Rationalizable implementation of social choice functions: complete characterization*

Siyang Xiong†

February 18, 2022

Abstract

We provide a complete answer regarding what social choice functions can be rationalizably implemented.

*I thank Federico Echenique, Anirudh Iyer, Ritesh Jain, Ville Korpela, Michele Lombardi and anonymous referees for insightful comments.

†Department of Economics, University of California, Riverside, United States, siyang.xiong@ucr.edu

1 Introduction

A social choice function describes a socially desirable outcome for each possible state. Given a solution concept (e.g., Nash equilibrium, rationalizability), we say that a social choice function can be fully implemented, if there exists a game (or equivalently, a mechanism) such that, at any state θ, the outcome induced by *any* solution in the game matches the outcome dictated by the social choice function at θ.

What social choice functions can be fully implemented? This question has been studied extensively in the literature (see, e.g., Jackson (2001) and Maskin and Sjöström (2002) for surveys), and most papers adopt the solution concept of Nash equilibria. However, we adopt a different solution concept in this paper, and study what social choice functions can be fully implemented in rationalizable strategies.

Nash equilibrium imposes two requirements: (1) (common knowledge of) players taking best strategies to their beliefs regarding other players' strategies, and (2) players' beliefs being correct. If we impose only the first requirement, we get the solution concept of rationalizability. Compared to Nash equilibrium, rationalizability has two advantages. First, though Nash equilibrium has a simpler definition than rationalizability, the epistemic foundation of the former is more complicated than the latter (see Aumann and Brandenburger (1995)). As a result, from an epistemic view, the interpretation of the results of rationalizable implementation is more clear than that of Nash implementation. Second, if players do not have common knowledge of primitives, a mechanism designer should require *robust mechanism design*.[1] Recent papers (e.g., Bergemann and Morris (2009), Bergemann and Morris (2011), Oury and Tercieux (2012)) have shown that robust mechanism design usually leads to requiring rationalizable implementation. Thus, compared to Nash implementation, rationalizable implementation better helps us understand robust mechanism design.

Rationalizable implementation is first studied in Bergemann, Morris, and Tercieux (2011) (hereafter, BMT). Focusing on social choice functions, BMT show that strict Maskin monotonicity is necessary for rationalizable implementation, and furthermore, given two additional technical conditions, strict Maskin monotonicity is also sufficient. In three subsequent papers, Kunimoto and Serrano (2019), Jain (2021) and Xiong (2018) study ra-

[1] Conceptually, robust mechanism design means that we implement a social choice function not only at each state, but also on a neighbourhood of each state.

tionalizable implementation of social choice correspondences.[2] In this paper, we focus on social choice functions, and our goal is to fully characterize rationalizable implementation when no technical condition is imposed.

Previous characterization of rationalizable implementation (BMT, Jain (2021), Kunimoto and Serrano (2019)) hinges critically on the following two assumptions.

- No Worst Alternatives (Definition 2, hereafter NWA);[3]
- Responsiveness (Definition 3).

NWA means that the targeted social choice function cannot choose a worst outcome of any agent at any state. Responsiveness means that the targeted social choice function is injective. There are plenty of examples in which either responsiveness or NWA fails. For instance, when the number of states is larger than the number of social outcomes, responsiveness fails automatically. Technically, this paper develop tools to dispense with responsiveness and NWA, and use these tools to fully characterize rationalizable implementation when neither assumption is imposed.

Besides the technical contribution, our results also clarify two conceptual puzzles implied by the results in BMT. Even though Nash equilibria and rationalizability are two very different solution concepts, the characterization of full implementation in these two solution concepts are surprisingly similar: Maskin monotonicity for the former and strict Maskin monotonicity for the latter. Our results identify the source of such a coincidence: NWA. When NWA is relaxed, rationalizable implementation is fully characterized by *strict event monotonicity* (Definition 9), which embeds an argument of iterated deletion of never-best replies—a feature that distinguishes rationalizability from Nash equilibria. When NWA holds, we still have iterated deletion, but the order of deletion does not matter, and hence, strict event monotonicity reduces to strict Maskin monotonicity.

Second, given NWA, BMT also aim to fully characterize rationalizable implementation when responsiveness is relaxed: they show that strict Maskin monotonicity* (Definition 7) suffices for rationalizable implementation, and given an additional assumption

[2] Chen, Kunimoto, Sun, and Xiong (2021) also study rationalizable implementation, but they allow for monetary transfers.

[3] NWA is first introduced in Cabrales and Serrano (2011).

called "the best-response property," strict Maskin monotonicity* is also necessary. However, "the best-response property" is not defined on primitives, and hence, it is not clear whether "the best-response property" suffers loss of generality.

Given NWA, we prove that strict Maskin monotonicity** (Definition 8) fully characterizes rationalizable implementation. In Xiong (2022), we provide an example in which NWA and strict Maskin monotonicity** hold, while strict Maskin monotonicity* does not. Therefore, it suffers loss of generality to impose "the best-response property."

The remainder of the paper proceeds as follows: we describe the model in Section 2; we provide motivating examples in Section 3; we illustrate canonical mechanisms in Section 4; we deal with responsiveness and NWA in Sections 5 and 6, respectively; we provide the full characterization in Section 7.

2 Model

The model consists of

$$\left\langle \mathcal{I} = \{i_1, .., i_I\}, \ \Theta = \{\theta_1, .., \theta_n\}, \ Z, f : \Theta \longrightarrow Z, \ (u_i : Z \times \Theta \longrightarrow \mathbb{R})_{i \in I} \right\rangle,$$

where $\mathcal{I}$ is a finite set of I agents with $I \geq 3$, Θ a finite set of n states, Z a countable set of pure social outcomes, f a social choice function (hereafter, SCF) which maps each state in Θ to an outcome in Z, u_i the Bernoulli utility function of agent i.

For notational ease, we write Y for $\triangle(Z)$ (i.e., $Y \equiv \triangle(Z)$). We assume that agents are expected utility maximizer, and with slight abuse of notation, we also use u_i to denote agent i's expected utility function, i.e.,

$$\begin{aligned} u_i &: Y \times \Theta \longrightarrow \mathbb{R}, \\ u_i(y, \theta) &= \sum_{z \in Z} y_z u_i(z, \theta), \end{aligned}$$

where y_z denotes the probability of z under y. Throughout the paper, we use $-i$ to denote $\mathcal{I} \backslash \{i\}$ and assume $|f(\Theta)| \geq 2$.[4] Define lower and upper contour sets as follows.

$$\begin{aligned} \mathcal{L}_i(y, \theta) &= \{y' \in Y : u_i(y, \theta) \geq u_i(y', \theta)\}, \forall y \in Y, \\ \mathcal{L}_i^\circ(y, \theta) &= \{y' \in Y : u_i(y, \theta) > u_i(y', \theta)\}, \forall y \in Y, \\ \mathcal{U}_i^\circ(y, \theta) &= \{y' \in Y : u_i(y, \theta) < u_i(y', \theta)\}, \forall y \in Y. \end{aligned}$$

[4] If $|f(\Theta)| = 1$, the implementation problem can be solved trivially.

A mechanism is a tuple $\mathcal{M} = \langle M \equiv \times_{i\in\mathcal{I}} M_i,\ g : M \longrightarrow Y \rangle$, where each M_i is a countable set, and it denotes the set of strategies for agent i in $\mathcal{M}$. We now define rationalizability and rationalizable implementation. For every $i \in \mathcal{I}$, define $\mathcal{S}_i \equiv 2^{M_i}$ and $\mathcal{S} \equiv \times_{i\in\mathcal{I}} \mathcal{S}_i$. Given any $(\mathcal{M}, \theta)$, consider an operator $b^{\mathcal{M},\theta} : \mathcal{S} \longrightarrow \mathcal{S}$ with $b^{\mathcal{M},\theta} \equiv \left[b_i^{\mathcal{M},\theta} : \mathcal{S} \longrightarrow \mathcal{S}_i \right]_{i\in\mathcal{I}}$, where each $b_i^{\mathcal{M},\theta}$ is defined as follows. For every $S \in \mathcal{S}$,

$$b_i^{\mathcal{M},\theta}(S) = \left\{ m_i \in M_i : \begin{array}{c} \exists \lambda_{-i} \in \triangle(M_{-i}) \text{ such that} \\ (1)\ \lambda_{-i}(m_{-i}) > 0 \text{ implies } m_{-i} \in S_{-i}, \text{ and} \\ (2)\ m_i \in \arg\max_{m_i' \in M_i} \Sigma_{m_{-i}\in M_{-i}} \lambda_{-i}(m_{-i})\, u_i\left(g\left(m_i', m_{-i}\right), \theta\right) \end{array} \right\}.$$

Clearly, $\mathcal{S}$ is a lattice with the order of "set inclusion," and $b^{\mathcal{M},\theta}$ is increasing, i.e.,

$$S \subseteq S' \implies b^{\mathcal{M},\theta}(S) \subseteq b^{\mathcal{M},\theta}(S').$$

By Tarski's fixed point theorem, a largest fixed point of $b^{\mathcal{M},\theta}$ exists, and we denote it by $S^{\mathcal{M},\theta} \equiv \left(S_i^{\mathcal{M},\theta}\right)_{i\in\mathcal{I}}$.[5] We say $m_i \in M_i$ is rationalizable in $\mathcal{M}$ at θ if and only if $m_i \in S_i^{\mathcal{M},\theta}$, i.e., $S_i^{\mathcal{M},\theta}$ is the set of rationalizable strategies of agent i in $\mathcal{M}$ at θ, and $S^{\mathcal{M},\theta}$ is the set of rationalizable strategy profiles. We say that $S \in \mathcal{S}$ satisfies *the best-reply property* in $\mathcal{M}$ at θ if and only if $S \subset b^{\mathcal{M},\theta}(S)$. Clearly, $S \subset S^{\mathcal{M},\theta}$, if S satisfies the best-reply property.[6]

Definition 1 *An SCF $f : \Theta \longrightarrow Z$ is rationalizably implemented by a mechanism $\mathcal{M}$ if*

$$g\left[S^{\mathcal{M},\theta}\right] = \{f(\theta)\}, \forall \theta \in \Theta.$$

f is rationalizably implementable if there exists a mechanism that rationalizably implements f.

3 Motivating examples

The characterization of rationalizable implementation in BMT hinges critically on the two conditions defined as follows.

Definition 2 (No worst alternative) *An SCF $f : \Theta \longrightarrow Z$ satisfies "no worst alternative" (NWA) if, for each $(i, \theta) \in \mathcal{I} \times \Theta$, there exists $z \in Z$ such that*

$$u_i(f(\theta), \theta) > u_i(z, \theta). \tag{1}$$

[5] That is, $S^{\mathcal{M},\theta} = b^{\mathcal{M},\theta}\left(S^{\mathcal{M},\theta}\right)$ and for any $S \in \mathcal{S}$ with $S = b^{\mathcal{M},\theta}(S)$, we have $S \subset S^{\mathcal{M},\theta}$.

[6] Suppose S satisfies the best-reply property. Inductively define $\left(b^{\mathcal{M},\theta}\right)^n(S) = b^{\mathcal{M},\theta}\left[\left(b^{\mathcal{M},\theta}\right)^{n-1}(S)\right]$. Then, $\cup_{n=1}^{\infty}\left(b^{\mathcal{M},\theta}\right)^n(S)$ is a fixed point. As a result, $S \subset \cup_{n=1}^{\infty}\left(b^{\mathcal{M},\theta}\right)^n(S) \subset S^{\mathcal{M},\theta}$.

Definition 3 (responsiveness) *An SCF $f : \Theta \longrightarrow Z$ is responsive if*

$$f(\theta) = f(\theta') \implies \theta = \theta', \forall \theta, \theta' \in \Theta.$$

In this section, we provide examples, showing that we can still characterize rationalizable implementation, even if any of the two conditions fails.

3.1 Violation of NWA

First, we provide an example in which NWA fails in a particular way, and we show that BMT's characterization can still be immediately applied, subject to some modification. Second, we argue that a similar logic can be applied to any general violation of NWA.

Example 1.a: $\mathcal{I} = \{i_1, i_2, i_3, i_4\}$, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$	$i = i_4$
$z = a$	2	0
$z = b$	1	1
$z = c$	0	2

state θ_2 with $f(\theta_2) = b$,

$u_i(z, \theta_2)$	$i = i_1, i_2, i_3$	$i = i_4$
$z = a$	0	2
$z = b$	2	0
$z = c$	1	1

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_2, i_3$	$i = i_4$
$z = a$	0	1
$z = b$	1	2
$z = c$	2	0

The tables record agents' utility of each social outcome at each state. For example, agent 1 has utility of 2 for the outcome a at state θ_1.

In Example 1.a, strict Maskin monotonicity and responsiveness hold, but NWA is violated (for agent i_4 at every state). Consider a modified version of Example 1.a, described as follows—the only difference is that agent 4 is eliminated in Example 1.b.

Example 1.b: $\mathcal{I} = \{i_1, i_2, i_3\}$, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$
$z = a$	2
$z = b$	1
$z = c$	0

state θ_2 with $f(\theta_2) = b$,

$u_i(z, \theta_2)$	$i = i_1, i_2, i_3$
$z = a$	0
$z = b$	2
$z = c$	1

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_2, i_3$
$z = a$	0
$z = b$	1
$z = c$	2

In this modified example, strict Maskin monotonicity, responsiveness and NWA hold, i.e., the sufficient condition in BMT is satisfied. As a result, we can achieve rationalizable implementation in Example 1.b, which further implies that we can achieve rationalizable implementation in Example 1.a (by using the same mechanism for Example 1.b and ignore agent 4's reports).—That is, literally, BMT's characterization cannot be applied in Example 1.a, but it can be applied in Example 1.b., which indirectly provides characterization in Example 1.a.

Examples 1.a and 1.b shed light on rationalizable implementation when NWA fails. We say agent i is *inactive* at state θ if and only if NWA is violated for agent i at θ.[7] The intuition is that, if agent i is inactive at θ, we can (and should) ignore agent i at state θ—this intuition is formalized in Lemma 2 in Section 6.2. In this simple example, agent 4 is inactive at every state, and hence, we can ignore her totally.

However, in more general cases, it may happen that i is inactive at some state θ, while j $(\neq i)$ is inactive at some other state θ', as described in the following example.

Example 2: $\mathcal{I} = \{i_1, i_2, i_3, i_4\}$, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$	$i = i_4$
$z = a$	2	0
$z = b$	1	1
$z = c$	0	2

state θ_2 with $f(\theta_2) = b$,

$u_i(z, \theta_2)$	$i = i_1, i_2, i_4$	$i = i_3$
$z = a$	0	2
$z = b$	2	0
$z = c$	1	1

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_3, i_4$	$i = i_2$
$z = a$	0	1
$z = b$	1	2
$z = c$	2	0

We will show that the same logic as above applies, i.e., we need to ignore the inactive agent i_4 at state θ_1, the inactive agent i_3 at state θ_2, the inactive agent i_2 at state θ_3. —This immediately leads to a technical difficulty: since the mechanism designer does not know the true state, and hence, *a priori*, cannot tell when to ignore which agent. Thus, the goal of the mechanism designer is to build a game in which the reports of all agents collectively determine the true state, which guides him regarding when to ignore which agent. We will build a new canonical mechanism which achieves this goal. For instance, it is intuitive that our canonical mechanism would dictate the following.

$$\begin{bmatrix} \text{if agents } i_1,\, i_2 \text{ an } i_3 \text{ report state } \theta_1, \text{ we would ignore agent } i_4, \text{ and implement } f(\theta_1), \\ \text{if agents } i_1,\, i_2 \text{ an } i_4 \text{ report state } \theta_2, \text{ we would ignore agent } i_3, \text{ and implement } f(\theta_2), \\ \text{if agents } i_1,\, i_3 \text{ an } i_4 \text{ report state } \theta_3, \text{ we would ignore agent } i_2, \text{ and implement } f(\theta_3). \end{bmatrix}$$

[7] Rigorously, agent i is inactive at state θ if and only if $u_i(f(\theta), \theta) \leq u_i(z, \theta)$ for every $z \in Z$.

