

Potentials and Implementation: A Generic Impossibility Theorem

Philippe Jehiel, Moritz Meyer-ter-Vehn and Benny Moldovanu*

July 16, 2002

Abstract

We introduce the concept of potentials (analogous to Monderer and Shapley (1996)) for mechanism design problems and relate potentials to implementation in ex-post equilibria. We then proceed to show that in generic settings with interdependent valuations only constant choice rules are ex-post implementable.

1 Introduction

The idea of aligning the interests of several heterogenous strategic agents that jointly control a decision is central to mechanism design and implementation. By attaching different monetary transfers to different social alternatives, the designer can affect the agents' preferences over these alternatives so that, ultimately, all agents agree about the preferred alternative (and hence so that all agents find it in their own strategic interest to behave in a way that leads to the commonly preferred outcome). Probably the most famous example of successful alignment is offered by the Vickrey-Clarke-Groves mechanisms (see Vickrey (1961), Clarke (1971) Groves (1973)) for private values environments with quasi-linear utility. There an agent receives a transfer equal to the sum of valuations of the other agents in the chosen social alternative. With such transfers all individual payoff maximization decision problems coincide with the maximization of social surplus, yielding the well known dominant strategy implementability of the efficient choice rule. More generally, whenever an

*We wish to thank Martin Hellwig for several helpful remarks. Jehiel: UCL London and CERAS, Paris; jehiel@enpc.fr; Meyer-ter-Vehn and Moldovanu: University of Mannheim; meyerter@pool.uni-mannheim.de, mold@pool.uni-mannheim.de

alignment of interests is possible for some desirable outcome from the point of view of the designer, we say that the outcome can be "implemented".

A seemingly unrelated notion of interest alignment is implicit in the definition of potentials for normal form games, due to Monderer and Shapley (1996). Roughly speaking, a game admits a potential if there exists a function (common to all players) from strategy profiles to the set of real numbers such that, for any player, changes in utility resulting from changes in own strategy (while keeping fixed others' strategies) are reflected in appropriate changes in the value of the common potential function. A main result is that a strategy profile is a Nash equilibrium of the original game if and only if it is a Nash equilibrium in the artificial game where each player's utility function is replaced by the common potential. Thus, the equilibria of strategic interaction in a potential game are mirrored in a much simpler game where all players' interests are identical.

In the first part of the present paper we introduce several notions of potentials for mechanism design problems with interdependent values, discuss their properties, and establish relations between these notions and implementability in ex-post equilibria. A mechanism design problem with a given set of valuation functions admits a potential if there exist monetary transfers such that the maximization problem of each agent coincides with the problem of maximizing a single function, common to all agents¹.

Ex-post implementation implies that, even after learning the signals of others, agents are not willing to change their strategy. This notion has received a lot of attention recently², essentially because it ensures that agents need not know the distribution from which others' signals are drawn in order to play³

Whereas in Monderer and Shapley's general setting a potential need not have an intuitive "economic" interpretation related to the game's features (see for example their derivation of a potential for a Cournot oligopoly), a potential in our mechanism design problem is, obviously, closely related to

¹The potentials introduced here should not be confused with the **individual** potential functions arising as expected equilibrium utility functions in Bayes-Nash implementation (see for example Jehiel, Moldovanu and Stacchetti, 1999). The gradient of each of those functions refers to features of a specific agent, whereas here the gradient connects all agents- see below. The common name reflects certain properties about path integrals first analyzed in the physical sciences (e.g., energy conservation).

²See, among others, Dasgupta and Maskin (2000), Jehiel and Moldovanu (2001), Bergemann and Valimaki (2002), Perry and Reny (2002), Chung and Ely (2002), Meyer-ter-Vehn and Moldovanu (2002).

³See Bergemann and Morris (2002) for a formal treatment of this issue, which is also connected to Wilson's concern about detail-free mechanisms (but, see McLean-Postlewaite (2002) for a challenging view on this point).

the outcome function that is being implemented through the alignment of interests induced by that potential. Thus, it becomes possible to deduce results about implementation by studying the properties of potential functions (and vice-versa).

Following Monderer and Shapley, we shall distinguish between ordinal and cardinal potentials. The former weaker concept roughly says that the potential function and each agent's payoff function agree on the best alternative, whereas the latter stronger concept says that the quantitative differences between alternatives are equal for the potential function and each agent's payoff function. In general, ex-post implementability is related to the existence of a corresponding ordinal potential⁴. While cardinal potentials are easily characterized via properties of second differences of valuation functions, ordinal potentials are much more fickle. Nevertheless, we give several examples of environments where all implementable outcome functions must correspond to a cardinal potential: in a private values setting with multi-dimensional signals, Roberts (1979) has shown that all deterministic dominant strategy implementable social choice rules can be characterized as maximizers of affine functions of the agents' signals. These maximizers are readily seen to be (weighted) cardinal potentials. Meyer-ter-Vehn and Moldovanu (2002) extended Roberts' result to a setting with interdependent values where the impact of one agent's information on his own valuation can be separated from the impact of information which is private to others. In particular, they show that the efficient allocation cannot be generically implemented in this framework⁵.

In the second part of the paper we apply the gained insights in order to obtain a general impossibility result for implementation in ex-post equilibria with interdependent quasi-linear values and multi-dimensional signals: generically⁶, only trivial outcome functions (which choose one alternative irrespective of the available information) can be implemented. This finding, which can be seen as a kind of counterpart to Gibbard and Satterthwaite's Theorem (Gibbard (1973) and Satterthwaite (1975)) about dominant strategy implementation, is the main result in the paper. We first illustrate the idea of the proof in a less general setting where valuations are linear in sig-

⁴It is not necessarily the case that a mechanism implementing a given social choice function gives rise to a potential game (even if we take the agent-normal form of the game with incomplete information)

⁵This holds even for Bayes-Nash implementation, as shown by Jehiel and Moldovanu (2001). With one-dimensional signals there are positive results. A good survey of the recent literature on efficient auctions with interdependent valuations is Maskin (2001).

