Efficient Full Implementation via Transfers: Uniqueness and Sensitivity in Symmetric Environments

By Mariann Ollár and Antonio Penta*

The problem of multiplicity is a key concern for the design of real-world mechanisms and institutions. Following the seminal work of Maskin (1999), the implementation literature has often addressed this concern via the design of complicated mechanisms and often relied on strong assumptions of common knowledge. Both of these features have been the object of famous critiques: on the one hand, Jackson (1992) called for a greater *relevance* of full implementation theory, initiating an agenda based on mechanisms with more economically appealing structure and properties; on the other, the so called *Wilson's doctrine* stressed the importance of weakening the reliance on common knowledge assumptions "[...] to conduct useful analyses of practical problems [...]" (Wilson (1987))

In this short paper we illustrate how novel insights gained from the robustness literature may be put to work to address both 'critiques' at once. We show this in a standard efficient implementation problem, with quasi-linear preferences and interdependent values, in environments that are *symmetric* in a twofold sense: (i) the total level of preference interdependence is constant across agents (*symmetric total preference interdependence*); (ii) types are drawn from distributions with identical means (*identical-means beliefs*).

To address the first of the critiques above, we pursue full implementation via the design of simple transfer schemes, that only elicit agents' payoff relevant information. To address the second, we only assume that agents commonly believe that others' types are drawn from distributions with an identical mean, but the actual distribution and its mean are unknown to both the agents and the designer.

Despite the weakness of these common knowledge assumptions and the demanding notion of implementation under restricted mechanisms, we identify surprisingly permissive conditions in such symmetric environments. Our main results characterize the conditions on agents' preferences under which full efficient implementation is possible in our sense, and identify a transfer scheme – the *equal-externality transfers* – that achieves full efficient implementation whenever possible. We further show that these transfers are also optimal in the sense that, among the set of all transfers that achieve full implementation, they uniquely minimize the sensitivity of the mechanism with respect to the presence of 'faulty' players (*Robustness to Mistakes in Play*).

I. Framework

A. *Symmetric Environments and Beliefs*

We consider public good environments with transferable utility. There is a finite set of agents $I = \{1, ..., n\}$, and we let $x \in \mathbb{R}_+$ denote the quantity of public good, which can be produced at cost $c(x) = \frac{1}{2}x^2$. The payoff type of agent i is $\theta_i \in \Theta_i := [0, 1]$, and we let $\theta_{-i} \in \Theta_{-i} = \times_{j \neq i} \Theta_j$ and $\theta \in \Theta = \times_{i \in I} \Theta_i$. Agent i's valuation for the public good is equal to $v_i(x, \theta) = (\theta_i + \sum_{j \neq i} \gamma_{ij} \theta_j)x$. Letting t_i denote the monetary transfer to agent i, the overall utility of agent i is:

$$u_i(x, \theta, t_i) = (\theta_i + \sum_{j \neq i} \gamma_{ij} \theta_j)x + t_i,$$

* Ollár: University of Edinburgh. Penta: ICREA, UPF, BSE and TSE. We thank M. Halac and the session organizers for spurring us to write this paper, and B. Golub for the insightful discussion. The BSE received the financial support of the Spanish Ministry of Econ. and Comp., through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). A. Penta acknowledges the financial support of the European Research Council (ERC-Starting Grant #759424).

The *efficient allocation rule* is thus $d : \Theta \to X$, such that $d(\theta) = \sum_{i\in I} \left(1 + \sum_{j\neq i} \gamma_{ji}\right)\theta_i$.

We maintain the following assumptions on preferences: first, the allocation rule is increasing in types $1+\sum_{j\neq i} \gamma_{ji} > 0$ for each i; second, valuations are *symmetric* in the sense that $\sum_{j\neq i} \gamma_{ij} = \xi$ for each i.

