2-2024

# Compellingness in Nash implementation

Shurojit CHATTERJI
*Singapore Management University*, shurojitc@smu.edu.sg

Takashi KUNIMOTO
*Singapore Management University*, tkunimoto@smu.edu.sg

Paulo Daniel Salles RAMOS
*Singapore Management University*, pauloramos.2018@phdecons.smu.edu.sg

SMU ECONOMICS & STATISTICS

SMU
SINGAPORE MANAGEMENT UNIVERSITY

# Compellingness in Nash Implementation

Shurojit Chatterji, Takashi Kunimoto, and Paulo Ramos

February 2024

Paper No. 02-2024

# Compellingness in Nash Implementation*

Shurojit Chatterji, Takashi Kunimoto, and Paulo Ramos†

This Version: February 2024

**Abstract**

A social choice function (SCF) is said to be Nash implementable if there exists a mechanism in which every Nash equilibrium induces outcomes specified by the SCF. The main objective of this paper is to assess the impact of considering mixed strategy equilibria in Nash implementation. We call a mixed strategy equilibrium "uncompelling" if its outcome is strictly Pareto dominated by that induced by the SCF. We show that if the finite environment and the SCF to be implemented jointly satisfy what we call *Condition COM*, then we can construct a finite mechanism which Nash implements the SCF in pure strategies with the property that any mixed strategy Nash equilibrium outcome is either uncompelling or consistent with the SCF. Our mechanism has several desirable features: transfers are completely dispensable; only finite mechanisms are considered; integer games are not invoked; and agents' attitudes toward risk do not matter. These features make our result quite distinct from prior attempts to handle mixed strategy equilibria in the theory of implementation.

*JEL Classification:* C72, D78, D82.

*Keywords:* implementation, ordinality, mixed strategies, Nash equilibrium, uncompelling mixed strategy equilibria.

*This paper subsumes Chatterji, Kunimoto, and Ramos (2022), which was restricted to two-person environments. We thank Bhaskar Dutta, Matthew Jackson, Atsushi Kajii, Jiangtao Li, Peng Liu, Antonio Penta, Stephen Morris, Ching-Jen Sun, Jingyi Xue, Hamid Sabourian, Yves Sprumont, and Takuro Yamashita as well as the conference/seminar audience at various places for helpful comments. This research is supported by the Ministry of Education, Singapore under MOE Academic Research Fund Tier 2 (MOE-T2EP402A20-0007). All remaining errors are our own.

†School of Economics, Singapore Management University, Singapore; shurojitc@smu.edu.sg, tkunimoto@smu.edu.sg, paulodsramos@smu.edu.sg

# 1 Introduction

The theory of implementation attempts to answer two questions. First, can one design a mechanism that successfully structures the interactions of agents in such a way that, in each state of the world, they always choose actions which result in the socially desirable outcomes for that state? Second, if agents possess information about the state and interact through a given mechanism, what properties do the resulting outcomes, viewed as a map from states to outcomes (and called social choice functions - henceforth, SCFs), possess? In answering these, the consequences of a given mechanism are predicted through the application of game theoretic solution concepts.[1]

In this paper we adopt Nash equilibrium as a solution concept, consider complete information environments, and ask if a given SCF is implementable, i.e., when can we design a mechanism in which "every" Nash equilibrium induces outcomes consistent with the SCF. Although the literature claims to care about all equilibria, it often ignores mixed strategy equilibria and only focuses on pure strategy equilibria. Jackson (1992) provides the most forceful argument for why the omission of mixed strategy equilibria brings about a serious consequence. In his Example 4, Jackson (1992) constructs a two-person environment and an SCF such that (i) there is a finite mechanism that pure Nash implements the SCF; and (ii) every finite pure Nash implementing mechanism always has a mixed strategy equilibrium that gives a lottery that is preferred by both agents to the outcome of the SCF. Thus, if we insist on using finite mechanisms, which is to be anticipated in an environment with finite number of alternatives and agents, we must question why agents would limit themselves to playing only pure strategies, particularly when there is a mixed strategy equilibrium that would be strictly preferred by both of them than any pure strategy equilibrium. In this paper, we revisit Jackson's example in Section 3.

To obtain the main result of the paper, we consider a finite environment with respect to an SCF on which we impose *Condition COM*, which delineates a set of conditions where it is always possible to construct a finite, pure Nash implementing mechanism such that every mixed strategy equilibrium outcome is either socially desirable or *uncompelling* in the sense that it is strictly Pareto dominated by the socially desirable outcome.[2] We call such a notion of implementation *compelling*

[1]See Jackson (2001), Maskin and Sjöström (2002), and Serrano (2004) for the survey of implementation theory.

[2]Note that Moore and Repullo (1990) identify Condition $\mu$ as a necessary and sufficient condition for pure strategy Nash implementation when there are at least three agents. In addition, Dutta and Sen (1991) and Moore and Repullo (1990) identify Condition $\beta$ and Condition $\mu 2$, respectively, as a necessary and sufficient condition for pure Nash implementation when there are only two agents.

*implementation.* Importantly, compelling implementation allows the implementing mechanism to admit mixed strategy equilibria that result in outcomes not consistent with the ones induced by the SCF, provided these mixed equilibria are uncompelling.[3] Hence, compelling implementation is considered a compromise between pure Nash implementation where only pure strategies are considered and mixed Nash implementation where all mixed strategy equilibria are fully considered.

To locate our contribution in a broader context, we first acknowledge that every prior work cited in the table below exploits some combination of the following five ingredients to handle mixed strategy equilibria in complete information environments: (i) infinite mechanisms; (ii) rationalizability as a stronger requirement than Nash equilibrium;[4] (iii) refinements of Nash equilibrium, such as subgame perfect equilibrium and undominated Nash equilibrium; (iv) environments with transfers or ones similar to separable environments of Jackson, Palfrey, and Srivstava (1994); and (v) cardinal utilities.[5]

| Combination of ingredients used | Previous works which handle mixed strategy equilibria in complete information environments |
|---|---|
| (i) | Kartik and Tercieux (2012), Maskin (1999), Maskin and Sjöström (2002), Mezzetti and Renou (2012a) |
| (i) × (v) | Kunimoto (2019), Serrano and Vohra (2010) |
| (i) × (ii) × (v) | Bergemann, Morris, and Tercieux (2011), Jain (2021), Kunimoto and Serrano (2019), Xiong (2022) |
| (ii) × (iv) × (v) | Abreu and Matsushima (1992), Chen, Kunimoto, Sun, and Xiong (2021) |
| (iii) × (iv) | Goltsman (2011), Jackson, Palfrey, and Srivastava (1994), Moore and Repullo (1988), Sjöström (1994) |
| (iii) × (iv) × (v) | Abreu and Matsushima (1994) |
| (iv) | Mezzetti and Renou (2012b) |
| (iv) × (v) | Chen, Kunimoto, Sun, and Xiong (2022) |

Table 1: The list of prior works handling mixed strategy equilibria in complete information environments.

We next emphasize that we obtain the main result of the paper without using any of the five ingredients used in the previous works. Our result is built on the following two requirements: (i) our implementing mechanism might admit mixed strategy equilibria which are uncompelling; and (ii) the planner is aware that each

[3]Our compelling implementation is similar to the notion of repeated implementation adopted by Lee and Sabourian (2015). They design a sequence of simple, finite mechanisms such that every pure strategy subgame perfect equilibrium "repeatedly" implements the efficient social choice function, while every mixed strategy subgame perfect equilibrium is strictly Pareto dominated by the pure equilibrium.

[4]Rationalizability is a more demanding requirement than Nash equilibrium because every action played with positive probability in a mixed strategy Nash equilibrium is rationalizable.

[5]This table, by no means, exhausts all related papers.

agent's cardinal utility from the socially desirable outcome is higher than that from the punishment lotteries. As long as this utility difference is positive, no matter how small it is, we can construct a mechanism that compellingly implements the SCF. In this sense, while compelling implementation is not completely ordinal, it can be made as ordinal as it can possibly be. We consider Nash implementation as the right notion of implementation if we insist on the robustness to information perturbations. This is so because Aghion, Fudenberg, Holden, Kunimoto, and Tercieux (2012) and Chung and Ely (2003) both show that Maskin monotonicity, a necessary condition for Nash implementation, is also necessary if we want implementation using refinements of Nash equilibria to be robust to information perturbations.[6] Our mechanism is finite so that it does not use the *integer games* which are often considered a questionable device in the literature.[7] The use of transfers can be dispensed with completely, which allows us to apply our result to an important class of environments including the models of voting and matching in which monetary transfers are simply unavailable.

We finally take up Korpela (2016) which is perhaps the closest to our paper.[8] Korpela (2016) uses a weaker notion of implementation than our compelling implementation in the sense that his notion of implementation ignores all uncompelling Nash equilibria, "regardless of whether they are pure or mixed." Therefore, Korpela's (2016) notion of implementation does not necessarily imply pure strategy Nash implementation, whereas our compelling implementation does. We believe that considering all pure strategy Nash equilibria is important because there is a well-known game with two Pareto-ranked strict Nash equilibria (e.g., the Stag-Hunt game), such that the Pareto-inferior (i.e., uncompelling) Nash equilibrium can sometimes prevail over the Pareto-superior Nash equilibrium.

We organize the rest of the paper as follows: Section 2 presents the environment, notation, mechanisms and solution concepts. Section 3 revisits Example 4 of Jackson (1992), which motivates our inquiry. Section 4 slightly modifies Example 4 of Jackson (1992) and presents an illustration of this paper's main result. Section 5 proposes Condition $COM$, provides an environment that satisfies Condition $COM$, and constructs a canonical mechanism that can achieve compelling implementation under Condition $COM$ when there are at least three agents. Section 6 argues that Condition $COM$ is indispensable for compelling implementation, in

[6] Aghion, Fudenberg, Holden, Kunimoto, and Tercieux (2012) and Chung and Ely (2003) adopt subgame perfect equilibrium and undominated Nash equilibrium as the solution concept, respectively. Note that permissive implementation results via refinements of Nash equilibria are usually obtained by weakening Maskin monotonicity.

[7] In the integer game, each agent announces some integer and the person who announces the highest integer gets to name his favorite outcome.

[8] This paper has been developed independently of Korpela (2016) and we only became aware of it after we completed Chatterji, Kunimoto, and Ramos (2022).

the sense that our mechanism fails to achieve compelling implementation if at least one property of Condition *COM* is violated. Section 7 extends the main result of Section 5 to the case of two agents. In Section 8, we compare part of Condition *COM* with Condition $\mu$ of Moore and Repullo (1990), which they show to be a necessary and sufficient condition for pure Nash implementation when there are at least three agents. Section 9 concludes the paper, and the Appendix contains the proofs omitted from the main body of the paper.

# 2 Preliminaries

There is a finite set of agents, denoted by $I = \{0, \ldots, n-1\}$ where we assume $n \geq 2$. Let $\Theta$ be the finite set of states. Let $A$ denote the set of social alternatives, which are assumed to be independent of the information state. We shall assume that $A$ is finite, and denote by $\Delta(A)$ the set of probability distributions over $A$. Associated with each state $\theta$ is a preference profile $\succeq^{\theta} = (\succeq_i^{\theta})_{i \in I}$ where $\succeq_i^{\theta}$ is agent $i$'s preference relation over $A$ at $\theta$. We write $a \succeq_i^{\theta} a^{'}$ when agent $i$ weakly prefers $a$ to $a^{'}$ in state $\theta$. We also write $a \succ_i^{\theta} a^{'}$ if agent $i$ strictly prefers $a$ to $a^{'}$ in state $\theta$ and $a \sim_i^{\theta} a^{'}$ if agent $i$ is indifferent between $a$ and $a^{'}$ in state $\theta$.

We assume that any preference relation $\succeq_i^{\theta}$ is extended to preferences over $\Delta(A)$ by a von Neumann-Morgenstern utility function $u_i(\cdot, \theta) : \Delta(A) \to \mathbb{R}$. We say that $u_i(\cdot, \theta)$ is consistent with $\succeq_i^{\theta}$ if, for any $a, a^{'} \in A$, $u_i(a, \theta) \geq u_i(a^{'}, \theta) \Leftrightarrow a \succeq_i^{\theta} a^{'}$. We denote by $\mathcal{U}_i^{\theta}$ the set of all possible cardinal utility functions $u_i(\cdot, \theta)$ that are consistent with $\succeq_i^{\theta}$. We formally define $\mathcal{U}_i^{\theta}$ as follows:

$$\mathcal{U}_i^{\theta} = \left\{ u_i(\cdot, \theta) \in [0,1]^{|A|} | \ u_i(\cdot, \theta) \text{ is consistent with } \succeq_i^{\theta} \right\},$$

where $|A|$ denotes the cardinality of $A$. Let $\mathcal{U}^{\theta} \equiv \times_{i \in I} \mathcal{U}_i^{\theta}$ and $\mathcal{U} \equiv \times_{\theta \in \Theta} \mathcal{U}^{\theta}$. We denote any subset of $\mathcal{U}_i^{\theta}$ by $\hat{\mathcal{U}}_i^{\theta}$, any subset of $\mathcal{U}^{\theta}$ by $\hat{\mathcal{U}}^{\theta}$, any subset of $\mathcal{U}$ by $\hat{\mathcal{U}}$, respectively.

We can now define an *environment* as $\mathcal{E} = (I, A, \Theta, (\succeq_i^{\theta})_{i \in I, \theta \in \Theta}, \hat{\mathcal{U}})$, which is implicitly understood to be commonly certain among agents. This paper introduces a *planner* who takes $\mathcal{E}$ as the primitive of her model and a *social choice function*—(henceforth, SCF) $f : \Theta \to \Delta(A)$ as her objective. The planner's objective takes only ordinal information about the state $\theta$ as input, but is allowed to have lotteries as outputs.[9] This implies that the planner is agnostic about agents' cardinal utility functions. Finally, we adopt the complete information assumption. That is, the underlying state $\theta \in \Theta$ and cardinal utility functions $u \in \hat{\mathcal{U}}$ together are commonly certain among agents.

[9]Although many papers deal with multi-valued social choice correspondences in the literature of Nash implementation, we focus only on single-valued SCFs.

## 2.1 Compelling Implementation

Let $\Gamma = ((M_i)_{i\in I}, g)$ be a finite mechanism where $M_i$ is a nonempty *finite* set of messages available to agent $i$; $g : M \to \Delta(A)$ (where $M \equiv \times_{i\in I} M_i$) is the *outcome function*. At each state $\theta \in \Theta$ and profile of cardinal utility functions $u \in \hat{\mathcal{U}}$, the environment and the mechanism together constitute a *game with complete information* which we denote by $\Gamma(\theta, u)$. Note that the restriction of $M_i$ to a finite set rules out the use of integer games (See, for example, Maskin (1999)).

Let $\sigma_i \in \Delta(M_i)$ be a mixed *strategy* of agent $i$ in the game $\Gamma(\theta, u)$. A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n) \in \times_{i\in I}\Delta(M_i)$ is said to be a mixed-strategy *Nash equilibrium* of the game $\Gamma(\theta, u)$ if, for all agents $i \in I$ and all messages $m_i \in \text{supp}\,(\sigma_i)$ and $m_i' \in M_i$, we have

$$\sum_{m_{-i}\in M_{-i}} \prod_{j\neq i} \sigma_j(m_j) u_i(g(m_i, m_{-i}), \theta) \geq \sum_{m_{-i}\in M_{-i}} \prod_{j\neq i} \sigma_j(m_j) u_i(g(m_i', m_{-i}), \theta).$$

A *pure-strategy* Nash equilibrium is a mixed-strategy Nash equilibrium $\sigma$ such that each agent $i$'s mixed-strategy $\sigma_i$ assigns probability one to some $m_i \in M_i$. Let $NE(\Gamma(\theta, u))$ denote the set of mixed-strategy Nash equilibria of the game $\Gamma(\theta, u)$. By $\Gamma(\theta)$ we mean the game in which the preference profile $(\succeq_i^\theta)_{i\in N}$ is commonly certain among the agents, while any cardinal utility function $u \in \hat{\mathcal{U}}$ is admissible. Since pure strategy equilibria do not depend on agents' cardinal utilities, we denote by $pureNE(\Gamma(\theta))$ the set of pure strategy Nash equilibria of the game $\Gamma(\theta)$. As far as we are only concerned with pure strategy equilibria, we only need ordinal preferences so that we can write $pureNE(\Gamma(\theta))$. We say that an SCF $f$ is *pure Nash implementable* if there exists a mechanism $\Gamma = (M, g)$ such that for any state $\theta$, the following two conditions hold: (i) $pureNE(\Gamma(\theta)) \neq \emptyset$; and (ii) $m \in pureNash(\Gamma(\theta)) \Rightarrow g(m) = f(\theta)$.

We strengthen the notion of pure Nash implementation by requiring that any mixed equilibrium outcome, if exists, be either socially desirable or "uncompelling" in the sense that it is strictly Pareto dominated by the socially desirable outcome. For every mixed strategy profile $\sigma \in \times_{i\in I}\Delta(M_i)$, we write

$$g(\sigma) \equiv \sum_{m\in M} \sigma(m) g(m).$$

Our notion of implementation can then be formally defined as follows:

**Definition 1** *An SCF $f$ is* ***compellingly implementable with respect to*** $\hat{\mathcal{U}}$ *if there exists a finite mechanism $\Gamma = (M, g)$ such that for every state $\theta \in \Theta$, (i) $pureNE(\Gamma(\theta)) \neq \emptyset$; (ii) $m \in pureNE(\Gamma(\theta)) \Rightarrow g(m) = f(\theta)$; and (iii) for any $u \in \hat{\mathcal{U}}^\theta$ and $\sigma \in NE(\Gamma(\theta, u))$, $g(\sigma) \neq f(\theta) \Rightarrow u_i(f(\theta), \theta) > u_i(g(\sigma), \theta)$ for all $i \in I$.*

**Remark 1** *The implementing mechanism may have two types of mixed strategy equilibria. We call the first type of it a "good" mixed strategy equilibrium in the sense that its outcome is socially desirable and call the second type of it a "bad" mixed strategy equilibrium in the sense that its outcome is strictly worse for all agents than the socially desirable outcome. Our notion of compelling implementation says that the planner should ignore bad mixed strategy equilibria in the mechanism.*

# 3 The Relevance of Mixed Strategy Equilibria in Nash Implementation

In this section, we articulate a compelling reason why we need to be worried about mixed strategy equilibria in Nash implementation. To do so, we revisit Example 4 of Jackson (1992), which shows that the omission of mixed strategy equilibria brings about a serious blow to Nash implementation.

We revisit Example 4 of Jackson (1992). Suppose that there are two agents $I = \{0, 1\}$; four alternatives $A = \{a, b, c, d\}$; and two states $\Theta = \{\theta, \theta'\}$. Suppose that agent 0 has the state-independent preference $a \succ_0 b \succ_0 c \sim_0 d$ and agent 1 has the preference $a \succ_1^{\theta} b \succ_1^{\theta} d \succ_1^{\theta} c$ at state $\theta$ and preference $b \succ_1^{\theta'} a \succ_1^{\theta'} c \sim_1^{\theta'} d$ at state $\theta'$. Consider the SCF $f$ such that $f(\theta) = a$ and $f(\theta') = c$.

First, Jackson (1992) constructs a finite mechanism $\Gamma = (M, g)$ (described in Table 2) that implements the SCF $f$ in pure-strategy Nash equilibria:

| $g(m)$ | | Agent 1 | | |
|---|---|---|---|---|
| | | $m_1^1$ | $m_1^2$ | $m_1^3$ |
| Agent 0 | $m_0^1$ | $c$ | $d$ | $d$ |
| | $m_0^2$ | $d$ | $a$ | $b$ |
| | $m_0^3$ | $d$ | $b$ | $a$ |

Table 2: The mechanism introduced in Example 4 of Jackson (1992).

There are two pure strategy Nash equilibria, $(m_0^2, m_1^2)$ and $(m_0^3, m_1^3)$, in the game $\Gamma(\theta)$, both of which result in outcome $a$. In the game $\Gamma(\theta')$, the unique pure-strategy Nash equilibrium is $(m_0^1, m_1^1)$, which results in outcome $c$. Thus, the SCF $f$ is implementable by the above finite mechanism in pure-strategy Nash equilibria. Due to the necessity of Maskin monotonicity for Nash implementation, we know that the SCF $f$ satisfies Maskin monotonicity. However, in the game $\Gamma(\theta')$, there is a mixed-strategy Nash equilibrium, where each agent $i$ plays $m_i^2$ and

$m_i^3$ with equal probability, which results in outcomes $a$ and $b$, each with probability 1/2. Both agents strictly prefer any outcome of the mixed-strategy equilibrium to the outcome of the pure-strategy equilibrium. Thus, this mixed strategy Nash equilibrium is Pareto-superior. Note that there is a conflict of interests between the two agents over $a$ and $b$ in state $\theta^{'}$, i.e., while agent 0 prefers $a$ to $b$, agent 1 prefers $b$ to $a$. This conflict of interests allows us to have the unique pure strategy Nash equilibrium in the game $\Gamma(\theta^{'})$, which results in outcome $c$. At the same time, this logic for the uniqueness of the pure-strategy equilibrium is extremely dubious because outcomes $a$ and $b$ are strictly better for both agents than outcome $c$.