⁶This means that the valuation functions for different agents and/or for different alternatives are not related by algebraic conditions.

nals (this allows for terms where two signals belonging to different agents are multiplied). In this framework cardinal potentials are readily shown to exist only if an algebraic condition (that cannot hold generically) is satisfied by the constants attached to the multiplicative terms. We then proceed to show how the same algebraic condition is also necessary for the existence of an ordinal potential, and hence for implementation. Finally, we extend the proof to general valuation functions.

Our result implies that, for any direct or indirect mechanism, if an ex-post equilibrium exists, it necessarily involves complete pooling for all agents. The other possibility is that an ex-post equilibrium does not exist at all - since this concept plays in frameworks with interdependent valuations a role akin to that of dominant strategy implementation with private values, it is also quite demanding. Thus, in general frameworks, one must come back to Bayes-Nash implementation in order to achieve positive results.

The rest of the paper is organized as follows: In Section 2 we briefly review Monderer and Shapley's definitions of potential for normal form games. In Section 3 we describe the mechanism design problem with interdependent values, and we define several notions of potential for this setting. The focus of this section is on establishing links between potentials and implementation in ex-post equilibria. In Section 4 we prove the generic impossibility result about implementation in ex-post equilibria. Several proofs appear in an Appendix.

2 Games and Potentials

We briefly review the original definitions of potentials for normal form games of complete information, due to Monderer and Shapley (1996). Let $\Gamma = \Gamma(u^1, u^2, \dots, u^n)$ be a game in strategic form played by the agents in a finite set N . The strategy set of player i is denoted by Y^i , and the payoff function of i is $u^i : Y \rightarrow \mathbb{R}$, where $Y = Y^1 \times Y^2 \times \dots \times Y^n$.

Definition 1 1. A function $P : Y \rightarrow \mathbb{R}$ is an **exact** (or **cardinal**) **potential** for Γ if for every $i \in N$ and for every $y^{-i} \in Y^{-i}$

$$u^i(y^{-i}, x) - u^i(y^{-i}, z) = P(y^{-i}, x) - P(y^{-i}, z)$$

for every $x, z \in Y^i$.

2. Given agent specific weights $\alpha = (\alpha^i)_{i \in N} \in \mathbb{R}_+^n$, a function $P : Y \rightarrow \mathbb{R}$ is an **α -potential** for Γ if for every $i \in N$ and for every $y^{-i} \in Y^{-i}$

$$\alpha^i(u^i(y^{-i}, x) - u^i(y^{-i}, z)) = (P(y^{-i}, x) - P(y^{-i}, z))$$

for every $x, z \in Y^i$. When we are not interested in the particular weights, we shall speak about **weighted (cardinal) potentials**.

3. A function $P : Y \rightarrow \mathbb{R}$ is an **ordinal potential** for Γ if for every $i \in N$ and for every $y^{-i} \in Y^{-i}$

$$u^i(y^{-i}, x) - u^i(y^{-i}, z) > 0 \Leftrightarrow P(y^{-i}, x) - P(y^{-i}, z) > 0$$

for every $x, z \in Y^i$.

A cardinal potential for Γ is a fortiori a weighted cardinal potential, and a weighted cardinal potential a fortiori an ordinal potential. The following result motivates the above definitions and shows how potentials align the agents' interests:

Proposition 1 (*Monderer and Shapley, 1996*): *Let P be an ordinal potential for $\Gamma(u^1, u^2, \dots, u^n)$. Then a strategy profile is a Nash equilibrium of $\Gamma(u^1, u^2, \dots, u^n)$ if and only if it is a Nash equilibrium of $\Gamma(P, P, \dots, P)$.*

Proof. From the definition of ordinal potentials it is immediate that $\Gamma(u^1, u^2, \dots, u^n)$ and $\Gamma(P, P, \dots, P)$ have the same best-response sets, and hence the same Nash equilibria. ■

3 Potentials for Mechanism Design Problems

We consider a model where a decision $k \in K$ must be taken by a set of agents N . Agent's i utility $u^i = v_k^i + t^i$ depends on i 's valuation of the chosen alternative k , v_k^i , and on a monetary transfer t^i . Her valuation $v_k^i = v_k^i(s)$ in turn depends on signals $s = (s^j)_{j \in N}$ drawn from some signal space $S := \prod_{j \in N} S^j$, where $s^j \in S^j$ is private information to agent j . While this structure is already sufficient in order to define potentials, we shall assume in some examples and in Section 4 that S^i is a subset of R^K , and that each signal $s^i = (s_k^i)_{k \in K}$ consists of K one-dimensional components s_k^i such that an agent's valuation of alternative k only depends on the k 'th components of every agent's signal $v_k^i(s) = v_k^i(s_k)$ where $s_k := (s_k^j)_{j \in N}$.

We are interested in deterministic choice rules $\psi : S \rightarrow K$ with the property that there are transfers functions $t^i : S \rightarrow \mathbb{R}$ such that truth-telling is an **ex-post equilibrium** in the incomplete information game induced by the direct revelation mechanism $(\psi, (t^i)_{i \in N})$, i.e.

$$v_{\psi(s)}^i(s) + t^i(s) \geq v_{\psi(\tilde{s}^i, s^{-i})}^i(s) + t^i(\tilde{s}^i, s^{-i}) \quad (1)$$

for all $s^i, \tilde{s}^i \in S^i$ and $s^{-i} \in S^{-i}$, where $s := (s^i, s^{-i})$. We shall call such ψ **implementable**.

It is easy to see that in an incentive compatible mechanism that implements ψ , the payments $t^i(s^i, s^{-i})$ may depend on agent i 's reported signal only via the chosen alternative. This allows us to write $t^i(s) = t_{\psi(s)}^i(s^{-i})$.⁷

Inspired by the definitions in the last Section, we have:

Definition 2 1. A family $(P_k)_{k \in K}$ of functions $P_k : S \rightarrow \mathbb{R}$ is a **cardinal potential** for v if there are payments $t_k^i(s^{-i})$ such that:

$$v_k^i(s) + t_k^i(s^{-i}) = P_k(s)$$

for all agents i , alternatives k and signals $s = (s^i, s^{-i})$.

2. Given agent specific weights $\alpha = (\alpha^i)_{i \in N} \in \mathbb{R}_+^N$, a family $(P_k)_{k \in K}$ of functions $P_k : S \rightarrow \mathbb{R}$ is an **α -potential** for v if there are payments $t_k^i(s^{-i})$ such that:

$$\alpha^i (v_k^i(s) + t_k^i(s^{-i})) = P_k(s)$$

for all agents i , alternatives k and signals s . We shall call any α -potential a **weighted cardinal potential** when we are not interested in the value of α .