As for the beliefs, we maintain common knowledge that agents believe the types of others to be distributed with identical means, but they do not necessarily agree on the actual distribution. Hence, the designer regards many beliefs $B^{IM}_{\theta_i} \subseteq \Delta(\Theta_{-i})$ as possible for any given type θ_i, namely all those which are consistent with the idea that the opponents' types come from distributions with identical means. This is formally represented by belief restrictions

$$\mathcal{B}^{IM} = ((B^{IM}_{\theta_i})_{\theta_i\in\Theta_i})_{i\in I},$$

such that for all i and θ_i, $B^{IM}_{\theta_i}$ is the set that contains all distributions $b_{\theta_i} \in \Delta(\Theta_{-i})$ which satisfy $\mathbb{E}^{b_{\theta_i}}(\theta_j) = \mathbb{E}^{b_{\theta_i}}(\theta_k)$ for all $j, k \neq i$.[1]

B. Direct Mechanisms

We consider *direct mechanisms* in which agents report their payoff types, the allocation is chosen according to d, and a transfer scheme $t = (t_i)_{i\in I}$, $t_i : M \to \mathbb{R}$ specifies the transfer to each agent i, for all profiles of reports $m \in \Theta$. (To distinguish the report from the state, we maintain the notation M_i even though the message spaces are $M_i = \Theta_i$.) We let $U^t_i(m;\theta) = v_i(d(m),\theta) + t_i(m)$ denote the payoff function of the game induced by transfer scheme $(t_i)_{i\in I}$, and let $\partial^2_{ij}U^t_i := \partial^2 U^t_i/\partial m_i\partial m_j$.

For every $\theta_i \in \Theta_i$, $\mu \in \Delta(M_{-i}\times\Theta_{-i})$ and $m_i \in M_i$, we let $EU^{\mu}_{\theta_i}(m_i) = \int_{M_{-i}\times\Theta_{-i}} U_i(m_i, m_{-i}; \theta_i, \theta_{-i})\, d\mu$ denote i's expected payoff from message m_i, given his type θ_i and conjectures μ, and let $BR_{\theta_i}(\mu) := \arg\max_{m_i\in M_i} EU^{\mu}_{\theta_i}(m_i)$. For conjectures that assign probability one to the opponents reporting truthfully, we let $\mathbb{E}^{b_{\theta_i}}(U_i(m_i, \theta_{-i}; \theta_i, \theta_{-i}))) := \int_{\Theta_{-i}} U_i(m_i, \theta_{-i}; \theta_i, \theta_{-i})\, db_{\theta_i}$.

C. Implementation Concepts

We first introduce two notions of incentive compatibility:

DEFINITION 1: *A direct mechanism is* ex-post incentive compatible (ep-IC) *if,* $U_i(\theta;\theta) \geq U_i(\theta'_i, \theta_{-i}; \theta)$ *for all* θ *and* θ'_i.

A direct mechanism is $\mathcal{B}^{IM}$*-incentive compatible (*$\mathcal{B}^{IM}$*-IC) if for all* $i \in I$*, for all* θ_i, θ'_i*, and for all* $b_{\theta_i} \in B^{IM}_{\theta_i}$*,* $\mathbb{E}^{b_{\theta_i}}(U_i(\theta;\theta)) \geq \mathbb{E}^{b_{\theta_i}}(U_i(\theta'_i, \theta_{-i}; \theta))$.

As it is well-known, ep-IC characterizes partial implementability when the designer has no information about agents' beliefs, since it requires truthful revelation to be a mutual best-reply for all beliefs in $\Delta(\Theta_{-i})$ (Bergemann and Morris (2005)). $\mathcal{B}^{IM}$-IC is less demanding than ep-IC, since it only requires truthful revelation to be a mutual best-reply for all beliefs in the set $B^{IM}_{\theta_i}$, but it is still stronger than interim incentive compatibility, in which truthful revelation is required to be a mutual best response only for the single beliefs that each type may have in a standard Bayesian setting.