Jackson (1992) further shows that his argument applies to *any* finite implementing mechanism. That is, for any finite mechanism which implements the SCF $f$ in pure-strategy Nash equilibria, there must also exist a Pareto-superior mixed-strategy Nash equilibrium at state $\theta'$ inducing a lottery different from $c$, which is the socially desirable outcome by the SCF $f$ at state $\theta^{'}$. Therefore, the SCF $f$ is "not" compellingly implementable with respect to $\mathcal{U}$, which is the set of "all" cardinal utility functions, or any of its subsets. It thus follows that the identified Pareto-superior mixed strategy equilibrium persists independently of any cardinal utility functions.

# 4 Illustration of the Main Result

The main objective of this paper is to identify a class of environments where the issue of mixed strategy equilibria can be avoided by carefully designing an implementing mechanism. In this section, we illustrate how we resolve this issue in the slightly modified version of Example 4 of Jackson (1992).

One crucial feature Jackson's Example 4 has is that its argument seems to rely heavily on the extreme inefficiency of the SCF, i.e., the SCF $f$ assigns the common worst outcome in state $\theta^{'}$.[10] To investigate how robust Jackson's argument is, we only make the following modification: both agents now strictly prefer $c$ to $d$ in state $\theta^{'}$, i.e., $c \succ_i^{\theta^{'}} d$ for each $i = 0, 1$.

We summarize the basic setup. Agent 0 has the state-independent preference $a \succ_0 b \succ_0 c \succ_0 d$ and agent 1 has the preference $a \succ_1^{\theta} b \succ_1^{\theta} d \succ_1^{\theta} c$ at state $\theta$ and preference $b \succ_1^{\theta'} a \succ_1^{\theta'} c \succ_1^{\theta'} d$ at state $\theta'$. Consider the same SCF $f$ such that $f(\theta) = a$ and $f(\theta') = c$. This way the SCF never assigns the worst outcome for any agent in either state (a feature that will also be implied by our sufficient condition).

With this modification, we are able to construct a mechanism that not only implements the SCF in pure-strategy Nash equilibrium but also guarantees that

[10] Jackson (1992, p.770) is well aware of this point.

all mixed-strategy equilibria of the constructed mechanism give each agent the expected payoff arbitrarily close to that of $d$, which is worse than that of $c$, the outcome induced by the SCF $f$ at state $\theta'$. Hence, we essentially overturn the implication of Jackson's Example 4 by assuming that there is a uniform bound for the utility difference.[11]

For each integer $k \geq 2$, we define $\Gamma^k = (M^k, g^k)$ as a mechanism with the following properties: (i) for each $i \in I$, $M_i^k = \{0, 1, \ldots, k\}$ and (ii) the outcome function $g^k : M^k \to A$ is given by the following rules: for each $m \in M^k$,

- If $m = (k, k)$, then $g^k(m) = c$;
- If there exists an integer $h$ with $0 \leq h \leq k-1$ such that $m = (h, h)$, then $g^k(m) = a$;
- If there exists an integer $h$ with $0 \leq h \leq k-1$ such that $m = (h, (h + 1 \bmod k))$, then $g^k(m) = b$; and
- Otherwise, $g^k(m) = d$.

We illustrate this mechanism in Table 3:

| $g^k(m)$ | | Agent 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $k$ | $k-1$ | $k-2$ | $k-3$ | $\cdots$ | 3 | 2 | 1 | 0 |
| Agent 0 | $k$ | $c$ | $d$ | $d$ | $d$ | $\cdots$ | $d$ | $d$ | $d$ | $d$ |
| | $k-1$ | $d$ | $a$ | $d$ | $d$ | $\cdots$ | $d$ | $d$ | $d$ | $b$ |
| | $k-2$ | $d$ | $b$ | $a$ | $d$ | $\cdots$ | $d$ | $d$ | $d$ | $d$ |
| | $k-3$ | $d$ | $d$ | $b$ | $a$ | $\cdots$ | $d$ | $d$ | $d$ | $d$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| | 3 | $d$ | $d$ | $d$ | $d$ | $\cdots$ | $a$ | $d$ | $d$ | $d$ |
| | 2 | $d$ | $d$ | $d$ | $d$ | $\cdots$ | $b$ | $a$ | $d$ | $d$ |
| | 1 | $d$ | $d$ | $d$ | $d$ | $\cdots$ | $d$ | $b$ | $a$ | $d$ |
| | 0 | $d$ | $d$ | $d$ | $d$ | $\cdots$ | $d$ | $d$ | $b$ | $a$ |

Table 3: $\Gamma^k = (M^k, g^k)$ where $k \geq 3$.

When $k = 2$, the above mechanism is reduced to the one introduced by Jackson (1992) where we set $m_i^1 = 2; m_i^2 = 1$; and $m_i^3 = 0$ for each $i \in \{0, 1\}$, as seen in Table 4.

[11] The mechanism presented here differs slightly from the canonical mechanism introduced in Section 5. Specifically, the mechanism here has been tailored to the particular example at hand and simplified for ease of exposition. Nevertheless, the main insights of this paper are still obtained in this illustration section.

| $g(m)$ | | Agent 1 | | |
|---|---|---|---|---|
| | | 2 | 1 | 0 |
| Agent 0 | 2 | $c$ | $d$ | $d$ |
| | 1 | $d$ | $a$ | $b$ |
| | 0 | $d$ | $b$ | $a$ |

Table 4: $\Gamma^k = (M^k, g^k)$ where $k = 2$.

For each $\theta \in \Theta$, $i \in \{0,1\}$, and $\varepsilon > 0$, we define $\mathcal{U}_i^{\theta,\varepsilon}$ as a subset of $\mathcal{U}_i^{\theta}$ as follows:

$$\mathcal{U}_i^{\theta,\varepsilon} = \left\{ u_i \in \mathcal{U}_i^{\theta} \middle| \ |u_i(a,\theta) - u_i(a',\theta)| \geq \varepsilon, \ \forall a \in A, \forall a' \in A \backslash \{a\}, \ \forall \theta \in \Theta \right\}.$$

Let $\mathcal{U}^{\theta,\varepsilon} \equiv \times_{i \in N} \mathcal{U}_i^{\theta,\varepsilon}$ and $\mathcal{U}^{\varepsilon} \equiv \times_{\theta \in \Theta} \mathcal{U}^{\theta,\varepsilon}$. We observe that $\mathcal{U}^{\varepsilon}$ possesses the following monotonicity:

$$\varepsilon > \varepsilon' > 0 \Rightarrow \mathcal{U}^{\varepsilon} \subsetneq \mathcal{U}^{\varepsilon'} \subseteq \mathcal{U} \subseteq \mathcal{U}^0.$$

Loosely speaking, if we choose $\varepsilon > 0$ small enough, we can approximate $\mathcal{U}$ by $\mathcal{U}^{\varepsilon}$ to an arbitrary degree. We are now ready to state the main result of this section.

**Proposition 1** *For any $\varepsilon > 0$, there exists $K \in \mathbb{N}$ large enough such that the SCF $f$ is compellingly implementable with respect to $\mathcal{U}^{\varepsilon}$ by the mechanism $\Gamma^K$.*

**Proof**: The proof is completed by a series of lemmas. For the moment, we fix $k$ in the proof and we ignore the dependence of the mechanism on $k$. We first establish pure Nash implementation by the mechanism $\Gamma^k$.

**Lemma 1** *The mechanism $\Gamma^k$ implements the SCF in pure-strategy Nash equilibrium.*

**Proof:** The message profile $(1,1)$ is a Nash equilibrium of the game $\Gamma^k(\theta)$, as it yields $a$ which is their most preferred outcome for both agents so that no agent can find a profitable deviation. We claim that $a$ is the unique pure Nash equilibrium outcome of the game $\Gamma^k(\theta)$. Let $m$ be a message profile such that $g(m) \neq a$. We will show that $m$ is "not" a Nash equilibrium in the game $\Gamma^k(\theta)$:

- If $g(m) = b$, there exists an integer $h$ with $0 \leq h \leq k-1$ such that $m = (h, (h+1 \text{ mod } k))$. Then agent 0 has an incentive to send a message $h + 1 \text{ mod } k$ so that outcome $a$ is induced.

- If $g(m) = c$, then $m = (k,k)$. Then, agent 1 has an incentive to send any message other than $k$ so that outcome $d$ is induced, as he strictly prefers outcome $d$ to outcome $c$ at state $\theta$.

- If $g(m) = d$, then we have $m = (m_0, m_1)$ where $m_0 \neq m_1$. If $m_0 > m_1$ then, agent 0 has an incentive to deviate from $m_0$ to $m_1$ so that outcome $a$ is induced. Conversely, if $m_1 > m_0$, then agent 0 has an incentive to deviate from $m_1$ to $m_0$, so that outcome $a$ is induced.

We next claim that $(k, k)$ is a Nash equilibrium of the game $\Gamma^k(\theta')$ because any unilateral deviation from $(k, k)$ yields $d$, which is inferior to $c$ induced by $(k, k)$ for both agents. Moreover, no other outcome can be induced by a Nash equilibrium in this game: every message profile $m = (m_0, m_1)$ where $m_1 < k$ and $g(m) \neq a$ has a profitable deviation for agent 0 at $m'_0 = m_1$, while every message profile $m = (m_0, m_1)$ where $m_0 < k$ and $g(m) \neq b$ has a profitable deviation for agent 1 at $m'_1 = m_0 + 1 \pmod k$. Since $g(m) = a$ implies $m_0 < k$ and $g(m) = b$ implies $m_1 < k$, we have that $(m_0, m_1)$ is not a Nash equilibrium if either $m_0 < k$ or $m_1 < k$. Thus, the only possible Nash equilibrium in pure strategies in the game $\Gamma^k(\theta')$ is $(k, k)$. ■

The following lemma is our key result, characterizing the set of Nash equilibria of the mechanism $\Gamma^k$ in state $\theta'$.

**Lemma 2** *For each $i \in \{0, 1\}$, let $\sigma_i = (\sigma_i(0), \sigma_i(1), ..., \sigma_i(k))$ denote agent $i$'s strategy and for each $x \in \{0, 1, \ldots, k\}$, let $\sigma_i(x)$ denote the probability that agent $i$ chooses $x$. If $\sigma = (\sigma_0, \sigma_1)$ is a Nash equilibrium in the game $\Gamma^k(\theta')$, then, for each $i \in \{0, 1\}$, there is a number $p^i \in [0, 1]$ such that $\sigma_i(x) = p^i/k$ for each $x \in \{0, \ldots, k-1\}$. Moreover, $p^0 = 0$ if and only if $p^1 = 0$.*

**Proof:** Recall that we set $u_i(d; \theta') = 0$ for each $u_i \in \mathcal{U}_i^{\theta'}$ and $i \in \{0, 1\}$. Let $\sigma$ be a Nash equilibrium of the game $\Gamma^k(\theta')$. If $\sigma_i(k) = 1$ for each $i \in \{0, 1\}$, such $p^i$ in the lemma is guaranteed to exist by setting $p^i = 0$. Thus, we assume that there exists $i \in \{0, 1\}$ for whom $\sigma_i(k) < 1$. We divide the proof into a series of steps. The proof of each step is in the Appendix.

**Step 1a**: If there exists $x \in \{0, \ldots, k-1\}$ such that $\sigma_0(x) > 0$, then $\sigma_1(x) > 0$.

**Step 1b**: If there exists $x \in \{1, \ldots, k-1\}$ such that $\sigma_1(x) > 0$, then $\sigma_0(x-1) > 0$. Moreover, if $\sigma_1(0) > 0$, then $\sigma_0(k-1) > 0$.

**Step 1c**: If there exist $i \in \{0, 1\}$ and $x' \in \{0, \ldots, k-1\}$ for whom $\sigma_i(x') > 0$, then $\sigma_0(x) > 0$ and $\sigma_1(x) > 0$ for all $x \in \{0, \ldots, k-1\}$.

**Step 2**: If there exist $i \in \{0, 1\}$ and $x, x' \in \{0, \ldots, k-1\}$ such that $\sigma_i(x) > 0$ and $\sigma_i(x') > 0$, then $\sigma_i(x) = \sigma_i(x')$.

It follows from both Steps 1c and 2 that $\sigma_i(x) = \sigma_i(x^{'})$ for every $x, x^{'} \in \{0, \ldots, k-1\}$ and $i \in \{0,1\}$. Thus, we can set $p^i = \sum_{x=0}^{k-1} \sigma_i(x)$ for each $i \in \{0,1\}$. Since we assume $\sigma_i(k) < 1$ for each $i \in \{0,1\}$, we have $p^i > 0$. This completes the proof of Lemma 2. ■

As we can easily see in the proof of Lemma 1,every mixed strategy Nash equilibria of the game $\Gamma^k(\theta)$ (if any) is uncompelling, because, in state $\theta$, the unique pure Nash equilibrium outcome is $a$, which is the best outcome for both agents. It thus remains to prove that every mixed Nash equilibrium of the game $\Gamma^k(\theta^{'})$ is also uncompelling.

If $k \geq 3$, we let $\sigma^k$ be a nontrivial mixed-strategy Nash equilibrium in the game $\Gamma^k(\theta^{'})$. Then, the resulting outcome distribution induced by $\sigma^k$ is given by

$$g \circ \sigma^k = \begin{cases} c & \text{w.p. } (1-p^0)(1-p^1) \\ a & \text{w.p. } (p^0 p^1)/k \\ b & \text{w.p. } (p^0 p^1)/k \\ d & \text{w.p. } ((k - 2p^0 p^1)/k) - ((1-p^0)(1-p^1)), \end{cases}$$

where $p^0, p^1 \in (0,1]$ and $p^i = \sum_{x=0}^{k-1} \sigma_i(x)$ for each $i \in \{0,1\}$. Recall the following pieces of notation:

$$\mathcal{U}_0^{\theta^{'}} = \left\{ u_0(\cdot;\theta^{'}) \in [0,1]^A \middle| \, 1 = u_0(a;\theta^{'}) > u_0(b;\theta^{'}) > u_0(c;\theta^{'}) > u_0(d;\theta^{'}) = 0 \right\};$$
$$\mathcal{U}_1^{\theta^{'}} = \left\{ u_1(\cdot;\theta^{'}) \in [0,1]^A \middle| \, 1 = u_1(b;\theta^{'}) > u_1(a;\theta^{'}) > u_1(c;\theta^{'}) > u_1(d;\theta^{'}) = 0 \right\}.$$

Let $\mathcal{U}^{\theta^{'}} \equiv \mathcal{U}_0^{\theta^{'}} \times \mathcal{U}_1^{\theta^{'}}$. For each $\varepsilon \in (0,1)$, we have

$$\mathcal{U}_0^{\theta^{'},\varepsilon} = \left\{ u_0(\cdot;\theta^{'}) \in \mathcal{U}_0^{\theta^{'}} \middle| \, u_0(c;\theta^{'}) \geq \varepsilon \right\};$$
$$\mathcal{U}_1^{\theta^{'},\varepsilon} = \left\{ u_1(\cdot;\theta^{'}) \in \mathcal{U}_1^{\theta^{'}} \middle| \, u_1(c;\theta^{'}) \geq \varepsilon \right\}.$$

Similarly, let $\mathcal{U}^{\theta^{'},\varepsilon} \equiv \mathcal{U}_0^{\theta^{'},\varepsilon} \times \mathcal{U}_1^{\theta^{'},\varepsilon}$.

By the lemma below, we show that for each $\varepsilon > 0$, there exists $K \in \mathbb{N}$ large enough so that, for any $u \in \mathcal{U}^{\theta^{'},\varepsilon}$, any non-trivial mixed strategy equilibrium of the game $\Gamma^K(\theta^{'}, u)$ is uncompelling.

**Lemma 3** *For each $\varepsilon > 0$, there exists an integer $K \in \mathbb{N}$ large enough so that for any $k \geq K$, $i \in \{0,1\}$, and $(u_0(\cdot,\theta^{'}), u_1(\cdot;\theta^{'})) \in \mathcal{U}^{\theta^{'},\varepsilon}$,*

$$u_i(g(\sigma^k);\theta^{'}) < u_i(c;\theta^{'}).$$

**Proof**: Fix $\varepsilon > 0$ and $i \in \{0, 1\}$. We compute

$$u_i(g(\sigma^k); \theta') = \frac{p^0 p^1}{k}[u_i(a; \theta') + u_i(b; \theta')] + (1 - p^0)(1 - p^1) u_i(c; \theta').$$

For each $(p^0, p^1) \in [0, 1]^2$, we define

$$k(p^0, p^1) = \frac{u_i(a; \theta') + u_i(b; \theta')}{u_i(c; \theta')} \left[\frac{1}{p^0} + \frac{1}{p^1} - 1\right]^{-1}.$$

In the rest of the proof, we make use of the following properties of $k(p^0, p^1)$:

- $k(\cdot, \cdot)$ is strictly increasing in both arguments over $[0, 1]^2$.
- $k(p^0_h, p^1_h)$ converges to zero no matter how the sequence $\{(p^0_h, p^1_h)\}_{h=1}^{\infty}$ approaches $(0, 0)$. Thus, $k(0, 0) \equiv \lim_{(p^0, p^1) \to (0,0)} k(p^0, p^1) = 0$.
- $k(1, 1) = [u_i(a; \theta') + u_i(b; \theta')]/u_i(c; \theta') = \max_{(p^0, p^1) \in [0,1]^2} k(p^0, p^1)$.
- We can conveniently rewrite $k(p^0, p^1)$ as

$$k(p^0, p^1) = \frac{u_i(a; \theta') + u_i(b; \theta')}{u_i(c; \theta')} \frac{p^0 p^1}{[1 - (1 - p^0)(1 - p^1)]}.$$

We set $K = \min\{k \in \mathbb{N} | k \geq 2/\varepsilon\}$. As $2/\varepsilon \geq [u_i(a; \theta') + u_i(b; \theta')]/u_i(c; \theta')$ for any $u_i(\cdot; \theta') \in \mathcal{U}_i^{\theta'}[\varepsilon]$, we have that $K \geq k(1, 1)$. Due to the strict monotonicity of $k(p^0, p^1)$ with respect to $p^0$ and $p^1$, we have that $K \geq k(p^0, p^1)$ for any $(p^0, p^1) \in [0, 1]^2$. Hence, for any $k \geq K$:

$$\begin{aligned}
u_i(g(\sigma^k); \theta') &= \frac{p^0 p^1}{k}[u_i(a; \theta') + u_i(b; \theta')] + (1 - p^0)(1 - p^1) u_i(c; \theta') \\
&\leq \frac{p^0 p^1}{k(p^0, p^1)}[u_i(a; \theta') + u_i(b; \theta')] + (1 - p^0)(1 - p^1) u_i(c; \theta') \\
&\quad (\because\ k \geq K \geq k(p^0, p^1)\ \forall (p^0, p^1) \in [0, 1]^2) \\
&= u_i(c; \theta')[1 - (1 - p^0)(1 - p^1)] + (1 - p^0)(1 - p^1) u_i(c; \theta') \\
&= u_i(c; \theta').
\end{aligned}$$

This completes the proof of Lemma 3. ■

Combining Lemmas 1, 2, and 3 together, we complete the proof of Proposition 1. ■

# 5 The Main Result When $n \geq 3$

Throughout this section, we assume that there are at least three agents, i.e., $n \geq 3$. We refer the reader to Section 7 where we address the result for the case of two agents, i.e., $n = 2$.

## 5.1 Acceptability and Forums

Let $\mathcal{G}$ be a pair of agents in $I$. We call $C : \mathcal{G} \rightrightarrows \Delta(A)$ a *choice correspondence* if it maps each agent in $\mathcal{G}$ into a nonempty, finite subset of lotteries in $\Delta(A)$. Given a pair of agents $\mathcal{G}$ and a choice correspondence $C$, we define the concept of $C$-acceptability:

**Definition 2** *Let $\mathcal{G} \subseteq I$ be a pair of agents, $C$ a choice correspondence, $\theta \in \Theta$ a state, and $u \in \hat{\mathcal{U}}$ a cardinal utility function. We say that lottery $x \in \Delta(A)$ is* $C$***-acceptable at*** $(\theta, u)$ *if the following two conditions are satisfied:*

- $x \in \bigcup_{i \in \mathcal{G}} C(i)$;
- *For every $i \in \mathcal{G}$ and $y \in C(i)$, $u_i(x, \theta) \geq u_i(y, \theta)$.*

**Remark**: Strictly speaking, the definition of $C$-acceptability depends upon $\mathcal{G}$. However, since such $\mathcal{G}$ is always clear from the context whenever we discuss $C$-acceptability, we omit $C$'s dependence on $\mathcal{G}$. Thus, we simply say $C$-acceptability without mentioning $\mathcal{G}$. If there are only two agents, i.e., $n = 2$, then there is no ambiguity about $\mathcal{G}$ so that we always take $\mathcal{G} = \{0, 1\}$ in the definition of $C$-acceptability.

We say that $\mathcal{F} = (\mathcal{G}, w, C, z)$ constitutes a *forum* if it satisfies the following properties:

1. $\mathcal{G}$ is a pair of agents in $I$;
2. $w : \{0, 1\} \to \mathcal{G}$ is a bijection, where we denote $w^{-1}$ by its inverse function so that $w(w^{-1}(i)) = i$;
3. $C : \mathcal{G} \rightrightarrows \Delta(A)$ is a choice correspondence; and
4. $z \in \Delta(A)$ is a lottery such that $z \in \bigcap_{i \in \mathcal{G}} C(i)$.