We discuss more intuition of our canonical mechanism in Section 4.5.

3.2 Violation of responsiveness

Consider the following degenerate example, in which responsiveness fails.

Example 3.a: $\mathcal{I} = \{i_1, i_2, i_3\}$, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$
$z = a$	1
$z = b$	2
$z = c$	0

state θ_2 with $f(\theta_2) = a$,

$u_i(z, \theta_2)$	$i = i_1, i_2, i_3$
$z = a$	1
$z = b$	2
$z = c$	0

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_2, i_3$
$z = a$	0
$z = b$	1
$z = c$	2

In this degenerate example, states θ_1 and θ_2 are the "same" in the sense that all agents' preferences do not change in the two states. Consider the following slightly modified version of Example 3.a, in which state θ_2 is eliminated.

Example 3.b: $\mathcal{I} = \{i_1, i_2, i_3\}$, $\Theta = \{\theta_1, \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$
$z = a$	1
$z = b$	2
$z = c$	0

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_2, i_3$
$z = a$	0
$z = b$	1
$z = c$	2

Clearly, responsiveness holds in Example 3.b, and BMT's result shows that we can achieve rationalizable implementation, which further implies that we can also achieve rationalizable implementation in Example 3.a (by using the same mechanism for Example 3.b).

Example 3.a is a degenerate case of violation of responsiveness. The following example shows that we can still achieve rationalizable implementation in more general cases. Example 3.c differs from Example 3.b only at state θ_2'.

Example 3.c: $\mathcal{I} = \{i_1, i_2, i_3\}$, $\Theta = \{\theta_1, \theta_2', \theta_3\}$, $Z = \{a, b, c\}$,

state θ_1 with $f(\theta_1) = a$,

$u_i(z, \theta_1)$	$i = i_1, i_2, i_3$
$z = a$	1
$z = b$	2
$z = c$	0

state θ_2 with $f(\theta_2') = a$,

$u_i(z, \theta_2')$	$i = i_1, i_2, i_3$
$z = a$	2
$z = b$	1
$z = c$	0

state θ_3 with $f(\theta_3) = c$,

$u_i(z, \theta_3)$	$i = i_1, i_2, i_3$
$z = a$	0
$z = b$	1
$z = c$	2

In Example 3.c, we have

$$\mathcal{L}_i(f(\theta_1), \theta_1) \subset \mathcal{L}_i(f(\theta_1), \theta_2'), \forall i \in \mathcal{I}, \tag{2}$$

which immediately implies $S^{\mathcal{M}, \theta_1} = S^{\mathcal{M}, \theta_2'}$, where $\mathcal{M}$ is the *canonical* mechanism used to implement f in Example 3.b. Hence, we can also achieve rationalizable implementation in Example 3.c (by using the same mechanism for Example 3.b).

In fact, given violation of responsiveness, we can achieve rationalizable implementation when a much weaker condition than (2) holds (see Definition 8 and Theorem 1).

4 Illustration of the canonical mechanisms

In this section, we describe different canonical mechanisms that are used to achieve implementation in Maskin (1999), BMT, and this paper.

4.1 The modified revelation principle

It is well-known that the revelation principle fails in full implementation. Nevertheless, a modified version holds. To see this, suppose that a mechanism $g : \times_{i \in \mathcal{I}} M_i \longrightarrow Z$ achieves full implementation in some solution concept (e.g., Nash equilibrium, rationalizability). Pick any one of the solution $((\phi_i : \Theta \longrightarrow M_i)_{i \in \mathcal{I}})$ of g,[8] and we re-label each "$\phi_i(\theta)$" to a new message "θ." Furthermore, denote $\widetilde{M}_i \equiv M_i \diagdown \phi_i(\Theta)$, and hence $M_i = [\phi_i(\Theta)] \cup \widetilde{M}_i$.

[8] That is, for Nash implementatin, $[\phi_i(\theta)]_{i \in \mathcal{I}}$ is one Nash equilibrium at state θ in g, and for rationalizable $[\phi_i(\theta)]_{i \in \mathcal{I}}$ is one rationalizable strategy profile at state θ in g.

Then, the original mechanism g can be "rephrased" to $\widetilde{g} : \times_{i \in \mathcal{I}} \left(\Theta \cup \widetilde{M}_i \right) \longrightarrow Z$ by relabeling $\phi_i(\theta)$ to θ, and $\widetilde{g}$ achieves full implementation.[9] Let us call such $\widetilde{g}$ an *augmented direct mechanism* (with the augmented messages in $\widetilde{M}_i$ for agent i). Therefore, this establishes a modified revelation principle for full implementation: it suffers no loss of generality to consider augmented direct mechanisms.[10]

This modified revelation principle provides the basis for canonical mechanisms in full implementation. First, all agents truthfully reporting θ at state θ is always a solution in the augmented direct mechanism. Second, the implementation problem is reduced to identifying the augmented messages in $\widetilde{M}_i$ in order to achieve full implementation. Most papers in the literature of full implementation follow this idea.

4.2 The canonical mechanisms

There is a generic form for most mechanisms in full implementation, which is described as follows.

> Agents are invited to report the true state. There are three cases for agents' reports: (1) agreement, i.e., all agents report the same state θ; (2) unilateral deviation, i.e., all agents except agent j report θ; (3) multi-lateral deviation, i.e., this includes all other scenarios. In Case (1), the canonical mechanism picks $f(\theta)$. In Case (2), agent j is allowed to choose any outcome in $\mathcal{L}_j(f(\theta), \theta)$, which ensures that truthful reporting is a Nash equilibrium (and hence, rationalizable). In Case (3), we first let all agents compete by submitting a positive integer. The agent who submits the largest integer wins (subject to any tie-breaking rule), and we let the winner choose any outcome in Z.

[9] Besides the solution $[\phi_i(\theta)]_{\theta \in \Theta, i \in \mathcal{I}}$ chosen in g, there may be other solutions involving messages in $[\times_{i \in \mathcal{I}} M_i] \diagdown \{\theta \in \Theta : [\phi_i(\theta)]_{i \in \mathcal{I}}\}$. Such solutions correspond to Case (2) and Case (3) in canonical mechanisms discussed below. Thus, our goal is to choose each $\widetilde{M}_i$ carefully so that the solutions involving Case (2) and Case (3) in canonical mechanisms still achieve full implementation.

[10] However, this revelation principle is much weaker than the original revelation principal for partial implemenation. For the latter, a direct mechanism is precisely defined, but for the former, this is not true for an augamented direct mechanism. That is, *a priori*, it is not clear what $\widetilde{M}_i$ should be. Finding $\widetilde{M}_i$ is one of the goals when we solve a full implementation problem.

At the true state θ, "all agents reporting θ" is a "good" equilibrium (or solution), which induces $f(\theta)$. In order to achieve full implementation, we have to further make sure that there is no "bad" equilibrium in any of Cases (1), (2) and (3).

For different environments and/or solution concepts, we may have to modify the canonical mechanism above slightly, which is illustrated below.

4.3 Illustration of Maskin (1999)

Maskin (1999) adopts the canonical mechanism in Section 4.2, and uses Maskin monotonicity to eliminate "bad" equilibria in Cases (1), and uses no-veto power to eliminate "bad" equilibria in Cases (2) and (3).

Definition 4 (Maskin monotonicity) *An SCF f satisfies Maskin monotonicity if*

$$f(\theta) \neq f(\theta') \Longrightarrow \left(\begin{array}{c} \exists j \in \mathcal{I}, \\ \mathcal{L}_j(f(\theta), \theta) \cap \mathcal{U}_j^{\circ}(f(\theta), \theta') \neq \varnothing \end{array} \right), \forall (\theta, \theta') \in \Theta \times \Theta. \quad (3)$$

Definition 5 (no-veto power) *An SCF f satisfies no-veto power if*

$$\left| \left\{ i \in \mathcal{I} : a \in \arg\max_{z \in Z} u_i(z, \theta) \right\} \right| \geq |\mathcal{I}| - 1 \Longrightarrow a \in f(\theta), \forall (\theta, a) \in \Theta \times Z.$$

Suppose the true state is θ'. A "bad" equilibrium in Case (1) means that all agents report θ with $f(\theta) \neq f(\theta')$. Given Maskin monotonicity, such a strategy profile cannot be an equilibrium, because (3) implies agent j has a profitable deviation to Case (2), i.e., j can pick $y \in \mathcal{L}_j(f(\theta), \theta) \cap \mathcal{U}_j^{\circ}(f(\theta), \theta')$.—When this happens, j is called a whistle-blower, and $y \in \mathcal{L}_j(f(\theta), \theta) \cap \mathcal{U}_j^{\circ}(f(\theta), \theta')$ is called j's blocking plan.[11]

A "bad" equilibrium in Case (2) or Case (3) is that agent j deviates from Case (1), and it induces $c \neq f(\theta)$. Given no-veto power, such a strategy profile cannot be an equilibrium, because the other $|\mathcal{I}| - 1$ agents (i.e., agents $-j$) can further deviate to Case

[11] That is, j uses $y \in \mathcal{L}_j(f(\theta), \theta) \cap \mathcal{U}_j^{\circ}(f(\theta), \theta')$ to inform the mechanism designer that agents $-j$ are lying by reporting θ. "$y \in \mathcal{L}_j(f(\theta), \theta)$" ensures that j's information is credible, because if agents $-j$ were not lying, y would be inferior to j at θ. "$y \in \mathcal{U}_j^{\circ}(f(\theta), \theta')$" ensures that y is indeed a profitable deviation to j at the true state θ'.

(3) , and induce their top outcomes in Z (by submitting a largest integer). If it were an equilibrium, c would be a top outcome for the other $|\mathcal{I}|-1$ agents, which, together with no-veto power, implies $c=f(\theta)$, contradicting $c \neq f(\theta)$.

4.4 Illustration of BMT

To achieve rationalizable implementation, BMT uses strict Maskin monotonicity to eliminate bad solutions in Case (1), when responsiveness holds.[12] The intuition is the same as above.

Definition 6 (strict Maskin monotonicity) *An SCF f satisfies strict Maskin monotonicity if*

$$f(\theta) \neq f(\theta') \Longrightarrow \left(\begin{array}{c} \exists j \in \mathcal{I}, \\ \mathcal{L}_j^{\circ}(f(\theta), \theta) \cap \mathcal{U}_j^{\circ}(f(\theta), \theta') \neq \varnothing \end{array} \right), \forall (\theta, \theta') \in \Theta \times \Theta. \tag{4}$$

BMT uses NWA to eliminate bad solutions in Cases (2) and (3). NWA implies existence of $\underline{y} \in Y$ such that

$$\underline{y} \notin \bigcup_{\theta \in \Theta} \bigcup_{i \in \mathcal{I}} \arg\max_{y \in Y} u_i(y, \theta), \tag{5}$$

i.e., $\underline{y}$ is never a top outcome for any agent at any state. Furthermore, BMT prove that NWA and strict Maskin monotonicity implies existence of $z_i(\theta, \theta) \in \mathcal{L}_i(f(\theta), \theta)$ for each $(\theta, i) \in \Theta \times \mathcal{I}$ such that

$$\max_{y \in \mathcal{L}_i(f(\theta), \theta)} u_i(y, \theta^*) > u_i(z_i(\theta, \theta), \theta^*), \forall \theta^* \in \Theta. \tag{6}$$

Furthermore, BMT modify Cases (2) and (3) in the canonical mechanism as follows. In Case (2), agent j is allowed to choose any $y \in \mathcal{L}_j(f(\theta), \theta)$ and any positive integer n, and the mechanism picks

$$\frac{n-1}{n} \times y + \frac{1}{n} \times z_i(\theta, \theta) \in \mathcal{L}_j(f(\theta), \theta) \tag{7}$$

[12] Suppose the true state in θ^*. More precisely, strict Maskin montonicity implies existence of a whistleblower whenever agents reach an agreement on θ with $f(\theta) \neq f(\theta^*)$. Furthermore, given responsiveness, "$f(\theta) \neq f(\theta^*)$" is equivalent to "$\theta \neq \theta^*$." Therefore, strict Maskin monotonicity eliminates any bad solutions in Case (1), i.e., agreement on θ with $\theta \neq \theta^*$.

In Case (3), the agent who submits the largest integer is allowed to choose any $z \in Z$ and any positive integer n, and the mechanism picks

$$\frac{n-1}{n} \times z + \frac{1}{n} \times \underline{y}. \tag{8}$$

(5), (6), (7) and (8) imply that all agents can never have a best reply in Cases (2) and (3), and hence, no "bad" solution.

4.5 Illustration of our canonical mechanism: NWA

When neither NWA nor responsiveness is imposed, the full characterization of rationalizable implementation is complicated. For expositional ease, we treat the two technical problems separately, and develop tools to fully characterize rationalizable implementation when only one of the two assumptions holds.

Suppose responsiveness holds, while NWA does not. The innovation of our canonical mechanism is introducing the notion of "active agents."

$$\mathcal{I}^{\theta} = \{ i \in \mathcal{I} : \exists z \in Z \text{ such that } u_i \left(f\left(\theta\right), \theta \right) > u_i \left(z, \theta \right) \}, \forall \theta \in \Theta,$$
$$\text{and } \mathcal{I}^{E} = \bigcap_{\theta \in E} \mathcal{I}^{\theta}, \forall E \in \left[2^{\Theta} \backslash \{\varnothing\} \right].$$

That is, $\mathcal{I}^{\theta}$ is the set of agents who can make condition (1) in NWA hold at state θ. We call agents in $\mathcal{I}^{\theta}$ *active agents* at state θ. Clearly, NWA is equivalent to requiring $\mathcal{I}^{\Theta} = \mathcal{I}$.

Without NWA, what goes wrong in BMT's canonical mechanism? Precisely, conditions (5) and (6) do not hold, and BMT's proof breaks down. However, by incorporating the notion of active agents, modified versions of (5) and (6) hold.

$$\underline{y} \notin \bigcup_{\theta \in \Theta} \bigcup_{i \in \mathcal{I}^{\theta}} \arg\max_{y \in Y} u_i \left(y, \theta \right). \tag{9}$$

$$\max_{y \in \mathcal{L}_i(f(\theta), \theta)} u_i \left(y, \theta^* \right) > u_i \left(z_i \left(\theta, \theta \right), \theta^* \right), \forall \left(\theta, \theta^* \right) \in \Theta \times \Theta, \forall i \in \mathcal{I}^{\theta^*}. \tag{10}$$

I.e., (5) and (6) hold only for active agents at each state.

With this being sorted out, there is only one major difference between BMT's canonical mechanism in Section 4.4 and ours: how is an agreement defined in case (1)?

Specifically, we modify the canonical mechanism as follows. In case (1), all agents in $\mathcal{I}^{\theta}$ report the same state θ,[13] and the mechanism picks $f(\theta)$. The rest of the canonical mechanism remains the same as above.

We will use strict event monotonicity and dictator monotonicity (see Definitions 9 and 10) to eliminate "bad" solutions in Case (1). And as above, all agents never have a best reply in Cases (2) and (3), i.e., no "bad" solution.