Our notion of full implementation requires truthful implementation to be the only strategy consistent with players' common belief in rationality and in the $\mathcal{B}^{IM}$-restrictions. Formally, for every i and θ_i, the set of conjectures that are consistent with common belief in identicality is defined as $C^{IM}_{\theta_i} := \left\{\mu_i \in \Delta(M_{-i}\times\Theta_{-i}) : marg_{\Theta_{-i}}\mu_i \in B^{IM}_{\theta_i}\right\}$. Then, given a transfer scheme t, for each $i \in I$, let $R^{IM,0}_i = \Theta_i \times M_i$ and for each $k = 1, 2, ...,$ let $R^{IM,k-1}_{-i} = \times_{j\neq i} R^{IM,k-1}_j$, where $R^{IM,k}_i$ is the set that contains all pairs (θ_i, m_i) such that $m_i \in BR_{\theta_i}(\mu_i)$ for some $\mu_i \in C^{IM}_{\theta_i} \cap \Delta\left(R^{id,k-1}_{-i}\right)$. In the limit, $R^{IM}_i = \bigcap_{k\geq 0} R^{IM,k}_i$. The set of $\mathcal{B}^{IM}$-rationalizable messages for type θ_i is defined as $R^{IM}_i(\theta_i) := \{m_i : (\theta_i, m_i) \in R^{IM}_i\}$.

Belief-free rationalizability, $R^{BF}_i(\theta_i)$, is defined similarly, replacing the set $C^{IM}_{\theta_i}$ with $\Delta(M_{-i}\times\Theta_{-i})$.

[1] These belief restrictions are a special case of the general notion of *moment conditions* introduced by Ollár and Penta (2017).

DEFINITION 2: *The transfer scheme* $t = (t_i)_{i\in I}$ fully $\mathcal{B}^{IM}$-implements d *if* $R_i^{IM}(\theta_i) = \{\theta_i\}$ *for all* θ_i *and all* i. *Allocation rule* d *is* fully $\mathcal{B}^{IM}$-implementable *if there exist some transfers that fully* $\mathcal{B}^{IM}$*-implement it.* Full belief-free implementation *obtains if* $R_i^{BF}(\theta_i) = \{\theta_i\}$ *for all* θ_i *and* i.

It is immediate that $\mathcal{B}^{IM}$-IC and ep-IC are necessary, respectively, for full $\mathcal{B}^{IM}$- and belief-free implementation.

II. The Equal-Externality Transfers

Since d is the efficient allocation rule, ep-IC is obtained by the *generalized VCG transfers*, which in this setting are:

$$t_i^*(m) = -\frac{\partial d}{\partial \theta_i}\left(\frac{1}{2}m_i^2 + \sum_{j\neq i}\gamma_{ij}m_jm_i\right).$$

Being ep-IC, the VCG transfers are obviously also $\mathcal{B}^{IM}$-IC. The following *equal-externality transfers*, instead, are $\mathcal{B}^{IM}$-IC but not ep-IC:

$$t_i^e(m) = -\frac{\partial d}{\partial \theta_i}\left(\frac{1}{2}m_i^2 + \xi\frac{\sum_{j\neq i}m_jm_i}{n-1}\right)$$

Aside from incentive compatibility, however, full implementation depends crucially on the properties of the *strategic externalities* that are induced by a mechanism (see Ollár and Penta (2017, 2021a)), that is on how players' best responses are affected by marginal misreports of the other players. Letting U_i^* and U_i^e denote, respectively, the payoff functions induced by t^* and t^e, such *strategic externalities* are conveniently captured by the second-order derivatives of these functions with respect to m_i and m_j.

It is easy to check that $\partial^2_{ij}U_i^* = -\gamma_{ij}$, and hence strong preference interdependence (i.e., large $((|\gamma_{ij}|)_{j\neq i})_{i\in I}$) immediately translates into strong strategic externalities in the VCG mechanism, which may induce multiplicity and hence a failure of full implementation.[2] This is the reason for Bergemann and Morris (2009a)'s negative result, for which belief-free full implementation is possible if and only if the preference interdependence is small enough. But if the designer has information about agents' beliefs, and specifically in the form of *moment conditions*, then incentive compatible transfers may be designed that have weak strategic externalities, and hence achieve uniqueness, even with strong preference interdependence (cf. Ollár and Penta (2017)).