3. A family $(P_k)_{k \in K}$ of functions $P_k : S \rightarrow \mathbb{R}$ is an **ordinal potential** for v if there are payments $t_k^i(s^{-i})$ such that:

$$\arg \max_k \{P_k(s)\} \subseteq \arg \max_k \{v_k^i(s) + t_k^i(s^{-i})\}$$

for all agents i and signals s .⁸

A cardinal potential for valuations v is a fortiori a weighted cardinal potential for v , and a weighted cardinal potential for v is a fortiori an ordinal potential. Also a cardinal potential for v is unique up to an additive constant in each P_k , and the same is true for an α -potential given α .

We now establish a close connection between potentials and implementable choice rules: For this, let \mathcal{S} be the set of all social choice rules $\psi : S \rightarrow K$, and let \mathcal{P} be the set of all potentials $P = (P_k)_{k \in K} : S \rightarrow \mathbb{R}^K$. Define two functions $\Phi : \mathcal{S} \rightarrow \mathcal{P}$ and $\Xi : \mathcal{P} \rightarrow \mathcal{S}$ as follows:

⁷This is related to the so called "taxation principle". If transfers $t' < t$ were possible in alternative k , agent i would never pick an announcement inducing the lower transfer.

⁸It would be unreasonable to require "=" instead of " \subseteq " in this definition, as this would require equality of the argmax sets between agents. As ordinal potentials are defined to translate into implementable choice rules (see Proposition 2), we only want to require, that there be one alternative which is in the argmax set of every agent, instead of these sets being equal.

1. The potential $\Phi(\psi)$ is given by $(\Phi(\psi))_k(s) := \begin{cases} 1 & \text{if } \psi(s) = k \\ 0 & \text{else} \end{cases}$
2. The social choice rule $\Xi(P)$ is given by $(\Xi(P))(s) := \arg \max_{k \in K} P_k(s)$

It is immediate that Ξ is a left inverse of Φ , i.e.: $\Xi(\Phi(\psi)) \equiv \psi$. We cannot expect the converse to be true, as the choice rule $\Xi(P)$ only contains the information about which of K numbers is the greatest for any given s . We shall call two potentials P, P' **equivalent**, if they represent the same choice rule in the sense that $\Xi(P) \equiv \Xi(P')$. The next Proposition is the counterpart to Proposition 1 for design problems:

Proposition 2 *If $P = (P_k)_{k \in K}$ is an ordinal potential for valuations v then $\Xi(P)$ is ex-post implementable. Conversely if a SCR ψ is ex-post implementable then $\Phi(\psi)$ is an ordinal potential for v .*

Proof. If P is an ordinal potential for v fix functions $t_k^i(s^{-i})$ for all i, k as in the definition of an ordinal potential. These transfers together with the choice rule $\psi := \Xi(P)$ induce truth-telling as an ex-post equilibrium in the associated information revelation game: By the definition of an ordinal potential $\psi(s) \in \arg \max_k \{P_k(s)\} \subseteq \arg \max_k \{v_k^i(s) + t_k^i(s^{-i})\}$, so no agent i has an incentive to misrepresent s^i .

Conversely, let ψ be implemented by transfers $t_k^i(s^{-i})$. To check that $P := \Phi(\psi)$ is an ordinal potential for v , we must show that $\arg \max_k \{(\Phi(\psi))_k(s)\} = \psi(s)$ is in $\arg \max_k \{v_k^i(s) + t_k^i(s^{-i})\}$ for every agent i . As (ψ, t) is incentive compatible this is satisfied after changing $t_k^i(s^{-i})$ to $-\infty$ (or a sufficiently negative number) for all s^{-i} such that there is no signal s^i of agent i that leads to the choice of $\psi(s^i, s^{-i}) = k$. ■

Remark 1 *It is worth stressing here a main difference between the notion of potential games and the notion of potentials for mechanism design. In particular, consider the direct revelation game associated with the SCR ψ which is assumed to be ex-post implementable. Every agent i (with signal s^i) is asked to report a signal \tilde{s}^i , the alternative $\psi(\tilde{s}^i, \tilde{s}^{-i})$ is chosen and agent i receives a transfer $t_{\psi(\tilde{s}^i, \tilde{s}^{-i})}(\tilde{s}^{-i})$. Proposition 2 shows that $P := \Phi(\psi)$ is a potential for v . However, the revelation game is not a potential game in the sense defined in Section 2. An obvious reason is that this is an incomplete information game. But, even taking the complete information projection of this game where the type of agent i is known to be s^i , it is readily verified that this game is not a potential game*

⁹in case $\arg \max_{k \in K} P_k(s)$ is not a singleton we choose the smallest k in this set, with respect to some order on K that is fixed throughout this paper.

Proposition 2 allows us to translate facts about implementable choice rules in the language of ordinal potentials. From a practical point of view, finding (weighted) cardinal potentials for given valuations v is more straightforward than finding ordinal potentials, since the former are more tightly connected with the valuations (see also Proposition 3 at the end of this section for straightforward characterizations of cardinal potentials). In contrast, as noticed already by Monderer and Shapley in their setting, ordinal potentials are not easily characterized. Since Proposition 2 connects implementability and ordinal potentials, it seems that there is not much gain by using the potential approach. Nevertheless, the next example (and the next Section) suggest that, in the mechanism design setting, ordinal and cardinal potentials are quite tightly linked.

Example 1 Consider a setting with private valuations where we can set $v_k^i(s) := s_k^i$. In order to establish whether the efficient SCR is implementable, we need to find an ordinal potential that corresponds to it. But note that the transfer functions $t_k^i(s^{-i}) = \sum_{j \neq i} s_k^j$ immediately yield that the family $(P_k)_{k \in K}$, where $P_k(s_k) = \sum_j s_k^j$, is a cardinal potential for v . More generally, under several technical conditions, it has been shown by Roberts (1979) that any dominant strategy implementable social choice rule must be of the form $\psi(s) \in \arg \max_{k \in K} (\sum_i \alpha^i s_k^i + \lambda_k)$, where $(\alpha^i)_{i \in N}$ are agent specific weights, and where $(\lambda_k)_{k \in K}$ are alternative specific weights. By setting transfers $t_k^i(s^{-i}) = \frac{1}{\alpha^i} (\sum_{j \neq i} s_k^j + \lambda_k)$ we obtain that the family of functions given by $P_k(s) = \sum_i \alpha^i s_k^i + \lambda_k$ is a weighted cardinal potential for valuations v . Thus Roberts' result can be read as saying that, for every ordinal potential for v , there is an equivalent weighted cardinal potential.