For each $\hat{\theta} \in \Theta$, we write $\mathcal{F}_{\hat{\theta}} = (\mathcal{G}_{\hat{\theta}}, w_{\hat{\theta}}, C_{\hat{\theta}}, z_{\hat{\theta}})$ as a forum indexed by the state $\hat{\theta} \in \Theta$. In the forum $\mathcal{F}_{\hat{\theta}}$, we have $z_{\hat{\theta}} \in C_{\hat{\theta}}(i)$ for each $i \in \mathcal{G}_{\hat{\theta}}$. If the forum $\mathcal{F}_{\hat{\theta}}$ is used when $\theta$ is the true state, we define

$$C^*_{\hat{\theta}}(j, \theta) \equiv \arg \max_{y \in C_{\hat{\theta}}(j)} u_j(y, \theta)$$

as the set of agent $j$'s best lotteries in state $\theta$ within $C_{\hat{\theta}}(j)$.

## 5.2 Condition *COM*

**Definition 3** *The environment $\mathcal{E} = (I, A, \Theta, (\succeq_i^\theta)_{i \in I, \theta \in \Theta}, \hat{\mathcal{U}})$ satisfies **Condition** COM with respect to the SCF $f$ if there exists a collection of forums $\{\mathcal{F}_{\tilde{\theta}}\}_{\tilde{\theta} \in \Theta} = \{\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}}\}_{\tilde{\theta} \in \Theta}$ with the following properties:*

1. *This property has two parts:*

   1-i. *For every $\theta \in \Theta$ and $u \in \hat{\mathcal{U}}$, $f(\theta)$ is $C_\theta$-acceptable at $(\theta, u)$.*

   1-ii. *For every $\theta, \hat{\theta} \in \Theta$, $f(\theta) \in C_{\hat{\theta}}(i) \Leftrightarrow f(\theta) \in C_{\hat{\theta}}(j)$, where $i = w_{\hat{\theta}}(0)$ and $j = w_{\hat{\theta}}(1)$. When $\hat{\theta} = \theta$, we have $f(\theta) \in C_\theta(i) \cap C_\theta(j)$ for all $\theta \in \Theta$.*

2. *For every $\theta, \hat{\theta} \in \Theta$ and $u \in \hat{\mathcal{U}}$, if $x \in \Delta(A)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then $x = f(\theta)$.*

3. *There exists $\varepsilon > 0$ such that for each $\theta \in \Theta$, $u_i(f(\theta), \theta) - u_i(z, \theta) \geq \varepsilon$ for all $i \in I$, all $u \in \hat{\mathcal{U}}$, and all $z \in \bigcup_{\tilde{\theta} \in \Theta} \{z_{\tilde{\theta}}\}$.*

4. *For all $\theta, \hat{\theta} \in \Theta$, $u \in \hat{\mathcal{U}}$, $i \in \mathcal{G}_{\hat{\theta}}$, if $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then $u_i(x, \theta) = u_i(f(\theta), \theta)$ implies $x = f(\theta)$ for all $x \in C_{\hat{\theta}}(i)$.*

**Remark 2** *There is a redundancy in Property 1. We do not need to assume $f(\theta) \in C_\theta(i) \cap C_\theta(j)$ for all $\theta \in \Theta$ in Property 1-ii because it follows from (i) $f(\theta)$ is $C_\theta$-acceptable at $(\theta, u)$ in Property 1-i and (ii) $f(\theta) \in C_\theta(i) \Leftrightarrow f(\theta) \in C_\theta(j)$ when we set $\hat{\theta} = \theta$ in Property 1-i. Nevertheless, for the convenience of writing the proofs in the paper, we explicitly add this property to part of Property 1-ii.*

*Property 2 admits a possibility that agent $i \in \mathcal{G}_{\hat{\theta}}$ is indifferent between $f(\theta)$ and some $x \neq f(\theta)$, while agent $j \in \mathcal{G}_{\hat{\theta}} \backslash \{i\}$ prefers $f(\theta)$ to $x$. However, Property 4 excludes this very possibility. This shows that Property 2 does not imply Property 4.*

Property 1-i says that when $\theta$ is the true state, $f(\theta)$ is the best outcome for both agents in the forum the forum $\theta$ (i.e., the two agents in $\mathcal{G}_\theta$). Property 1-ii says that, if $\theta$ is the true state, then for every forum $\hat{\theta} \neq \theta$ either $f(\theta)$ is in the choice sets for both agents in $\mathcal{G}_{\hat{\theta}}$ or it is not in the choice sets for either agent in $\mathcal{G}_{\hat{\theta}}$. Moreover, if $\hat{\theta} = \theta$, $f(\theta)$ must be in the choice sets for both agents in $\mathcal{G}_{\hat{\theta}}$. Property 2 says that, for every forum $\hat{\theta}$, if the true state is $\theta$ and $x$ is the best outcome within the choice sets for both agents in $\mathcal{G}_{\hat{\theta}}$, then $x$ must be $f(\theta)$. Property 3 says that there exists the smallest unit $\varepsilon > 0$ such that each agent's cardinal utility from $f(\theta)$ is at least as high as that from any punishment lottery $z$ by $\varepsilon$. Property 4 says that, for every forum $\hat{\theta}$ if $\theta$ is the true state and $f(\theta)$ is the best outcome within the choice sets for both agents in $\mathcal{G}_{\hat{\theta}}$, then, $f(\theta)$ is the only lottery with such property.

We also provide an intuitive account for why Condition $COM$ gets us compelling implementation. It is well known that Maskin monotonicity is necessary for Nash implementation. It requires that, for any $\theta, \hat{\theta} \in \Theta$, whenever $f(\theta) \neq f(\hat{\theta})$, then there be agent $i \in I$ and lottery $x$ such that $f(\theta) \succeq_i^{\theta} x$, while $x \succ_i^{\hat{\theta}} f(\theta)$. Properties 1-i, 1-ii, 2 and 4 of Condition $COM$ induces a stronger version of Maskin monotonicity: agent $i$ must be one of the two agents in $\mathcal{G}_{\theta}$ and $x$ must be among the lotteries in $C_{\theta}(i)$. Property 3, on the other hand, is used to construct a set of punishment lotteries which are guaranteed to cause a loss of utility for each agent of at least $\varepsilon$ when compared to their utility from the socially desirable alternative. The stronger form of Maskin monotonicity is used to control how the set of Nash Equilibria appear in our canonical mechanism, while the punishments are used to ensure that any mixed equilibrium either yields a socially desirable outcome or is uncompelling.

We now provide an example with three agents in which Condition $COM$ is satisfied.

**Example 1** *There are three agents, i.e.,* $I = \{0, 1, 2\}$. *Let* $A = \{a, b, c, d, z\}$ *be the set of pure alternatives,* $\Theta = \{\theta_a, \theta_b, \theta_c, \theta_d\}$ *be the set of states, and* $f$ *be the SCF such that* $f(\theta_x) = x$ *for any* $x \in A \backslash \{z\}$. *Agents' preferences over* $A \backslash \{z\} = \{a, b, c, d\}$ *are summarized in Table 5. In addition, for any* $x \in A \backslash \{z\}$, *any* $i \in I$, *and any* $\theta \in \Theta$, *we assume* $x \succ_i^{\theta} z$, *so* $z$ *is the common worst outcome across states.*[12]

*Fix* $\varepsilon > 0$ *as an arbitrary small number. For each* $i \in I$ *and* $\tilde{\theta} \in \Theta$, *we define*

$$\mathcal{U}_i^{\hat{\theta}, \varepsilon} = \left\{ u_i(\cdot, \tilde{\theta}) \in [0,1]^5 \middle| \; u_i(z, \tilde{\theta}) = 0, \max_{\tilde{a} \in A} u_i(\tilde{a}, \tilde{\theta}) = 1, \text{ and } u_i(\tilde{a}, \tilde{\theta}) \geq \varepsilon, \forall \tilde{a} \in A \backslash \{z\} \right\},$$

[12] It is worthwhile to mention that we do not need the common worst outcome across states, such as $z$, in this example. In particular, $z$ can be state dependent and it does not have to be the worst outcome for any agent. What we really need is that every state-dependent $z_{\theta}$ is worse than the social choice outcome in any state for any agent and $z_{\theta}$ needs to be in both agents' choice sets in the forum $\mathcal{G}_{\theta}$

| State | Agent 0 | Agent 1 | Agent 2 |
|---|---|---|---|
| $\theta_a$ | $a \succ_0^{\theta_a} b \succ_0^{\theta_a} d \succ_0^{\theta_a} c$ | $a \succ_1^{\theta_a} b \succ_1^{\theta_a} d \succ_1^{\theta_a} c$ | $a \succ_2^{\theta_a} b \succ_2^{\theta_a} c \succ_2^{\theta_a} d$ |
| $\theta_b$ | $a \succ_0^{\theta_b} b \succ_0^{\theta_b} c \succ_0^{\theta_b} d$ | $b \succ_1^{\theta_b} a \succ_1^{\theta_b} c \succ_1^{\theta_b} d$ | $b \succ_2^{\theta_b} a \succ_2^{\theta_b} d \succ_2^{\theta_b} c$ |
| $\theta_c$ | $a \succ_0^{\theta_c} b \succ_0^{\theta_c} c \succ_0^{\theta_c} d$ | $b \succ_1^{\theta_c} a \succ_1^{\theta_c} c \succ_1^{\theta_c} d$ | $a \succ_2^{\theta_c} b \succ_2^{\theta_c} c \succ_2^{\theta_c} d$ |
| $\theta_d$ | $d \succ_0^{\theta_d} a \succ_0^{\theta_d} b \succ_0^{\theta_d} c$ | $b \succ_1^{\theta_d} a \succ_1^{\theta_d} c \succ_1^{\theta_d} d$ | $a \succ_2^{\theta_d} b \succ_2^{\theta_d} d \succ_2^{\theta_d} c$ |

Table 5: Agents' Preferences over $A\backslash\{z\}$

*as the set of all possible cardinal utility functions $u_i(\cdot,\tilde{\theta})$ that are consistent with ordinal preferences $\succeq_i^{\tilde{\theta}}$ given in Table 5. Let $\mathcal{U}^{\tilde{\theta},\varepsilon} \equiv \times_{i\in I}\mathcal{U}_i^{\tilde{\theta}}$ and $\mathcal{U}^{\varepsilon} \equiv \times_{\tilde{\theta}\in\Theta}\mathcal{U}^{\tilde{\theta},\varepsilon}$. We construct the following collection of forums $\{\mathcal{F}_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta} = \{\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$ with the following properties:*

- *for each $\theta \in \Theta$, we set $z_\theta = z$.*
- *At state $\theta_a$ we set $\mathcal{G}_{\theta_a} = \{0,1\}$, $w_{\theta_a}(0) = 0$, $w_{\theta_a}(1) = 1$, and $C_{\theta_a}(0) = C_{\theta_a}(1) = \{a,b,c,d,z\}$.*
- *At state $\theta_b$, we set $\mathcal{G}_{\theta_b} = \{1,2\}$, $w_{\theta_b}(0) = 1$, $w_{\theta_b}(1) = 2$, and $C_{\theta_b}(1) = C_{\theta_b}(2) = \{a,b,c,d,z\}$.*
- *At state $\theta_c$, we set $\mathcal{G}_{\theta_c} = \{1,2\}$, $w_{\theta_c}(0) = 1$, $w_{\theta_c}(1) = 2$, and $C_{\theta_c}(1) = C_{\theta_c}(2) = \{c,d,z\}$.*
- *Finally, at state $\theta_d$, we set $\mathcal{G}_{\theta_d} = \{0,2\}$, $w_{\theta_d}(0) = 0$, $w_{\theta_d}(1) = 2$, and $C_{\theta_d}(0) = C_{\theta_c}(2) = \{c,d,z\}$.*

*We show in the Appendix that all properties of Condition $COM$ are satisfied.*

## 5.3 The Canonical Mechanism

Condition $COM$ is utilized to construct our canonical mechanism that achieves compelling implementation. By Condition $COM$, we can fix a collection of forums $\{\mathcal{F}_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$ that satisfies all the properties in Condition $COM$. Recall that we assume that there are at least three agents, i.e., $n \geq 3$. Fix $\theta^* \in \Theta$ and $k \geq 2$ as a state and an integer, respectively. We write $\Gamma^k = (M^k, g^k)$ as a mechanism. We define $M_i^k \equiv M_i^1 \times M_i^2 \times M_i^3$ as agent $i$'s message space in the mechanism $\Gamma^k$. Let $m_i = (m_i^1, m_i^2, m_i^3) \in M_i$ be agent $i$'s generic message such that (i) $m_i^1 \in M_i^1 = \Theta$; (ii) $m_i^2 = (m_i^2[\tilde{\theta}])_{\tilde{\theta}\in\Theta} \in M_i^2 = \times_{\tilde{\theta}\in\Theta} M_i^2[\tilde{\theta}]$ where $m_i^2[\tilde{\theta}] \in \{0,\ldots,k-1\}$; and (iii) $m_i^3 = (m_i^3[\tilde{\theta}])_{\tilde{\theta}\in\Theta} \in M_i^3 \equiv \times_{\tilde{\theta}\in\Theta} M_i^3[\tilde{\theta}]$ where, for all $i \in I$, $M_i^3[\tilde{\theta}] = C_{\tilde{\theta}}(i)$ if $i \in \mathcal{G}_{\tilde{\theta}}$ and $M_i^3[\tilde{\theta}] = \{\emptyset\}$ if $i \notin \mathcal{G}_{\tilde{\theta}}$. In words, each agent $i$ announces a state, a collection of state-contingent integers between 0 and $k-1$, and a collection of state-contingent outcomes such that each outcome in state $\tilde{\theta}$ is required to be chosen from $C_{\tilde{\theta}}(i)$ only if $i \in \mathcal{G}_{\tilde{\theta}}$. We thus define $M = \times_{i \in I} M_i$ as the set of message profiles in the mechanism $\Gamma^k$. For any $m \in M$, we define $\theta^m \in \Theta$ as follows:

$$\theta^m = \begin{cases} \theta' & \text{if there exists } \theta' \in \Theta \text{ such that } |\{j \in I | m_j^1 = \theta'\}| > n/2, \\ \theta^* & \text{otherwise.} \end{cases}$$

Note that $\theta^m$ is well defined because we assume $n \geq 3$. Thus, $\theta^m$ is determined by the simple majority rule.

For any $m \in M$, $g^k(m)$ induces the following two rules:

**Rule 1**: If there exists $q \in \{0,1\}$ such that $q = \left[\sum_{j \in \mathcal{G}_{\theta^m}} m_j^2[\theta^m] \pmod{k}\right]$, then

$$g^k(m) = m_{i^*}^3[\theta^m],$$

where $i^* = w_{\theta^m}(q)$. Rule 1 says that after $\theta^m$ is selected, the two agents in $\mathcal{G}_{\theta^m}$ will play a modified version of the modulo game using the integer in the second component of their messages that corresponds to state $\theta^m$. In this modulo game, if the modulo sum given by $q$ is equal to either 0 or 1, there is a winner, given by $i^* = w_{\theta^m}(q)$, which then dictates the outcome as one of the lotteries chosen from his choice set. However, if the modulo sum results in any other values, then there are no winners and we move to Rule 2 below.

**Rule 2**: If $\left[\sum_{j \in \mathcal{G}_{\theta^m}} m_j^2[\theta^m] \pmod{k}\right] > 1$, then

$$g^k(m) = z_{\theta^m}.$$

Rule 2 says that when two agents in $\mathcal{G}_{\theta^m}$ play the modified modulo game and there are no winners, the mechanism induces lottery $z_{\theta^m}$.

Our canonical mechanism can be understood as the following two-step procedure: the first step is to identify the state $\theta^m$ by the simple majority rule, where each agent $i$'s vote is represented by the first component of his message, $m_i^1$. The second step is to pick $\mathcal{G}_{\theta^m}$ and ask each agent $i \in \mathcal{G}_{\theta^m}$ to pick a lottery from his choice set $C_{\theta^m}(i)$. If $\theta^m$ corresponds to the true state $\theta$, then Condition $COM$ ensures that not only $f(\theta)$ will be in their choice sets, but also that both agents will select it in the third component of their message $m_i^3[\theta^m]$, and furthermore that they have the right incentive to coordinate the second component of their message $m_i^2[\theta^m]$ so that a winner is chosen in the modified modulo game. However, if $\theta^m \neq \theta$, then Condition $COM$ ensures that there is a mixed strategy equilibrium in which both agents in $\mathcal{G}_{\theta^m}$ pick different lotteries. This undermines their incentives to coordinate on the modulo game, resulting in Rule 2 being triggered with high probability, so that, if we choose $k$ large enough, we can choose an arbitrarily high probability that the punishment lotteries will be chosen. Then, Condition $COM$ further ensures that each agent's expected utility from this mixed strategy equilibrium is worse than if they induce $\theta^m = \theta$, making the focused mixed strategy equilibrium uncompelling.

## 5.4 Main Theorem

**Theorem 1** *Let $f$ be an SCF. Suppose that the finite environment $\mathcal{E} = \left(I, A, \Theta, (\succeq_i^{\theta})_{i\in I,\theta\in\Theta}\right)$ satisfies Condition $COM$ with respect to $f$ and $\hat{\mathcal{U}}$. Then, the SCF $f$ is compellingly implementable with respect to $\hat{\mathcal{U}}$.*

**Proof**: Suppose that $\mathcal{E}$ satisfies Condition $COM$ with respect to the SCF $f$ and $\hat{\mathcal{U}}$. Therefore, throughout the proof of the theorem, we fix a collection of forums $(\mathcal{F}_{\tilde{\theta}})_{\tilde{\theta}\in\Theta} = (\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}})_{\tilde{\theta}\in\Theta}$ that satisfies Properties 1 through 4 of Condition $COM$ with respect to $f$ and $\hat{\mathcal{U}}$. We prove this theorem through a series of steps.

**Step 1**: For any $k \geq 2$, the SCF $f$ is pure Nash implementable by the mechanism $\Gamma^k$.

**Proof of Step 1**: Let $\theta \in \Theta$ be a true state. Fix $u \in \hat{\mathcal{U}}$ arbitrarily. It follows from Property 1 of Condition $COM$ that there exists $i = w_\theta(0)$ for whom $f(\theta) \in C_\theta(i)$. Let $m \in M$ be a message profile with the following properties:

- $m_j^1 = \theta$ for all $j \in I$;
- $m_j^2[\theta] = 0$ for all $j \in \mathcal{G}_\theta$;
- $m_i^3[\theta] = f(\theta)$.

Since $i = w_\theta(0)$ where $\sum_{j\in\mathcal{G}_\theta} m_j^2[\theta] \pmod{k} = 0$ under Rule 1, it follows that $g(m) = f(\theta)$. We next claim that $m$ is a pure strategy Nash equilibrium in the game $\Gamma^k(\theta)$. First, since there are at least three agents (i.e., $n \geq 3$), no agent can unilaterally change $\theta^m = \theta$. Thus, every agent $i$ cannot find any profitable deviation from $m$ when restricting her deviation strategy to $M_i^1$. Hence, any profitable unilateral deviation of agent $i$ from $m$, if any, must involve the change of her message in either $M_i^2$, $M_i^3$, or both. By construction of the mechanism $\Gamma^k$, the only agents who can unilaterally change the outcome from $g^k(m)$ are those who are in $\mathcal{G}_\theta$. We also know that each agent $j \in \mathcal{G}_\theta$ can only induce outcomes within $C_\theta(j)$ by her unilateral deviation from $m$. Property 1-i of Condition $COM$ ensures that any agent $j \in \mathcal{G}_\theta$ finds $f(\theta)$ as her best outcome within $C_\theta(j)$ in state $\theta$.

Therefore, no agent $j \in \mathcal{G}_\theta$ can find any profitable unilateral deviation by inducing Rules 1 or 2. Thus, $m$ is indeed a pure strategy Nash equilibrium in the game $\Gamma^k(\theta)$.

Now we show that $m \in pureNE(\Gamma^k(\theta))$ implies $g(m) = f(\theta)$. We assume by way of contradiction that there exists $m \in pureNE(\Gamma(\theta))$ such that $g(m) \neq f(\theta)$.