In the present setting, however, due to the limited set of moment conditions that are available to the designer, such strategic externalities cannot be weakened without violating incentive compatibility: Similar to the case studied by Ollár and Penta (2021a), they can only be *redistributed*, and as it turns out, the implementation problem can be cast as one of optimally redistributing the strategic externalities induced by the VCG transfers, in order to induce a mechanism with contractive best replies.[3]

As it is easy to check, for the equal-externality transfers we have that $\partial^2_{ij}U_i^e = \frac{1}{n-1}\sum_{j\neq i}\partial^2_{ij}U_i^*$ for all i and $j\neq i$, and $\partial^2_{ii}U_i^e = \partial^2_{ii}U_i^*$ for all i. Hence, as required by the $\mathcal{B}^{IM}$-IC constraints, these transfers do preserve the *total* strategic externalities of the VCG mechanism. They only differ in that they redistribute them *evenly* across the opponents (hence their name). Such a redistribution does expand the possibility of achieving full implementation (i.e., there are environments in which $(\gamma_{ij})_{i,j\in I}$ are such that t^e achieves full implementation, but t^* does not), and in fact maximally so under the present symmetry restrictions on the environment. That is, as we show next, in our setting full implementation is achieved by some transfers if and only if it is achieved by the 'equal-externality' transfers:

PROPOSITION 1: *In symmetric environments, the following are equivalent:*

1) *Full* $\mathcal{B}^{IM}$*-Implementation is possible.*

[2]The technical condition is the following: the VCG transfers achieve full implementation if and only if the matrix SE^* in which the ij-th entry is equal to $|\gamma_{ij}|$ if $j\neq i$, and 0 otherwise, has a spectral radius smaller than one (see Ollár and Penta (2021a)).

[3]As discussed in Ollár and Penta (2021a), this is formally equivalent to an *optimal network design* problem.

2) *Full $\mathcal{B}^{IM}$-Implementation is achieved by transfers $(t_i^e)_{i\in I}$.*

3) *$|\xi|$ is less than* 1.

The improvement that t^e marks relative to t^*, in terms of possibility of achieving full implementation, is captured by the latter condition, which expresses an upper bound on the *total* level of preference interdependence. This is a less stringent requirement than the one that applies to the VCG transfers, which depends on the individual $(\gamma_{ij})_{i,j\in I}$ parameters (cf. footnote 2). Intuitively, while the condition for VCG transfers requires each γ_{ij} not to be 'too large' in absolute value, the t^e transfers enable one to accommodate settings in which some such parameters have large positive and large negative values, that 'cancel each other out' in terms of the *total* externalities condition in point (3) above.

III. Implementing Transfers: Multiplicity and Sensitivity

A. *Other Transfers for Full Implementation*

As discussed, in symmetric environments t^e achieve efficient full $\mathcal{B}^{IM}$-implementation whenever it is possible, but they are not the only transfers that do this. For instance, the *loading transfers* introduced by Ollár and Penta (2021a) can be shown to also achieve full $\mathcal{B}^{IM}$-implementation whenever t^e does. In the present setting, the loading transfers are defined as follows:

$$t_i^l(m) = -\frac{\partial d}{\partial \theta_i}\left(\frac{1}{2}m_i^2 + \xi m_1 m_i\right) \text{ if } i \neq 1$$

$$t_1^l(m) = -\frac{\partial d}{\partial \theta_1}\left(\frac{1}{2}m_1^2 + \xi m_2 m_1\right).$$

To appreciate the difference between t^e and t^l, it is best to consider the strategic externality matrices that they induce:

$$SE^e = \begin{bmatrix} 0 & \frac{\xi}{n-1} & \dots & \dots & \frac{\xi}{n-1} \\ \frac{\xi}{n-1} & 0 & \frac{\xi}{n-1} & \dots & \frac{\xi}{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\xi}{n-1} & \dots & \frac{\xi}{n-1} & 0 & \frac{\xi}{n-1} \\ \frac{\xi}{n-1} & \dots & \dots & \frac{\xi}{n-1} & 0 \end{bmatrix}, \text{ and}$$

$$SE^l = \begin{bmatrix} 0 & \xi & 0 & \dots & 0 \\ \xi & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi & 0 & 0 & \dots & 0 \end{bmatrix}$$