Meyer-ter-Vehn and Moldovanu (2002) generalized the above result by showing that, for additive, semi-separable, interdependent valuations where $v_k^i(s_k) := s_k^i + h_k^i(s_k^{-i})$, any ex-post implementable social choice rule must be of the form $\psi(s) \in \arg \max_{k \in K} (\sum_i \alpha^i s_k^i + \lambda_k)$. Again, we obtain that. Note also that the efficient SCR $\psi(s) \in \arg \max_{k \in K} (\sum_i v_k^i(s_k))$ cannot be (generically) implemented in this framework.

The next example shows that the above observation (i.e., for every ordinal potential there is an equivalent weighted cardinal potential) does not always hold.

Example 2 In their Proposition 5.1, Jehiel and Moldovanu (2001) study a model with one-dimensional signals s^i where agent i 's valuation of alternative k is given by $v_k^i(s) = \sum_{j \in N} a_{ki}^j s^j$. Their proof (which only deals with the possibility of implementing the efficient choice rule) can be extended to show that

any functions $P_k(s) = \sum_{j \in N} A_k^j s^j$ where $a_{kj}^j > a_{k'j}^j \Leftrightarrow A_k^j > A_{k'}^j$ for all j, k, k' represent ex-post implementable choice rules, and thus are ordinal potentials for v^{10} . Generically, no weighted cardinal potential $P'_k(s) = \sum_{j \in N} \alpha^j a_{kj}^j s^j$ can be equivalent to such an ordinal potential, because α^j would have to simultaneously satisfy the K equations $\alpha^j a_{kj}^j = A_k^j$.

The different insights offered by the two examples above are due to the differences in the dimension of each agent's signals: The slackness inherent in an ordinal potential can only be tightened (to obtain a cardinal potential) by the incentive constraints at signals which are on the border of at least three areas in the signal space (such that a different alternative is chosen in each area). But, for such signals to exist, the signal space must be at least two-dimensional (see also the characterization results in Meyer-ter-Vehn and Moldovanu (2002)).

We now display necessary and sufficient¹¹ conditions on the valuations v for (weighted) cardinal potentials to exist. Denote by $\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^i(s^{-i,j})$ the second difference of v_k^i :

$$\begin{aligned} \partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^i(s^{-i,j}) & : = (v_k^i(s^i, s^j, s^{-i,j}) - v_k^i(s^i, \widehat{s}^j, s^{-i,j})) \\ & - (v_k^i(\widehat{s}^i, s^j, s^{-i,j}) - v_k^i(\widehat{s}^i, \widehat{s}^j, s^{-i,j})) \end{aligned} \quad (2)$$

Proposition 3 1. *There exists a cardinal potential for v if and only if*

$$\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^i(s^{-i,j}) = \partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^j(s^{-i,j})$$

for all $i, j \in N, k \in K, s^{i,j}, \widehat{s}^{i,j} \in S^{i,j}$ and $s^{-ij} \in S^{-ij}$.

2. *There exists an α -potential for v if and only if:*

$$\alpha^i \partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^i(s^{-i,j}) = \alpha^j \partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^j(s^{-i,j})$$

for all $i, j \in N, k \in K, s^{i,j}, \widehat{s}^{i,j} \in S^{i,j}$ and $s^{-ij} \in S^{-ij}$. Thus, there is a weighted cardinal potential for v only if:

$$\begin{aligned} & (\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^j(s^{-i,j})) (\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_l^i(s^{-i,j})) \\ & = (\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_k^i(s^{-i,j})) (\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} v_l^j(s^{-i,j})) \end{aligned}$$

for all $i, j \in N, k, l \in K, s^{i,j}, \widehat{s}^{i,j} \in S^{i,j}$ and $s^{-ij} \in S^{-ij}$.

¹⁰A characterization of all implementable choice rules in this setting is available from the authors upon request.

¹¹Note that Monderer and Shapley's analogous sufficient conditions on cross derivatives use one dimensional strategy spaces and differentiability in order to apply Green's Theorem. Our discrete version is more general and applies to their setting as well.

Proof. The necessity of these conditions is immediate from the definitions. The sufficiency part is more delicate, and it is relegated to the appendix. For an intuition consider the case of two agents 1, 2. Given that the second difference of $v_k^1 - v_k^2$ is equal to zero, it must have a representation which is additively separable in its arguments: $v_k^1(s) - v_k^2(s) = t_k^2(s^1) - t_k^1(s^2)$. This yields the cardinal potential $P_k(s) = v_k^i(s) + t_k^i(s^{-i})$. ■

Note that the sufficiency part implies that there is an α -potential for any semi-separable valuation functions $v_k^i(s) = g_k^i(s^i) + h_k^i(s^{-i})$ ¹². We shall see in Proposition 4 that, as soon as we move away from the semi-separable case, non-trivial ordinal potentials cannot exist generically. Again, this indicates that ordinal and cardinal potentials are not so far apart as one may think.

We conclude this Section with a straightforward result displaying some operations that can be performed on valuation functions without affecting the existence of a potential. The Lemma will be used in the next Section.