4.6 Illustration of our canonical mechanism: responsiveness

Suppose NWA holds, while responsiveness does not. What goes wrong in BMT's canonical mechanism? To see this, consider a concrete example.

$$\Theta = \left\{\theta^1, \theta^2, \theta^3\right\} \text{ such that } f\left(\theta^1\right) \neq f\left(\theta^2\right) = f\left(\theta^3\right).$$

Suppose that the true state is θ^2. Strict Maskin monotonicity ensures that agreement on θ^1 in the canonical mechanism will not be a Nash equilibrium (or more precisely, will not be rationalizable). However, strict Maskin monotonicity does not preclude the possibility that agreement on θ^3 in the canonical mechanism is a Nash equilibrium.[14] Such a possibility does not destroy Nash implementation, because $f\left(\theta^2\right) = f\left(\theta^3\right)$. However, it destroys rationalizable implementation, which is due to a distinct feature of rationalizability. When both "all agents reporting θ^2" and "all agents reporting θ^3" are Nash equilibria, one rationalizable strategy profile could be: "odd-indexed agents reporting θ^2, and even-indexed agents reporting θ^3," which would trigger either case (2) or case (3), and induce undesired outcomes $z_i(\theta, \theta)$ or $\underline{y}$ with positive probability, i.e., rationalizable implementation is not achieved.

[13] To see why an agreement is determined by agents in $\mathcal{I}^{\theta}$ *only*, suppose that a canonical mechanism achieves rationalizable implementation. At state θ, consider any inactive agent $j \notin \mathcal{I}^{\theta}$. Pick any rationalizable strategy of agent j, and it is a best reply to a rationalizable conjecture, which induces the worst outcome $f(\theta)$ for j. This immediately implies any other strategy must also be a best reply to the same rationalizable conjecture, i.e., all the other strategies are rationalizable for j. Or equivalently, any of j's report of the true state is not informative. Therefore, an agreement is determined by agents in $\mathcal{I}^{\theta}$ only.

[14] Strict Maskin monotonicity kicks in only when $f\left(\theta^2\right) \neq f\left(\theta^3\right)$, but, here, we have $f\left(\theta^2\right) = f\left(\theta^3\right)$.

4.6.1 BMT's attempt

BMT provide a first attempt to characterize rationalizable implementation, when responsiveness is violated. Suppose the true state is θ^2. Strict Maskin monotonicity provides tools for a whistle-blower to block *only* "reporting θ" with $f(\theta) \neq f\left(\theta^2\right)$. The example above shows that the problem comes from θ' with $f(\theta') = f\left(\theta^2\right)$. Thus, if there is no whistle-blower who is able to block "reporting θ'" with $f(\theta') = f\left(\theta^2\right)$, we have to identify θ' and θ^2, in order to avoid (the undesired outcomes in) Cases (2) and (3).[15] Or equivalently, we must form a partition $\mathcal{P}^*$ on Θ such that

$$\mathcal{P}^*(\theta) = \mathcal{P}^*(\theta') \Longrightarrow f(\theta) = f(\theta'), \forall (\theta, \theta') \in \Theta \times \Theta,$$

and at any true state θ, reporting any state in $\mathcal{P}^*(\theta)$ must be rationalizable in the canonical mechanism for all agents at θ. This immediately leads to an additional requirement:

$$\mathcal{P}^*(\theta) \neq \mathcal{P}^*(\theta') \Longrightarrow \begin{pmatrix} \text{at the true state } \theta', \\ \text{there exists a whistle-blower } j \in \mathcal{I} \text{ such that} \\ j \text{ can block "agents } -j \text{ reporting } \widehat{\theta}\text{"} \\ \textit{simultaneously} \text{ for any } \widehat{\theta} \in \mathcal{P}^*(\theta) \end{pmatrix}, \forall (\theta, \theta') \in \Theta \times \Theta. \tag{11}$$

By reporting $\widehat{\theta} \in \mathcal{P}^*(\theta)$, agents $-j$ disclose that any state in $\mathcal{P}^*(\theta)$ might be the true state, and hence, the whistle-blower must block all of the false states in $\mathcal{P}^*(\theta)$ simultaneously.

One critical issue is how we should formalize "simultaneously" in (11). BMT adopt the following formalization, i.e., strict Maskin monotonicity* (Definition 7).

$$\mathcal{P}^*(\theta) \neq \mathcal{P}^*(\theta') \Longrightarrow \left(\begin{array}{c} \exists j \in \mathcal{I}, \\ \left[\bigcap_{\widehat{\theta} \in \mathcal{P}^*(\theta)} \mathcal{L}_j^\circ \left(f\left(\widehat{\theta}\right), \widehat{\theta} \right) \right] \cap \mathcal{U}_j^\circ \left(f(\theta), \theta' \right) \neq \varnothing \end{array} \right), \forall (\theta, \theta') \in \Theta \times \Theta.$$

That is, at the true state θ', when all agents report $\widehat{\theta} \in \mathcal{P}^*(\theta)$, there must exist a whistle-blower j with a blocking plan $y \in \left[\bigcap_{\widehat{\theta} \in \mathcal{P}^*(\theta)} \mathcal{L}_j^\circ \left(f\left(\widehat{\theta}\right), \widehat{\theta} \right) \right] \cap \mathcal{U}_j^\circ \left(f(\theta), \theta' \right)$, i.e., y is credible at *all states in* $\mathcal{P}^*(\theta)$, and is strictly profitable at θ'.

[15] That is, when all agents report θ' or θ^2, we regard it as Case (1), and the canonical mechanism picks $f(\theta') = f\left(\theta^2\right)$.

Specifically, BMT modify the canonical mechanism as follows. In case (1), all agents report some states in $\mathcal{P}^*(\theta)$, the mechanism picks $f(\theta)$. In case (2), all agents except agent j report some states in $\mathcal{P}^*(\theta)$, then we let agent j choose any outcome $y \in \left[\bigcap_{\widehat{\theta}\in\mathcal{P}^*(\theta)} \mathcal{L}_j^{\circ}\left(f\left(\widehat{\theta}\right),\widehat{\theta}\right)\right]$ and any positive integer n, and the mechanism picks $\frac{n-1}{n}\times y+\frac{1}{n}\times z_i(\theta,\theta)\in\mathcal{L}_i(f(\theta),\theta)$. The rest of the canonical mechanism remains the same.

BMT uses strict Maskin monotonicity* to eliminate "bad" solutions in Case (1). And as above, all agents never have a best reply in Cases (2) and (3), i.e., no "bad" solution.

4.6.2 Our canonical mechanism

We take a different formalization of "simultaneously" in (11) as follows (i.e., strict Maskin monotonicity** in Definition 8).

$$\mathcal{P}^*(\theta)\neq\mathcal{P}^*(\theta') \Longrightarrow \left(\begin{array}{c}\exists j\in\mathcal{I},\exists\phi:\Theta\to Y,\\ \phi\left(\widehat{\theta}\right)\in\mathcal{L}_j^{\circ}\left(f\left(\widehat{\theta}\right),\widehat{\theta}\right)\cap\mathcal{U}_j^{\circ}\left(f(\theta),\theta'\right)\neq\varnothing,\\ \forall\widehat{\theta}\in\mathcal{P}^*(\theta),\end{array}\right),\forall(\theta,\theta')\in\Theta\times\Theta. \tag{12}$$

Our innovation is that we allow a whistle-blower to adopt a *state-contingent blocking plan*, i.e., $\phi:\Theta\to Y$ in (12).

Specifically, we modify the canonical mechanism as follows. In case (1), all agents report some states in $\mathcal{P}^*(\theta)$, the mechanism picks $f(\theta)$. In case (2), all agents except agent j report some states in $\mathcal{P}^*(\theta)$, then we let agent j choose any $\phi:\Theta\to Y$ such that $\phi\left(\widehat{\theta}\right)\in\mathcal{L}_j^{\circ}\left(f\left(\widehat{\theta}\right),\widehat{\theta}\right)$ for every $\widehat{\theta}\in\mathcal{P}^*(\theta)$ and any positive integer n, and the mechanism picks $\frac{n-1}{n}\times\phi\left(\theta^{j+1}\right)+\frac{1}{n}\times z_i\left(\theta^{j+1},\theta^{j+1}\right)\in\mathcal{L}_i\left(f\left(\theta^{j+1}\right),\theta^{j+1}\right)$, where θ^{j+1} denotes the report of agent $(j+1)$ module I. The rest of the mechanism remains the same.

We use strict Maskin monotonicity** to eliminate "bad" solutions in Case (1). And as above, all agents never have a best reply in Cases (2) and (3), i.e., no "bad" solution.

5 How to deal with violation of responsiveness?

In this section, we drop responsiveness, and fully characterize rationalizable implementation when only NWA is assumed.

5.1 A summary of the full characterization

Let $\mathcal{P}_f$ denote the partition on Θ induced by f, which is defined as follows.

$$\mathcal{P}_f(\theta) = \left\{\theta' \in \Theta : f\left(\theta'\right) = f(\theta)\right\}, \forall \theta \in \Theta.$$

Given NWA, BMT show that strict Maskin monotonicity* defined below is sufficient for rationalizable implementation.

Definition 7 (strict Maskin monotonicity*) *An SCF $f : \Theta \longrightarrow Z$ satisfies strict Maskin monotonicity* if there exists a partition $\mathcal{P}$ on Θ finer than $\mathcal{P}_f$ such that for any $\left(\theta, \theta'\right) \in \Theta \times \Theta$,*

$$\theta' \in \mathcal{P}(\theta) \Longleftarrow \left[\begin{array}{c} \forall (y,i) \in Y \times \mathcal{I}, \\ \left(\begin{array}{c} u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y, \widehat{\theta}\right), \\ \forall \widehat{\theta} \in \mathcal{P}(\theta) \end{array}\right) \Longrightarrow u_i\left(f(\theta), \theta'\right) \geq u_i\left(y, \theta'\right) \end{array}\right],$$

or, equivalently,

$$\theta' \notin \mathcal{P}(\theta) \Longrightarrow \left[\begin{array}{c} \exists (y,i) \in Y \times \mathcal{I}, \\ \left(\begin{array}{c} u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y, \widehat{\theta}\right), \\ \forall \widehat{\theta} \in \mathcal{P}(\theta) \end{array}\right) \text{ and } u_i\left(y, \theta'\right) > u_i\left(f(\theta), \theta'\right) \end{array}\right]. \tag{13}$$

We propose a new axiom which is weaker than strict Maskin monotonicity*.

Definition 8 (strict Maskin monotonicity**) *An SCF $f : \Theta \longrightarrow Z$ satisfies strict Maskin monotonicity** if there exists a partition $\mathcal{P}$ on Θ finer than $\mathcal{P}_f$ such that for any $\left(\theta, \theta'\right) \in \Theta \times \Theta$,*

$$\theta' \in \mathcal{P}(\theta) \Longleftarrow \left[\begin{array}{c} \forall i \in \mathcal{I},\ \exists \widehat{\theta} \in \mathcal{P}(\theta), \forall y \in Y, \\ u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y, \widehat{\theta}\right) \Longrightarrow u_i\left(f(\theta), \theta'\right) \geq u_i\left(y, \theta'\right) \end{array}\right], \tag{14}$$

or, equivalently,

$$\theta' \notin \mathcal{P}(\theta) \implies \left[\begin{array}{c} \exists i \in \mathcal{I} \text{ such that } \forall \widehat{\theta} \in \mathcal{P}(\theta), \exists y^{\widehat{\theta}} \in Y, \\ u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y^{\widehat{\theta}}, \widehat{\theta}\right) \text{ and } u_i\left(y^{\widehat{\theta}}, \theta'\right) > u_i\left(f(\theta), \theta'\right). \end{array} \right]. \quad (15)$$

Conditions (13) and (15) speak out the difference between the two axioms: when $\theta' \notin \mathcal{P}(\theta)$ (i.e., the true state is θ' and all agents falsely report states in $\mathcal{P}(\theta)$), strict Maskin monotonicity* requires existence of a whistle-blower i and a *common* blocking plan y which works for every states $\widehat{\theta} \in \mathcal{P}(\theta)$, while strict Maskin monotonicity** requires existence of a whistle-blower i and a *state-contingent* blocking plan $y^{\widehat{\theta}}$ which works for each state $\widehat{\theta} \in \mathcal{P}(\theta)$. Clearly, strict Maskin monotonicity* implies strict Maskin monotonicity**, and the latter also suffices for rationalizable implementation.

Theorem 1 *Suppose that an SCF $f : \Theta \longrightarrow Z$ satisfies NWA. Then, f is rationalizably implementable if and only if f satisfies strict Maskin monotonicity**.*

The "only if" and "if" parts of Theorem 1 are proved in Sections 5.2 and A.3.

Given a condition called "the best-response property," BMT show that strict Maskin monotonicity* is necessary for rationalizable implementation. However, the best-response property is not defined on primitives, and hence it remains an open question regarding whether it suffers loss of generality to assume the best-response property? Xiong (2022) provides a negative answer for this question. Specifically, it constructs an example in which NWA and strict Maskin monotonicity** hold, but strict Maskin monotonicity* fails. By Theorem 1, we can achieve rationalizable implementation in this example, even though Maskin monotonicity* doses not hold.

5.2 The proof of the "only if" part of Theorem 1

Suppose that f is rationalizably implemented by a mechanism $\mathcal{M} = \langle M,\ g : M \longrightarrow Y \rangle$. Consider the partition defined as follows.

$$\mathcal{P}(\theta) = \left\{ \widetilde{\theta} \in \Theta : S^{\mathcal{M}, \widetilde{\theta}} = S^{\mathcal{M}, \theta} \right\}, \forall \theta \in \Theta.$$

Since f is rationalizably implemented by $\mathcal{M}$, the partition $\mathcal{P}$ defined above is finer than $\mathcal{P}_f$. Furthermore, for any $(\theta, \theta') \in \Theta \times \Theta$, suppose that the right-hand side of (14) holds and we aim to prove $\theta' \in \mathcal{P}(\theta)$ (i.e., strict Maskin monotonicity** holds). We need the following result, and its proof is relegated to Appendix A.1.

Lemma 1 *If an SCF f is rationalizably implemented by a mechanism $\mathcal{M}$, we have*

$$S^{\mathcal{M},\theta} \subset S^{\mathcal{M},\theta'} \implies S^{\mathcal{M},\theta} = S^{\mathcal{M},\theta'}, \forall (\theta, \theta') \in \Theta \times \Theta.$$

We will show that $S^{\mathcal{M},\theta}$ satisfies the best-reply property in $\mathcal{M}$ at state θ', i.e., $S^{\mathcal{M},\theta} \subset S^{\mathcal{M},\theta'}$. By Lemma 1, we have $S^{\mathcal{M},\theta} = S^{\mathcal{M},\theta'}$, and hence, $\theta' \in \mathcal{P}(\theta)$.