Similar to the equal-externality transfers, also the loading transfers preserve the *total* strategic externalities of the VCG mechanism (the row-sums of the two SE-matrices are the same, and equal to ξ). The two mechanisms only differ in the way in which the VCG-strategic externalities are *redistributed.* This is not coincidental: adapting arguments from Ollár and Penta (2021a), it can be shown that such redistribution is necessary for $\mathcal{B}^{IM}$-IC in this setting. Given this, designing transfers for full implementation boils down to a problem of 'optimally' redistributing the given total, on the entries of a strategic externality matrix, with the objective of ensuring 'most contractive' best replies. As shown in Ollár and Penta (2021a), this objective entails minimizing the spectral radius of the resulting SE-matrix. In the present setting, both t^l and t^e achieve this minimum, which is equal to $|\xi|$. For that reason, if $|\xi| < 1$, they are both "fully implementing" transfers, as any other transfer whose SE-matrix has the same spectral radius would be (there is a continuum of those).

B. *Sensitivity to Mistakes in Play*

In this section we explore the sensivity of transfers with respect to the possibility of small 'mistakes' by the agents. In words, the idea is that the designer does not know how many or which agents might be potentially faulty, and the criterion with which he/she assesses the robustness of the mechanism is the worst-case scenario across all possible configurations of sets of faulty agents. The measure of the fragility of the mechanism is therefore provided by the largest misreport consistent with $R_i^{F_\varepsilon}$, across all agents and all configurations of the set of faulty agents.

Formally, we consider mistakes in play made by a subset of agents F of size $f \neq 0$, whose choice in the mechanism is within $\varepsilon > 0$ from optimal. At each step of the iterative process, replacing the best reply sets of these agents by the set $BR^{\varepsilon}_{\theta_i}(\mu_i) = \{m_i : |m_i - m'_i| \leq \varepsilon \text{ and } m'_i \in BR_{\theta_i}(\mu_i)\}$ defines the set of F_ε-rationalizable messages, $R_i^{F_\varepsilon}(\theta_i)$. Then, the transfer scheme t is 'η-sensitive' to mistakes in play by f agents if for all F with $|F| = f$ and for all θ_i, $R_i^{IM,F_\varepsilon}(\theta_i) \subseteq [\theta_i \pm \eta\varepsilon]$. Setting $\eta_0^f(t)$ to be the infimum of such ηs, we have a measure of sensitivity to mistakes in play by a group of agents.

For the next result, we focuse on quadratic ts. This simplifies the proof, but the result holds with respect to all ts.

The next result shows that, under the maintained assumptions, t^e are *the* most robust transfers among all those which achieve full $\mathcal{B}^{IM}$-implementation:

PROPOSITION 2: *The equal-externality transfer scheme t^e is least sensitive to mistakes in play: $\eta_0^f(t^e) \leq \eta_0^f(t)$ for all t that fully $\mathcal{B}^{IM}$-implement d. Moreover if $\xi \neq 0$, then for all $t \neq t^e$ and for all $f < n$, $\eta_0^f(t^e) < \eta_0^f(t)$.*

The intuition behind this result is the following: as can be gathered from the SE^l-matrix, the loading transfers induce a very hierarchical strategic structure, in which the contractiveness of the mechanism is completely determined by the two agents with smallest preference interdependence. But loading all strategic externalities on these agents also makes the mechanism especially vulnerable to the possibility of these agents being faulty. To avoid this risk, and not knowing which set of agents may potentially be faulty, the safest solution for the designer is to redistribute the strategic externalities uniformly across all players, so that no player is especially critical for the mechanism. The same logic extends to other mechanims that achieve full implementation: as long as they induce uneven strategic externalities, the worst-case scenario of a set of faulty agent makes the mechanism less robust than the equal-externality tranfers.