- Lemma 1** 1. *If there is an ordinal [cardinal], [weighted cardinal] potential for $v = (v_k^i(s))_{i,k}$ then there is an ordinal [cardinal], [weighted cardinal] potential for $(v_k^i(s) + h_k^i(s^{-i}))_{i,k}$ for any functions h_k^i .*
2. *If there is a cardinal potential for $v = (v_k^i(s))_{i,k}$ then there is a cardinal potential for $(v_k^i(s) + w_k(s))_{i,k}$ for any functions w_k .*
3. *If there is a weighted cardinal potential for $v = (v_k^i(s))_{i,k}$ then there is a weighted cardinal potential for $(\lambda^i v_k^i(s))_{i,k}$ for any weights λ^i .*

Proof. Immediate by the definitions of potentials. ■

4 Impossibility Results

In order to obtain relatively straightforward and intuitive arguments for the results derived in this section, we impose that $S^i = [0, 1]^K$ and that the valuation function $v_k^i(s) = v_k^i(s_k)$ only depends on the k -th component of every agent's signal. We also restrict ourselves to two agents i, j and two alternatives k, l . As our main result is an impossibility one (see Proposition 5 below), the additional restrictions do not matter: more general settings only aggravate the problem of implementation and thus reinforce our result.

¹²The implementability of the corresponding SCRs $\psi(s) = \arg \max_{k \in K} \{ \sum_i \alpha^i g_k^i(s^i) + \lambda_k \}$ follows by a simple generalization of results in Meyer-ter-Vehn, Moldovanu (2002).

The functions $v_k^i : [0, 1]^N \rightarrow \mathbb{R}$ are assumed to be sufficiently smooth and to satisfy $\frac{\partial v_k^i}{\partial s_k^i} > 0$. Finally, we consider transfer functions $t_k^i(s^{-i})$ that are differentiable in s^{-i} .

Since the agents' preferences between alternatives k and l are only affected by differences in payments $t_k^i(s^{-i}) - t_l^i(s^{-i})$ we loose no generality by setting $t_l^i \equiv 0$ and denoting $t^i \equiv t_k^i$.

We denote the indifference set of a choice rule ψ by $I := \overline{\psi^{-1}(k)} \cap \overline{\psi^{-1}(l)}$, and call ψ non-trivial if $I \cap S^\circ \neq \emptyset$. For an implementable rule ψ , we obtain by the continuity of the functions t_k^i that

$$s \in I \Leftrightarrow v_k^1(s_k^1, s_k^2) + t^1(s^2) = v_l^1(s_l^1, s_l^2) \Leftrightarrow v_k^2(s_k^2, s_k^1) + t^2(s^1) = v_l^2(s_l^2, s_l^1)$$

This equation is the key for the impossibility results Propositions 4 and 5. We shall deduce from it implications on the transfer functions at $s \in I$. This, in turn, will yield non generic conditions on the valuation functions that must be satisfied for implementability of a non-trivial choice rule.

Before we elaborate on this, it is instructive to note why the conditions for a **cardinal** potential cannot be satisfied generically. Of course, this is not a sufficient argument, as ex-post implementation is equivalent to the existence of an **ordinal** potential (Proposition 2). But, the argument can be viewed as a first step toward the impossibility result.

Proposition 3 tells us that for a cardinal potential for v to exist it must be that for all $s^1, \hat{s}^1, s^2, \hat{s}^2$

$$\partial_{(s^1, \hat{s}^1)} \partial_{(s^2, \hat{s}^2)} v_k^1 = \partial_{(s^1, \hat{s}^1)} \partial_{(s^2, \hat{s}^2)} v_k^2$$

Assuming that the functions v_k^i , $i = 1, 2$ are twice continuously differentiable, this is equivalent to

$$\frac{\partial^2 v_k^1}{\partial s_k^1 \partial s_k^2}(s_k^1, s_k^2) \equiv \frac{\partial^2 v_k^2}{\partial s_k^1 \partial s_k^2}(s_k^1, s_k^2)$$

Clearly, such a condition cannot be satisfied generically¹³

The rest of the Section will in fact establish that the impossibility applies to all non-trivial choice rules, i.e. to ordinal potentials as well.

¹³In the case where the valuations are additively separable, then the two derivatives are zero and a cardinal potential exists. This is the situation studied by Meyer-ter-Vehn and Moldovanu (2002).

4.1 Linear Valuations

Much intuition about the reasons behind the general impossibility result can be gained by first restricting attention to valuations that are linear in every agent's signal. Assume then the following valuation structure:

$$\begin{aligned}
v_k^1(s_k) &: = a_k^1 s_k^1 + b_k^1 s_k^1 s_k^2 = s_k^1 (a_k^1 + b_k^1 s_k^2) \\
v_l^1(s_l) &: = a_l^1 s_l^1 + b_l^1 s_l^1 s_l^2 = s_l^1 (a_l^1 + b_l^1 s_l^2) \\
v_k^2(s_k) &: = a_k^2 s_k^2 + b_k^2 s_k^1 s_k^2 = s_k^2 (a_k^2 + b_k^2 s_k^1) \\
v_l^2(s_l) &: = a_l^2 s_l^2 + b_l^2 s_l^1 s_l^2 = s_l^2 (a_l^2 + b_l^2 s_l^1)
\end{aligned} \tag{3}$$

Note that leaving out terms such as $c_k^i + d_k^i s_k^j$ in $v_k^1(s_k)$ is without loss of generality by Lemma 1.

A key observation is offered by Proposition 3: it tells us that a cardinal potential for v exists if and only if $b_k^1 = b_k^2$, $b_l^1 = b_l^2$, and that a weighted cardinal potential exists if and only if $b_k^1 b_l^2 = b_l^1 b_k^2$. Thus, cardinal potentials exist only for non-generic sets of valuations, i.e., for valuations that are related by some algebraic equations. In Section 3 we have already seen that ordinal potentials (which correspond to implementable social choice rules) and (weighted) cardinal potential are often tightly related. The next Proposition offers another instance of this phenomenon:

Proposition 4 *Given the linear valuations specified in equations 3, a non-trivial deterministic SCR ψ is implementable if and only if*

$$b_l^1 b_k^2 - b_k^1 b_l^2 = 0. \tag{4}$$

4.2 General valuations

We now consider the case of general valuation functions. We use the same idea as above, namely we show that implementability of any non-trivial SCR necessarily implies some algebraic relations among valuations. We call a situation *non-generic* if there exists a non-null rational function H such that $\forall s, H(\{v_k^i(s), \frac{\partial v_k^i(s)}{\partial s_k^i}, \frac{\partial^2 v_k^i(s)}{\partial s_k^i \partial s_k^j}\}_{i,j,k}) = 0$ ¹⁴.