Consider any $i \in \mathcal{I}$, and pick any $m_i \in S_i^{\mathcal{M},\theta}$. By the right-hand side of (14), there exists $\exists \widehat{\theta} \in \mathcal{P}(\theta)$ such that

$$u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y, \widehat{\theta}\right) \implies u_i\left(f(\theta), \theta'\right) \geq u_i\left(y, \theta'\right), \forall y \in Y. \tag{16}$$

Since $\widehat{\theta} \in \mathcal{P}(\theta)$, we have $m_i \in S_i^{\mathcal{M},\theta} = S_i^{\mathcal{M},\widehat{\theta}}$, and hence, there exists $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\widehat{\theta}}\right)$ such that m_i is a best reply to λ_{-i} for agent i at State $\widehat{\theta}$, i.e.,

$$u_i\left(g(m_i, \lambda_{-i}), \widehat{\theta}\right) = u_i\left(f\left(\widehat{\theta}\right), \widehat{\theta}\right) = u_i\left(f(\theta), \widehat{\theta}\right) \geq u_i\left(g(\widetilde{m}_i, \lambda_{-i}), \widehat{\theta}\right), \forall \widetilde{m}_i \in M_i,$$

which further implies

$$u_i\left(g(m_i, \lambda_{-i}), \widehat{\theta}\right) = u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(g(\widetilde{m}_i, \lambda_{-i}), \widehat{\theta}\right), \forall \widetilde{m}_i \in M_i \diagdown S_i^{\mathcal{M},\widehat{\theta}}, \tag{17}$$

$$\text{and } g(m_i, \lambda_{-i}) = g(\widetilde{m}_i, \lambda_{-i}) = f\left(\widehat{\theta}\right), \forall \widetilde{m}_i \in S_i^{\mathcal{M},\widehat{\theta}}. \tag{18}$$

Then, (16) and (17) imply

$$u_i\left(g(m_i, \lambda_{-i}), \theta'\right) = u_i\left(f(\theta), \theta'\right) = u_i\left(f\left(\widehat{\theta}\right), \theta'\right) \geq u_i\left(g(\widetilde{m}_i, \lambda_{-i}), \theta'\right), \forall \widetilde{m}_i \in M_i \diagdown S_i^{\mathcal{M},\widehat{\theta}}. \tag{19}$$

Finally, (18) and (19) imply

$$u_i\left(g(m_i, \lambda_{-i}), \theta'\right) = u_i\left(f\left(\widehat{\theta}\right), \theta'\right) \geq u_i\left(g(\widetilde{m}_i, \lambda_{-i}), \theta'\right), \forall \widetilde{m}_i \in M_i,$$

i.e., m_i is a best reply to λ_{-i} for i at θ'. Thus, $S^{\mathcal{M},\theta}$ satisfies the best reply property at θ'.■

6 How to deal with violation of NWA?

In this section, we drop NWA, and fully characterize rationalizable implementation when only responsiveness is assumed.

6.1 A summary of the full characterization

We fully characterize rationalizable implementation by two new axioms. First, we propose an axiom called "strict event monotonicity," which strengthens strict Maskin monotonicity (Definition 6), and when NWA holds, the two notions coincide.

Definition 9 (strict event monotonicity) *An SCF $f : \Theta \longrightarrow Z$ satisfies strict event monotonicity if for every $(\theta', E) \in \Theta \times [2^\Theta \backslash \{\varnothing\}]$:*

$$\{f(\theta')\} = f(E) \text{ whenever}$$

$$u_i(f(\theta), \theta) > u_i(y, \theta) \implies u_i(f(\theta), \theta') \geq u_i(y, \theta'), \forall (\theta, y, i,) \in E \times Y \times \mathcal{I}^E, \quad (20)$$

or, equivalently,

$$\{f(\theta')\} \neq f(E) \text{ implies}$$

$$u_i(f(\theta), \theta) > u_i(y, \theta) \text{ and } u_i(y, \theta') > u_i(f(\theta), \theta'), \text{for some } (\theta, y, i,) \in E \times Y \times \mathcal{I}^E. \quad (21)$$

There are two subtle differences between strict event monotonicity and strict Maskin monotonicity. First, pairwise comparison between states (i.e., "θ' Vs θ") is conducted in strict Maskin monotonicity, while a state is compared to a group of states (i.e., "θ' Vs E") in strict event monotonicity. Second, as shown in conditions (4) and (21), the whistle-blower is required to be an active agent in $\mathcal{I}^E$ in strict event monotonicity, while he or she could be anyone in $\mathcal{I}$ in strict Maskin monotonicity. It is straightforward to show that, given NWA, strict event monotonicity is equivalent to strict Maskin monotonicity.

Definition 10 (dictator monotonicity) *Agent $i \in \mathcal{I}$ is a dictator if $\{i\} = \mathcal{I}^\theta$. An SCF $f : \Theta \longrightarrow Z$ satisfies dictator monotonicity if for every $(i, \theta, \theta', \theta'') \in \mathcal{I} \times \Theta \times \Theta \times \Theta$,*

$$\begin{bmatrix} \{i\} = \mathcal{I}^\theta \\ \text{and } f(\theta) \neq f(\theta') \end{bmatrix} \implies \begin{bmatrix} \exists y \in Y \text{ such that} \\ u_i(f(\theta''), \theta'') \geq u_i(y, \theta'') \text{ and } u_i(y, \theta') > u_i(f(\theta), \theta') \end{bmatrix}. \quad (22)$$

Theorem 2 *A responsive SCF $f : \Theta \longrightarrow Z$ is rationalizably implementable if and only if f satisfies strict event monotonicity and dictator monotonicity.*

We provide an intuition of Theorem 2 in Sections 6.3 and 6.4, and the proofs are presented in Sections 6.3.2, 6.4.2 and A.5.

6.2 A crucial observation

We first offer a crucial observation, which provides a powerful tool in establishing both the necessity and the sufficiency parts of Theorem 2.

Lemma 2 *Suppose that a social choice function $f : \Theta \longrightarrow Z$ is rationalizably implemented by a mechanism $\mathcal{M} = \langle M,\ g : M \longrightarrow Y \rangle$. Then, for every $(i, \theta) \in \mathcal{I} \times \Theta$,*

$$i \notin \mathcal{I}^{\theta} \implies \left[\begin{array}{c} S_i^{\mathcal{M},\theta} = M_i \text{ and} \\ g\left(m_i, m_{-i}\right) = f\left(\theta\right), \forall \left(m_i, m_{-i}\right) \in M_i \times S_{-i}^{\mathcal{M},\theta}. \end{array} \right]. \tag{23}$$

Given f being rationalizably implemented by $\mathcal{M}$, Lemma 2 says that at any state, all strategies are rationalizable for an inactive agent. The proof is straightforward: given $i \notin \mathcal{I}^{\theta}$, pick any $m_i \in S_i^{\mathcal{M},\theta}$, and there exists $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\theta}\right)$ such that m_i is a best reply to λ_{-i} and $g\left(m_i, \lambda_{-i}\right) = f\left(\theta\right)$. Given $i \notin \mathcal{I}^{\theta}$, $f\left(\theta\right)$ is a worst outcome for i at θ, and hence, any other strategy is a best reply to λ_{-i}. Therefore, $S_i^{\mathcal{M},\theta} = M_i$.

Lemma 2 sheds light on the canonical mechanism that rationalizably implements f: at true state θ, we should let active agents in $\mathcal{I}^{\theta}$ *only* to determine the outcome of the mechanism, and ignore agents in $\mathcal{I} \backslash \mathcal{I}^{\theta}$, because they are not informative.

6.3 The meaning of dictator monotonicity

6.3.1 The sufficiency part of dictator monotonicity: intuition

We first show why we need dictator monotonicity when we prove the sufficiency part of Theorem 2. If agent i is a dictator at θ and i reports θ in the canonical mechanism, by Lemma 2, we should trust i and ignore other agents' reports, and pick $f\left(\theta\right)$. In particular, consider the following scenario: at true state θ', we aim to implements $f\left(\theta'\right)$, but agent i with $\{i\} = \mathcal{I}^{\theta}$ reports θ with $f\left(\theta\right) \neq f\left(\theta'\right)$, while all the other agents report θ''—in this scenario, $f(\theta)$ "should" be chosen by Lemma 2. Since $f(\theta') \neq f(\theta)$, a whistle-blower must exist to block this false reporting, but who should this whistle-blower be? Recall Lemma 2, which says that we should ignore all the other agents' reports, when i is a dictator at θ and reports θ. As a result, the only possible whistle-blower must be agent i. To ensure that the whistle-blower i has a blocking plan, we need

$$\left[\begin{array}{c}\exists y \in Y \text{ such that} \\ u_i\left(f\left(\theta''\right), \theta''\right) \geq u_i\left(y, \theta''\right) \text{ and } u_i\left(y, \theta'\right) > u_i\left(f\left(\theta\right), \theta'\right)\end{array}\right]. \tag{24}$$

That is, this legitimate blocking plan for i must satisfy two conditions. First, given agents $-i$ reporting θ'', the blocking plan y must be credible,[16]

$$u_i\left(f\left(\theta''\right), \theta''\right) \geq u_i\left(y, \theta''\right). \tag{25}$$

Second, the blocking plan must be strictly profitable at θ', i.e.,

$$u_i\left(y, \theta'\right) > u_i\left(f\left(\theta\right), \theta'\right). \tag{26}$$

(25) and (26) imply (24), i.e., the dictator monotonicity in Definition 10 (precisely, (22)).

6.3.2 The necessity part of dictator monotonicity: proof

To prove the necessity of dictator monotonicity, we show a contrapositive statement of (22): for every for every $\left(i, \theta, \theta', \theta''\right) \in \mathcal{I} \times \Theta \times \Theta \times \Theta$,

$$\left[\begin{array}{c}\{i\} = \mathcal{I}^{\theta} \text{ and} \\ \left(\begin{array}{c}\forall y \in Y, \\ u_i\left(f\left(\theta''\right), \theta''\right) \geq u_i\left(y, \theta''\right) \Longrightarrow u_i\left(f\left(\theta\right), \theta'\right) \geq u_i\left(y, \theta'\right)\end{array}\right)\end{array}\right] \Longrightarrow f\left(\theta\right) = f\left(\theta'\right). \tag{27}$$

Suppose that f is rationalizably implemented by a mechanism $\mathcal{M} = \langle M, \ g : M \longrightarrow Y \rangle$, and that the left-hand-side of (27) holds. We aim to show $f(\theta) = f(\theta')$.

Pick any $\left(m_i, m_i''\right) \in S_i^{\mathcal{M}, \theta} \times S_i^{\mathcal{M}, \theta''}$, and there exists $\lambda_{-i}'' \in \triangle\left(S_{-i}^{\mathcal{M}, \theta''}\right)$ such that m_i'' is a best reply to λ_{-i}'' for i at θ'', i.e.,

$$u_i\left(g\left(m_i'', \lambda_{-i}''\right), \theta''\right) = u_i\left(f\left(\theta''\right), \theta''\right) \geq u_i\left(g\left(m_i, \lambda_{-i}''\right), \theta''\right), \forall m_i \in M_i. \tag{28}$$

[16] Agent i uses "y" to report that agents $-i$ are lying by reporting θ''. It is credible because if the true state were θ'', y would be inferior to $f\left(\theta''\right)$ for agent i.

By $\{i\} = \mathcal{I}^{\theta}$ and Lemma 2, we have

$$g\left(m_i, m_{-i}\right) = f\left(\theta\right) = g\left(m_i, \lambda''_{-i}\right), \forall m_{-i} \in M_{-i}, \tag{29}$$

which, together with (28) and the left-hand side (27), implies

$$u_i\left(g\left(m_i, \lambda''_{-i}\right), \theta'\right) = u_i\left(f\left(\theta\right), \theta'\right) \geq u_i\left(g\left(m_i, \lambda''_{-i}\right), \theta'\right), \forall m_i \in M_i. \tag{30}$$

For $\left(m_i, \lambda''_{-i}\right)$, (30) shows that i does not have a profitable deviation at θ', and (29) shows that agents $-i$ do not have a profitable deviation at θ'. Therefore, $\left(m_i, \lambda''_{-i}\right)$ is a Nash equilibrium at θ', which induces $f\left(\theta\right)$. Therefore, $f\left(\theta\right) = f\left(\theta'\right)$.■

6.4 The meaning of strict event monotonicity

6.4.1 The sufficiency part of strict event monotonicity: intuition

To illustrate strict event monotonicity, we consider an alternative and equivalent notion.

Definition 11 (strict iterated-elimination monotonicity) *An SCF $f : \Theta \longrightarrow Z$ satisfies strict iterated-elimination monotonicity if for every $\theta' \in \Theta$, there exists $\left(\theta^1, \theta^2, ..., \theta^n\right)$ such that*

$$\left\{\theta^1, \theta^2, ..., \theta^n\right\} = \Theta,$$
$$\theta^n = \theta',$$

and for every $k \in \{1, 2, ..., n-1\}$,

$$u_i\left(f\left(\theta^k\right), \theta^k\right) > u_i\left(y, \theta^k\right) \text{ and } u_i\left(y, \theta'\right) > u_i\left(f\left(\theta^k\right), \theta'\right), \text{for some } (y, i) \in Y \times \mathcal{I}^{\left\{\theta^k, \theta^{k+1}, ..., \theta^n\right\}}. \tag{31}$$

Proposition 1 *A responsive SCF $f : \Theta \longrightarrow Z$ satisfies strict event monotonicity if and only if f satisfies strict iterated-elimination monotonicity.*

The proof of Proposition 1 is relegated to Appendix A.4. To see the intuition of the "if" part of Theorem 2, we first recall the canonical mechanism in Sections 4.5. In this mechanism, we invite agents to report the true state, and there are three cases: (1) agreement, (2) unilateral deviation, and (3) multi-lateral deviation. A distinct feature of this mechanism is that agents do not have a best reply when Cases (2) and (3) are triggered. As a result, a strategy can be rationalized *only in Case (1)*. We now show that "truthful report" is the only rationalizable strategy in this mechanism. Suppose the true state is θ'. We start a iterative process of deletion with $\Theta = \left\{\theta^1, \theta^2, ..., \theta^n\right\}$. First, suppose all agents report θ^1. By (31) in Proposition 1, a whistle-blower $i \in \mathcal{I}^{\Theta}$ finds it strictly profitable to deviate to Case (2), and pick $y \in \mathcal{L}_i^{\circ}\left(f\left(\theta^1\right), \theta\right) \cap \mathcal{U}_i^{\circ}\left(f\left(\theta^1\right), \theta'\right)$. Thus, reporting θ^1 is not rationalizable for i, and as a result, θ^1 can be deleted from the rationalizable set of every agent in $\mathcal{I}^{\Theta}$.[17] Second, with $\Theta' = \left\{\theta^2, \theta^3, ..., \theta^n\right\}$, suppose all agents report θ^2. Similarly, by Proposition 1, reporting θ^2 is not rationalizable for some whistle-blower $i' \in \mathcal{I}^{\Theta'}$, and hence not rationalizable for all agents in $\mathcal{I}^{\Theta'}$....we continue this iterative process of deletion until we delete θ^{n-1}. As a result, only reporting $\theta^n = \theta'$ is rationalizable for agents in $\mathcal{I}^{\{\theta'\}}$, which induces $f\left(\theta'\right)$ at θ', i.e., we achieve rationalizable implementation.

6.4.2 The necessity part of strict event monotonicity: proof

Suppose that f is rationalizably implemented by a mechanism $\mathcal{M} = \langle M,\ g : M \longrightarrow Y\rangle$, and that (20) holds for some $(\theta', E) \in \Theta \times \left[2^{\Theta} \setminus \{\varnothing\}\right]$, i.e.,

$$u_i\left(f\left(\theta\right),\ \theta\right) > u_i\left(y,\ \theta\right) \implies u_i\left(f\left(\theta\right),\ \theta'\right) \geq u_i\left(y,\ \theta'\right), \forall\left(\theta, y, i,\right) \in E \times Y \times \mathcal{I}^{E}, \quad (32)$$

We aim to show $\left\{f\left(\theta'\right)\right\} = f\left(E\right)$, which establishes strict event monotonicity. Consider

$$S^{\mathcal{M},\,E} \equiv \left(S_i^{\mathcal{M},\,E} \equiv \left(\bigcup_{\theta \in E} S_i^{\mathcal{M},\,\theta}\right)\right)_{i \in \mathcal{I}}.$$

We will show that $S^{\mathcal{M},\,E}$ satisfies the best-reply property in $\mathcal{M}$ at state θ', which further implies $S^{\mathcal{M},\,E} \subset S^{\mathcal{M},\,\theta'}$, and hence, $\left\{f\left(\theta'\right)\right\} = f\left(E\right)$.