Since $g(m) \neq f(\theta)$, it follows from Property 2 that, for any $\tilde{\theta} \in \Theta$, $g(m)$ is "not" $C_{\tilde{\theta}}$-acceptable at state $\theta$. In particular, we have that $g(m)$ is not $C_{\theta^m}$-acceptable at state $\theta$. This implies that there exist $i \in \mathcal{G}_{\theta^m}$ and $x \in C_{\theta^m}(i)$ such that $u_i(x,\theta) > u_i(g(m),\theta)$. We define $\hat{m}_i$ to be identical to $m_i$ except that $\hat{m}_i^3[\theta^m] = x$ and $\hat{m}_i^2[\theta^m]$ such that agent $i$ is the modulo game winner, i.e.,

$$w_{\theta^m}\left(\left(\hat{m}_i^2[\theta^m] + m_j^2[\theta^m]\right) \mod k\right) = i.$$

Then, agent $i$ has a profitable unilateral deviation from $m$. This shows that $m$ is not a pure strategy Nash equilibrium in the game $\Gamma^k(\theta)$, which is a desired contradiction. Thus, $f$ is pure Nash implementable by mechanism $\Gamma^k$. ■

Throughout the proof, we denote by $\theta$ the true state and by $\hat{\theta}$ the state determined by the agents' announcement in the mechanism. Let $\Gamma^k = (M^k, g^k)$ be our canonical mechanism where $k \geq 3$. We define $C_{\hat{\theta}} \equiv \bigcup_{i \in I} C_{\hat{\theta}}(i)$ for each $\hat{\theta} \in \Theta$, and $C \equiv \bigcup_{\hat{\theta} \in \Theta} C_{\hat{\theta}}$. Note that $C_{\hat{\theta}}$ and $C$ are both finite. For each $\hat{\theta} \in \Theta$, $q \in \{0, \ldots, k-1\}$, $i \in I$, and $x \in C$, we define

$$M^*(\hat{\theta}, q, x) = \left\{ m \in M^k \,\middle|\, \theta^m = \hat{\theta},\ \sum_{j \in \mathcal{G}_{\hat{\theta}}} m_j^2[\hat{\theta}] \pmod{k} = q,\ g^k(m) = x \right\}$$

as a subset of $M^k$. Notice that as $M^*(\hat{\theta}, q, x)$ requires a lot structure on itself, it may well be empty for certain combinations of $\hat{\theta}$, $q$ and $x$; in particular, it is empty for all $(\hat{\theta}, q, x)$ with $q \geq 2$ and $x \neq z_{\hat{\theta}}$. By construction, we have

$$\bigcup_{\hat{\theta} \in \Theta} \bigcup_{q \in \{0,\ldots,k-1\}} \bigcup_{x \in C} M^*(\hat{\theta}, q, x) = M^k.$$

Let $\sigma$ be a mixed strategy profile in the mechanism $\Gamma^k$. For any $\hat{\theta} \in \Theta$, $q \in \{0, \ldots, k-1\}$, and $x \in C$, we define

$$P^\sigma(\hat{\theta}, q, x) \equiv \sum_{m \in M^*(\hat{\theta}, q, x)} \sigma(m).$$

For each $\hat{\theta} \in \Theta$ and $q \in \{0, \ldots, k-1\}$, we define

$$M^*(\hat{\theta}, q) = \bigcup_{x \in C} M^*(\hat{\theta}, q, x) \text{ and } P^\sigma(\hat{\theta}, q) = \sum_{m \in M^*(\hat{\theta}, q)} \sigma(m).$$

For any $\hat{\theta} \in \Theta$, we define

$$M^*(\hat{\theta}) = \bigcup_{q \in \{0,\dots,k-1\}} M^*(\hat{\theta}, q) \text{ and } P^\sigma(\hat{\theta}) = \sum_{m \in M^*(\hat{\theta})} \sigma(m).$$

We can now define conditional probabilities as well:

$$P^\sigma(q, x|\hat{\theta}) = \begin{cases} P^\sigma(\hat{\theta}, q, x)/P^\sigma(\hat{\theta}) & \text{if } P^\sigma(\hat{\theta}) > 0, \\ 0 & \text{if } P^\sigma(\hat{\theta}) = 0. \end{cases}$$

$$P^\sigma(q|\hat{\theta}) = \begin{cases} P^\sigma(\hat{\theta}, q)/P^\sigma(\hat{\theta}) & \text{if } P^\sigma(\hat{\theta}) > 0, \\ 0 & \text{if } P^\sigma(\hat{\theta}) = 0. \end{cases}$$

$$P^\sigma(x|q, \hat{\theta}) = \begin{cases} P^\sigma(q, x|\hat{\theta})/P^\sigma(q|\hat{\theta}) & \text{if } P^\sigma(q|\hat{\theta}) > 0, \\ 0 & \text{if } P^\sigma(q|\hat{\theta}) = 0. \end{cases}$$

We define the set of message profiles in which $\hat{\theta}$ is the agreed-upon state chosen by the mechanism and agent $i$ sends $m_i^2[\hat{\theta}] = q$:

$$M^*(\hat{\theta}, q, i) = \left\{ m \in M^k \,\middle|\, \theta^m = \hat{\theta},\ m_i^2[\hat{\theta}] = q \right\}.$$

Using this set, we define $P_i^\sigma(q|\hat{\theta})$ as the probability that agent $i$ sends $q$ under $\sigma$ conditional on $M^*(\hat{\theta}, q, i)$:

$$P_i^\sigma(q|\hat{\theta}) = \begin{cases} \sum_{m_i:(m_i, m_{-i}) \in M^*(\hat{\theta}, q, i)} \sigma_i(m_i)/P^\sigma(\hat{\theta}) & \text{if } P^\sigma(\hat{\theta}) > 0, \\ 0 & \text{if } P^\sigma(\hat{\theta}) = 0. \end{cases}$$

Lastly, for each $\hat{\theta} \in \Theta$ and each $q \in \{0, 1\}$, define the following lottery:

$$\ell_q^k(\hat{\theta}, \sigma) \equiv \sum_{x \in C} P^\sigma(x|q, \hat{\theta}) x.$$

**Step 2**: For any mixed strategy profile $\sigma$ in the mechanism $\Gamma^k = (M^k, g^k)$, $g^k(\sigma)$ can be represented by the following multiple forms:

$$\begin{aligned} g^k(\sigma) &= \sum_{\hat{\theta} \in \Theta} \sum_{x \in C} \left[ \sum_{q \in \{0,1\}} P^\sigma(\hat{\theta}, q, x) x + \sum_{q \in \{2,\dots,k-1\}} P^\sigma(\hat{\theta}, q, x) z_{\hat{\theta}} \right] \\ &= \sum_{\hat{\theta} \in \Theta} P^\sigma(\hat{\theta}) \sum_{x \in C} \left[ \sum_{q \in \{0,1\}} P^\sigma(q, x|\hat{\theta}) x + \sum_{q \in \{2,\dots,k-1\}} P^\sigma(q, x|\hat{\theta}) z_{\hat{\theta}} \right] \\ &= \sum_{\hat{\theta} \in \Theta} P^\sigma(\hat{\theta}) \ell^k(\hat{\theta}, \sigma), \end{aligned}$$

where, for each $\hat{\theta} \in \Theta$,

$$\begin{aligned}
\ell^k(\hat{\theta}, \sigma) &\equiv \sum_{x \in C} \left[ \sum_{q \in \{0,1\}} P^\sigma(q, x|\hat{\theta})x + \sum_{q \in \{2,\ldots,k-1\}} P^\sigma(q, x|\hat{\theta})z_{\hat{\theta}} \right] \\
&= \sum_{q \in \{0,1\}} \sum_{x \in C} P^\sigma(q, x|\hat{\theta})x + \sum_{q \in \{2,\ldots,k-1\}} \sum_{x \in C} P^\sigma(q, x|\hat{\theta})z_{\hat{\theta}} \\
&= \sum_{q \in \{0,1\}} P^\sigma(q|\hat{\theta})\ell_q^k(\hat{\theta}, \sigma) + \sum_{q \in \{2,\ldots,k-1\}} P^\sigma(q|\hat{\theta})z_{\hat{\theta}} \\
&= P^\sigma(q=0|\hat{\theta})\ell_0^k(\hat{\theta}, \sigma) + P^\sigma(q=1|\hat{\theta})\ell_1^k(\hat{\theta}, \sigma) + \left( 1 - \sum_{q \in \{0,1\}} P^\sigma(q|\hat{\theta}) \right) z_{\hat{\theta}} \\
&= P^\sigma(q=0|\hat{\theta})(\ell_0^k(\hat{\theta}, \sigma) - z_{\hat{\theta}}) + P^\sigma(q=1|\hat{\theta})(\ell_1^k(\hat{\theta}, \sigma) - z_{\hat{\theta}}) + z_{\hat{\theta}}
\end{aligned}$$

**Proof of Step 2:** This comes from the construction of our mechanism. ■

**Step 3**: Let $\sigma \in NE(\Gamma^k(\theta, u))$ for some $u \in \hat{\mathcal{U}}$. Then, for any $m \in \text{supp}(\sigma)$, $q \in \{0, 1\}$, and $\hat{\theta} \in \Theta$, if $m \in M^*(\hat{\theta}, q)$, then $m_i^3[\hat{\theta}] \in C^*_{\hat{\theta}}(i, \theta)$, where $i = w_{\hat{\theta}}(q)$.

**Proof of Step 3:** Fix $u \in \hat{\mathcal{U}}$. Let $\sigma \in NE(\Gamma^k(\theta, u))$. Fix $m \in \text{supp}(\sigma)$, $q \in \{0, 1\}$, and $\hat{\theta} \in \Theta$. Assume that $m \in M^*(\hat{\theta}, q)$, and let $i = w_{\hat{\theta}}(q)$. Suppose, by way of contradiction, that $m_i^3[\hat{\theta}] \notin C^*_{\hat{\theta}}(i, \theta)$. We define $\bar{m}_i$ to be identical to $m_i$ except that $\bar{m}_i^3[\hat{\theta}] = c^*_{\hat{\theta}}(i, \theta)$ for some $c^*_{\hat{\theta}}(i, \theta) \in C^*_{\hat{\theta}}(i, \theta)$. By construction, we know that $\bar{m}_i$ weakly dominates $m_i$. We next define $\bar{\sigma}_i$ to be the following deviation strategy: for any $\tilde{m}_i \in M_i$,

$$\bar{\sigma}_i(\tilde{m}_i) = \begin{cases} \sigma_i(\bar{m}_i) + \sigma_i(m_i) & \text{if } \tilde{m}_i = \bar{m}_i, \\ 0 & \text{if } \tilde{m}_i = m_i, \\ \sigma_i(\tilde{m}_i) & \text{otherwise.} \end{cases}$$

We compute the following utility difference:

$$\begin{aligned}
& u_i(g^k(\bar{\sigma}_i, \sigma_{-i}), \theta) - u_i(g^k(\sigma_i, \sigma_{-i}), \theta) \\
= & \sum_{\tilde{m}_{-i}} \sigma_i(m_i) \left[ u_i(g^k(\bar{m}_i, \tilde{m}_{-i}), \theta) - u_i(g^k(m_i, \tilde{m}_{-i}), \theta) \right] \\
= & \sigma(m) \left[ u_i(g^k(\bar{m}_i, m_{-i}), \theta) - u_i(g^k(m_i, m_{-i}), \theta) \right] \\
& + \sum_{\tilde{m}_{-i} \neq m_{-i}} \sigma(m_i, \tilde{m}_{-i}) \left[ u_i(g^k(\bar{m}_i, \tilde{m}_{-i}), \theta) - u_i(g^k(m_i, \tilde{m}_{-i}), \theta) \right] \\
\geq & \sigma(m) \left[ u_i(g^k(\bar{m}_i, m_{-i}), \theta) - u_i(g^k(m_i, m_{-i}), \theta) \right] \\
= & \sigma(m) \left[ u_i(c^*_{\hat{\theta}}(i, \theta), \theta) - u_i(m^3_i[\tilde{\theta}], \theta) \right] \\
> & 0,
\end{aligned}$$

where the weak inequality follows because $\bar{m}_i$ weakly dominates $m_i$, and the strict inequality follows because $\sigma(m) > 0$ and $u_i(c^*_{\hat{\theta}}(i, \theta)) > u_i(m^3_i[\hat{\theta}], \theta)$, as $c^*_{\hat{\theta}}(i, \theta) \in C^*_{\hat{\theta}}(i, \theta)$ and $m^3_i[\hat{\theta}] \notin C^*_{\hat{\theta}}(i, \theta)$. This shows that $\sigma$ is not a Nash equilibrium in the game $\Gamma^k(\theta, u)$, which is the desired contradiction. Thus, we complete the proof. ■

**Step 4**: Let $\sigma$ be a mixed strategy profile in the mechanism $\Gamma^K$, where we later choose $K \geq 3$ large enough, and fix $\hat{\theta} \in \Theta$ such that $P^\sigma(\hat{\theta}) > 0$. Assume that $P^\sigma(q = 0|\hat{\theta}) + P^\sigma(q = 1|\hat{\theta}) \leq 2/K$ . Then, there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$ and $u \in \hat{\mathcal{U}}$.

**Proof of Step 4:** Fix $\hat{\theta} \in \Theta$ with $P^\sigma(\hat{\theta}) > 0$, $u \in \hat{\mathcal{U}}$, and $i \in I$ arbitrarily. Since the utility achievable is bounded above from 1 and the compound lottery $\ell^K(\hat{\theta}, \sigma)$ induces $z_{\hat{\theta}}$ with probability equal to at least $1 - 2/K$, we have

$$\begin{aligned}
u_i(\ell^K(\hat{\theta}, \sigma), \theta) & \leq \frac{2}{K} \cdot 1 + \frac{K-2}{K} u_i(z_{\hat{\theta}}, \theta) \\
& = u_i(z_{\hat{\theta}}, \theta) + \frac{2\,(1 - u_i(z_{\hat{\theta}}, \theta))}{K} \\
& \leq u_i(z_{\hat{\theta}}, \theta) + \frac{2}{K}
\end{aligned}$$

By Property 3 of Condition $COM$, we have that, for any $i \in I$ and $u \in \hat{\mathcal{U}}$,

$$u_i(z_{\hat{\theta}}, \theta) \leq u_i(f(\theta), \theta) - \varepsilon.$$

Hence, we combine this with the previous inequality obtained so that

$$u_i(\ell^K(\hat{\theta}, \sigma), \theta) \leq u_i(f(\theta), \theta) - \varepsilon + \frac{2}{K}.$$

If we choose $K$ to be the smallest integer such that $K \geq 4/\varepsilon$, we have

$$u_i(\ell^K(\hat{\theta}, \sigma), \theta) \leq u_i(f(\theta), \theta) - \frac{\varepsilon}{2},$$

for any $i \in I$ and $u \in \hat{\mathcal{U}}$. Since $\varepsilon > 0$, we have that $u_i(\ell^K(\hat{\theta}, \theta)) < u_i(f(\theta), \theta)$ for any $i \in I$ and $u \in \hat{\mathcal{U}}$, as desired. This completes the proof. ■

From here till the end of the proof of Step 7, we fix $u \in \hat{\mathcal{U}}$, $\hat{\theta} \in \Theta$, and adopt the convention that $w_{\hat{\theta}}(0) = i$ and $w_{\hat{\theta}}(1) = j$. We focus now on the lotteries that emerge from the remaining components of a strategy profile $\sigma$ after a particular state $\hat{\theta}$ is selected by the mechanism, which we denote by $\ell^k(\hat{\theta}, \sigma)$. Steps 5, 6 and 7 are all used to show that if $\sigma$ is a Nash equilibrium of the game $\Gamma^K(\theta, u)$ and $f(\theta) \neq \ell^K(\hat{\theta}, \sigma)$, all agents must prefer $f(\theta)$ to $\ell^K(\hat{\theta}, \sigma)$ where we choose $K$ large enough.

Let $m_j \in \text{supp}(\sigma_j)$, $\bar{m}_j$ be agent $j$'s arbitrary message sent, and $\sigma_{-j}$ be other players' strategy profile. Let $(m_j, \sigma_{-j})$ denote the strategy profile in which agent $j$ plays $m_j$ and other agents play $\sigma_{-j}$, and $(\bar{m}_j, \sigma_{-j})$ be the strategy profile in which agent $j$ plays $\bar{m}_j$ and other agents play $\sigma_{-j}$. These two strategy profiles will induce different lotteries, $\ell^k(\hat{\theta}, (m_j, \sigma_{-j}))$ and $\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_{-j}))$, respectively. Using Step 2, we compute the difference in expected payoff for agent $j$ at state $\theta$ between these two lotteries as:

$$\begin{aligned}
& u_j(\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^k(\hat{\theta}, (m_j, \sigma_j)), \theta) \\
= \; & P^{(\bar{m}_j, \sigma_{-j})}(q = 0|\hat{\theta}) \left( u_j(\ell_0^k(\hat{\theta}, (m_j, \sigma_{-j})), \theta) - u_j(z_{\hat{\theta}}, \theta) \right) \\
+ \; & P^{(\bar{m}_j, \sigma_{-j})}(q = 1|\hat{\theta}) \, (\bar{u}_1 - u_j(z_{\hat{\theta}}, \theta)) + u_j(z_{\hat{\theta}}, \theta) \\
- \; & \left[ P^{(m_j, \sigma_{-j})}(q = 0|\hat{\theta}) \left( u_j(\ell_0^k(\hat{\theta}, (m_j, \sigma_{-j})), \theta) - u_j(z_{\hat{\theta}}, \theta) \right) \right. \\
+ \; & \left. P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) \left( u_j(m_j^3[\hat{\theta}], \theta) - u_j(z_{\hat{\theta}}, \theta) \right) + u_j(z_{\hat{\theta}}, \theta) \right]
\end{aligned}$$

To ease the notation, we will adopt the following set of conventions:

$$\begin{aligned}
P^{(m_j, \sigma_{-j})}(q = 0|\hat{\theta}) &= p_0; \; P^{(\bar{m}_j, \sigma_{-j})}(q = 0|\hat{\theta}) = \bar{p}_0; P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) = p_1; \\
P^{(\bar{m}_j, \sigma_{-j})}(q = 1|\hat{\theta}) &= \bar{p}_1; \; u_j(\ell_0^k(\hat{\theta}, (m_j, \sigma_{-j})), \theta) = u_0; \\
u_j(m_j^3[\hat{\theta}], \theta) &= u_1; \; u_j(\bar{m}_j^3[\hat{\theta}], \theta) = \bar{u}_1; u_j(z_{\hat{\theta}}, \theta) = u_z.
\end{aligned}$$

This introduced notation allows us to simply the previous expression to

$$\begin{aligned}
& u_j(\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^k(\hat{\theta}, (m_j, \sigma_j)), \theta) \\
= \; & \bar{p}_0(u_0 - u_z) + \bar{p}_1(\bar{u}_1 - u_z) + u_z - [p_0(u_0 - u_z) + p_1(u_1 - u_z) + u_z.]
\end{aligned}$$

The above expression will be extensively used in the rest of the proof. For any $x \in \{0,1\}$ and $q \in \{0,\dots,K-1\}$, we define $b_x(q) \in \{0,\dots,K-1\}$ such that $q + b_x(q) \pmod K = x$.

**Step 5**: Let $\sigma \in NE(\Gamma^K(\theta,u))$, where we later choose $K$ large enough. For any $\hat{\theta} \in \Theta$ with $P^\sigma(\hat{\theta}) > 0$, we assume that there exists $q \in \{0,1\}$ such that $P^\sigma(q|\hat{\theta}) = 0$. Then, there exists $K \in \mathbb{N}$ large enough so that, for any $\hat{\theta} \in \Theta$ with $P^\sigma(\hat{\theta}) > 0$, if $\ell^K(\hat{\theta},\sigma) \neq f(\theta)$, then $u_i(\ell^K(\hat{\theta},\sigma),\theta) < u_i(f(\theta),\theta)$ for all $i \in I$.

**Proof of Step 5:** Fix $\sigma \in NE(\Gamma^K(\theta,u))$, and $\hat{\theta} \in \Theta$ with $P^\sigma(\hat{\theta}) > 0$ arbitrarily. We assume that there exists $q \in \{0,1\}$ such that $P^\sigma(q|\hat{\theta}) = 0$. We further assume that $\ell^K(\hat{\theta},\sigma) \neq f(\theta)$. We divide the proof into each of the following two cases.

**Case 1**: $P^\sigma(q|\hat{\theta}) = 0$ for all $q \in \{0,1\}$

This implies that $\ell^K(\hat{\theta},\sigma)$ induces $z_{\hat{\theta}}$ with probability one. It follows from Property 3 of Condition $COM$ that $u_i(\ell^K(\hat{\theta},\sigma),\theta) < u_i(f(\theta),\theta)$ for any $i \in I$, $u \in \hat{\mathcal{U}}$, and $K \geq 3$.

**Case 2**: $P^\sigma(q=0|\hat{\theta}) > 0$ and $P^\sigma(q=1|\hat{\theta}) = 0$

The argument we provide below regarding Case 2 will ensure that we can handle the case that $P^\sigma(q=1|\hat{\theta}) > 0$ and $P^\sigma(q=0|\hat{\theta}) = 0$ in a similar fashion. So we omit this case. With the help of Step 2, we can write

$$\ell^K(\hat{\theta},\sigma) = P^\sigma(q=0|\hat{\theta})(\ell_0^K(\hat{\theta},\sigma) - z_{\hat{\theta}}) + z_{\hat{\theta}}.$$

Fix $c^*_{\hat{\theta}}(j,\theta) \in C^*_{\hat{\theta}}(j,\theta)$. We further divide Case 2 into four sub-cases:

**Case 2-1**: $u_0 \geq u_j(c^*_{\hat{\theta}}(j,\theta),\sigma),\theta) \geq u_z$

This implies that some $c^*_{\hat{\theta}}(i,\theta) \in \text{supp}\ (\ell_0^K(\hat{\theta},\sigma))$ is $C_{\hat{\theta}}$-acceptable at $(\theta,u)$. It then follows from Property 2 of Condition $COM$ that $c^*_{\hat{\theta}}(i,\theta) = f(\theta)$. From Step 4, we also have that $\text{supp}\ (\ell_0^k(\hat{\theta},\sigma)) \subseteq C_{\hat{\theta}}(i,\theta)$. Thus, $f(\theta) \in C^*_{\hat{\theta}}(i,\theta)$, which further implies that $u_i(c,\theta) = u_i(f(\theta),\theta)$ for all $c \in C^*_{\hat{\theta}}(i,\theta)$. Finally, by Property 4 of Condition $COM$, we have $C^*_{\hat{\theta}}(i,\theta) = \{f(\theta)\}$, which further implies $\ell_0^k(\hat{\theta},\sigma) = f(\theta)$. Since we assume that $\ell^K(\hat{\theta},\sigma) \neq f(\theta)$, we must have $P^\sigma(q=0|\hat{\theta}) < 1$. By Property 3 of Condition $COM$, there exists $\varepsilon > 0$ such that $u_j(f(\theta),\theta) - u_j(z_{\hat{\theta}},\theta) \geq \varepsilon$ for all $j \in I$ and $u \in \hat{\mathcal{U}}$. Due to the construction of $\ell^K(\hat{\theta},\theta)$ and the fact that $\ell_0^k(\hat{\theta},\sigma) = f(\theta)$, we conclude that $u_j(\ell^K(\hat{\theta},\sigma),\theta) < u_j(f(\theta),\theta)$ for all $j \in I$.