Proposition 5 *Assume that a non-trivial deterministic choice rule ψ is implementable. Then, the function $\frac{\partial v_l^2}{\partial s_l^2}$ is determined by the functions v_k^1, v_l^1, v_k^2 up to five parameters. In other words, for generic valuation functions, only trivial choice rules are implementable.*

¹⁴Note that the private values case and the case with interdependent but semi-separable valuations are non-generic in this sense.

It is instructive to compare the above impossibility result with the celebrated Gibbard-Satterthwaite Theorem (see Gibbard (1973), and Satterthwaite (1975)). These authors focused on dominant strategy implementation in a framework with private values where utility functions need not be quasi-linear (i.e, agent i 's utility function depends in an arbitrary way on i 's type and on a social alternative whose description may or may not include monetary transfers). Their main assumptions on the underlying model are: 1) There are at least three alternatives; 2) By varying the "type" of agent i , the obtained set of ordinal preference relations over the alternatives covers the entire set of rational ordinal preferences for i . Under these assumptions, they showed that any exhaustive¹⁵ and dominant strategy implementable SCR must be dictatorial. It is well-known that restricting attention to quasi-linear utility functions in the private values case yields positive results, such as the Vickrey-Clarke-Groves mechanisms (see Example 1 in Section 3). Moreover, with two alternatives, the simple majority rule is non-dictatorial and dominant strategy implementable (without restrictions on the utility functions).

In contrast, we only looked at quasi-linear utility functions (where positive results exist for private values), but we allowed for a general form of interdependence in valuations. Accordingly, we shifted attention to implementation in ex-post equilibria. Interdependence seems weaker than requirement 2 above, but plays a similar role: by varying the agents' signals we obtain a sufficiently rich environment. Even for settings with two alternatives, we showed that, generically, only constant social choice rule are implementable¹⁶. Recall also that in the less rich framework where valuations are quasi-linear, interdependent and semi-separable, Meyer-ter-Vehn and Moldovanu (2002) have obtained a positive result (see Example 1 in Section 3).

5 Summary

We have introduced several notions of potentials for mechanism design problems. Whenever a potential exists, individual interests can be aligned by suitable monetary transfers. This feature allowed us to connect the existence of potentials and the possibility to implement various social choice functions. We have used this connection in order to prove a generic impossibility result for ex-post implementation in frameworks with multi-dimensional signals and interdependent valuations. The insights and relations introduced in this paper should be useful in other mechanism design contexts.

¹⁵i.e, any alternative is chosen for some realization of types.

¹⁶Note that dictatorship is not well-defined with interdependent valuations.

Since our result implies that, for any generic social choice situation, direct or indirect mechanisms either do not have ex-post equilibria, or have only ex-post equilibria involving complete pooling, we identify an important role for Bayes-Nash implementation.

6 Appendix

For the proof of Proposition 3 we need the following:

Lemma 2 *Let $f : X \times Y \times Z \rightarrow \mathbb{R}$ (where X, Y, Z are arbitrary sets) be any function that satisfies: $\forall x, x', y, y', z$,*

$$\partial_{(x,x')} \partial_{(y,y')} f(\cdot, \cdot, z) := (f(x, y, z) - f(x, y', z)) - (f(x', y, z) - f(x', y', z)) = 0$$

Then f must be additive in x and y : There are functions $f_x(x, z)$, $f_y(y, z)$ such that $f(x, y, z) = f_x(x, z) + f_y(y, z)$.

Proof. Fix x^*, y^* and define $f_x(x, z) := f(x, y^*, z)$ and $f_y(y, z) := f(x^*, y, z) - f(x^*, y^*, z)$. Solving $\partial_{(x,x')} \partial_{(y,y')} f(\cdot, \cdot, z) = 0$ for $f(x, y, z)$ yields the result. ■

Proof of Proposition 3. As the alternative k does not matter for this proof we simplify notation by leaving away the subscripts. We exhibit the proof for exact potentials. The weighted case is analogous.

Lemma 2 gives the proof for the case of two agents i, j : $\partial_{(s^i, \widehat{s}^i)} \partial_{(s^j, \widehat{s}^j)} (v^i - v^j) \equiv 0$ yields $v^1(s) - v^2(s) = t^2(s^1) - t^1(s^2)$ and we can define the cardinal potential $P(s) = v^i(s) + t^i(s^{-i})$. We now use this as the initial step in an induction argument.

Assume that we have already defined a potential $P(s)$ and payments $t^j(s^{-j})$ such that $v^j(s) + t^j(s^{-j}) = P(s)$ for all agents j in a subset $J \subset N$. We now construct a potential $P'(s)$ and payments $t^{j'}(s^{-j})$ such that $v^j(s) + t^{j'}(s^{-j}) = P'(s)$ for all agents j in $J \cup \{i\}$ where $i \in N \setminus J$.

By the result in the two-agent case we can construct a potential $P^j(s)$ and payments $t^{j,i}(s^{-j}), t^{i,j}(s^{-i})$ for each agent $j \in J$ such that $v^j(s) + t^{j,i}(s^{-j}) = P^j(s) = v^i(s) + t^{i,j}(s^{-i})$. Using the potentials $P^j(s)$ we now change $P(s)$ (which is a potential for J) into $P'(s)$, a potential for $J \cup \{i\}$. Fix an agent $h \in J$. By the properties of the potentials P, P^j we have:

$$\begin{aligned} t^{h,i}(s^{-h}) - t^h(s^{-h}) &= P^h(s) - P(s) = P^h(s) - P^j(s) + P^j(s) - P(s) \\ &= t^{i,h}(s^{-i}) - t^{i,j}(s^{-i}) + t^{j,i}(s^{-j}) - t^j(s^{-j}), \forall j \in J \end{aligned}$$

Since $t^{i,h} - t^{i,j}$ does not depend on s^i and since $t^{j,i} - t^j$ does not depend on s^j , we get that $\partial_{(s^i, \tilde{s}^i)} \partial_{(s^j, \tilde{s}^j)} (t^{h,i} - t^h) \equiv 0$ for all $j \in J$. Lemma 2 yields now

$$t^{h,i}(s^{-h}) - t^h(s^{-h}) = t(s^i) - t^i(s^{-i,h}) \quad (5)$$

for some functions t and t^i .