First, consider any $i \notin \mathcal{I}^{E}$, i.e., $i \notin \mathcal{I}^{\theta}$ for some $\theta \in E$. By Lemma 2, we have $S_i^{\mathcal{M},\,\theta} = M_i$, and hence,

$$M_i = S_i^{\mathcal{M},\,\theta} \subset S_i^{\mathcal{M},\,E} \subset M_i,$$

i.e., $S_i^{\mathcal{M},\,E} = M_i$. Pick any $\widehat{m}_{-i} \in S_{-i}^{\mathcal{M},\,\theta}$, Lemma 2 implies

$$g\left(m_i, \widehat{m}_{-i}\right) = f\left(\theta\right), \forall m_i \in M_i,$$

[17] Given θ^1 being not rationalizable for agent i, agents in $\mathcal{I}^{\Theta} \diagdown \{i\}$ can rationalize "reporting θ^1" only in Cases (2) and (3), in which a best reply does not exist.

i.e., every $m_i \in M_i = S_i^{\mathcal{M},E}$ is a best reply to $\widehat{m}_{-i} \in S_{-i}^{\mathcal{M},\theta} \subset S_{-i}^{\mathcal{M},E}$ for agent i at state θ'.

Second, consider any $i \in \mathcal{I}^E$. Pick any $\theta \in E$ and any $m_i \in S_i^{\mathcal{M},\theta}$, and we will show that m_i is a best reply for agent i at state θ' to some $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},E}\right)$, which would establish the best-reply property of $S^{\mathcal{M},E}$ at state θ'.

Since $m_i \in S_i^{\mathcal{M},\theta}$, there exists $\widetilde{\lambda}_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\theta}\right)$ such that m_i is a best reply to $\widetilde{\lambda}_{-i}$ for i at θ, i.e.,

$$u_i\left(g\left(m_i,\widetilde{\lambda}_{-i}\right),\theta\right) = u_i\left(f(\theta),\theta\right) \geq u_i\left(g\left(\overline{m}_i,\widetilde{\lambda}_{-i}\right),\theta\right), \forall \overline{m}_i \in M_i,$$

and more precisely,

$$g\left(m_i,\widetilde{\lambda}_{-i}\right) = f(\theta) = g\left(\overline{m}_i,\widetilde{\lambda}_{-i}\right), \forall \overline{m}_i \in S_i^{\mathcal{M},\theta}, \tag{33}$$

$$u_i\left(g\left(m_i,\widetilde{\lambda}_{-i}\right),\theta\right) = u_i\left(f(\theta),\theta\right) > u_i\left(g\left(\overline{m}_i,\widetilde{\lambda}_{-i}\right),\theta\right), \forall \overline{m}_i \in M_i \diagdown S_i^{\mathcal{M},\theta}. \tag{34}$$

Thus, (33) implies

$$u_i\left(g\left(m_i,\widetilde{\lambda}_{-i}\right),\theta'\right) = u_i\left(f(\theta),\theta'\right) = u_i\left(g\left(\overline{m}_i,\widetilde{\lambda}_{-i}\right),\theta'\right), \forall \overline{m}_i \in S_i^{\mathcal{M},\theta}, \tag{35}$$

and (32) and (34) imply

$$u_i\left(g\left(m_i,\widetilde{\lambda}_{-i}\right),\theta'\right) = u_i\left(f(\theta),\theta'\right) \geq u_i\left(g\left(\overline{m}_i,\widetilde{\lambda}_{-i}\right),\theta'\right), \forall \overline{m}_i \in M_i \diagdown S_i^{\mathcal{M},\theta}. \tag{36}$$

(35) and (36) imply m_i is a best reply to $\widetilde{\lambda}_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\theta}\right) \subset \triangle\left(S_{-i}^{\mathcal{M},E}\right)$ for i at θ'.■

7 A full characterization of rationalizable implementation

In this section, we combine the techniques developed in Sections 5 and 6, and fully characterize rationalizable implementation when neither NWA nor responsiveness is imposed.

From the analysis in Section 5, we learn that a necessary and sufficient condition requires existence of a partition $\mathcal{P}$ finer than $\mathcal{P}_f$ such that for any two states, θ and θ', with $\mathcal{P}(\theta) \neq \mathcal{P}(\theta')$, there must exist a whistle-blower i, who, at true state θ', can always block any false state $\widehat{\theta} \in \mathcal{P}(\theta)$ reported by agent $-i$. Furthermore, from the analysis in Section 6, we learn that three additional modification should be imposed: (1) we should compare $E\,(\subset \Theta)$ with θ' (rather than "θ Vs θ'"), and (2) the whistle-blower must be active at any state in E, and (3) dictator monotonicity holds.

Definition 12 (strict event monotonicity**) *A SCF $f : \Theta \longrightarrow Y$ satisfies strict event monotonicity** if there exists a partition $\mathcal{P}$ of Θ finer than $\mathcal{P}_f$ such that the following two conditions hold.*

1. *(strict event monotonicity) for every $(\theta', E) \in \Theta \times [2^{\Theta} \backslash \{\varnothing\}]$,*

$$\mathcal{P}(\theta') = \bigcup_{\theta \in E} \mathcal{P}(\theta) \Longleftarrow \left[\begin{array}{c} \forall (\theta, i) \in E \times \mathcal{I}^{[\cup_{\theta \in E} \mathcal{P}(\theta)]},\ \exists \widehat{\theta} \in \mathcal{P}(\theta), \forall y \in Y, \\ u_i\left(f(\theta), \widehat{\theta}\right) > u_i\left(y, \widehat{\theta}\right) \Longrightarrow u_i\left(f(\theta), \theta'\right) \geq u_i\left(y, \theta'\right) \end{array} \right],$$

2. *(dictator monotonicity) for every $(i, \theta, \theta', \theta'') \in \mathcal{I} \times \Theta \times \Theta \times \Theta$,*

$$\left[\begin{array}{c} \{i\} = \mathcal{I}^{\mathcal{P}(\theta)} \\ \text{and } \mathcal{P}(\theta) \neq \mathcal{P}(\theta') \end{array} \right] \Longrightarrow \left[\begin{array}{c} \exists y \in Y \text{ such that} \\ u_i\left(f(\theta''), \theta''\right) \geq u_i\left(y, \theta''\right) \text{ and } u_i\left(y, \theta'\right) > u_i\left(f(\theta), \theta'\right) \end{array} \right].$$

Two points are worth noting. First, strict event monotonicity** combines strict event monotonicity and dictator monotonicity. Without responsiveness, all axioms must be based on a *common* partition on Θ. Because of this, we have to write both strict event monotonicity and dictator monotonicity into one single axiom, which is based on a common partition on Θ. With abuse of notation, we call this new axiom, strict event monotonicity**.

Second, $\mathcal{P}(\theta)$ represents the set of states which are indistinguishable from θ (regarding players' rationalizable strategies in canonical mechanisms). Parts 1 and 2 of Definition 12 are simply the corresponding versions of Definitions 9 and 10, respectively, incorporating this idea of equivalent class of states (induced by the partition $\mathcal{P}$).

Theorem 3 *An SCF f is rationalizably implementable if and only if f satisfies strict event monotonicity**.*

The proof of Theorem 3 is similar to those of Theorems 1 and 2, and we omit it.

A Proofs

A.1 Proof of Lemma 1

Suppose that f is rationalizably implemented by a mechanism $\mathcal{M} = \langle M,\ g : M \longrightarrow Y \rangle$ and that $S^{\mathcal{M},\theta} \subset S^{\mathcal{M},\theta'}$. We aim to show $S^{\mathcal{M},\theta} = S^{\mathcal{M},\theta'}$. Clearly, $S^{\mathcal{M},\theta} \subset S^{\mathcal{M},\theta'}$ implies

$$f(\theta) = f(\theta'). \tag{37}$$

We will show $S^{\mathcal{M},\theta'}$ satisfies the best-reply property at state θ, which would imply $S^{\mathcal{M},\theta'} \subset S^{\mathcal{M},\theta}$, and hence, also $S^{\mathcal{M},\theta} = S^{\mathcal{M},\theta'}$. For any $i \in \mathcal{I}$, pick any $m_i \in S_i^{\mathcal{M},\theta}$, and there exists $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\theta}\right) \subset \triangle\left(S_{-i}^{\mathcal{M},\theta'}\right)$ such that

$$u_i(g(m_i,\lambda_{-i}),\theta) = u_i(f(\theta),\theta) \geq u_i(g(\widetilde{m}_i,\lambda_{-i}),\theta), \forall \widetilde{m}_i \in M_i. \tag{38}$$

Pick any $m_i' \in S_i^{\mathcal{M},\theta'}$. Then, $S^{\mathcal{M},\theta} \subset S^{\mathcal{M},\theta'}$ and (37) imply

$$u_i\left(g\left(m_i',\lambda_{-i}\right),\theta\right) = u_i\left(f\left(\theta'\right),\theta\right) = u_i\left(f(\theta),\theta\right). \tag{39}$$

Thus, (38) and (39) imply

$$u_i\left(g\left(m_i',\lambda_{-i}\right),\theta\right) = u_i(f(\theta),\theta) \geq u_i(g(\widetilde{m}_i,\lambda_{-i}),\theta), \forall \widetilde{m}_i \in M_i,$$

i.e., m_i' is a best reply to λ_{-i} for i at θ. And, $S^{\mathcal{M},\theta'}$ satisfies the best-reply property at θ.■

A.2 A useful lemma

Following a similar construction as in BMT, we get the following lemma.

Lemma 3 *There exist lotteries*

$$\underline{y} \in Y,$$
$$\{y_i^*(\theta) \in Y : (\theta,i) \in \Theta \times \mathcal{I}\},$$
$$\left\{z_i\left(\theta,\theta'\right) \in Y : \left(\theta,\theta'\right) \in \Theta \times \Theta \text{ and } i \in \mathcal{I}\right\},$$

such that

$$u_i(y_i^*(\theta),\theta) > u_i\left(\underline{y},\theta\right), \forall \theta \in \Theta, \forall i \in \mathcal{I}^{\theta}, \tag{40}$$

$$u_i\left(f\left(\theta'\right),\theta'\right) > u_i\left(z_i\left(\theta,\theta'\right),\theta'\right), \forall \left(\theta,\theta'\right) \in \Theta \times \Theta, \forall i \in \mathcal{I}^{\theta'}, \tag{41}$$

$$u_i\left(z_i\left(\theta,\theta'\right),\theta\right) > u_i\left(z_i\left(\theta',\theta'\right),\theta\right), \forall \left(\theta,\theta'\right) \in \Theta \times \Theta \text{ with } \theta \neq \theta', \forall i \in \mathcal{I}^{\theta}. \tag{42}$$

If NWA is imposed, i.e., $\mathcal{I}^{\Theta} = \mathcal{I}$, Lemma 3 reduces to Lemma 2 in BMT.

Proof of Lemma 3: We can find a set $\left\{\underline{y}_i(\theta) : (\theta, i) \in \Theta \times \mathcal{I}\right\}$ such that

$$u_i(f(\theta), \theta) > u_i\left(\underline{y}_i(\theta), \theta\right), \forall \theta \in \Theta, \forall i \in \mathcal{I}^{\theta}. \tag{43}$$

Define

$$\underline{y}_i \triangleq \frac{1}{|\Theta|} \sum_{\theta \in \Theta} \underline{y}_i(\theta) \text{ and}$$

$$y_i(\theta) \triangleq \frac{1}{|\Theta|} \sum_{\widehat{\theta} \in \Theta \diagdown \{\theta\}} \underline{y}_i\left(\widehat{\theta}\right) + \frac{1}{|\Theta|} f(\theta), \forall (\theta, i) \in \Theta \times \mathcal{I},$$

which, together with (43), imply

$$u_i(y_i(\theta), \theta) > u_i\left(\underline{y}_i, \theta\right), \forall \theta \in \Theta, \forall i \in \mathcal{I}^{\theta}. \tag{44}$$

Furthermore, define

$$\underline{y} \triangleq \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \underline{y}_i \text{ and}$$

$$y_i^*(\theta) \triangleq \frac{1}{|\mathcal{I}|} \sum_{j \in \mathcal{I} \diagdown \{i\}} \underline{y}_j + \frac{1}{|\mathcal{I}|} y_i(\theta), \forall (\theta, i) \in \Theta \times \mathcal{I},$$

which, together with (44), imply

$$u_i(y_i^*(\theta), \theta) > u_i\left(\underline{y}, \theta\right), \forall \theta \in \Theta, \forall i \in \mathcal{I}^{\theta},$$

which shows (40) in Lemma 3. With $\varepsilon > 0$, define

$$z_i(\theta', \theta') \triangleq (1-\varepsilon)\underline{y}_i(\theta') + \varepsilon \underline{y}_i, \forall (\theta', i) \in \Theta \times \mathcal{I} \text{ and} \tag{45}$$

$$z_i(\theta, \theta') \triangleq (1-\varepsilon)\underline{y}_i(\theta') + \frac{\varepsilon}{|\Theta|}\left(\sum_{\widehat{\theta} \in \Theta \diagdown \{\theta\}} \underline{y}_i\left(\widehat{\theta}\right) + f(\theta)\right), \forall (\theta, \theta', i) \in \Theta \times \Theta \times \mathcal{I} \text{ with } \theta \neq \theta'. \tag{46}$$

By (43), we have

$$u_i(f(\theta'), \theta') > u_i\left(\underline{y}_i(\theta'), \theta'\right), \forall \theta' \in \Theta, \forall i \in \mathcal{I}^{\theta'}, \tag{47}$$

$$u_i(f(\theta), \theta) > u_i\left(\underline{y}_i(\theta), \theta\right), \forall \theta \in \Theta, \forall i \in \mathcal{I}^{\theta}. \tag{48}$$

Then, (45), (46) and (48) imply

$$u_i(z_i(\theta, \theta'), \theta) > u_i(z_i(\theta', \theta'), \theta), \forall (\theta, \theta') \in \Theta \times \Theta \text{ with } \theta \neq \theta', \forall i \in \mathcal{I}^{\theta},$$

which establish (42) in Lemma 3.

Furthermore, by choosing sufficient small $\varepsilon > 0$, (45), (46) and (47) imply

$$u_i\left(f\left(\theta'\right),\theta'\right) > u_i\left(z_i\left(\theta,\theta'\right),\theta'\right), \forall\left(\theta,\theta'\right) \in \Theta \times \Theta, \forall i \in \mathcal{I}^{\theta'},$$

which establish (41) in Lemma 3.■

A.3 Proof: the "if" part of Theorem 1

Suppose that NWA and strict Maskin monotonicity** hold. Thus, there exists a partition $\mathcal{P}$ of Θ finer than $\mathcal{P}_f$ such that for any $\left(\theta,\theta'\right) \in \Theta \times \Theta$,

$$\theta' \notin \mathcal{P}(\theta) \text{ implies } \left[\begin{array}{c} \exists i \in \mathcal{I} \text{ such that } \forall \widehat{\theta} \in \mathcal{P}(\theta), \exists y^{\widehat{\theta}} \in Y, \\ u_i\left(f\left(\widehat{\theta}\right), \widehat{\theta}\right) > u_i\left(y^{\widehat{\theta}}, \widehat{\theta}\right) \text{ and } u_i\left(y^{\widehat{\theta}}, \theta'\right) > u_i\left(f\left(\widehat{\theta}\right), \theta'\right). \end{array}\right] \tag{49}$$

That is, $y^{\widehat{\theta}}$ in (49) is the blocking plan for agent i, when the true state is θ' and all other agents report $\widehat{\theta}$. Let $\mathcal{B}$ denote the finite set of all such $y^{\widehat{\theta}}$. Since $\mathcal{I} \times \Theta$ is finite, there exists a finite set $\Sigma \subset Y$ such that

$$\mathcal{B} \cup f(\Theta) \cup \left\{z_i\left(\theta,\theta'\right) \in Y : \left(\theta,\theta'\right) \in \Theta \times \Theta\right\} \subset \Sigma \text{ and}$$
$$\theta' \notin \mathcal{P}(\theta) \text{ implies } \left[\begin{array}{c} \exists i \in \mathcal{I} \text{ such that } \forall \widehat{\theta} \in \mathcal{P}(\theta), \exists y^{\widehat{\theta}} \in \Sigma, \\ u_i\left(f\left(\widehat{\theta}\right), \widehat{\theta}\right) > u_i\left(y^{\widehat{\theta}}, \widehat{\theta}\right) \text{ and } u_i\left(y^{\widehat{\theta}}, \theta'\right) > u_i\left(f\left(\widehat{\theta}\right), \theta'\right). \end{array}\right] \tag{50}$$

where $z_i\left(\theta,\theta'\right)$ are defined in Lemma 3. That is, Σ is a finite set which contains all of $f(\theta)$, $z_i\left(\theta,\theta'\right)$ and potential blocking plans for all possible profiles of $\left(\theta', \widehat{\theta}\right)$.—Our canonical mechanism is required to be a countable-action game, and hence, we should focus on Σ.