**Case 2-2**: $u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_z > u_0$.

In this case, we will show that there is no Nash equilibrium $\sigma$ in the game $\Gamma^K(\theta,u)$ such that $P^\sigma(q=0|\hat{\theta}) > 0$ and $P^\sigma(q=1|\hat{\theta}) = 0$. Suppose, on the contrary, that such $\sigma$ constitutes a Nash equilibrium in the game $\Gamma^K(\theta,u)$.

Fix $m \in \text{supp}(\sigma)$ as a message profile such that $\theta^m = \hat{\theta}$. The existence of such $m$ is guaranteed because we have $P^\sigma(\hat{\theta}) > 0$. Note that $P^\sigma(q=1|\hat{\theta}) = 0$ implies that $P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta}) = p_1 = 0$. Our goal is to find a message $\bar{m}_j$, which together with $\sigma_{-j}$ induces a lottery $\ell^K(\hat{\theta},(\bar{m}_j,\sigma_{-j}))$ that first-order stochastically dominates the lottery induced by $(m_j,\sigma_{-j})$, showing that $\sigma$ is not a Nash equilibrium of the game $\Gamma^K(\theta,u)$. To achieve this, we need to find $\bar{m}^2_j[\hat{\theta}] \in \{0,\ldots,K-1\}$, which together with $\sigma_{-j}$ guarantees that $\bar{p}_1 > 0$ and $\bar{p}_0 = 0$. We shall propose an algorithm selecting such $\bar{m}^2_j[\hat{\theta}]$.

Recall the following notation:

$$M^*(\hat{\theta},q,i) = \left\{ m \in M^K \,\middle|\, \theta^m = \hat{\theta},\ m^2_i[\hat{\theta}] = q \right\}.$$

and

$$P^\sigma_i(q|\hat{\theta}) = \sum_{m_i:(m_i,m_{-i})\in M^*(\hat{\theta},q,i)} \sigma_i(m_i)/P^\sigma(\hat{\theta}).$$

We also use this notation in Case 2-3 later. Start the algorithm by setting $q_0 = b_1(m^2_j[\hat{\theta}])$, where, for any $q \in \{0,\ldots,K-1\}$, we define $b_1(q) \in \{0,\ldots,K-1\}$ such that $q + b_1(q) \pmod K = 1$. It follows from $P^\sigma(q=1|\hat{\theta}) = 0$ that $P^\sigma_i(q_0|\hat{\theta}) = 0$. Next, for any $h \in \{1,\cdots,K-1\}$ and $q_{h-1} \in \{0,\ldots,K-1\}$, we define

$$q_h = q_{h-1} + 1 \pmod K.$$

Since $\sum_{q=0}^{K-1} P^\sigma_i(q|\hat{\theta}) = 1$, we can choose $h \in \{1,\ldots,K-1\}$ uniquely in such a way that $P^\sigma_i(q_h|\hat{\theta}) > 0$ and $P^\sigma_i(q_{h'}|\hat{\theta}) = 0$ for all $h' \in \{0,\ldots,h-1\}$. Then, we set $\bar{m}^2_j[\hat{\theta}] = b_1(q_h)$. Define $\bar{m}^3_j[\hat{\theta}]$ as follows:

$$\bar{m}^3_j[\hat{\theta}] = \begin{cases} c^*_{\hat{\theta}}(j,\theta) & \text{if } m^3_j[\hat{\theta}] \notin C^*_{\hat{\theta}}(j,\theta), \\ m^3_j[\hat{\theta}] & \text{if } m^3_j[\hat{\theta}] \in C^*_{\hat{\theta}}(j,\theta). \end{cases}$$

Thus, we define $\bar{m}_j$ to be identical to $m_j$ except that $m^2_j[\hat{\theta}]$ is replaced by $\bar{m}^2_j[\hat{\theta}]$ and $m^3_j[\hat{\theta}]$ is replaced by $\bar{m}^3_j[\hat{\theta}]$. By the algorithm to find $q_h$ and construction of

$\bar{m}_j^2[\hat{\theta}]$, we have the following properties:

$$\begin{aligned}
\bar{p}_1 \equiv P^{(\bar{m}_j,\sigma_{-j})}(q=1|\hat{\theta}) &= P^{(\bar{m}_j,\sigma_{-j})}\Big(q_h + b_1(q_h) \pmod K \big|\hat{\theta}\Big) \\
&= P_i^{(\bar{m}_j,\sigma_{-j})}(q_h|\hat{\theta}) > 0, \\
\bar{p}_0 \equiv P^{(\bar{m}_j,\sigma_{-j})}(q=0|\hat{\theta}) &= P^{(\bar{m}_j,\sigma_{-j})}\Big(q_{h-1} + b_1(q_h) \pmod K \big|\hat{\theta}\Big) \\
&= P_i^{(\bar{m}_j,\sigma_{-j})}(q_{h-1}|\hat{\theta}) = 0.
\end{aligned}$$

It follows from $P^{\sigma}(q=1|\hat{\theta}) = 0$ that $p_1 \equiv P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta}) = 0$, implying $\bar{p}_1 - p_1 > 0$. Since $\bar{p}_0 = 0$, we also have

$$p_0 \equiv P^{(m_j,\sigma_{-j})}(q=0|\hat{\theta}) \geq P^{(\bar{m}_j,\sigma_j)}(q=0|\hat{\theta}) \equiv \bar{p}_0,$$

which implies $p_0 - \bar{p}_0 \geq 0$. Due to the construction of $\bar{m}_j^3[\hat{\theta}]$, we obtain the following inequalities:

$$\begin{aligned}
u_j(\bar{m}_j^3[\hat{\theta}],\theta) = u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_j(z_{\hat{\theta}},\theta) &\quad\Rightarrow\quad \bar{u}_1 - u_z > 0, \\
u_j(\bar{m}_j^3[\hat{\theta}],\theta) = u_j(c^*_{\hat{\theta}}(j,\theta),\theta) \geq u_j(m_j^3[\hat{\theta}],\theta) &\quad\Rightarrow\quad \bar{u}_1 - u_1 \geq 0.
\end{aligned}$$

Now, we claim that $u_j(g(\bar{m}_j,\sigma_j),\theta) - u_j(g(m_j,\sigma_j),\theta) > 0$. Since $P^{\sigma}(\hat{\theta}) > 0$, by Step 2 and the construction of $\bar{m}_j$, it suffices to show that $u_j(\ell^K(\hat{\theta},(\bar{m}_j,\sigma_j)),\theta) - u_j(\ell^K(\hat{\theta},(m_j,\sigma_j)),\theta) > 0$. Thus, we compute

$$\begin{aligned}
& u_j(\ell^K(\hat{\theta},(\bar{m}_j,\sigma_j)),\theta) - u_j(\ell^K(\hat{\theta},(m_j,\sigma_j)),\theta) \\
= \;& \bar{p}_0(u_0 - u_z) + \bar{p}_1(\bar{u}_1 - u_z) + u_z - [p_0(u_0 - u_z) + p_1(u_1 - u_z) + u_z] \\
= \;& (\bar{p}_1 - p_1)(\bar{u}_1 - u_z) + (p_0 - \bar{p}_0)(u_z - u_0) + p_1(\bar{u}_1 - u_1) \\
> \;& 0,
\end{aligned}$$

where the strict inequality follows because $(p_0 - \bar{p}_0)(u_z - u_0) \geq 0$, $p_1(\bar{u}_1 - u_1) \geq 0$, and $(\bar{p}_1 - p_1)(\bar{u}_1 - u_z) > 0$. This contradicts the hypothesis that $\sigma$ is a Nash equilibrium in the game $\Gamma^K(\theta,u)$. ■

**Case 2-3**: $u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_0 > u_z$

In this case, we will show that there is no Nash equilibrium $\sigma$ in the game $\Gamma^K(\theta,u)$ such that $P^{\sigma}(q=0|\hat{\theta}) > 0$ and $P^{\sigma}(q=1|\hat{\theta}) = 0$. Suppose, on the contrary, that such $\sigma$ constitutes a Nash equilibrium in the game $\Gamma^K(\theta,u)$.

Fix $m \in \text{supp}(\sigma)$ as a message profile such that $\theta^m = \hat{\theta}$. The existence of such $m$ is guaranteed because we have $P^{\sigma}(\hat{\theta}) > 0$. Note that $P^{\sigma}(q=1|\hat{\theta}) = 0$ implies that $P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta}) = p_1 = 0$. Our goal here is to find a message $\bar{m}_j$, which together with $\sigma_{-j}$ induces a lottery $\ell^K(\hat{\theta}, (\bar{m}_j, \sigma_{-j}))$ that first-order stochastically dominates the lottery induced by $(m_j, \sigma_{-j})$, showing that $\sigma$ is not a Nash equilibrium of the game $\Gamma^K(\theta, u)$. To achieve this, we need to find $\bar{m}_j^2[\hat{\theta}] \in \{0, \ldots, K-1\}$, which together with $\sigma_{-j}$ guarantees that $\bar{p}_1 > 0$ and $\bar{p}_0 + \bar{p}_1 > p_0 + p_1$. Such $\bar{m}_j^2[\hat{\theta}]$ will be found by the following algorithm:

Start the algorithm by setting $q_0 = b_1(m_j^2[\hat{\theta}])$. It follows from $P^\sigma(q = 1|\hat{\theta}) = 0$ that $P_i^\sigma(q_0|\hat{\theta}) = 0$. Next, for any $h \in \{1, \cdots, K-1\}$ and $q_{h-1} \in \{0, \ldots, K-1\}$, we define

$$q_h = q_{h-1} - 1 \pmod K.$$

Since $\sum_{q=0}^{K-1} P_i^\sigma(q|\hat{\theta}) = 1$, we can choose $h \in \{1, \ldots, K-1\}$ uniquely in such a way that $P_i^\sigma(q_h|\hat{\theta}) > 0$ and $P_i^\sigma(q_{h'}|\hat{\theta}) = 0$ for all $h' \in \{0, \ldots, h-1\}$. Then, we set $\bar{m}_j^2[\hat{\theta}] = b_1(q_h)$. Define also $\bar{m}_j^3[\hat{\theta}]$ as follows:

$$\bar{m}_j^3[\hat{\theta}] = \begin{cases} c^*_{\hat{\theta}}(j, \theta) & \text{if } m_j^3[\hat{\theta}] \notin C^*_{\hat{\theta}}(j, \theta), \\ m_j^3[\hat{\theta}] & \text{if } m_j^3[\hat{\theta}] \in C^*_{\hat{\theta}}(j, \theta), \end{cases}$$

Thus, we define $\bar{m}_j$ to be identical to $m_j$ except that $m_j^2[\hat{\theta}]$ is replaced by $\bar{m}_j^2[\hat{\theta}]$ and $m_j^3[\hat{\theta}]$ is replaced by $\bar{m}_j^3[\hat{\theta}]$.

By the algorithm to find $q_h$ and construction of $\bar{m}_j^2[\hat{\theta}]$, we have the following properties:

$$\begin{aligned} \bar{p}_1 \equiv P^{(\bar{m}_j, \sigma_{-j})}(q = 1|\hat{\theta}) &= P^{(\bar{m}_j, \sigma_{-j})}(q_h + b_1(q_h) \pmod K|\hat{\theta}) \\ &= P_i^{(\bar{m}_j, \sigma_{-j})}(q_h|\hat{\theta}) > 0, \\ &\geq P_i^{(\bar{m}_j, \sigma_{-j})}(q_1|\hat{\theta}) \\ &= P^{(m_j, \sigma_{-j})}(q_0 - 1 + m_j^2[\hat{\theta}] \pmod K|\hat{\theta}) \\ &= P^{(m_j, \sigma_{-j})}(q = 0|\hat{\theta}) \equiv p_0 \ (\because q_0 + m_j^2[\hat{\theta}] \pmod K = 1) \\ P^{(\bar{m}_j, \sigma_{-j})}(q = 2|\hat{\theta}) &= P^{(\bar{m}_j, \sigma_{-j})}(q_{h-1} + b_1(q_h) \pmod K|\hat{\theta}) \\ &= P_i^{(\bar{m}_j, \sigma_{-j})}(q_{h-1}|\hat{\theta}) = 0, \end{aligned}$$

which implies $\bar{p}_1 - p_0 \geq 0$. Since $p_1 \equiv P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) = 0$ and $P^{(\bar{m}_j, \sigma_{-j})}(q = 0|\hat{\theta}) \geq 0$, we have $P^{(\bar{m}_j, \sigma_{-j})}(q = 0|\hat{\theta}) \geq P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta})$, which implies that $\bar{p}_0 - p_1 \geq 0$. In addition, since $p_1 \equiv P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) = 0$, we also have $P^{(\bar{m}_j, \sigma_{-j})}(q = 1|\hat{\theta}) > P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta})$, which implies that $\bar{p}_1 - p_1 > 0$.

Since $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) > u_0 > u_z$, due to the construction of $\bar{m}_j^3[\hat{\theta}]$, we have the following inequality:

$$\bar{u}_1 \equiv u_j(\bar{m}_j^3[\hat{\theta}], \theta) = u_j(c^*_{\hat{\theta}}(j, \theta), \theta) > u_j(\ell_0^K(\hat{\theta}, (m_j, \sigma_{-j})), \theta) \equiv u_0$$

Thus, $\bar{u}_1 - u_0 > 0$.

We claim now that $u_j(g(\bar{m}_j, \sigma_j), \theta) - u_j(g(m_j, \sigma_j), \theta) > 0$. Since $P^\sigma(\hat{\theta}) > 0$, by Step 2, it suffices to show that $u_j(\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^k(\hat{\theta}, (m_j, \sigma_j)), \theta) > 0$. Thus, we compute the following:

$$\begin{aligned} & u_j(\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^k(\hat{\theta}, (m_j, \sigma_j)), \theta) \\ = \ & \bar{p}_0(u_0 - u_z) + \bar{p}_1(\bar{u}_1 - u_z) + u_z - [p_0(u_0 - u_z) + p_1(u_1 - u_z) + u_z] \\ = \ & (\bar{p}_1 - p_1)(\bar{u}_1 - u_0) + (\bar{p}_1 - p_0)(u_0 - u_z) + (\bar{p}_0 - p_1)(u_0 - u_z) + p_1(\bar{u}_1 - u_1) \\ > \ & 0, \end{aligned}$$

where the strict inequality follows because $(\bar{p}_1 - p_0)(u_0 - u_z) \geq 0$, $(\bar{p}_0 - p_1)(u_0 - u_z) \geq 0$, $p_1(\bar{u}_1 - u_1) \geq 0$, and $(\bar{p}_1 - p_1)(\bar{u}_1 - u_0) > 0$. This contradicts the hypothesis that $\sigma$ is a Nash equilibrium in the game $\Gamma^K(\theta, u)$. ■

**Case 2-4**: $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) = u_z > u_0$

We will show that $P^\sigma(q = 0|\hat{\theta}) \leq 1/K$. Suppose not, that is, $P^\sigma(q = 0|\hat{\theta}) > 1/K$. We construct $\bar{\sigma}_j$ to be identical to $\sigma_j$ except that $P_i^{(\bar{\sigma}_j, \sigma_{-j})}(q|\hat{\theta}) = 1/K$ for all $q \in \{0, \cdots, K-1\}$ and $\bar{m}_j^3[\hat{\theta}] = c^*_{\hat{\theta}}(j, \theta)$ for all $m_j \in \text{supp}\ (\bar{\sigma}_j)$. Then, since we have $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) = u_z$, we compute agent $j$'s payoff difference between $(\bar{\sigma}_j, \sigma_{-j})$ and $\sigma$:

$$\begin{aligned} & u_j(g^K(\bar{\sigma}_j, \sigma_{-j}), \theta) - u_j(g^K(\sigma), \theta) \\ = \ & P^\sigma(\hat{\theta}) \left[ \frac{K-1}{K} u_z + \frac{1}{K} u_0 \right] \\ & -P^\sigma(\hat{\theta}) \left[ P^\sigma(q = 0|\hat{\theta}) u_0 + (1 - P^\sigma(q = 0|\hat{\theta})) u_z \right] \\ = \ & P^\sigma(\hat{\theta}) \left( P^\sigma(q = 0|\hat{\theta}) - \frac{1}{K} \right) [u_z - u_0] \\ > \ & 0, \end{aligned}$$

where the strict inequality follows because $P^\sigma(\hat{\theta}) > 0$; $P^\sigma(q = 0|\hat{\theta}) > 1/K$; and $u_z > u_0$. This implies that $\sigma$ is not a Nash equilibrium in the game $\Gamma^K(\theta, u)$, which is the desired contradiction. Thus, we have $P^\sigma(q = 0|\hat{\theta}) + P^\sigma(q = 0|\hat{\theta}) \leq 1/K$. We can then use Step 4 to conclude that there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$. ■

**Step 6**: Let $\sigma \in NE(\Gamma^K(\theta, u))$, where we later choose $K$ large enough. Assume that there exists $\hat{\theta} \in \Theta$ with $P^\sigma(\hat{\theta}) > 0$ such that $P^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. If $f(\theta)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

**Proof of Step 6**: Fix $\sigma \in NE(\Gamma^K(\theta, u))$ and $\hat{\theta} \in \Theta$ such that $P^\sigma(\hat{\theta}) > 0$ and $P^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. We take the contrapositive statement of Property 2 of Condition $COM$: for any $x \in \bigcup_{i \in \mathcal{G}_{\hat{\theta}}} C_{\hat{\theta}}(i)$, if $x \neq f(\theta)$, $x$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$. Let $\mathcal{G}_{\hat{\theta}} = \{i, j\}$ in the rest of the proof. In addition, we know that either

$$f(\theta) \in C_{\hat{\theta}}(i) \cup C_{\hat{\theta}}(j) \text{ or } f(\theta) \notin C_{\hat{\theta}}(i) \cup C_{\hat{\theta}}(j)$$

holds. Since $f(\theta)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, we therefore have the following property: for any $x \in C_{\hat{\theta}}(i) \cup C_{\hat{\theta}}(j)$, $x$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$. This implies that for any $c^*_{\hat{\theta}}(i, \theta) \in C^*_{\hat{\theta}}(i, \theta)$ and any $c^*_{\hat{\theta}}(j, \theta) \in C^*_{\hat{\theta}}(j, \theta)$, we have $u_i(c^*_{\hat{\theta}}(i, \theta), \theta) > u_i(c^*_{\hat{\theta}}(j, \theta), \theta)$ and $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) > u_j(c^*_{\hat{\theta}}(i, \theta), \theta)$.

If either $u_i(c^*_{\hat{\theta}}(i, \theta), \theta) = u_i(z_{\hat{\theta}}, \theta)$ or $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) = u_j(z_{\hat{\theta}}, \theta)$ holds, then we can appeal to an argument identical to the one employed in Case 2-4 of Step 5 to conclude that there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

Therefore, we can assume that both $u_i(c^*_{\hat{\theta}}(i, \theta), \theta) > u_i(z_{\hat{\theta}}, \theta)$ and $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) > u_j(z_{\hat{\theta}}, \theta) \equiv u_z$ hold. We define $\pi_i^{min}$ and $\pi_i^{max}$ as follows:

$$\begin{aligned} \pi_i^{min} &= \min\{P_i^\sigma(q|\hat{\theta}) \in [0,1] | q \in \{0, \dots, K-1\}\}, \\ \text{and } \pi_i^{max} &= \max\{P_i^\sigma(q|\hat{\theta}) \in [0,1] | q \in \{0, \dots, K-1\}\}. \end{aligned}$$

We can also define $\pi_j^{min}$ and $\pi_j^{max}$ in a similar fashion. We will show that $\pi_i^{min} = \pi_i^{max}$ and $\pi_j^{min} = \pi_j^{max}$, which implies that both agents randomize uniformly in their choice of integer. We shall prove this through Steps 6.a and 6.b below.

**Step 6.a**: For any $q^* \in \{0, \dots, K-1\}$, if $P_i^\sigma(q^*|\hat{\theta}) = \pi_i^{min}$ and $\pi_i^{min} < \pi_i^{max}$, then $P_j^\sigma(b_1(q^*)|\hat{\theta}) = 0$, where $b_1(q^*) \in \{0, \dots, K-1\}$ such that $q^* + b_1(q^*) \pmod K = 1$.