For $j \in J$ set

$$P'(s) := P(s) + t(s^i) = P^h(s) + t^i(s^{-i,h}), t^{j'}(s^{-j}) := t^j(s^{-j}) + t(s^i) \quad (6)$$

Together with

$$t^{i'}(s^{-i}) := t^{i,h}(s^{-i}) + t^i(s^{-i,h}) \quad (7)$$

we get that $v^j(s) + t^{j'}(s^{-j}) = P'(s)$ for all agents j in $J \cup \{i\}$. ■

Proof of Proposition 4. Let $\psi : S \rightarrow \{k, l\}$ be a non-trivial choice rule and fix a signal $s = (s^1, s^2) \in I \cap S^\circ$. We shall vary s^2 in a way that leaves the signal in the indifference set, and study the implications for $t^1(s^2)$. The signals \tilde{s}^2 such that $s = (s^1, \tilde{s}^2) \in I$ are characterized by agent 2's indifference condition:

$$v_k^2(\tilde{s}_k) + t^2(s^1) = v_l^2(\tilde{s}_l) \quad (8)$$

The above equation is satisfied if and only if

$$\tilde{s}^2 = s^2(\varepsilon) := (s_k^2(\varepsilon), s_l^2(\varepsilon)) := (s_k^2 + \varepsilon(a_l^2 + b_l^2 s_l^1), s_l^2 + \varepsilon(a_k^2 + b_k^2 s_k^1)) \quad (9)$$

for some $\varepsilon \in \mathbb{R}$. Agent 1 must also be indifferent for $(s^1, s^2(\varepsilon)) \in I$, yielding:

$$t^1(s^2(\varepsilon)) = v_l^1(s_l^1, s_l^2(\varepsilon)) - v_k^1(s_k^1, s_k^2(\varepsilon)) \quad (10)$$

Thus:

$$\begin{aligned} t^1(s^2(\varepsilon)) - t^1(s^2) &= \varepsilon (s_l^1 b_l^1 (a_k^2 + b_k^2 s_k^1) - s_k^1 b_k^1 (a_l^2 + b_l^2 s_l^1)) \\ &= \varepsilon (a_k^2 b_l^1 s_l^1 - a_l^2 b_k^1 s_k^1 + (b_l^1 b_k^2 - b_k^1 b_l^2) s_k^1 s_l^1) \end{aligned} \quad (11)$$

The above equation defines $t^1 : S^2 \rightarrow \mathbb{R}$ on a ray through s^2 . This ray can now be rotated by changing s^1 . Let

$$s^1(\delta) := (s_k^1(\delta), s_l^1(\delta)) := (s_k^1 + \delta(a_l^1 + b_l^1 s_l^2), s_l^1 + \delta(a_k^1 + b_k^1 s_k^2)) \quad (12)$$

As $(s^1(\delta), s^2) \in I$ we can define a ray

$$(s_k^2(\varepsilon, \delta), s_l^2(\varepsilon, \delta)) := (s_k^2 + \varepsilon(a_l^2 + b_l^2 s_l^1(\delta)), s_l^2 + \varepsilon(a_k^2 + b_k^2 s_k^1(\delta))) \quad (13)$$

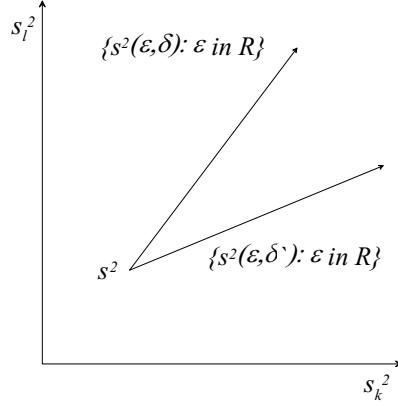


Figure 1:

such that $(s^1(\delta), s^2(\varepsilon, \delta)) \in I$. The set $\{s^2(\varepsilon, \delta) : \varepsilon, \delta \in \mathbb{R}_+\}$ defines a cone in S^2 anchored at s^2 . This cone is depicted in figure 1.

Equation 13 yields

$$t^1(s^2(\varepsilon, \delta)) - t^1(s^2) = \varepsilon (a_k^2 b_l^1 s_l^1(\delta) - a_l^2 b_k^1 s_k^1(\delta) + (b_l^1 b_k^2 - b_k^1 b_l^2) s_k^1(\delta) s_l^1(\delta)) \quad (14)$$

As a function of ε in the space S^2 , the slope of the ray $s^2(\varepsilon, \delta)$ is given by

$$\frac{a_k^2 + b_k^2 (s_k^1 + \delta (a_l^1 + b_l^1 s_l^2))}{a_l^2 + b_l^2 (s_l^1 + \delta (a_k^1 + b_k^1 s_k^2))} = \frac{a_k^2 + b_k^2 s_k^1 + \delta b_k^2 (a_l^1 + b_l^1 s_l^2)}{a_l^2 + b_l^2 s_l^1 + \delta b_l^2 (a_k^1 + b_k^1 s_k^2)} \quad (15)$$

We consider now two cases:

Case 1: The slope does not depend on δ . In this case the cone $\{s^2(\varepsilon, \delta) : \varepsilon, \delta \in \mathbb{R}_+\}$ degenerates to a single ray. Then $s^2(\varepsilon, \delta)$ is an affine function of δ , for fixed $\varepsilon \neq 0$. On the other hand, equation 14 shows that $t^1(s^2)$ is an affine function of s^2 when restricted to the ray $\{s^2(\varepsilon, \delta) : \varepsilon, \delta \in \mathbb{R}_+\}$. This makes $t^1(s^2(\varepsilon, \delta))$ an affine function of δ . Thus, we obtain that $b_l^1 b_k^2 - b_k^1 b_l^2$, the coefficient of the quadratic term in equation 14, must be equal to zero.