We use $\mathcal{M} = \left\langle M = (M_i)_{i\in\mathcal{I}},\ g : M \longrightarrow Y\right\rangle$ defined as follows to implement f. In this mechanism, each agent i sends a message $m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4\right] \in M_i$, where

$$\begin{aligned} m_i^1 &\in \Theta, \\ m_i^2 &\in \mathbb{N}, \\ m_i^3 &\in \Sigma^{\Theta}, \\ m_i^4 &\in Z. \end{aligned}$$

The innovation is that $m_i^3 \in \Sigma^\Theta$ is a state-contingent blocking plan. As usual, we partition M into three sets: agreement, unilateral deviation and multi-lateral deviation.

$$M' = \left\{ \left(m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4 \right] \right)_{i \in \mathcal{I}} \in M : \exists \theta \in \Theta, m_i^1 \in \mathcal{P}(\theta) \text{ and } m_i^2 = 1, \forall i \in \mathcal{I} \right\},$$

$$M'' = \left\{ m_i \in M \diagdown M' : \exists (\theta, i) \in \Theta \times \mathcal{I}, m_j^1 \in \mathcal{P}(\theta) \text{ and } m_j^2 = 1, \forall j \in \mathcal{I} \diagdown \{i\} \right\},$$

$$M''' = M \diagdown \left(M' \cup M'' \right).$$

Then, g is defined by the following rules.

Rule 1 (agreement): when $m \in M'$: there exists $\theta \in \Theta$, $m_i^1 \in \mathcal{P}(\theta)$ and $m_i^2 = 1$ for every $i \in \mathcal{I}$. In particular, $f(\theta)$ is unique, and g picks $f(\theta)$;

Rule 2 (unilateral deviation): when $m \in M''$: there exists $(\theta, i) \in \Theta \times \mathcal{I}$ such that $m_j^1 \in \mathcal{P}(\theta)$and $m_j^2 = 1$ for every $j \in \mathcal{I} \diagdown \{i\}$. In particular, such $f(\theta)$ is unique. For notational ease, consider agent $i+1$ (module I)[18] and set $\widehat{\theta} \equiv m_{i+1}^1$ — the interpretation is: even though agents $-i$ may report different states in $\mathcal{P}(\theta)$, we "hypothetically" regard they all reporting $\widehat{\theta}$, when agent i is the whistle-blower. We further distinguish two sub-cases:

Rule (2.a): if $u_i\left(f\left(\widehat{\theta}\right), \widehat{\theta}\right) \geq u_i\left(m_i^3\left(\widehat{\theta}\right), \widehat{\theta}\right)$, then g picks $m_i^3\left(\widehat{\theta}\right)$ with probability $1 - \frac{1}{m_i^2+1}$ and g picks $z_i\left(\widehat{\theta}, \widehat{\theta}\right)$ with probability $\frac{1}{m_i^2+1}$;

Rule (2.b): otherwise, g picks $z_i\left(\widehat{\theta}, \widehat{\theta}\right)$;

Rule 3 (multi-lateral deviation): when $m \in M'''$: consider agent $j = max\left[\arg\max_{h \in \mathcal{I}} m_h^2\right]$, i.e., j is the largest-numbered agent who reports the largest integer in the second dimension. Then, g picks m_j^4 with probability $1 - \frac{1}{m_j^2+1}$ and g picks $\underline{y}$ with probability $\frac{1}{m_j^2+1}$.

From now on, let us assume that the true state is θ^*, and we will show that reporting states in $\mathcal{P}(\theta^*)$ are the only rationalizable strategies for all agents. Before starting our proof, we first show that there exist best challenging schemes in Rules 2 and 3. Fix any

$$\widehat{m}_i^3(\theta) \in \arg \max_{y \in \{\widetilde{y} \in \Sigma : u_i(f(\theta), \theta) \geq u_i(\widetilde{y}, \theta)\}} u_i(y, \theta^*), \forall \theta \in \Theta, \tag{51}$$

$$\widehat{m}_i^3 \equiv \left[\widehat{m}_i^3(\theta)\right]_{\theta \in \Theta},$$

$$\widehat{m}_i^4 \in \arg\max_{z \in Z} u_i(z, \theta^*). \tag{52}$$

[18] That is, agent $(I+1)$ is agent 1.

That is, $\widehat{m}_i^3 \equiv \left[\widehat{m}_i^3(\theta)\right]_{\theta\in\Theta}$ and $\widehat{m}_i^4$ are the best options for agent i if Rules 2 and 3 are triggered, respectively. Specifically, we have

$$u_i\left(\widehat{m}_i^4, \theta^*\right) \geq u_i\left(y_i(\theta^*), \theta^*\right) > u_i\left(\underline{y}, \theta^*\right), \forall\theta\in\Theta, \quad (53)$$

$$u_i\left(\widehat{m}_i^4, \theta^*\right) \geq u_i\left(\widehat{m}_i^3(\theta), \theta^*\right) > u_i\left(z_i(\theta,\theta), \theta^*\right), \forall\theta\in\Theta, \quad (54)$$

where $y_i(\theta^*)$, $\underline{y}$ and $z_i(\theta,\theta)$ are defined in Lemma 3. In particular, the weak inequalities in (53) and (54) follow from (52), and the strict inequality in (53) from (40) in Lemma 3. To see strict inequality in (54), first suppose $\theta = \theta^*$, and we have

$$u_i\left(\widehat{m}_i^3(\theta^*), \theta^*\right) \geq u_i\left(f(\theta^*), \theta^*\right) > u_i\left(z_i(\theta^*, \theta^*), \theta^*\right), \quad (55)$$

where the weak inequality follows from $f(\theta^*) \in \{\widetilde{y}\in\Sigma : u_i(f(\theta^*), \theta^*) \geq u_i(\widetilde{y}, \theta^*)\}$ and the strict inequality follows from (41) in Lemma 3. Second, suppose $\theta \neq \theta^*$, and we have

$$u_i\left(\widehat{m}_i^3(\theta), \theta^*\right) \geq u_i\left(z_i(\theta^*, \theta), \theta^*\right) > u_i\left(z_i(\theta,\theta), \theta^*\right), \quad (56)$$

where the weak inequality follows from $z_i(\theta^*, \theta) \in \{\widetilde{y}\in\Sigma : u_i(f(\theta), \theta) \geq u_i(\widetilde{y}, \theta)\}$ due to (41) in Lemma 3 and the strict inequality follows from (42) in Lemma 3. Hence, (55) and (56) imply the strict inequality in (54).

When either Rule 2 or Rule 3 is triggered, the induced payoffs are listed as follows.

$$\text{Rule 2:} \quad \left[\begin{array}{c} \frac{(n-1)\times u_i(y,\theta^*)+u_i(z_i(\theta,\theta),\theta^*)}{n} \text{for some } (n,\theta,y)\in\mathbb{N}\times\Theta\times\Sigma \\ \text{such that } u_i(f(\theta),\theta)\geq u_i(y,\theta) \end{array}\right],$$

$$\text{Rule 3}: \quad \left[\frac{n\times u_i(z,\theta^*)+u_i\left(\underline{y},\theta^*\right)}{n+1} \text{ for some } (n,z)\in\mathbb{N}\times Z\right],$$

Then, (54) implies

$$u_i\left(\widehat{m}_i^4, \theta^*\right) \geq u_i\left(\widehat{m}_i^3(\theta), \theta^*\right) > \frac{(n-1)\times u_i(y,\theta^*)+u_i\left(z_i(\theta,\theta),\theta^*\right)}{n}, \quad (57)$$
$$\forall (n,\theta,y)\in\mathbb{N}\times\Theta\times\Sigma \text{ such that } u_i(f(\theta),\theta)\geq u_i(y,\theta),$$

and (53) implies

$$u_i\left(\widehat{m}_i^4, \theta^*\right) > \frac{n\times u_i(z,\theta^*)+u_i\left(\underline{y},\theta^*\right)}{n+1}, \quad (58)$$
$$\forall (n,z)\in\mathbb{N}\times Z,$$

i.e., $\widehat{m}_i^3$ and $\widehat{m}_i^4$ are strictly better than any induced payoffs in Rule 2 and 3, respectively.

In five steps, we now prove that $\mathcal{M}$ rationalizably implements f.

Step 1: at true state $\theta^* \in \Theta$, any $[m_i^1, m_i^2, m_i^3, m_i^4]_{i\in\mathcal{I}}$ with $(m_i^1, m_i^2) = (\theta^*, 1)$ for every $i \in I$ is a Nash equilibrium, which induces $f(\theta^*)$ as dictated by Rule 1.

For every $i \in \mathcal{I}$, any deviation of i would either stay in Rule 1 and induce the same outcome $f(\theta^*)$, or trigger Rule 2, which induces either $z_i(\theta^*, \theta^*)$ or a mixture of $z_i(\theta^*, \theta^*)$ and $m_i^3(\theta^*)$ with $u_i(f(\theta^*), \theta^*) \geq u_i(m_i^3(\theta^*), \theta^*)$. Clearly, $z_i(\theta^*, \theta^*)$ is worse than $f(\theta^*)$ by Lemma 3 (precisely, (41)). Therefore, any deviation of i is not a profitable.

Step 2: at true state $\theta^* \in \Theta$, for every $i \in \mathcal{I}$, if any $m_i \in M_i$ is a best reply to $\lambda_{-i} \in \triangle(M_{-i})$, then (m_i, λ_{-i}) induces Rules 2 and 3 with probability 0.

We prove this by contradiction. Suppose (m_i, λ_{-i}) induces Rules 2 or 3 with a positive probability. We thus partition M_{-i} as follows.

$$M_{-i} = \left(\bigcup_{\theta\in\Theta} M_{-i}^{\theta}\right) \cup \left(\bigcup_{(n,\theta,y)\in\mathbb{N}\times\Theta\times\Sigma} M_{-i}^{(n,\theta,y)}\right) \cup \left(\bigcup_{(n,z)\in\mathbb{N}\times Z} M_{-i}^{(n,z)}\right),$$

$$\begin{aligned}
\text{where } M_{-i}^{\theta} &\equiv \left\{ m_{-i} \in M_{-i} : \begin{array}{l} (m_i, m_{-i}) \text{ triggers Rule 1 and induces payoff} \\ u_i(f(\theta), \theta^*) \text{ for agent } i \end{array} \right\}, \\
\text{where } M_{-i}^{(n,\theta,y)} &\equiv \left\{ m_{-i} \in M_{-i} : \begin{array}{l} u_i(f(\theta), \theta) \geq u_i(y, \theta) \text{ and} \\ (m_i, m_{-i}) \text{ triggers Rule 2 and induces payoff} \\ \frac{(n-1)\times u_i(y,\theta^*) + u_i(z_i(\theta,\theta),\theta^*)}{n} \text{ for agent } i \end{array} \right\}, \\
\text{and } M_{-i}^{(n,z)} &\equiv \left\{ m_{-i} \in M_{-i} : \begin{array}{l} (m_i, m_{-i}) \text{ triggers Rule 3 and induces payoff} \\ \frac{n\times u_i(z,\theta^*) + u_i(\underline{y},\theta^*)}{n+1} \text{ for agent } i \end{array} \right\}.
\end{aligned}$$

Suppose agent i deviates from m_i to $[m_i^1, \widetilde{m}_i^2, \widehat{m}_i^3, \widehat{m}_i^4]$ with $\widetilde{m}_i^2 \to \infty$, where $\widehat{m}_i^3$ and $\widehat{m}_i^4$ are defined in (51) and (52). The payoff changes are listed in the following table.

as $\widetilde{m}_i^2 \to \infty$	payoff under m_i		the supremum of payoffs under $\left[m_i^1, \widetilde{m}_i^2, \widehat{m}_i^3, \widehat{m}_i^4\right]$
$m_{-i} \in M_{-i}^{\theta}$	$u_i\left(f\left(\theta\right), \theta^*\right)$	$\leq$	$u_i\left(\widehat{m}_i^3\left(\theta\right), \theta^*\right)$
$m_{-i} \in M_{-i}^{(n,\theta,y)} \neq \varnothing$	$\frac{(n-1)\times u_i(y,\theta^*)+u_i(z_i(\theta,\theta),\theta^*)}{n}$	$<$	$u_i\left(\widehat{m}_i^3\left(\theta\right), \theta^*\right)$
$m_{-i} \in M_{-i}^{(n,z)}$	$\frac{n\times u_i(y,\theta^*)+u_i\left(\underline{y},\theta^*\right)}{n+1}$	$<$	$u_i\left(\widehat{m}_i^4, \theta^*\right)$

where the strict inequality follows from (57) and (58). Hence, for any mixed strategy $\lambda_{-i} \in \triangle\left(M_{-i}\right)$ which induces Rules 2 and 3 with a positive probability, we have

$$u_i\left(g\left(\left[m_i^1, m_i^2, m_i^3, m_i^4\right], \lambda_{-i}\right), \theta^*\right) < \lim_{\widetilde{m}_i^2 \to \infty} u_i\left(g\left(\left[m_i^1, \widetilde{m}_i^2, \widehat{m}_i^3, \widehat{m}_i^4\right], \lambda_{-i}\right), \theta^*\right)$$

and as a result, there exists $\widehat{m}_i^2 \in \mathbb{N}$ such that

$$u_i\left(g\left(\left[m_i^1, m_i^2, m_i^3, m_i^4\right], \lambda_{-i}\right), \theta^*\right) < u_i\left(g\left(\left[m_i^1, \widehat{m}_i^2, \widehat{m}_i^3, \widehat{m}_i^4\right], \lambda_{-i}\right), \theta^*\right),$$

contradicting m_i being a best reply to λ_{-i}.

Step 3: at true state $\theta^* \in \Theta$, for every $i \in \mathcal{I}$, any strategy $m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4\right]$ with $m_i^2 > 1$ is not rationalizable for agent i.

This follows from Step 2, because $m_i^2 > 1$ induces rules 2 or 3 with probability 1.

Step 4: at true state $\theta^* \in \Theta$, for any $\overline{\theta} \in \Theta \diagdown \mathcal{P}\left(\theta^*\right)$ (or equivalently, $\theta^* \notin \mathcal{P}\left(\overline{\theta}\right)$), there exists $j \in \mathcal{I}$ such that any strategy $m_j = \left[m_j^1, m_j^2, m_j^3, m_j^4\right]$ with $m_j^1 \in \mathcal{P}\left(\overline{\theta}\right)$ is not rationalizable for agent j.