IV. Proofs

For the Proof of Propositon 1, extending the proof of Lemma 1 in Ollár and Penta (2021a), we use the following Lemma:

LEMMA 1: *If (d, t) is $\mathcal{B}^{IM}$-IC, and if the spectral radius of $(|SE^t|)$ is less than one, then t ensures full $\mathcal{B}^{IM}$-implementation.*

PROOF OF PROPOSITION 1. Clearly, (2) ⇒ (1). To see that (1) ⇒ (3), note that $\mathcal{B}^{id} \subset \mathcal{B}^{IM}$ and hence $R_i^{id,k} \subset R_i^{IM,k}$ $\forall k$. Thus, if some t achieves full $\mathcal{B}^{IM}$-implementation, then it also achieves full $\mathcal{B}^{id}$-implementation, which by Theorem 2 in Ollár and Penta (2021a) is possible iff $|\xi_1\xi_2| < 1$. In a symmetric environment, $|\xi_1\xi_2| = \xi^2$, and thus $|\xi| < 1$. To see that (3) ⇒ (2), note that if $|\xi| < 1$, Gershgorin circle theorem implies that all eigenvalues of SE^e are less than 1 in absolute value. Full $\mathcal{B}^{IM}$-implementation follows from Lemma 1. ■

For the Proof of Proposition 2, we first prove the following Lemma:

LEMMA 2: *Fix $\varepsilon > 0$ and $F \subseteq I$. If t achieves $\mathcal{B}$-implementation and is s.t. $\partial^2_{ij}U_i^t$ is constant in m for all i, j, then the largest set of reports in $R_i^{F_\varepsilon}$ is the largest element of the vector $(I - |SE^t|)^{-1}\boldsymbol{\varepsilon}^F$, where $\boldsymbol{\varepsilon}^F \in \mathbb{R}^n$ is s.t. $\varepsilon_i = \varepsilon$ if $i \in F$ and $\varepsilon_i = 0$ if $i \notin F$.*

PROOF OF LEMMA 2. Let $\boldsymbol{\varepsilon} \in \mathbb{R}^n_+$ be an arbitrary non-negative vector. The characterization of the best replies in the Proof of Lemma 1 in Ollár and Penta (2021a) extends from $\mathcal{B}^{id}$ to $\mathcal{B}^{IM}$. Given the assumptions on t and U^t, letting $l := \overline{\theta} - \underline{\theta}$ and $a_1 := |SE^t|\mathbf{1}l + \boldsymbol{\varepsilon}$, it follows $\forall\theta_i$:

$$R_i^{\mathcal{B}_{\boldsymbol{\varepsilon}},1}(\theta_i) = [\theta_i \pm [a_1]_i] \cap [\underline{\theta}, \overline{\theta}]$$

In the second round, $\forall m_i \in R_i^{\mathcal{B}_{\boldsymbol{\varepsilon}},2}(\theta_i)$,

$$|\theta_i - m_i| \leq \sum_{j \neq i} \frac{|\partial^2_{ij}U_i|}{|\partial^2_{ii}U_i|}\left[\sum_{m \neq j} \frac{|\partial^2_{jm}U_j|l}{|\partial^2_{jj}U_j|} + \varepsilon_j\right] + \varepsilon_i.$$

Moreover, applying again the bounds in Lemma 1 Ollár and Penta (2021a), letting $a_2 := |SE^t|^2\mathbf{1}l + |SE^t|\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}$, for each θ_i:

$$R_i^{\mathcal{B}_{\boldsymbol{\varepsilon}},2}(\theta_i) = [\theta_i \pm [a_2]_i] \cap [\underline{\theta}, \overline{\theta}]$$

By induction, letting $a_k := |SE^t|^k \mathbf{1}l + |SE^t|^{k-1}\boldsymbol{\varepsilon} + \ldots + |SE^t|\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}$, we obtain:

$$R_i^{\mathcal{B}_{\varepsilon},k}(\theta_i) = [\theta_i \pm [a_k]_i] \cap [\underline{\theta}, \overline{\theta}]$$

Taking limits as $k \to \infty$ (while assuming that $\rho(|SE^t|) < 1$), we have that for all i and θ_i, the rationalizable messages for all θ_i are

$$R_i^{\mathcal{B}_{\varepsilon}}(\theta_i) = \left[\theta_i \pm \left[\left(I - |SE^t|\right)^{-1} \boldsymbol{\varepsilon}\right]_i\right] \cap [\underline{\theta}, \overline{\theta}].$$