**Proof of Step 6.a**: Fix $q^* \in \{0, \dots, K-1\}$ such that $P_i^\sigma(q^*|\hat{\theta}) = \pi_i^{min} < \pi_i^{max}$. Assume, by way of contradiction, that $P_j^\sigma(b_1(q^*)|\hat{\theta}) > 0$. This implies that there exists $m \in \text{supp }(\sigma)$ such that $\theta^m = \hat{\theta}$ and $m_j^2[\hat{\theta}] = b_1(q^*)$. This further implies

$$p_1 \equiv P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) = P^{(m_j, \sigma_{-j})}(q^* + b_1(q^*)|\hat{\theta}) = P_i^{(m_j, \sigma_{-j})}(q^*|\hat{\theta}) = \pi_i^{min}.$$

We claim that $m_j$ is not a best response to $\sigma_{-j}$, contradicting the hypothesis that $\sigma$ is a Nash equilibrium in the game $\Gamma^K(\theta, u)$. To prove this, we consider two possible cases:

**Case 1**: $u_j(c^*_{\hat{\theta}}(j, \theta), \theta) > u_z > u_0$.

Our goal here is to find a message $\bar{m}_j$, which together with $\sigma_{-j}$ induces a lottery $\ell^K(\hat{\theta}, (\bar{m}_j, \sigma_{-j}))$ that first-order stochastically dominates the lottery induced by

$(m_j, \sigma_{-j})$, hence showing that $\sigma$ is not a Nash equilibrium of the game $\Gamma^K(\theta, u)$. To achieve this, we need to find $\bar{m}_j^2[\hat{\theta}] \in \{0, \ldots, K-1\}$, which together with $\sigma_{-j}$ guarantees that $\bar{p}_1 > \pi_i^{min}$ and $\bar{p}_0 = \pi_i^{min}$. We shall propose an algorithm selecting such $\bar{m}_j^2[\hat{\theta}]$.

Recall the following notation:

$$M^*(\hat{\theta}, q, i) = \left\{ m \in M^K \,\middle|\, \theta^m = \hat{\theta},\ m_i^2[\hat{\theta}] = q \right\},$$

and

$$P_i^\sigma(q|\hat{\theta}) = \sum_{m_i:(m_i, m_{-i}) \in M^*(\hat{\theta}, q, i)} \sigma_i(m_i)/P^\sigma(\hat{\theta}).$$

We also use this notation in Case 2 later. Start the algorithm by setting $q_0 = q^*$. Next, for any $h \in \{1, \cdots, K-1\}$ and $q_{h-1} \in \{0, \ldots, K-1\}$, we define

$$q_h = q_{h-1} + 1 \pmod{K}.$$

Since $\pi_i^{min} < \pi_i^{max}$, we can choose $h \in \{1, \ldots, K-1\}$ uniquely in such a way that $P_i^\sigma(q_h|\hat{\theta}) > \pi_i^{min}$ and $P_i^\sigma(q_{h'}|\hat{\theta}) = \pi_i^{min}$ for all $h' \in \{0, \ldots, h-1\}$. Then, we set $\bar{m}_j^2[\hat{\theta}] = b_1(q_h)$. Define $\bar{m}_j^3[\hat{\theta}]$ as follows:

$$\bar{m}_j^3[\hat{\theta}] = \begin{cases} c_{\hat{\theta}}^*(j, \theta) & \text{if } m_j^3[\hat{\theta}] \notin C_{\hat{\theta}}^*(j, \theta), \\ m_j^3[\hat{\theta}] & \text{if } m_j^3[\hat{\theta}] \in C_{\hat{\theta}}^*(j, \theta). \end{cases}$$

Thus, we define $\bar{m}_j$ to be identical to $m_j$ except that $m_j^2[\hat{\theta}]$ is replaced by $\bar{m}_j^2[\hat{\theta}]$ and $m_j^3[\hat{\theta}]$ is replaced by $\bar{m}_j^3[\hat{\theta}]$. By the algorithm selecting $q_h$ and construction of $\bar{m}_j^2[\hat{\theta}]$, we have the following properties:

$$\begin{aligned} \bar{p}_1 \equiv P^{(\bar{m}_j, \sigma_{-j})}(q = 1|\hat{\theta}) &= P^{(\bar{m}_j, \sigma_{-j})}\left(q_h + b_1(q_h) \pmod{K} \middle| \hat{\theta}\right) \\ &= P_i^{(\bar{m}_j, \sigma_{-j})}(q_h|\hat{\theta}) > \pi_i^{min}, \\ \bar{p}_0 \equiv P^{(\bar{m}_j, \sigma_{-j})}(q = 0|\hat{\theta}) &= P^{(\bar{m}_j, \sigma_{-j})}\left(q_{h-1} + b_1(q_h) \pmod{K} \middle| \hat{\theta}\right) \\ &= P_i^{(\bar{m}_j, \sigma_{-j})}(q_{h-1}|\hat{\theta}) = \pi_i^{min}. \end{aligned}$$

It follows from $p_1 \equiv P^{(m_j, \sigma_{-j})}(q = 1|\hat{\theta}) = \pi_i^{min}$ that $\bar{p}_1 - p_1 > 0$. Since $\bar{p}_0 = \pi_i^{min}$, we also have

$$p_0 \equiv P^{(m_j, \sigma_{-j})}(q = 0|\hat{\theta}) \geq P^{(\bar{m}_j, \sigma_j)}(q = 0|\hat{\theta}) \equiv \bar{p}_0,$$

which implies $p_0 - \bar{p}_0 \geq 0$. Due to the construction of $\bar{m}_j^3[\hat{\theta}]$, we obtain the following inequalities:

$$u_j(\bar{m}_j^3[\hat{\theta}], \theta) = u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_j(z_{\hat{\theta}}, \theta) \quad \Rightarrow \quad \bar{u}_1 - u_z > 0,$$
$$u_j(\bar{m}_j^3[\hat{\theta}], \theta) = u_j(c^*_{\hat{\theta}}(j,\theta),\theta) \geq u_j(m_j^3[\hat{\theta}], \theta) \quad \Rightarrow \quad \bar{u}_1 - u_1 \geq 0.$$

We claim now that $u_j(g(\bar{m}_j, \sigma_j), \theta) - u_j(g(m_j, \sigma_j), \theta) > 0$. Since $P^\sigma(\hat{\theta}) > 0$, by Step 2 and the construction of $\bar{m}_j$, it suffices to show that $u_j(\ell^K(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^K(\hat{\theta}, (m_j, \sigma_j)), \theta) > 0$. Thus, we compute

$$\begin{aligned} & u_j(\ell^K(\hat{\theta}, (\bar{m}_j, \sigma_j)), \theta) - u_j(\ell^K(\hat{\theta}, (m_j, \sigma_j)), \theta) \\ = \; & \bar{p}_0(u_0 - u_z) + \bar{p}_1(\bar{u}_1 - u_z) + u_z - [p_0(u_0 - u_z) + p_1(u_1 - u_z) + u_z] \\ = \; & (\bar{p}_1 - p_1)(\bar{u}_1 - u_z) + (p_0 - \bar{p}_0)(u_z - u_0) + p_1(\bar{u}_1 - u_1) \\ > \; & 0, \end{aligned}$$

where the strict inequality follows because $(p_0 - \bar{p}_0)(u_z - u_0) \geq 0$, $p_1(\bar{u}_1 - u_1) \geq 0$, and $(\bar{p}_1 - p_1)(\bar{u}_1 - u_z) > 0$. This contradicts the hypothesis that $\sigma$ is a Nash equilibrium in the game $\Gamma^K(\theta, u)$. ■

**Case 2**: $u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_0 > u_z$.

Our goal here is to find a message $\bar{m}_j$, which together with $\sigma_{-j}$ induces a lottery $\ell^k(\hat{\theta}, (\bar{m}_j, \sigma_{-j}))$ that first-order stochastically dominates the lottery induced by $(m_j, \sigma_{-j})$, showing that $\sigma$ is not a Nash equilibrium of the game $\Gamma^K(\theta, u)$. To achieve this, we need to find $\bar{m}_j^2[\hat{\theta}] \in \{0, \ldots, K-1\}$, which together with $\sigma_{-j}$ guarantees that $\bar{p}_1 > \pi_i^{min}$ and $\bar{p}_0 + \bar{p}_1 > p_0 + p_1$. Such $\bar{m}_j^2[\hat{\theta}]$ will be found by the following algorithm:

Start the algorithm by setting $q_0 = q^*$. Next, for any $h \in \{1, \cdots, K-1\}$ and $q_{h-1} \in \{0, \ldots, K-1\}$, we define

$$q_h = q_{h-1} - 1 \pmod{K}.$$

Since $\pi_i^{min} < \pi_i^{max}$, we can choose $h \in \{1, \ldots, K-1\}$ uniquely in such a way that $P_i^\sigma(q_h|\hat{\theta}) > \pi_i^{min}$ and $P_i^\sigma(q_{h'}|\hat{\theta}) = \pi_i^{min}$ for all $h' \in \{0, \ldots, h-1\}$. Then, we set $\bar{m}_j^2[\hat{\theta}] = b_1(q_h)$. Define also $\bar{m}_j^3[\hat{\theta}]$ as follows:

$$\bar{m}_j^3[\hat{\theta}] = \begin{cases} c^*_{\hat{\theta}}(j,\theta) & \text{if } m_j^3[\hat{\theta}] \notin C^*_{\hat{\theta}}(j,\theta), \\ m_j^3[\hat{\theta}] & \text{if } m_j^3[\hat{\theta}] \in C^*_{\hat{\theta}}(j,\theta). \end{cases}$$

Thus, we define $\bar{m}_j$ to be identical to $m_j$ except that $m_j^2[\hat{\theta}]$ is replaced by $\bar{m}_j^2[\hat{\theta}]$ and $m_j^3[\hat{\theta}]$ is replaced by $\bar{m}_j^3[\hat{\theta}]$.

By the algorithm selection $q_h$ and construction of $\bar{m}_j^2[\hat{\theta}]$, we have the following properties:

$$\begin{aligned}
\bar{p}_1 \equiv P^{(\bar{m}_j,\sigma_{-j})}(q=1|\hat{\theta}) &= P^{(\bar{m}_j,\sigma_{-j})}(q_h + b_1(q_h) \pmod K|\hat{\theta}) \\
&= P_i^{(\bar{m}_j,\sigma_{-j})}(q_h|\hat{\theta}) > \pi_i^{min}, \\
&\geq P_i^{(\bar{m}_j,\sigma_{-j})}(q_1|\hat{\theta}) \\
&= P^{(m_j,\sigma_{-j})}(q_0 - 1 + m_j^2[\hat{\theta}] \pmod K|\hat{\theta}) \\
&= P^{(m_j,\sigma_{-j})}(q=0|\hat{\theta}) \equiv p_0 \ (\because q_0 + m_j^2[\hat{\theta}] \pmod K = 1) \\
P^{(\bar{m}_j,\sigma_{-j})}(q=2|\hat{\theta}) &= P^{(\bar{m}_j,\sigma_{-j})}(q_{h-1} + b_1(q_h) \pmod K|\hat{\theta}) \\
&= P_i^{(\bar{m}_j,\sigma_{-j})}(q_{h-1}|\hat{\theta}) = \pi_i^{min},
\end{aligned}$$

which implies $\bar{p}_1 - p_0 \geq 0$. Since $p_1 \equiv P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta}) = \pi_i^{min}$ and $P^{(\bar{m}_j,\sigma_{-j})}(q = 0|\hat{\theta}) \geq \pi_i^{min}$, we have $P^{(\bar{m}_j,\sigma_{-j})}(q=0|\hat{\theta}) \geq P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta})$, which implies that $\bar{p}_0 - p_1 \geq 0$. In addition, since $p_1 \equiv P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta}) = \pi_i^{min}$, we also have $P^{(\bar{m}_j,\sigma_{-j})}(q=1|\hat{\theta}) > P^{(m_j,\sigma_{-j})}(q=1|\hat{\theta})$, which implies that $\bar{p}_1 - p_1 > 0$.

Since we have $u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_0 > u_z$, due to the construction of $\bar{m}_j^3[\hat{\theta}]$, we have the following inequality:

$$\bar{u}_1 \equiv u_j(\bar{m}_j^3[\hat{\theta}],\theta) = u_j(c^*_{\hat{\theta}}(j,\theta),\theta) > u_j(\ell_0^K(\hat{\theta},(m_j,\sigma_{-j})),\theta) \equiv u_0$$

Thus, $\bar{u}_1 - u_0 > 0$.

Now, we claim that $u_j(g(\bar{m}_j,\sigma_j),\theta) - u_j(g(m_j,\sigma_j),\theta) > 0$. Since $P^\sigma(\hat{\theta}) > 0$, by Step 2, it suffices to show that $u_j(\ell^k(\hat{\theta},(\bar{m}_j,\sigma_j)),\theta) - u_j(\ell^k(\hat{\theta},(m_j,\sigma_j)),\theta) > 0$. Thus, we compute the following:

$$\begin{aligned}
& u_j(\ell^k(\hat{\theta},(\bar{m}_j,\sigma_j)),\theta) - u_j(\ell^k(\hat{\theta},(m_j,\sigma_j)),\theta) \\
= \ & \bar{p}_0(u_0 - u_z) + \bar{p}_1(\bar{u}_1 - u_z) + u_z - [p_0(u_0 - u_z) + p_1(u_1 - u_z) + u_z] \\
= \ & (\bar{p}_1 - p_1)(\bar{u}_1 - u_0) + (\bar{p}_1 - p_0)(u_0 - u_z) + (\bar{p}_0 - p_1)(u_0 - u_z) + p_1(\bar{u}_1 - u_1) \\
> \ & 0,
\end{aligned}$$

where the strict inequality follows because $(\bar{p}_1 - p_0)(u_0 - u_z) \geq 0$, $(\bar{p}_0 - p_1)(u_0 - u_z) \geq 0$, $p_1(\bar{u}_1 - u_1) \geq 0$, and $(\bar{p}_1 - p_1)(\bar{u}_1 - u_0) > 0$. This contradicts the hypothesis that $\sigma$ is a Nash equilibrium in the game $\Gamma^K(\theta,u)$. ■

**Step 6.b**: $\pi_i^{min} = \pi_i^{\max}$.

**Proof of Step 6.b**: Assume, by way of contradiction, that $\pi_i^{min} < \pi_i^{\max}$. We then use Step 6.a to conclude that, for each $q \in \{0,\cdots,K-1\}$, $P_i^\sigma(q|\hat{\theta}) = \pi_i^{min}$ implies $P_j^\sigma(b_1(q)|\hat{\theta}) = 0$. This implies $\pi_j^{min} = 0$ so that $\pi_j^{min} < \pi_j^{\max}$. We then establish the counterpart of Step 6.a by swapping the roles of $i$ and $j$ and replacing the function $b_1(q)$ with the function $b_0(q)$, where we define $b_0(q) \in \{0, \ldots, K-1\}$ such that $q + b_0(q) \pmod{K} = 0$. Therefore, we conclude that, for each $q \in \{0, \cdots, K-1\}$, if $P_j^\sigma(q|\hat{\theta}) = \pi_j^{min}$, then $P_i^\sigma(b_0(q)|\hat{\theta}) = 0$. Hence, $\pi_i^{min} = 0$.

Then, we can set $q \in \{0, \ldots, K-1\}$ such that $P_i^\sigma(q|\hat{\theta}) = 0$. By Step 6.a, we have

$$P_j^\sigma(b_1(q)|\hat{\theta}) = P_j^\sigma(1 - q \pmod{K}|\hat{\theta}) = 0.$$

By Step 6.a, we also have

$$P_i^\sigma(b_0(b_1(q))|\hat{\theta}) = P_i^\sigma(q - 1 \pmod{K}|\hat{\theta}) = 0.$$

We use Step 6.a repeatedly to conclude that $P_i^\sigma(q|\hat{\theta}) = P_j^\sigma(q|\hat{\theta}) = 0$ for each $q \in \{0, \ldots, K-1\}$. However, this is simply impossible, as we have $\sum_{q=0}^{K-1} P_i^\sigma(q|\hat{\theta}) = \sum_{q=0}^{K-1} P_j^\sigma(q|\hat{\theta}) = 1$. Therefore, we must have $\pi_i^{min} = \pi_i^{\max}$. ■

Since we can replace the role of agent $i$ with that of agent $j$ in the entire argument, we conclude that $\pi_i^{min} = \pi_i^{max}$ and $\pi_j^{min} = \pi_j^{max}$. Thus, we have $P_i^\sigma(q|\hat{\theta}) = P_j^\sigma(q|\hat{\theta}) = 1/K$ for each $q \in \{0, \ldots, K-1\}$. This implies that $P^\sigma(q = 0|\hat{\theta}) + P^\sigma(q = 1|\hat{\theta}) \leq 2/K$. By Step 4, we conclude that there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$. This completes the proof. ■

**Step 7**: Let $\sigma \in NE(\Gamma^K(\theta, u))$, where we later choose $K$ large enough. Assume that there exists $\hat{\theta} \in \Theta$ such that $P^\sigma(\hat{\theta}) > 0$ and $P^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. If $\ell^K(\hat{\theta}, \sigma) \neq f(\theta)$ and $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then there exists $K \in \mathbb{N}$ large enough so that $u_i(\ell^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

**Proof of Step 7**: Fix $\sigma \in NE(\Gamma^K(\theta, u))$. Assume that there exists $\hat{\theta} \in \Theta$ such that $P^\sigma(\hat{\theta}) > 0$ and $P^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. Assume further that $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$. Then, there exist $q \in \{0, 1\}$ and $i \in \mathcal{G}_{\hat{\theta}}$ such that $f(\theta) \in C_{\hat{\theta}}(i)$, where $i = w_{\hat{\theta}}(q)$. We write $\mathcal{G}_{\hat{\theta}} = \{i, j\}$ in the rest of the proof. By Property 1-ii of Condition $COM$, we have $f(\theta) \in C_{\hat{\theta}}(i) \cap C_{\hat{\theta}}(j)$.[13] By Property 4 of Condition $COM$, $f(\theta)$ is agent $i$'s unique maximal element in $C_{\hat{\theta}}(i)$ so that $C_{\hat{\theta}}^*(i, \theta) = \{f(\theta)\}$. Similarly, by Property 4 of Condition $COM$, $f(\theta)$ is also agent $j$'s unique maximal element in $C_{\hat{\theta}}(j)$ so that $C_{\hat{\theta}}^*(j, \theta) = \{f(\theta)\}$. Step 3 implies that

[13] When there are only two agents, we do not need to assume $f(\theta) \in C_{\hat{\theta}}(i) \cap C_{\hat{\theta}}(j)$. The hypothesis that $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$ and $f(\theta) \neq \ell^K(\hat{\theta}, \sigma)$ implies that both agents must strictly prefer $f(\theta)$ to lottery $\ell^K(\hat{\theta}, \sigma)$. When there are only two agents, this implies that $\ell^K(\hat{\theta}, \sigma)$ is strictly Pareto dominated by $f(\theta)$. Hence, we can conclude that we can ignore $\sigma$ because it is "uncompelling" without the help of Property 1-ii of Condition $COM$.

for each $q \in \{0,1\}$, we have $\text{supp}(\ell_q^K(\hat{\theta},\sigma)) \in C^*_{\hat{\theta}}(w_{\hat{\theta}}(q),\theta)$, hence we can conclude that both $\ell_0^K(\hat{\theta},\sigma) = f(\theta)$ and $\ell_1^K(\hat{\theta},\sigma) = f(\theta)$. Using the notation developed in Step 2, we can write

$$\begin{aligned} \ell^K(\hat{\theta},\sigma) &= P^\sigma(q=0|\hat{\theta})(\ell_0^K(\hat{\theta},\sigma) - z_{\hat{\theta}}) + P^\sigma(q=1|\hat{\theta})(\ell_1^K(\hat{\theta},\sigma) - z_{\hat{\theta}}) + z_{\hat{\theta}} \\ &= \left(P^\sigma(q=0|\hat{\theta}) + P^\sigma(q=1|\hat{\theta})\right)(f(\theta) - z_{\hat{\theta}}) + z_{\hat{\theta}}. \end{aligned}$$

Moreover, since we assume $\ell^K(\hat{\theta},\sigma) \neq f(\theta)$, $z_{\hat{\theta}}$ is induced with positive probability. By Property 3 of Condition $COM$, we have $u_i(f(\theta),\theta) - u_i(z_{\hat{\theta}},\theta) \geq \varepsilon$ for all $i \in I$ and $u \in \hat{\mathcal{U}}$. By the continuity of expected payoff, this implies that $u_i(\ell^K(\hat{\theta},\sigma),\theta) < u_i(f(\theta),\theta)$ for all $i \in I$ and $u \in \hat{\mathcal{U}}$. ■

By Step 2, we have the following expressions:

$$g^K(\sigma) = \sum_{\hat{\theta}\in\Theta} P^\sigma(\hat{\theta})\ell^K(\hat{\theta},\sigma) \text{ and } u_i(g^K(\sigma),\theta) = \sum_{\hat{\theta}\in\Theta} P^\sigma(\hat{\theta})u_i(\ell^K(\hat{\theta},\sigma),\theta).$$

Steps 5, 6 and 7 show that, for a fixed $u \in \hat{\mathcal{U}}$, we can find a value of $K$ such that, for every $\theta, \hat{\theta} \in \Theta$ and every $i \in I$, if $P^\sigma(\hat{\theta}) > 0$ and $\ell^K(\hat{\theta},\sigma) \neq f(\theta)$, then we have $u_i(\ell^K(\hat{\theta},\sigma),\theta) < u_i(f(\theta,\theta)$. By Property 3 of Condition $COM$, we have that there exists $\varepsilon > 0$ such that $u_i(f(\theta),\theta) - u_i(z_{\hat{\theta}},\theta) \geq \varepsilon$ for all possible cardinal utility functions $u \in \hat{\mathcal{U}}$. Therefore, we can choose $K \in \mathbb{N}$ large enough such that $u_i(\ell^K(\hat{\theta},\sigma),\theta) < u_i(f(\theta),\theta)$ hold for all $u \in \hat{\mathcal{U}}$ (and all $\theta, \hat{\theta} \in \Theta$, $i \in I$). We summarize this into the following step:

**Step 8**: There exists $K \in \mathbb{N}$ large enough such that, for any $u \in \hat{\mathcal{U}}$ and $\sigma \in NE(\Gamma^K(\theta,u))$, it follows that either $g^K(\sigma) = f(\theta)$ or $u_i(g^K(\sigma),\theta) < u_i(f(\theta),\theta)$ for all $i \in I$.