Case 2: The slope depends on δ . In this case the slope of each ray $\{s^2(\varepsilon, \delta) : \varepsilon \in \mathbb{R}_+\}$ is a continuous and strictly monotone function of δ . Equation 14 again forces t^1 to be an affine function on any ray in the cone $\{s^2(\varepsilon, \delta) : \varepsilon, \delta \in \mathbb{R}_+\}$. As the above argument can be repeated for any s^2 in the cone, t^1 must be an affine function on the whole cone. As before we see by equation 13 that $s^2(\varepsilon, \delta)$ is an affine function of δ and, putting the pieces together, that $t^1(s^2(\varepsilon, \delta))$ is an affine function of δ . Again we can conclude that $b_l^1 b_k^2 - b_k^1 b_l^2 = 0$. ■

Proof of Proposition 5. Take $(s^{1*}, s^{2*}) \in I$. For $s \in I$ one has (recall that the transfer for l is null):

$$v_k^1(s_k^1, s_k^2) + t^1(s^2) = v_l^1(s_l^1, s_l^2) \quad (16)$$

and similarly,

$$v_k^2(s_k^2, s_k^1) + t^2(s^1) = v_l^2(s_l^2, s_l^1) \quad (17)$$

Change now slightly s^{2*} into $s^{2*,\varepsilon}$ around s^* so that equation 17 still holds, i.e.

$$s^{2*,\varepsilon} = (s_k^{2*} + \varepsilon \frac{\partial v_l^2}{\partial s_l^2}, s_l^{2*} + \varepsilon \frac{\partial v_k^2}{\partial s_k^2}) \quad (18)$$

Equation 16 should also hold for $(s^{1*}, s^{2*,\varepsilon})$. Assuming that t^1 is locally differentiable, we obtain that

$$\begin{aligned} & -\frac{\partial v_k^1}{\partial s_k^2}(s_k^*) \frac{\partial v_l^2}{\partial s_l^2}(s_l^*) + \frac{\partial v_l^1}{\partial s_l^2}(s_l^*) \frac{\partial v_k^2}{\partial s_k^2}(s_k^*) \\ = & \frac{\partial t^1}{\partial s_k^2}(s^{2*}) \frac{\partial v_l^2}{\partial s_l^2}(s_l^*) + \frac{\partial t^1}{\partial s_l^2}(s^{2*}) \frac{\partial v_k^2}{\partial s_k^2}(s_k^*) \end{aligned} \quad (19)$$

The next step is to realize that equation 19 should hold for all $s = (s^1, s^{2*})$ such that

$$v_k^1(s_k^1, s_k^{2*}) - v_l^1(s_l^1, s_l^{2*}) = v_k^1(s_k^{1*}, s_k^{2*}) - v_l^1(s_l^{1*}, s_l^{2*}) \quad (20)$$

Now keep s^{2*} constant and re-write all functions assuming $s^2 = s^{2*}$. For some constants A, B, C one gets:

$$-\frac{\partial v_k^1}{\partial s_k^2}(s_k^1) \frac{\partial v_l^2}{\partial s_l^2}(s_l^1) + \frac{\partial v_l^1}{\partial s_l^2}(s_l^1) \frac{\partial v_k^2}{\partial s_k^2}(s_k^1) = A \frac{\partial v_l^2}{\partial s_l^2}(s_l^1) + B \frac{\partial v_k^2}{\partial s_k^2}(s_k^1) \quad (21)$$

whenever

$$v_k^1(s_k^1) - v_l^1(s_l^1) = C \quad (22)$$

Writing $s_k^1 = s_k^1(s_l^1)$ via equation 22 we can rewrite equation 21 to get:

$$\frac{\partial v_l^2}{\partial s_l^2}(s_l^1) = \frac{\frac{\partial v_k^2}{\partial s_k^2}(s_k^1) \left(\frac{\partial v_l^1}{\partial s_l^2}(s_l^1) - B \right)}{A + \frac{\partial v_k^1}{\partial s_k^2}(s_k^1)} \quad (23)$$

Given v^1, v_k^2 , (23) is a condition on $\frac{\partial v_l^2}{\partial s_l^2}$ with five free parameters $A, B, C, s_k^{2*}, s_l^{2*}$ where the last three are implicit. Thus for generic v_l^2 , we get the impossibility result. ■

References

- [1] Bergemann, D. and S. Morris (2002): "Robust Mechanism Design", discussion paper, Yale University.
- [2] Bergemann, D. and J. Valimäki (2002): "Information Acquisition and Efficient Mechanism Design", *Econometrica* **70**, 1007-10035.
- [3] Chung, K-S. and J. Ely (2001): "Ex-post Incentive Compatible Mechanism Design", discussion paper, Northwestern University.
- [4] Clarke E. (1971): "Multipart Pricing of Public Goods", *Public Choice* **8**, 19-33.
- [5] Dasgupta, P., and Maskin, E. (2000): "Efficient Auctions", *Quarterly Journal of Economics* **115**, 341-388.
- [6] Gibbard, A. (1973): "Manipulation of Voting Schemes", *Econometrica* **41**, 587-601.
- [7] Groves T. (1973): "Incentives in Teams", *Econometrica* **41**, 617-631.
- [8] Jehiel, P., Moldovanu, B. and Stacchetti, E. (1999) "Multidimensional Mechanism Design for Auctions with Externalities", *Journal of Economic Theory* **85**, 814-829.
- [9] Jehiel, P., and Moldovanu, B. (2001) "Efficient Design with Interdependent Valuations", *Econometrica* **69**, 1237-1259.
- [10] Maskin, E. (2001): "Auctions and Efficiency", discussion paper, Institute of Advanced Studies, Princeton.
- [11] McLean, R. and A. Postlewaite (2001): "Efficient Auction Mechanisms with Multidimensional Signals", discussion paper, University of Pennsylvania
- [12] Meyer-ter-Vehn, M., and Moldovanu, B. (2002): "Ex-Post Implementation with Interdependent Valuations", discussion paper, University of Mannheim.
- [13] Monderer D., and Shapley L. (1996): "Potential Games", *Games and Economic Behavior* **14**, 124-143.
- [14] Perry, M. and P. Reny (2002): "An Efficient Auction", *Econometrica* **70**, 1199-1213.

- [15] Roberts, K. (1979): "The Characterization of Implementable Choice Rules" in *Aggregation and Revelation of Preferences*, ed. Laffont J.J., North Holland.
- [16] Satterthwaite, M. (1975): "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting procedures and Social Welfare Functions", *Journal of Economic Theory* **10**, 187-217.
- [17] Vickrey, W. (1961): "Counterspeculation, Auctions and Competitive Sealed Tenders", *Journal of Finance* **16**, 8-37.