By our construction of Σ above and (50), we have

$$\forall \widetilde{\theta} \in \mathcal{P}\left(\overline{\theta}\right), \exists y^{\widetilde{\theta}} \in \Sigma,$$
$$u_j\left(f\left(\widetilde{\theta}\right), \widetilde{\theta}\right) > u_j\left(y^{\widetilde{\theta}}, \widetilde{\theta}\right) \text{ and } u_j\left(y^{\widetilde{\theta}}, \theta^*\right) > u_j\left(f\left(\widetilde{\theta}\right), \theta^*\right).$$

and furthermore, by our definition of $\widehat{m}_j^3\left(\theta\right)$ above, we have

$$\forall \widetilde{\theta} \in \mathcal{P}\left(\overline{\theta}\right),$$
$$u_j\left(f\left(\widetilde{\theta}\right), \widetilde{\theta}\right) \geq u_j\left(\widehat{m}_j^3\left(\widetilde{\theta}\right), \widetilde{\theta}\right) \text{ and } u_j\left(\widehat{m}_j^3\left(\widetilde{\theta}\right), \theta^*\right) > u_j\left(f\left(\widetilde{\theta}\right), \theta^*\right). \tag{59}$$

For any $m_j = \left[m_j^1, m_j^2, m_j^3, m_j^4\right]$ with $m_j^1 \in \mathcal{P}\left(\overline{\theta}\right)$, we prove by contradiction that it is not rationalizable for j. Suppose otherwise, i.e., m_j is a best reply to some $\lambda_{-j} \in \triangle\left(S_{-j}^{\mathcal{M},\theta^*}\right)$. By Step 2, $\left(m_j, \lambda_{-j}\right)$ must induce Rule 1 with probability 1. Thus, every agent i report $m_i^1 \in \mathcal{P}\left(\overline{\theta}\right)$ and $m_i^2 = 1$, which induces $f\left(\overline{\theta}\right)$. As a result, agent j's payoff is $u_j\left(f\left(\overline{\theta}\right), \theta^*\right)$. Then, agent j would like to deviate from m_j to $\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right]$ with $\widetilde{m}_j^2 \to \infty$, which would induce Rule 2 and

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\overline{\theta}\right), \theta^*\right) < \lim_{\widetilde{m}_j^2 \to \infty} u_i\left(g\left(\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

where the inequality follows from (59). Hence, there exists $\widehat{m}_j^2 \in \mathbb{N}$ such that

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\overline{\theta}\right), \theta^*\right) < u_i\left(g\left(\left[m_j^1, \widehat{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

contradicting m_j being a best reply to λ_{-j}.

Step 5: at true state $\theta^* \in \Theta$, for any $\left(i, \overline{\theta}\right) \in \mathcal{I} \times \left[\Theta \diagdown \mathcal{P}\left(\theta^*\right)\right]$, any $m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4\right]$ with $m_i^1 \in \mathcal{P}\left(\overline{\theta}\right)$ is not rationalizable for agent i.

First, this is true for the agent j identified in Step 4. Second, consider any $i \neq j$. We prove it by contradiction. Suppose $m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4\right]$ with $m_i^1 \in P\left(\overline{\theta}\right)$ is rationalizable for agent i. Then, m_i is a best reply to some rationalizable conjecture $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M},\theta^*}\right)$. By Step 2, $\left(m_i, \lambda_{-i}\right)$ must induce Rule 1 with probability 1, or equivalently, with probability 1, every agent $h \in \mathcal{I}$ reports $m_h^1 \in \mathcal{P}\left(\overline{\theta}\right)$, including agent j, contradicting Step 4.

To sum, Step 1 shows

$$S^{\mathcal{M},\theta^*} \supset \prod_{i \in \mathcal{I}} \left\{\left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in M_i : \left(m_i^1, m_i^2\right) = \left(\theta^*, 1\right)\right\},$$

and Steps 2-5 show

$$S^{\mathcal{M},\theta^*} \subset \prod_{i \in \mathcal{I}} \left\{\left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in M_i : m_i^1 \in \mathcal{P}\left(\theta^*\right) \text{ and } m_i^2 = 1\right\}.$$

Thus, every $m \in S^{\mathcal{M},\theta^*}$ triggers Rule 1 and induces $f\left(\theta^*\right)$, i.e., $g\left(S^{\mathcal{M},\theta^*}\right) = \left\{f\left(\theta^*\right)\right\}$.■

A.4 Proofs of Proposition 1

Consider any responsive SCF $f : \Theta \longrightarrow Z$. First, suppose strict event monotonicity, and we show strict iterated-elimination monotonicity. Fix any $\theta' \in \Theta$, and we will define a sequence $\left(\theta^1,\theta^2,...,\theta^n\right)$ inductively. Define $\theta^n=\theta'$, and apply strict event monotonicity on $E=\Theta$. Given responsiveness, we have $\{f(\theta')\}\neq f(E)$ and hence, strict event monotonicity implies

$$u_i\left(f\left(\theta^1\right),\,\theta^1\right)>u_i\left(y,\,\theta^1\right)\text{ and }u_i\left(y,\,\theta'\right)>u_i\left(f\left(\theta^1\right),\,\theta'\right),$$
$$\text{for some }\left(\theta^1,y,i\right)\in\Theta\times Y\times\mathcal{I}^{\Theta}.$$

Inductively, for each $k\in\{2,...,n-1\}$, apply strict group monotonicity on

$$E=\Theta\diagdown\left\{\theta^1,...,\theta^{k-1}\right\},$$

and we get

$$u_i\left(f\left(\theta^k\right),\,\theta^k\right)>u_i\left(y,\,\theta^k\right)\text{ and }u_i\left(y,\,\theta'\right)>u_i\left(f\left(\theta^k\right),\,\theta'\right),$$
$$\text{for some }\left(\theta^k,y,i\right)\in\left[\Theta\diagdown\left\{\theta^1,...,\theta^{k-1}\right\}\right]\times Y\times\mathcal{I}^{\Theta\diagdown\left\{\theta^1,...,\theta^{k-1}\right\}},$$

i.e., strict iterated-elimination monotonicity holds.

Second, suppose strict iterated-elimination monotonicity, and we show strict event monotonicity. For any (θ',E) with $\{f(\theta')\}\neq f(E)$, and we aim to show

$$u_i\left(f(\theta),\,\theta\right)>u_i\left(y,\,\theta\right)\text{ and }u_i\left(y,\,\theta\right)>u_i\left(f(\theta),\,\theta'\right),\text{ for some }(\theta,y,i,)\in E\times Y\times\mathcal{I}^E. \tag{60}$$

Given strict iterated-elimination monotonicity, there exists $\left(\theta^1,\theta^2,...,\theta^n\right)$ such that

$$\left\{\theta^1,\theta^2,...,\theta^n\right\}=\Theta,$$
$$\theta^n=\theta',$$

and for every $k\in\{1,2,...,n-1\}$,

$$u_i\left(f\left(\theta^k\right),\,\theta^k\right)>u_i\left(y,\,\theta^k\right)\text{ and }u_i\left(y,\,\theta'\right)>u_i\left(f\left(\theta^k\right),\,\theta'\right),\text{ for some }(y,i)\in Y\times\mathcal{I}^{\left\{\theta^k,\theta^{k+1},...,\theta^n\right\}}. \tag{61}$$

We write $E=\left\{\theta^{k_1},...\theta^{k_n}\right\}\subset\Theta$ with $k_1<...<k_n$ and $k_1<n$ due to $\{f(\theta')\}\neq f(E)$. Then, (61) implies

$$u_i\left(f\left(\theta^{k_1}\right),\,\theta^{k_1}\right)>u_i\left(y,\,\theta^{k_1}\right)\text{ and }u_i\left(y,\,\theta'\right)>u_i\left(f\left(\theta^{k_1}\right),\,\theta'\right), \tag{62}$$
$$\text{for some }(y,i)\in Y\times\mathcal{I}^{\left\{\theta^{k_1},\theta^{k_1+1},...,\theta^n\right\}}.$$

Note that $\theta^{k_1}\in E$ and $E=\left\{\theta^{k_1},...\theta^{k_n}\right\}\subset\left\{\theta^{k_1},\theta^{k_1+1},...,\theta^n\right\}$ and hence, $\mathcal{I}^{\left\{\theta^{k_1},\theta^{k_1+1},...,\theta^n\right\}}\subset\mathcal{I}^E$. As a result, (62) implies (60).■

A.5 Proof: the "if" part of Theorem 2

Suppose that f satisfies responsiveness, strict event monotonicity and dictator monotonicity. As argued above, since $\mathcal{I} \times \Theta$ is finite, there exists a finite set $\Sigma \subset Y$ such that

$$f(\Theta) \cup \{z_i(\theta, \theta') \in Y : (\theta, \theta') \in \Theta \times \Theta\} \subset \Sigma,$$

$$\text{and } \forall (\theta', E) \in \Theta \times \left[2^{\Theta} \setminus \{\varnothing\}\right], \tag{63}$$

$$\{f(\theta')\} \neq f(E) \text{ implies } \left[\begin{array}{c} u_i(f(\theta), \theta) > u_i(y, \theta) \text{ and } u_i(y, \theta') > u_i(f(\theta), \theta'), \\ \text{for some } (\theta, y, i,) \in E \times \Sigma \times \mathcal{I}^E, \end{array}\right],$$

$$\text{and } \forall (i, \theta, \theta', \theta'') \in \mathcal{I} \times \Theta \times \Theta \times \Theta,$$

$$\left[\begin{array}{c} \{i\} = \mathcal{I}^{\theta} \\ \text{and } f(\theta) \neq f(\theta') \end{array}\right] \Longrightarrow \left[\begin{array}{c} \exists y \in \Sigma \text{ such that} \\ u_i(f(\theta''), \theta'') \geq u_i(y, \theta'') \text{ and } u_i(y, \theta') > u_i(f(\theta), \theta') \end{array}\right], \tag{64}$$

$$\left[\begin{array}{c} \{i\} = \mathcal{I}^{\theta} \\ \text{and } f(\theta) \neq f(\theta') \end{array}\right] \Longrightarrow u_i(f(\theta'), \theta') > u_i(f(\theta), \theta'), \tag{65}$$

where $z_i(\theta, \theta')$ are defined in Lemma 3, and (63) follows from strict event monotonicity, and (64) from dictator monotonicity, and (65) from (64) by considering $\theta' = \theta''$. That is, Σ is a finite set which contains all of $f(\theta)$, $z_i(\theta, \theta')$ and all potential blocking plans. In particular, when we take $E = \Theta$, we get $\{f(\theta')\} \neq f(E)$, and (63) implies

$$\mathcal{I}^{\Theta} \neq \varnothing. \tag{66}$$

We use $\mathcal{M} = \langle M = (M_i)_{i \in \mathcal{I}}, \, g : M \longrightarrow Y \rangle$ defined as follows to implement f. In this mechanism, each agent i sends a message $m_i = [m_i^1, m_i^2, m_i^3, m_i^4] \in M_i$, where

$$\begin{aligned} m_i^1 &\in \Theta, \\ m_i^2 &\in \mathbb{N}, \\ m_i^3 &\in \Sigma^{\Theta}, \\ m_i^4 &\in Z. \end{aligned}$$

We partition M as follows: agreement, unilateral deviation and multi-lateral deviation.

$$M' = \left\{ \left(m_i = \left[m_i^1, m_i^2, m_i^3, m_i^4\right]\right)_{i \in \mathcal{I}} \in M : \exists \theta \in \Theta, \, \left(m_i^1, m_i^2\right) = (\theta, 1), \forall i \in \mathcal{I}^{\theta} \right\},$$

$$M'' = \left\{ m_i \in M \diagdown M' : \exists (\theta, i) \in \Theta \times \mathcal{I}, \left(m_j^1, m_j^2 \right) = (\theta, 1), \forall j \in \mathcal{I} \diagdown \{i\} \right\},$$
$$M''' = M \diagdown \left(M' \cup M'' \right).$$

It is worth noting that $\mathcal{I}^\theta$ is used in the definition of M', and $\mathcal{I}$ is used in the definition of M''. Specifically, agreement is defined as all agents in $\mathcal{I}^\theta$ reporting $(\theta, 1)$ in the first two dimensions. Furthermore, unilateral deviation refers to the unilateral deviation from all agents in $\mathcal{I}$ reporting the same $(\theta, 1)$ for some θ. Thus, a unilateral deviation from a message profile in M' may induce a message profile in M'''.

Then, g is defined by the following rules.

Rule 1 (agreement): when $m \in M'$: there exists $\theta \in \Theta$ such that $\left(m_i^1, m_i^2 \right) = (\theta, 1)$ for every $i \in \mathcal{I}^\theta$. By (66), such θ is unique. Then, g picks $f(\theta)$;

Rule 2 (unilateral deviation): when $m \in M''$: there exists $(\theta, i) \in \Theta \times \mathcal{I}$ such that $\left(m_j^1, m_j^2 \right) = (\theta, 1)$ for every $j \in \mathcal{I} \diagdown \{i\}$, and such (θ, i) is unique due to $|\mathcal{I}| \geq 3$. We further distinguish two sub-cases:

Rule (2.a): if $u_i (f(\theta), \theta) \geq u_i \left(m_i^3 (\theta), \theta \right)$, then g picks $m_i^3 (\theta)$ with probability $1 - \frac{1}{m_i^2 + 1}$ and g picks $z_i (\theta, \theta)$ with probability $\frac{1}{m_i^2 + 1}$;

Rule (2.b): if $u_i (f(\theta), \theta) < u_i \left(m_i^3 (\theta), \theta \right)$, then g picks $z_i (\theta, \theta)$;

Rule 3 (multi-lateral deviation): when $m \in M'''$: consider agent $j = max \left[\arg\max_{h \in \mathcal{I}} m_h^2 \right]$, i.e., j is the largest-numbered agent who report the largest integer in the second dimension. Then, g picks m_j^4 with probability $1 - \frac{1}{m_j^2 + 1}$ and g picks $\underline{y}$ with probability $\frac{1}{m_j^2 + 1}$.

From now on, fix any true state is $\theta^* \in \Theta$. As above, fix any

$$\widehat{m}_i^3 (\theta) \in \arg \max_{y \in \{\widetilde{y} \in \Sigma \,:\, u_i(f(\theta), \theta) \geq u_i(\widetilde{y}, \theta)\}} u_i (y, \theta^*), \forall \theta \in \Theta,$$
$$\widehat{m}_i^3 \equiv \left[\widehat{m}_i^3 (\theta) \right]_{\theta \in \Theta},$$
$$\widehat{m}_i^4 \in \arg\max_{z \in Z} u_i (z, \theta^*).$$

Using the same argument as in Appendix A.3, we can show

$$u_i \left(\widehat{m}_i^4, \theta^* \right) \geq u_i \left(\widehat{m}_i^3 (\theta), \theta^* \right) > \frac{(n-1) \times u_i (y, \theta^*) + u_i (z_i (\theta, \theta), \theta^*)}{n}, \tag{67}$$
$$\forall i \in \mathcal{I}^{\theta^*}, \forall (n, \theta, y) \in \mathbb{N} \times \Theta \times \Sigma \text{ such that } u_i (f(\theta), \theta) \geq u_i (y, \theta),$$

$$\text{and } u_i\left(\widehat{m}_i^4,\theta^*\right) > \frac{n\times u_i\left(z,\theta^*\right)+u_i\left(\underline{y},\theta^*\right)}{n+1}, \tag{68}$$
$$\forall i\in\mathcal{I}^{\theta^*}, \forall\left(n,z\right)\in\mathbb{N}\times Z.$$

In five steps, we now prove that $\mathcal{M}$ rationalizably implements f.