Combining Steps 1 and 8, we conclude that there exists $K \in \mathbb{N}$ large enough such that the SCF $f$ is compellingly implementable with respect to $\hat{\mathcal{U}}$ by the mechanism $\Gamma^K$. This completes the proof of the theorem. ■

# 6 Indispensability of Condition *COM*

In this section, we show that Condition $COM$ is indispensable for our Theorem 1. We show this by arguing that Properties 1-i and 2 in Condition $COM$ are in fact implied by the very requirement of our compelling implementation, while for each of Properties 1-ii, 3 and 4, we will provide an example that satisfies all but one of Condition $COM$ in which our canonical mechanism fails to achieve compelling implementation.

## 6.1 Properties 1-i and 2 of Condition $COM$

We show that if the SCF $f$ is compellingly implementable by our canonical mechanism $\Gamma^k$, Properties 1-i and 2 of Condition $COM$ are satisfied.

**Proposition 2** *Let $\mathcal{E} = (I, A, \Theta, (\succeq_i^\theta)_{i\in I,\theta\in\Theta}, \hat{\mathcal{U}})$ be a finite environment. If the SCF $f$ is compellingly implementable with respect to $\hat{\mathcal{U}}$ by the mechanism $\Gamma^k$, then the finite environment $\mathcal{E}$ satisfies Properties 1-i and 2 of Condition COM with respect to $f$ and $\hat{\mathcal{U}}$.*

**Proof:** By Condition $COM$, we first define a collection of forums, $(\mathcal{F}_{\tilde{\theta}})_{\tilde{\theta}\in\Theta} = \{\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$.

Assume, on the contrary, that Property 1-i is violated. That is, there exist $\theta \in \Theta$ and $u \in \hat{\mathcal{U}}$ such that $f(\theta)$ is not $C_\theta$-acceptable at $(\theta, u)$. This implies that there exist a player $i \in \mathcal{G}_\theta$ and a lottery $l \in C_\theta(i)$ such that $u_i(l, \theta) > u_i(f(\theta), \theta)$. Since the SCF $f$ is compellingly implementable by the mechanism $\Gamma^k$, there exists $m \in pureNE(\Gamma^k(\theta))$ such that $g(m) = f(\theta)$ and $m_j^1 = \theta$ for each $j \in I$ (i.e., all agents agree upon the true state $\theta$). We define $\bar{m}_i^2[\theta] \in \{0, \dots, k-1\}$ such that

$$\bar{m}_i^2[\theta] + m_j^2[\theta] \pmod{k} = w_\theta^{-1}(i),$$

where $j \in \mathcal{G}_\theta \backslash \{i\}$. We also define $\bar{m}_i^3[\theta] = l$. Then, we define $\bar{m}_i$ to be identical to $m_i$ except that $m_i^2[\theta]$ is replaced by $\bar{m}_i^2[\theta]$ and $m_i^3[\theta]$ is replaced by $\bar{m}_i^3[\theta]$. It then follows from the construction of $m$ and $\bar{m}_i$ that $g(m) = f(\theta)$ and $g(\bar{m}_i, m_{-i}) = l$. Therefore, $\bar{m}_i$ is a profitable deviation from $m$ in the game $\Gamma^k(\theta)$, contradicting the hypothesis that $m \in pureNE(\Gamma^k(\theta))$. This proves that Property 1-i holds if the mechanism $\Gamma^k$ $C$-implements $f$.

Next, assume, on the contrary, that Property 2 is violated. That is, there exist $\theta, \hat{\theta} \in \Theta$, $u \in \hat{\mathcal{U}}$, and a lottery $l \neq f(\theta)$ such that $l$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$. This implies that there exists $i^* \in \mathcal{G}_{\hat{\theta}}$ such that $l \in C_{\hat{\theta}}(i^*)$. We assume, without loss of generality, that $i^* = w_{\hat{\theta}}(0)$. For all $i \in I$, we define

- $\hat{m}_i^1 = \hat{\theta}$;
- $\hat{m}_i^2[\tilde{\theta}] = 0$ for all $\tilde{\theta} \in \Theta$;
- $\hat{m}_i^3[\tilde{\theta}] = l$ for all $\tilde{\theta} \in \Theta$.

So, we define $\hat{m}_i = (\hat{m}_i^1, (\hat{m}_i^2[\tilde{\theta}])_{\tilde{\theta}\in\Theta}, (\hat{m}_i^3[\tilde{\theta}])_{\tilde{\theta}\in\Theta})$ so that we can write $\hat{m} = (\hat{m}_i)_{i\in I}$ as a message profile. We claim that $\hat{m} \in pureNE(\Gamma^k(\theta))$. By construction of the mechanism and the fact that $n \geq 3$, no agent can unilaterally change the agreed upon state $\hat{\theta}$ by changing the first component of their message. By construction of the mechanism, $\mathcal{G}_{\hat{\theta}}$ contains only two agents who can affect the outcome of the mechanism. This implies that all the other agents are left with no profitable deviations. Moreover, by construction of the mechanism, the only other outcomes attainable by a unilateral deviation of any agent in $\mathcal{G}_{\hat{\theta}}$ are the ones in the image of $C_{\hat{\theta}}$. Since the lottery $l$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$ and $g(\hat{m}) = l$, neither of the two agents in $\mathcal{G}_{\hat{\theta}}$ can find a profitable deviation from $\bar{m}$. Hence, $\hat{m} \in pureNE(\Gamma^k(\theta))$, which contradicts the hypothesis that the mechanism $\Gamma^k$ $C$-implements $f$. Thus, Property 2 holds. ■

## 6.2 Property 3 of Condition $COM$

We argue by means of an example that Property 3 of Condition $COM$ is indispensable if we are to achieve compelling implementation by our canonical mechanism. To do this, we revisit Example 1 in Section 5.2 but we set $\hat{\mathcal{U}} = \mathcal{U}$, which is the set of *all* possible cardinal utility functions consistent with the ordinal preferences in the example. It is easy to see that the collection of forums $\mathcal{F}_{\tilde{\theta}\in\Theta}$ we constructed in Section 5.2 continue to satisfy Properties 1, 2, and 4 of Condition $COM$ even if we replace $\mathcal{U}^{\varepsilon}$ with $\mathcal{U}$.

We claim that for any integer $K \in \mathbb{N}$ with $K \geq 3$, there exist $u \in \mathcal{U}$ and strategy profile $\sigma \in NE(\Gamma^k(\theta_c, u))$ such that $u_i(g(\sigma), \theta_c) > u_i(f(\theta_c), \theta_c)$ for all $i \in I$. We define $\sigma$ as follows:

- $\sigma_i^1(\theta_a) = 1$ for all $i \in I$, i.e., all agents commonly announce $\theta_a$.
- $\sigma_i^2[\tilde{\theta}](\tilde{k}) = 1/K$ for each $i \in I$, $\tilde{\theta} \in \Theta$, and $\tilde{k} \in \{0, \ldots, K-1\}$, each agent announces every integer in $\{0, \ldots, K-1\}$ with equal probability.
- $\sigma_0^3[\theta_a](a) = \sigma_0^3[\theta_b](\emptyset) = \sigma_0^3[\theta_c](\emptyset) = \sigma_0^3[\theta_d](c) = 1$. Here, agent 0 always propose $a$ when they agree on $\theta_a$.
- $\sigma_1^3[\theta_a](b) = \sigma_1^3[\theta_b](b) = \sigma_1^3[\theta_c](c) = \sigma_1^3[\theta_d](\emptyset) = 1$. Here, agent 1 always proposes $b$ when they agree on $\theta_a$.
- $\sigma_2^3[\theta_a](\emptyset) = \sigma_2^3[\theta_b](a) = \sigma_2^3[\theta_c](c) = \sigma_2^3[\theta_d](c) = 1$. As agent 2 is not a part of forum $\theta_a$, he cannot propose any alternative when agents agree on $\theta_a$.

By construction of the mechanism and the fact that $n = 3$, no agent can unilaterally change the agreed upon state $\theta_a$ by changing the first component of their message. By construction of the mechanism, $\mathcal{G}_{\theta_a}$ contains only agents 0 and 1 who can affect the outcome of the mechanism. This implies that agent 2 is left with no profitable deviations. Moreover, by construction of the mechanism, the only other outcomes attainable by a unilateral deviation of either agents 0 or 1 in $\mathcal{G}_{\theta_a}$ are the ones in the image of $C_{\theta_a}$.

We consider the game $\Gamma^K(\theta_c, u)$. In this game, the constructed $\sigma$ dictates that agent 0 already chooses his best outcome $a$ in state $\theta_c$, while agent 1 already chooses his best outcome $b$ in state $\theta_c$. Therefore, agents 0 and 1 have no profitable deviations by changing the third component of their message. It thus remains to show that agents 0 and 1 have no profitable deviations by changing the second component of their message. Given $\sigma_1$ and $\sigma_2$, any deviation $m_0$ restricted to the second component of agent 0's message induces the following lottery:

$$\ell^K(\theta_a, (m_0, \sigma_1, \sigma_2)) = \frac{1}{K}a + \frac{1}{K}b + \left(1 - \frac{2}{K}\right) z = \ell^K(\theta_a, \sigma),$$

where $\ell^K(\theta_a, \sigma)$ denotes the lottery induced by $\sigma$ in which the agreed-upon state is $\theta_a$. Similarly, given $\sigma_0$ and $\sigma_2$, any deviation $m_1$ restricted to the second component of agent 1's message induces the following lottery:

$$\ell^K(\theta_a, (m_1, \sigma_0, \sigma_2)) = \frac{1}{K}a + \frac{1}{K}b + \left(1 - \frac{2}{K}\right) z = \ell^K(\theta_a, \sigma).$$

So, agents 0 and 1 cannot change the induced lottery by their unilateral deviation. Thus, $\sigma \in NE(\Gamma^K(\theta_c, u))$ for any $u \in \mathcal{U}$.

For any $K \in \mathbb{N}$, we can select $\varepsilon_1, \varepsilon_2 \in (0, 1)$ such that

$$\frac{2(1 - \varepsilon_1)}{K} > \varepsilon_2,$$

Since $\mathcal{U}$ includes *all* possible cardinal utility functions consistent with the underlying ordinal preferences, we can select $u \in \mathcal{U}$ such that $u_0(a, \theta_c) - u_0(b, \theta_c) = u_1(b, \theta_c) - u_1(a, \theta_c) = u_2(a, \theta_c) - u_2(b, \theta_c) = \varepsilon_1$ and $u_i(c, \theta_c) = \varepsilon_2$ for all $i \in I$. By Step 2, we have that $u_i(g^K(\sigma), \theta_c) = u_i(\ell^K(\theta_a, \sigma), \theta_c)$ for all $i \in I$ and $P^\sigma(q = 0|\theta_a) = P^\sigma(q = 1|\theta_a) = 1/K$. Therefore, for any $i \in I$, we have

$$\begin{aligned} u_i(g^K(\sigma), \theta_c) &= \frac{2(1 - \varepsilon_1)}{K}, \\ u_i(f(\theta_c), \theta_c) = u_i(c, \theta_c) &= \varepsilon_2. \end{aligned}$$

This implies that, for any $K \in \mathbb{N}$, there exists $u \in \mathcal{U}$ such that $u_i(g^K(\sigma), \theta_c) > u_i(f(\theta_c), \theta_c)$ for all $i \in I$. Thus, the SCF $f$ is not compellingly implementable in this example. ■

## 6.3 Property 4 of Condition *COM*

In what follows, we will first present an environment in which Condition *COM* holds, then modify it in a specific way so that only a specific property is violated.

Let $\mathcal{E}^* = \{I, A, \Theta, (\succeq_i^\theta)_{i\in I,\theta\in\Theta}\}$ be an environment such that $I = \{0, 1, 2, 3\}$, $\Theta = \{\theta, \theta'\}$, $A = \{a, b, c\}$, and preferences $(\succeq_i^\theta)$ are directly given by their cardinal utilities in Table 6. We therefore assume $\hat{\mathcal{U}} = \{u\}$ - a singleton set. We consider this restriction to the singleton set inconsequential because the same argument we provide below goes through with any $\hat{\mathcal{U}}$ as long as it contains $\{u\}$.

| | $u_i(a,\theta), u_i(a,\theta')$ | $u_i(b,\theta), u_i(b,\theta')$ | $u_i(c,\theta), u_i(c,\theta')$ |
|---|---|---|---|
| $i = 0$ | 1 , 0.8 | 0.8 , 0 | 0 , 1 |
| $i = 1$ | 0 , 1 | 1 , 0 | 0.8 , 0.8 |
| $i = 2$ | 0 , 0.8 | 0.8 , 1 | 1 , 0 |
| $i = 3$ | 0 , 1 | 1 , 0 | 0.8 , 0.8 |

Table 6: The environment $\mathcal{E}^*$

The SCF $f$ is specified as follows: $f(\theta) = (0.1, 0.8, 0.1)$, in which $(0.1, 0.8, 0.1)$ denotes the lottery which induces $a$ with probability 0.1, $b$ with probability 0.8, and $c$ with probability 0.1. Using the same notation, we have $f(\theta') = (0.8, 0.1, 0.1)$.

This environment satisfies Condition $COM$ with the following two forums:

- At state $\theta$, we have $\mathcal{G}_\theta = \{1, 2\}$, $w_\theta(0) = 1$, $w_\theta(1) = 2$, $C_\theta(1) = \{z_\theta, a, c, f(\theta)\}$ and $C_\theta(2) = \{z_\theta, f(\theta)\}$, where $z_\theta = (1/3, 1/3, 1/3)$.
- At state $\theta'$, we have $\mathcal{G}_{\theta'} = \{1, 0\}$, $w_{\theta'}(0) = 1$, $w_{\theta'}(1) = 0$, $C_{\theta'}(1) = \{z_{\theta'}, b, c, f(\theta')\}$ and $C_{\theta'}(0) = \{z_{\theta'}, f(\theta')\}$, where $z_{\theta'} = (1/3, 1/3, 1/3)$.

Using the utility values in Table 6, we compute $u_0(f(\theta'), \theta) = u_1(f(\theta), \theta) = u_1(f(\theta'), \theta') = u_2(f(\theta), \theta') = 0.88$, $u_0(f(\theta), \theta) = u_0(f(\theta'), \theta') = u_2(f(\theta), \theta) = u_2(f(\theta'), \theta') = 0.74$, $u_1(f(\theta), \theta') = u_1(f(\theta'), \theta) = 0.18$, and $u_i(z, \theta) = u_i(z, \theta') = 0.6$ for all $i \in I$. We shall verify that the environment $\mathcal{E}^*$ satisfies Condition $COM$ with respect to $f$ and $\mathcal{U}$:

- Property 1-i

  $f(\theta)$ is $C_\theta$-acceptable at $(\theta, u)$ because $f(\theta)$ is the best outcome within $C_\theta(1) = \{z_\theta, a, c, f(\theta)\}$ for agent 1 and the best outcome within $C_\theta(2) = \{z_\theta, f(\theta)\}$ for agent 2. $f(\theta')$ is $C_{\theta'}$-acceptable at $(\theta', u)$ because $f(\theta')$ is the best outcome within $C_{\theta'}(1) = \{z_{\theta'}, b, c, f(\theta')\}$ for agent 1 and the best outcome within $C_{\theta'}(0) = \{z_{\theta'}, f(\theta')\}$ for agent 0.

- Property 1-ii

  We have $f(\theta) \in \{z_\theta, a, c, f(\theta)\} \cap \{z_\theta, f(\theta)\} = C_\theta(1) \cap C_\theta(2)$ as well as $f(\theta') \in \{z_{\theta'}, b, c, f(\theta')\} \cap \{z_{\theta'}, f(\theta')\} = C_{\theta'}(1) \cap C_\theta(0)$. Likewise, $f(\theta') \notin \{z_\theta, a, c, f(\theta)\} \cap \{z_\theta, f(\theta)\} = C_\theta(1) \cap C_\theta(2)$ as well as $f(\theta) \notin \{z_{\theta'}, b, c, f(\theta')\} \cap \{z_{\theta'}, f(\theta')\} = C_{\theta'}(1) \cap C_\theta(0)$.

- Property 2

  $f(\theta)$ is the only outcome that is $C_\theta$-acceptable at $(\theta, u)$. Since $u_1(b, \theta) > u_1(f(\theta'), \theta)$ while $u_0(f(\theta'), \theta) > u_0(b, \theta)$, there are no outcomes that are $C_{\theta'}$-acceptable $(\theta, u)$. Likewise, $f(\theta')$ is the only outcome that is $C_{\theta'}$-acceptable at $(\theta^{'}, u)$. Since $u_1(a, \theta') > u_1(f(\theta), \theta')$ while $u_2(f(\theta), \theta') > u_2(a, \theta)$, there are no outcomes that are $C_\theta$-acceptable at $(\theta, u)$.

- Property 3

  Since we have $u_i(z_\theta, \theta) = u_i(z_{\theta'}, \theta') = 0.6$ for any $i \in I$ and $\min_{i\in I} u_i(f(\theta), \theta) = \min_{i\in I} u_i(f(\theta'), \theta') = 0.74$, Property 3 holds.

- Property 4

  As we argue above, there are no outcomes that are either $C_\theta$-acceptable at $(\theta^{'}, u)$ or $C_{\theta'}$-acceptable at $(\theta, u)$. Moreover, at state $\theta$, $f(\theta)$ is the unique best outcome in $C_\theta(1)$ for agent 1 and the unique best outcome in $C_\theta(2)$ for agent 2. Likewise, at state $\theta'$, $f(\theta')$ is the unique best outcome in $C_{\theta'}(1)$ for agent 1 and the unique best outcome in $C_{\theta'}(0)$ for agent 0. Hence, Property 4 holds.

To construct an environment in which only Property 4 is violated, we modify the environment by adding the lottery $y = (0.16, 0.5, 0.34)$ to $C_\theta(2)$. Under cardinal utility function $u$, we have that $u_2((0.16, 0.5, 0.34), \theta) = 0.74 = u_2(f(\theta), \theta)$, hence violating Property 4. Let $\sigma$ be a strategy profile satisfying the following properties:

- $\sigma_i^1(\theta) = 1$ for all $i \in I$;

- $\sigma_0^2[\tilde{\theta}](0) = \sigma_1^2[\tilde{\theta}](0) = \sigma_3^2[\tilde{\theta}](0) = 1$ and $\sigma_2^2[\tilde{\theta}](0) = \sigma_2^2[\tilde{\theta}](1) = 1/2$ for all $\tilde{\theta} \in \Theta$, meaning that when state $\theta$ is selected half of the time agent 1 wins the modulo game and the other half the winner is agent 2;

- $\sigma_1^3[\tilde{\theta}](f(\tilde{\theta})) = 1$ for all $\tilde{\theta} \in \Theta$, $\sigma_2^3[\theta](f(\theta)) = 0.99$, $\sigma_2^3[\theta](y) = 0.01$, $\sigma_2^3[\theta'](\emptyset) = 1$, $\sigma_3^3[\tilde{\theta}](\emptyset) = 1$ and $\sigma_0^3[\theta](\emptyset) = 1$ and $\sigma_0^3[\theta^{'}](f(\theta^{'})) = 1$, meaning that at state $\theta$, when agent 1 wins the modulo game, he always selects $f(\theta)$; and when agent 2 is the winner, he chooses $f(\theta)$ with probability 0.99 and $y$ with the remaining probability.

The key reason why $\sigma$ is a Nash equilibrium in the game $\Gamma^k(\theta, u)$ is that agent 2 is indifferent between $y$ and $f(\theta)$, while agent 1 has no profitable deviations because sending any other integer increases the probability that outcome $z_\theta$ is realized. It thus follows that $\sigma \in NE(\Gamma^k(\theta, u))$ so that $u_2(g(\sigma), \theta)) = u_2(f(\theta), \theta))$. This implies that that compelling implementation fails. ■

## 6.4 Property 1-ii of Condition $COM$

We first modify the environment $\mathcal{E}^*$ in the previous section by removing $f(\theta)$ from $C_\theta(1)$, thus violating Property 1-ii. We next eliminate $c$ from $C_\theta(1)$ and instead add the following lottery $y^{'} = (2/9, 7/18, 7/18)$ to $C_\theta(1)$ so that $u_1(y^{'}, \theta) = u_2(y^{'}, \theta) = 0.7$. Since both $f(\theta)$ and $c$ are no longer in $C_\theta(1)$, $y^{'}$ is the new unique best outcome for agent 1 in $C_\theta(1)$ at state $\theta$. Finally, we modify agent 3's utilities at state $\theta$ as follows: $u_3(a, \theta) = 0$, $u_3(b, \theta) = 0.51$, and $u_3(c, \theta) = 1$. Under these new utility values, we note that $u_3(f(\theta), \theta) = 0.508$, $u_3(z_\theta, \theta) \approx 0.503$ and $u_3(y, \theta) \approx 0.587$. We also note that none of these modifications violate any of the other properties in Condition $COM$. Let $\sigma$ be a strategy profile satisfying the following characteristics:

- $\sigma_i^1(\theta) = 1$ for all $i \in I$;
- $\sigma_1^2[\tilde{\theta}](0) = 7/9, \sigma_1^2[\tilde{\theta}](1) = 2/9, \sigma_2^2[\tilde{\theta}](0) = 14/23, \sigma_2^2[\tilde{\theta}](1) = 9/23$, and $\sigma_0^2[\tilde{\theta}](0) = \sigma_3^2[\tilde{\theta}](0) = 1$ for all $\tilde{\theta} \in \Theta$. This implies that $P^\sigma(q = 0|\theta) = 98/207$, $P^\sigma(q = 1|\theta) = 91/207$, and $P^\sigma(q = 2|\theta) = 18/207$;
- $\sigma_3^3[\tilde{\theta}](\emptyset) = 1$ for all $\tilde{\theta} \in \Theta$, $\sigma_1^3[\theta](y^{'}) = \sigma_1^3[\theta'](b) = 1$ $\sigma_2^3[\theta](f(\theta)) = 1$, $\sigma_0^3[\theta^{'}](f(\theta^{'})) = 1$, $\sigma_2^3[\theta^{'}](\emptyset) = 1$, and $\sigma_0^3[\theta](\emptyset) = 1$. This implies that at state $\theta$ whenever agent 1 wins the modulo game, he always selects $y^{'}$, and when agent 2 is the winner, he always chooses $f(\theta)$.