Applying this formula to ε-faulty agents with F_ε completes the proof. ■

PROOF OF PROPOSITION 2. Fix $f < n$. Lemma 2 implies that, for any t, the sensitivity to mistakes in play is equal to $\eta_0^f(t) := g_f(SE^t) := \max_{F:|F|=f} \max_i \left[\left(I - |SE^t|\right)^{-1} \boldsymbol{\varepsilon}_F\right]_i$. Minimizng $\eta_0^f(t)$ in t is equivalent to:

$$\min_{SE^t} \; g_f\left(SE^t\right)$$

$$\text{s.t. } \rho\left(|SE^t|\right) < 1$$

$$\sum_{j \neq i} SE_{ij}^t = \xi \text{ and } SE_{ii}^t = 0 \text{ for all } i \in I.$$

Let t be s.t. $SE^t \neq SE^e$. If SE^t is feasible, then so is $\pi\ SE^t$, for every permutation π of the ordered set of agents $\{1, 2, \ldots, n\}$. Moreover, $g_f(SE^t) = g_f(\pi\ SE^t)$. For $F_t^* \in \operatorname{argmax}_{F:|F|=f} \max_i \left[\left(I - |SE^t|\right)^{-1} \boldsymbol{\varepsilon}_F\right]_i$,

$$g_f\left(SE^t\right) > \frac{1}{f} \boldsymbol{\varepsilon}_{F_t^*}^T \left(I - |SE^t|\right)^{-1} \boldsymbol{\varepsilon}_{F_t^*}$$

$$= \frac{1}{f} \boldsymbol{\varepsilon}_{F_t^*}^T \sum_{k=0}^{\infty} |SE^t|^k \boldsymbol{\varepsilon}_{F_t^*}.$$

The inequality is strict because SE_{ij}^t is not uniform across all i, j and $f < n$ (the latter implies that $\boldsymbol{\varepsilon}_{F_t^*}$ has at least one 0).

Moreover, for every permutation π,

$$g_f\left(\pi\ SE^t\right) > \frac{1}{f} \boldsymbol{\varepsilon}_{F_t^*}^T \sum_{k=0}^{\infty} |\pi\ SE^t|^k \boldsymbol{\varepsilon}_{F_t^*},$$

otherwise the equality of $g_f(\pi\ SE^t)$s is contradicted. Now, adding up the previous inequalities for all π, we get that

$$g_f\left(SE^t\right) > \frac{1}{n!} \sum_{\pi} \frac{1}{f} \boldsymbol{\varepsilon}_{F_t^*}^T \sum_{k=0}^{\infty} |\pi\ SE^t|^k \boldsymbol{\varepsilon}_{F_t^*}.$$

Exchanging the order of summations on the RHS, we obtain:

$$g_f\left(SE^t\right) > \frac{1}{f} \boldsymbol{\varepsilon}_{F_t^*}^T \left(I - |SE^e|\right)^{-1} \boldsymbol{\varepsilon}_{F_t^*}$$

$$= g_f\left(SE^e\right),$$

which completes the proof. ■

REFERENCES

BERGEMANN, D. AND S. MORRIS, *Robust Mechanism Design*, Econometrica, 73(6), 1771–1813, 2005.

BERGEMANN, D. AND S. MORRIS, *Robust Implementation in Direct Mechanisms*, Review of Economic Studies, 76, 1175–1204, 2009.

JACKSON, M. O., *Implementation in Undominated Strategies: A Look at Bounded Mechanisms*, Review of Economic Studies, 59(4), 757–75, 1992.

MASKIN, E., *Nash Equilibrium and Welfare Optimality*, The Review of Economic Studies, 66(1), 23–38, 1999.

OLLÁR, M. AND A. PENTA, *Full Implementation and Belief Restrictions*, American Economic Review, August, 2017.

OLLÁR, M. AND A. PENTA, *A Network Solution to Robust Implementation: the Case of Identical but Unknown Distributions*, Working Paper, 2021.

OLLÁR, M. AND A. PENTA, *Incentive Compatibility and Belief Restrictions*, mimeo, 2021.

WILSON, R., *Game-Theoretic Analysis of Trading Processes*, Advances in Economic Theory, ed. by Bewley, Cambridge University Press, 1987.