Step 1: at true state $\theta^* \in \Theta$, any $\left[m_i^1, m_i^2, m_i^3, m_i^4\right]_{i\in\mathcal{I}}$ with $\left(m_i^1, m_i^2\right) = \left(\theta^*, 1\right)$ for every $i \in I$ is a Nash equilibrium, which induces $f\left(\theta^*\right)$ as dictated by Rule 1.

By following this strategy profile, agent i gets payoff $u_i\left(f\left(\theta^*\right), \theta^*\right)$. For any $i \notin \mathcal{I}^{\theta^*}$, any deviation of i would still induce $f\left(\theta^*\right)$, i.e., not a profitable deviation. For any $i \in \mathcal{I}^{\theta^*}$, consider any of i's deviation $\left(\widehat{m}_i^1, \widehat{m}_i^2, \widehat{m}_i^3, \widehat{m}_i^4\right)$ that changes the outcome chosen by g. This deviation would either trigger Rule 1, when $\left(\widehat{m}_i^1, \widehat{m}_i^2\right) = \left(\widehat{\theta}, 1\right)$ with $\widehat{\theta} \neq \theta^*$ and $\{i\} = \mathcal{I}^{\widehat{\theta}}$, or trigger Rule 2 otherwise. For the former case, this is not profitable because of dictator monotonicity (precisely, (65)). In the latter case, g picks either $z_i\left(\theta^*, \theta^*\right)$, or a mixture of $z_i\left(\theta^*, \theta^*\right)$ and $\widehat{m}_i^3\left(\theta^*\right)$ with $u_i\left(f\left(\theta^*\right), \theta^*\right) \geq u_i\left(m_i^3\left(\theta^*\right), \theta^*\right)$, all of which are worse than $f\left(\theta^*\right)$ for i at θ^* by (41) in Lemma 3, i.e., not a profitable deviation.

Step 2: at true state $\theta^* \in \Theta$, for every $i \in \mathcal{I}^{\theta^*}$, if any $m_i \in M_i$ is a best reply to $\lambda_{-i} \in \triangle\left(M_{-i}\right)$, then $\left(m_i, \lambda_{-i}\right)$ induces Rules 2 and 3 with probability 0.

The proof is the same as Step 2 in Appendix A.3, and we omit it.

Step 3: at true state $\theta^* \in \Theta$, for any $\overline{\theta} \in \Theta \diagdown \{\theta^*\}$, there exists $j \in \mathcal{I}^{\overline{\theta}} \cap \mathcal{I}^{\theta^*}$ such that any $m_j = \left[m_j^1, m_j^2, m_j^3, m_j^4\right]$ with $\left(m_j^1, m_j^2\right) = \left(\overline{\theta}, 1\right)$ is not rationalizable for agent j.

By Proposition 1, strict event monotonicity is equivalent to strict iterated-elimination monotonicity, which means that there exists $\left(\theta^1, \theta^2, ..., \theta^n\right)$ such that

$$\left\{\theta^1, \theta^2, ..., \theta^n\right\} = \Theta,$$
$$\theta^n = \theta^*,$$

and for every $k \in \{1, 2, ..., n-1\}$,

$$u_j\left(f\left(\theta^k\right), \theta^k\right) > u_j\left(y, \theta^k\right) \text{ and } u_j\left(y, \theta^*\right) > u_j\left(f\left(\theta^k\right), \theta^*\right), \text{ for some } (y, j) \in \Sigma \times \mathcal{I}^{\left\{\theta^k, \theta^{k+1}, ..., \theta^n\right\}}. \tag{69}$$

Then, inductively, for each $k \in \{1, 2, ..., n-1\}$, we will show that it is not rationalizable for agent j (identified in (69)) to report $\left(\theta^k, 1\right)$ in the first two dimensions.

Clearly, $j \in \mathcal{I}^{\left\{\theta^k, \theta^{k+1}, \ldots, \theta^n\right\}} \subset \mathcal{I}^{\theta^n} = \mathcal{I}^{\theta^*}$. Furthermore, (69) immediately implies

$$u_j\left(\widehat{m}_j^4,\, \theta^*\right) \geq u_j\left(\widehat{m}_j^3\left(\theta^k\right),\, \theta^*\right) \geq u_j\left(y,\, \theta^*\right) > u_j\left(f\left(\theta^k\right),\, \theta^*\right), \tag{70}$$

where the strict inequality follows from (69), the first weak inequality from (67), the second weak inequality from the definition of $\widehat{m}_j^3\left(\theta^k\right)$ and $y \in \left\{\widetilde{y} \in \Sigma : u_i\left(f\left(\theta^k\right),\, \theta^k\right) \geq u_i\left(\widetilde{y}, \theta^k\right)\right\}$.

We now consider two cases: (a) $\{j\} = \mathcal{I}^{\theta^k}$, (b) $\{j\} \neq \mathcal{I}^{\theta^k}$.

Step 3.a: when $\{j\} = \mathcal{I}^{\theta^k}$, any $m_j = \left(m_j^1, m_j^2, m_j^3, m_j^4\right)$ with $\left(m_j^1, m_j^2\right) = \left(\theta^k, 1\right)$ is not rationalizable for j.

We prove it by contradiction. Given $\{j\} = \mathcal{I}^{\theta^k}$, suppose $m_j = \left(m_j^1, m_j^2, m_j^3, m_j^4\right)$ with $\left(m_j^1, m_j^2\right) = \left(\theta^k, 1\right)$ is rationalizable for agent j. Then, m_j is a best reply to some $\lambda_{-j} \in \triangle\left(S_{-j}^{\mathcal{M},\,\theta^*}\right)$. Since $\{j\} = \mathcal{I}^{\theta^k}$, we reach agreement (i.e., Rule 1) under the strategy profile $\left(m_j, \lambda_{-j}\right)$, which induces the outcome $f\left(\theta^k\right)$.

Given $\theta^k \neq \theta^*$ and responsiveness, dictator monotonicity (i.e., (64)) implies

$$\forall \theta'' \in \Theta,\ \exists y \in \Sigma \text{ such that}$$
$$u_i\left(f\left(\theta''\right),\, \theta''\right) \geq u_i\left(y,\, \theta''\right) \text{ and } u_i\left(y,\, \theta^*\right) > u_i\left(f\left(\theta^k\right),\, \theta^*\right),$$

which further implies

$$u_j\left(\widehat{m}_j^3\left(\theta''\right),\, \theta^*\right) > u_j\left(f\left(\theta^k\right),\, \theta^*\right), \forall \theta'' \in \Theta, \tag{71}$$

i.e., agent j always finds it profitable to use $\widehat{m}_j^3$ to deviate to Rule 2. Also, (70) implies

$$u_j\left(\widehat{m}_j^4,\, \theta^*\right) > u_j\left(f\left(\theta^k\right),\, \theta^*\right), \tag{72}$$

i.e., whenever possible, agent j always finds it profitable to use the blocking plan $\widehat{m}_j^4$ to deviate to Rule 3. Then, agent j would like to deviate from m_j to $\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right]$ with $\widetilde{m}_j^2 \to \infty$, which would induce either Rule 2 or Rule 3 and

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\theta^k\right),\, \theta^*\right) < \lim_{\widetilde{m}_j^2 \to \infty} u_i\left(g\left(\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

where the inequality follows from (71) and (72). Thus, there exists $\widehat{m}_j^2 \in \mathbb{N}$ such that

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\theta^k\right),\, \theta^*\right) < u_i\left(g\left(\left[m_j^1, \widehat{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

contradicting m_j being a best reply to λ_{-j}.

Step 3.b: when $\{j\} \neq \mathcal{I}^{\theta^k}$, any strategy $m_j = \left(m_j^1, m_j^2, m_j^3, m_j^4\right)$ with $\left(m_j^1, m_j^2\right) = \left(\theta^k, 1\right)$ is not rationalizable for agent j.

We prove it by contradiction. Suppose $m_j = \left(m_j^1, m_j^2, m_j^3, m_j^4\right)$ with $\left(m_j^1, m_j^2\right) = \left(\theta^k, 1\right)$ is rationalizable for agent j. Then, m_j is a best reply to some $\lambda_{-j} \in \triangle\left(S_{-j}^{\mathcal{M}, \theta^*}\right)$.

Recall $j \in \mathcal{I}^{\left\{\theta^k, \theta^{k+1}, ..., \theta^l\right\}} \subset \mathcal{I}^{\theta^k} \cap \mathcal{I}^{\theta^*}$. By Step 2, the strategy profile (m_j, λ_{-j}) induces Rule 1 with probability 1. Given $\left(m_j^1, m_j^2\right) = \left(\theta^k, 1\right)$, all agents in $\mathcal{I}^{\theta^k}$ must report $\left(\theta^k, 1\right)$ under (m_j, λ_{-j}).[19] With $\mathcal{I}^{\theta^k} \diagdown \{j\} \neq \varnothing$, we consider two sub-cases (1) agents $-j$ all report $\left(\theta^k, 1\right)$ and (2) otherwise. In Case (1) agent j can deviate to Rule 2 and in Case (2) agent j can deviate to Rule 3. Note that (70) implies

$$u_j\left(\widehat{m}_j^3\left(\theta^k\right), \theta^*\right) > u_j\left(f\left(\theta^k\right), \theta^*\right), \tag{73}$$
$$u_j\left(\widehat{m}_j^4, \theta^*\right) > u_j\left(f\left(\theta^k\right), \theta^*\right), \tag{74}$$

i.e., (73) says that, in Case (1), agent j always finds it profitable to use the blocking plan $\widehat{m}_j^3$ to deviate to Rule 2; (74) says that, in Case (2), agent j always finds it profitable to use the blocking plan $\widehat{m}_j^4$ to deviate to Rule 3. Thus, agent j would like to deviate from m_j to $\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right]$ with $\widetilde{m}_j^2 \to \infty$, which would induce either Rule 2 or Rule 3 and

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\theta^k\right), \theta^*\right) < \lim_{\widetilde{m}_j^2 \to \infty} u_i\left(g\left(\left[m_j^1, \widetilde{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

where the inequality follows from (73) and (74). Thus, there exists $\widehat{m}_j^2 \in \mathbb{N}$ such that

$$u_j\left(g\left(m_j, \lambda_{-j}\right), \theta^*\right) = u_j\left(f\left(\theta^k\right), \theta^*\right) < u_i\left(g\left(\left[m_j^1, \widehat{m}_j^2, \widehat{m}_j^3, \widehat{m}_j^4\right], \lambda_{-j}\right), \theta^*\right)$$

contradicting m_j being a best reply to λ_{-j}.

Step 4: at true state $\theta^* \in \Theta$, for every $i \in \mathcal{I}^{\Theta}$, any strategy $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right)$ with $\left(m_i^1, m_i^2\right) \neq \left(\theta^*, 1\right)$ is not rationalizable for agent i.

For any $i \in \mathcal{I}^{\Theta}$, pick any $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in S_i^{\mathcal{M}, \theta^*}$. Then, m_i is a best reply to some $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M}, \theta^*}\right)$. By Step 2, (m_i, λ_{-i}) induces Rule 1 with probability 1. However,

[19] By the induction hypothesis, we cannot reach agreement on any state in $\left\{\theta^1, \theta^2, ..., \theta^{k-1}\right\}$.

by Step 3, any agreement on $\overline{\theta} \in \Theta \diagdown \{\theta^*\}$ cannot be reached. Thus, the only possible is agreement on θ^*, i.e. all agents in $\mathcal{I}^{\theta^*}$ report $(\theta^*, 1)$. As a result, only $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right)$ with $\left(m_i^1, m_i^2\right) = (\theta^*, 1)$ is rationalizable for $i \in \mathcal{I}^{\Theta}$ at θ^*.

Step 5: at true state $\theta^* \in \Theta$, for every $i \in \mathcal{I}^{\theta^*}$, any strategy $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right)$ with $\left(m_i^1, m_i^2\right) \neq (\theta^*, 1)$ is not rationalizable for agent i.

For any $i \in \mathcal{I}^{\theta^*}$, pick any $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in S_i^{\mathcal{M}, \theta^*}$. Then, m_i is a best reply to some $\lambda_{-i} \in \triangle\left(S_{-i}^{\mathcal{M}, \theta^*}\right)$. By Step 2, (m_i, λ_{-i}) induces Rule 1 with probability 1. Since $\mathcal{I}^{\Theta} \neq \varnothing$, by Step 4, it must be the case that all agents in $\mathcal{I}^{\Theta}$ report $(\theta^*, 1)$. Thus, the only possibility is agreement on θ^*, i.e. all agents in $\mathcal{I}^{\theta^*}$ report $(\theta^*, 1)$. Therefore, only $m_i = \left(m_i^1, m_i^2, m_i^3, m_i^4\right)$ with $\left(m_i^1, m_i^2\right) = (\theta^*, 1)$ is rationalizable for agent $i \in \mathcal{I}^{\theta^*}$ at θ^*.

To sum, Step 1 shows

$$S^{\mathcal{M}, \theta^*} \supset \prod_{i \in \mathcal{I}} \left\{ \left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in M_i : \left(m_i^1, m_i^2\right) = (\theta^*, 1) \right\},$$

and Steps 2-5 show

$$S^{\mathcal{M}, \theta^*} \subset \left(\prod_{i \in \mathcal{I}^{\theta^*}} \left\{ \left(m_i^1, m_i^2, m_i^3, m_i^4\right) \in M_i : \left(m_i^1, m_i^2\right) = (\theta^*, 1) \right\} \right) \times \left(\prod_{i \in \mathcal{I} \diagdown \mathcal{I}^{\theta^*}} M_j \right).$$

As a result, we have $g\left(S^{\mathcal{M}, \theta^*}\right) = \{f(\theta^*)\}$, i.e., rationalizable implementation is achieved.■

References

AUMANN, R. J., AND A. BRANDENBURGER (1995): "Epistemic conditions for Nash equilibrium," *Econometrica*, 63, 1161–1180.

BERGEMANN, D., AND S. MORRIS (2009): "Robust implementation in direct mechanims," *Review of Economic Studies*, 76, 1175–1206.

——— (2011): "Robust implementation in general mechanism," *Games and Economic Behavior*, 71, 261–281.

BERGEMANN, D., S. MORRIS, AND O. TERCIEUX (2011): "Rationalizable Implementation," *Journal of Economic Theory*, 146, 1253–1274.

CABRALES, A., AND R. SERRANO (2011): "Implementation in adaptive better-response dynamics: Towards a general theory of bounded rationality in mechanisms," *Games and Economic Behavior*, 73, 360–374.

CHEN, Y.-C., T. KUNIMOTO, Y. SUN, AND S. XIONG (2021): "Rationalizable Implementation in Finite Mechanisms," *Games and Economic Behavior*, 129, 181–197.

JACKSON, M. O. (2001): "A crash course in implementation theory," *Social Choice Welfare*, 18, 655–708.

JAIN, R. (2021): "Rationalizable Implementation of Social Choice Correspondences," *Games and Economic Behavior*, 127, 47–66.

KUNIMOTO, T., AND R. SERRANO (2019): "Rationalizable Implementation of Correspondences," *Mathematics of Operations Research*, 44, 1145–1509.

MASKIN, E. (1999): "Nash equilibria and welfare optimality," *Review of Economic Studies*, 66, 23–38.

MASKIN, E., AND SJÖSTRÖM (2002): *Implementation Theory*, Handbook of Social Choice and Welfare. North-Holland, Amsterdam.

OURY, M., AND O. TERCIEUX (2012): "Continuous implementation," *Econometrica*, 80, 1605–1637.

XIONG, S. (2018): "Rationalizable Implementation: Social Choice Correspondences," *mimeo*.

——— (2022): "Rationalizable Implementation of Social Choice Functions: Complete Characterization," *working paper*, https://arxiv.org/abs/2202.04885.