We claim $\sigma \in NE(\Gamma^k(\theta, u))$. As we discuss in other examples, no agent can find profitable deviations by changing the first component of his messages because all agents announce $\theta$. This implies that $P^{(m_i, \sigma_{-i})}(\theta) = 1$ for any $i \in I$ and $m_i \in M_i$. By construction of the mechanism, $\mathcal{G}_\theta$ contains only agents 1 and 2 who can affect the outcome of the mechanism. This implies that agents 0 and 3 are left with no profitable deviations when they agree on $\theta$. Since agents 1 and 2 choose their unique best outcome in the third component of their message under $\sigma$, they find no profitable deviations by changing the third component of their message. Using the notation developed in the proof of Theorem 1, we have

$$g^k((m_i, \sigma_{-i})) = P^{(m_i, \sigma_{-i})}(q = 0|\theta)(y - z_\theta) + P^{(m_i, \sigma_{-i})}(q = 1|\theta)(f(\theta) - z_\theta) + z_\theta.$$

This shows that the only deviations with which we are concerned are the ones changing the second component of their message, which induces a change in $P^{(m_i, \sigma_{-i})}(q = 0|\theta)$ and $P^{(m_i, \sigma_{-i})}(q = 1|\theta)$. Moreover, for such a deviation to be profitable, it must either increase $P^{(m_i, \sigma_{-i})}(q = 0|\theta) + P^{(m_i, \sigma_{-i})}(q = 1|\theta)$ or increase $P^{(m_i, \sigma_{-i})}(q = 1|\theta)$ without decreasing $P^{(m_i, \sigma_{-i})}(q = 0|\theta) + P^{(m_i, \sigma_{-i})}(q = 1|\theta)$. This rules out any message $m_1$ with $m_1^2[\theta] \notin \{0, 1\}$, since these messages would increase the probability that $z_\theta$ is realized and decrease the probability that $f(\theta)$ is realized. Similarly, any message $m_2$ with $m_2^2[\theta] \notin \{0, 1\}$ is an unprofitable deviation. Hence, the only profitable deviation from $\sigma$, if any, is to change $P_1^{\sigma}(q=0|\theta)$, $P_1^{\sigma}(q=1|\theta)$, $P_2^{\sigma}(q=0|\theta)$, or $P_2^{\sigma}(q=1|\theta)$, all of which are the probabilities that agents 1 and 2 announce integers 0 and 1. For each $q \in \{0,1\}$, let $m_2^q \in M_2$ be a deviation message such that $m_2^{1,q} = \theta$; $m_2^{2,q}[\tilde{\theta}] = q$ for each $\tilde{\theta} \in \Theta$; $m_2^{3,q}[\theta] = f(\theta)$; $m_2^{3,q}[\theta'] = \emptyset$. However, given the strategies of the other agent, both agent 1 and 2 are indifferent between sending each of these two integers:

$$\begin{aligned} & P_1^{(m_2^0,\sigma_{-2})}(q=0|\theta)u_2(y,\theta) + P_1^{(m_2^0,\sigma_{-2})}(q=1|\theta)u_2(f(\theta),\theta) \\ = \; & P_1^{(m_2^1,\sigma_{-2})}(q=0|\theta)u_2(f(\theta),\theta) + P_1^{(m_2^1,\sigma_{-2})}(q=1|\theta)u_2(z_\theta,\theta) \end{aligned}$$

In the equation above, the left hand side represents the expected payoff for agent 2 to play $m_2^0$ given $\sigma_{-2}$, while the right hand side represents the expected payoff for agent 2 to play $m_2^1$ given $\sigma_{-2}$. For each $q \in \{0,1\}$, let $m_1^q \in M_1$ be a deviation message such that $m_1^{1,q} = \theta$; $m_1^{2,q}[\tilde{\theta}] = q$ for each $\tilde{\theta} \in \Theta$; $m_1^{3,q}[\theta] = m_1^{3,q}[\theta'] = y'$; $m_2^{3,q}[\theta'] = \emptyset$.

$$\begin{aligned} & P_2^{(m_1^0,\sigma_{-1})}(q=0|\theta)u_1(y',\theta) + P_2^{(m_1^0,\sigma_1)}(q=1|\theta)u_1(f(\theta),\theta) \\ = \; & P_2^{(m_1^1,\sigma_{-1})}(q=0|\theta)u_1(f(\theta),\theta) + P_2^{(m_1^1,\sigma_{-1})}(q=1|\theta)u_1(z_\theta,\theta). \end{aligned}$$

In the equation above, the left hand side represents the expected payoff for agent 1 to play $m_1^0$ given $\sigma_{-1}$, while the right hand side represents the expected payoff for agent 1 to play $m_1^1$ given $\sigma_{-1}$. This shows that there are no profitable deviations from $\sigma$ so that $\sigma \in NE(\Gamma^k(\theta,u))$.

In this equilibrium $\sigma$, although both agents want agent 2 to be the winner of the modulo game, the lack of coordination causes agent 1 to win the modulo game with positive probability. Given the strategy profile of other agents, both agents 1 and 2 are indifferent between sending either 0 or 1 in the modulo game, which induces a positive probability that agent 1 is the winner. When that happens, agent 1 chooses $y'$ as his best outcome within $C_\theta(1)$ in state $\theta$, while agent 3 strictly prefers $y'$ to $f(\theta)$ in state $\theta$. Thus, we have $u_3(g(\sigma),\theta) \approx 0.545 > 0.508 = u_3(f(\theta),\theta)$. This implies that compelling implementation fails.

# 7 The Main Result When $n = 2$

Condition $COM$ guarantees that each forum essentially induces a subgame with only two agents. Therefore, it is natural to ask if Condition $COM$ is also sufficient to achieve compelling implementation when there are only two agents. It is well known that two-person pure Nash implementation is much more demanding than when there are at least three agents.[14] This same difficulty persists in

[14]Dutta and Sen (1991) and Moore and Repullo (1990) independently identify a necessary and sufficient condition (called Condition $\beta$ and Condition $\mu 2$, respectively) for two-person pure Nash

compelling implementation because compelling implementation implies pure Nash implementation. This difficulty boils down to how they agree on the state $\hat{\theta}$ in the mechanism. When there are three or more agents, if all agents announce the same state $\hat{\theta}$, then no agent can change $\hat{\theta}$. This, in turn, limits the set of possible outcomes that an agent can induce by a unilateral deviation to the outcomes in $C_{\hat{\theta}}$ alone. However, this argument fails because a two-person environment allows a single agent to induce a change in $\hat{\theta}$ by changing his own message. As a result, the set of outcomes that agent can achieve through a unilateral deviation becomes (potentially) larger than $C_{\hat{\theta}}$. Thus, we need to modify both our canonical mechanism and the notions of forums.

Our first step is to modify the notion of forums so that our canonical mechanism and Condition $COM$ are both redefined. With $I = \{0, 1\}$, both $\mathcal{G}_{\hat{\theta}}$ and $w_{\hat{\theta}}$ become redundant in the definition of Condition $COM$. We also define a punishment function $z$ to depend on both agents' state announcements. Hence, we modify our definition of forums from the one in Section 5.

We say that $\mathcal{F}^2 = (\{C_\theta\}_{\theta\in\Theta}, z)$ is a *forum-2* if it satisfies the following properties:

1. $C : \{0, 1\} \rightrightarrows \Delta(A)$ is a choice correspondence; and
2. $z : \Theta \times \Theta \to \Delta(A)$ is a function such that, for all $\theta_0, \theta_1 \in \Theta$, $z(\theta_0, \theta_1) \in C_{\theta_1}(0) \cap C_{\theta_0}(1)$.

We now introduce the two-person counterpart of Condition $COM$, which we call Condition $COM_2$.

**Definition 4** *The environment* $\mathcal{E} = \left(\{0, 1\}, A, \Theta, (\succeq_i^\theta)_{i\in\{0,1\},\theta\in\Theta}, \hat{\mathcal{U}}\right)$ *satisfies* ***Condition*** $COM_2$ *with respect to the SCF* $f$ *if there exists a forum-2* $\mathcal{F}^2 = (\{C_\theta\}_{\theta\in\Theta}, z)$ *such that:*

1. *For every* $\theta \in \Theta$ *and* $u \in \hat{\mathcal{U}}$*,* $f(\theta)$ *is* $C_\theta$*-acceptable at* $(\theta, u)$*.*
2. *For every* $\theta, \hat{\theta} \in \Theta$ *and* $u \in \hat{\mathcal{U}}$*, if* $x \in A$ *is is* $C_{\hat{\theta}}$*-acceptable* $(\theta, u)$*, then* $x = f(\theta)$*.*
3. *There exists* $\varepsilon > 0$ *such that, for all* $\theta, \theta_0, \theta_1 \in \Theta$*,* $u_i(f(\theta), \theta) - u_i(z(\theta_0, \theta_1), \theta) \geq \varepsilon$ *for all* $i \in \{0, 1\}$ *and all* $u \in \hat{\mathcal{U}}$*.*
4. *For all* $\theta, \hat{\theta} \in \Theta$*,* $u \in \hat{\mathcal{U}}$*,* $i \in \{0, 1\}$*, and* $x \in C_{\hat{\theta}}$*, if* $f(\theta)$ *is* $C_{\hat{\theta}}$*-acceptable at* $(\theta, u)$*, then* $u_i(x, \theta) = u_i(f(\theta), \theta)$ *implies* $x = f(\theta)$*.*

---

implementation.

5. *For each* $\theta, \theta_0, \theta_1 \in \Theta$ *and each* $u \in \hat{\mathcal{U}}$*, there exists either* $a_{(\theta_0,\theta_1)} \in C_{\theta_1}(0)$ *such that* $u_0(a_{(\theta_0,\theta_1)}, \theta) > u_0(z(\theta_0, \theta_1), \theta)$ *or* $b_{(\theta_0,\theta_1)} \in C_{\theta_0}(1)$ *such that* $u_1(b_{(\theta_0,\theta_1)}, \theta) > u_1(z(\theta_0, \theta_1), \theta)$*;*

For each $k \geq 2$, we define a new canonical mechanism, $\tilde{\Gamma}^k = (M^k, \tilde{g}^k)$. For each $i \in \{0, 1\}$, agent $i$'s message space is given by $M_i^k \equiv M_i^1 \times M_i^2 \times M_i^3$. Each component of the message space is identical to the one introduced to prove Theorem 1. Let $m_i = (m_i^1, m_i^2, m_i^3) \in M_i^k$ be agent $i$'s generic message such that (i) $m_i^1 \in M_i^1 = \Theta$; (ii) $m_i^2 = (m_i^2[\tilde{\theta}])_{\tilde{\theta}\in\Theta} \in M_i^2 = \times_{\tilde{\theta}\in\Theta} M_i^2[\tilde{\theta}]$ where $m_i^2[\tilde{\theta}] \in \{0, \ldots, k-1\}$; and (iii) $m_i^3 = (m_i^3[\tilde{\theta}])_{\tilde{\theta}\in\Theta} \in M_i^3 \equiv \times_{\tilde{\theta}\in\Theta} M_i^3[\tilde{\theta}]$ where $M_i^3[\tilde{\theta}] = C_{\tilde{\theta}}(i)$. Throughout this section, we write $\tilde{\Gamma}^k$ and $\tilde{g}^k$ as the two-person counterparts of $\Gamma^k$ and $g^k$, respectively, used for Theorem 1.

The outcome function $\tilde{g}^k$ induces the following rules in a similar way $g^k$ used for Theorem 1 is defined: for each $m \in M^k$,

**Rule 1:** If there exists $\theta^m \in \Theta$ such that $m_0^1 = m_1^1 = \theta^m$ and $i^* \in \{0, 1\}$ such that $[m_0^2[\theta^m] + m_1^2[\theta^m] \pmod{k}] = i^*$, then

$$\tilde{g}^k(m) = m_{i^*}^3[\theta^m].$$

In words, Rule 1 says that if two agents agree on the state $\theta^m$, they play the modified modulo game so that the modulo game winner - if one is selected - dictates the lottery which needs to be chosen in his choice set.

**Rule 2:** Otherwise,

$$\tilde{g}^k(m) = z(m_0^1, m_1^1).$$

In words, Rule 2 says that if either the two agents do not agree on the state or if they agree upon a state but no modulo winner is selected, then their distinct state announcements $m_0^1$ and $m_1^1$ jointly determine the lottery by $z(m_0^1, m_1^1)$.

Before we prove our main theorem in this section, we first establish the following result.

**Lemma 4** *If a finite environment* $\mathcal{E} = (\{0, 1\}, A, \Theta, (\succeq_i^\theta)_{i\in\{0,1\},\theta\in\Theta}, \hat{\mathcal{U}})$ *satisfies Condition* $COM_2$ *with respect to* $f$*, then there exists a collection of forums* $\{\mathcal{F}_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta} = \{\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$ *which satisfies Properties 1-i, 2, 3 and 4 of Condition COM with respect to* $f$*.*

**Proof:** For each $\theta \in \Theta$, we set $\mathcal{G}_\theta = \{0, 1\}$ and $w_\theta(i) = i$ for all $i \in \{0, 1\}$. By our hypothesis, there is a collection of choice sets $\{C_\theta\}_{\theta\in\Theta}$ and a punishment function $z$ that satisfy Properties 1 through 4 of Condition $COM_2$. For each $\theta \in \Theta$, we set $z_\theta = z(\theta, \theta)$. We first claim that $\{\mathcal{F}_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta} = \{\mathcal{G}_{\tilde{\theta}}, w_{\tilde{\theta}}, C_{\tilde{\theta}}, z_{\tilde{\theta}}\}_{\tilde{\theta}\in\Theta}$ is a collection of forums. It is immediate to see that the constructed collection of forums satisfies

Properties 1 through 3 of the definition of forums. Property 4 of a forum follows from Property 2 of a forum-2, which implies $z(\theta,\theta) \in C_\theta(0) \cap C_\theta(1)$.

It remains to show that all properties except Property 1-ii of Condition $COM$ are satisfied. First, we observe that Property 1 of Condition $COM_2$ implies Property 1-i of Condition $COM$. Second, we observe that Properties 2 and 4 of Condition $COM_2$ imply Properties 2 and 4 of Condition $COM$. Finally, we observe that Property 3 of Condition $COM_2$ implies Property 3 of Condition $COM$ because it ensures that $u_i(f(\theta),\theta) - u_i(z(\hat{\theta},\hat{\theta}),\theta) \geq \varepsilon$ for all $\theta, \hat{\theta} \in \Theta$, $u \in \hat{\mathcal{U}}$, and $i \in I$. ■

We state our second theorem for the case of two agents.

**Theorem 2** *Suppose that the two-person finite environment* $\mathcal{E} = \left(\{0,1\}, A, \Theta, (\succeq_i^\theta)_{i\in I,\theta\in\Theta}, \hat{\mathcal{U}}\right)$ *satisfies Condition* $COM_2$ *with respect to* $f$. *Then, the SCF* $f$ *is compellingly implementable with respect to* $\hat{\mathcal{U}}$.

**Proof**: The proof is in the Appendix. ■

# 8 Relation with pure Nash Implementation

In Step 1 in the proof of Theorem 1, we show that our canonical mechanism achieves pure Nash implementation under Properties 1-i, 1-ii and 2 of Condition $COM$. These properties are, therefore, sufficient for pure Nash implementation by our mechanism. It is then natural to ask how they compare to the necessary and sufficient conditions in Moore and Repullo (1990).

It turns out that Properties 1-i and 2 are more restrictive than Condition $\mu$ of Moore and Repullo (1990). The reason for this is simply that our mechanism is different from the mechanism used in Moore and Repullo (1990). In sum, Properties 1-i and 2 of Condition $COM$ allow our canonical mechanism to achieve compelling implementation, while they impose more stringent requirements on environments than those only necessary to achieve pure Nash implementation by the mechanism proposed by Moore and Repullo (1990).

We propose an environment that does not satisfy Properties 1-i and 2 of Condition $COM$, but in which pure Nash implementation is still possible:

**Example 2** *Let* $I = \{0,1,2\}$ *be the set of agents,* $A = \{a,b,c,d\}$ *be the set of social alternatives,* $\Theta = \{\theta,\theta'\}$ *be the set of states. Let* $f$ *be an SCF such that* $f(\theta) = a$, $f(\theta') = c$. *Agent 0's state dependent preferences over A are given by*

$$\succ_0^\theta = a \succ_0^\theta b \succ_0^\theta d \succ_0^\theta c \text{ and } \succ_0^{\theta'} = a \succ_0^{\theta'} b \succ_0^{\theta'} c \succ_0^{\theta'} d.$$

*Agent 1's state dependent preferences over A are given by*

$$\succ_1^\theta = a \succ_1^\theta \succ_2^\theta b \succ_1^\theta d \succ_1^\theta c \text{ and } \succ_1^{\theta'} = b \succ_1^{\theta'} a \succ_1^{\theta'} c \succ_1^{\theta'} d.$$

*Finally, agent 2's preferences over A at state $\theta$ are identical to those by agent 0 in that state, while agent 2's preferences over A at state $\theta^{'}$ are identical to those by agent 1 in that state. We summarize this environment as $\mathcal{E}^{**} = (I, A, \Theta, (\succeq_i^\theta )_{i\in I,\theta\in\Theta})$*

Assume that the environment $\mathcal{E}^{**}$ admits a collection of forums $\{\mathcal{F}_{\tilde\theta}\}_{\tilde\theta\in\Theta} = \{\mathcal{G}_{\tilde\theta}, w_{\tilde\theta}, C_{\tilde\theta}, z_{\tilde\theta}\}_{\tilde\theta\in\Theta}$. We also set $\hat{\mathcal{U}} = \mathcal{U}$, which is the set of all possible cardinal utility functions consistent with $(\succeq_i^{\tilde\theta})_{\tilde\theta\in\Theta}$. In the rest of the argument, we fix $u \in \hat{\mathcal{U}}$ arbitrarily. We claim that, if $\mathcal{F}_{\theta'}$ satisfies Property 1-i of Condition $COM$, it violates Property 2 of Condition $COM$. Notice that, for any $i \in \mathcal{G}_{\theta'}$, Property 1-i requires that $C_{\theta'}(i)$ be a subset of agent $i$'s lower contour set of $c$ at state $\theta'$. For all three agents, this lower contour set is equal to $\{c, d\}$. Moreover, at state $\theta$, all three agents have identical preferences over $\{c, d\}$. This implies that either $c$ or $d$ must be $C_{\theta'}$-acceptable at $(\theta, u)$.[15] This violates Property 2, as $f(\theta) = a \neq c, d$. In particular, if the SCF $f$ is compellingly implementable by our canonical mechanism $\Gamma^k$, there necessarily exists $m \in pureNE(\Gamma^k(\theta))$ such that $g(m) \in \{c, d\}$.

Despite the failure of compelling implementation in this example, pure Nash implementation is still possible if we adopt a different mechanism than our canonical mechanism. Consider the 2-agent mechanism depicted in Table 7, in which only the messages from agents 0 and 1 are used to determine the outcome and agent 2 is excluded from the mechanism:

| $g(m)$ | | Agent 1 | | |
|---|---|---|---|---|
| | | $m_1^1$ | $m_1^2$ | $m_1^3$ |
| Agent 0 | $m_0^1$ | $c$ | $d$ | $d$ |
| | $m_0^2$ | $d$ | $a$ | $b$ |
| | $m_0^3$ | $d$ | $b$ | $a$ |

Table 7: The mechanism used by Example 4 of Jackson (1992).

This is the same mechanism featured in Example 4 of Jackson (1992) with the same SCF as in the current example. In addition, the preferences in this example are very similar to the ones in Jackson's (1992) example. In this case, we have $\mathcal{G}_\theta = \mathcal{G}_{\theta'} = \{0, 1\}$. The main reason why pure Nash implementation

[15] We can show this easily. If $d \in C_{\theta'}(i)$ for some $i \in \mathcal{G}_{\theta'}$, it follows that $d$ is $C_{\theta'}$-acceptable at $(\theta, u)$. Otherwise, we have $C_{\theta'}(i) = c$ for all $i \in \mathcal{G}_{\theta'}$. This implies that $c$ is $C_{\theta'}$-acceptable at $(\theta, u)$.

is possible, while Property 2 is violated, is that in the above mechanism, the set of outcomes each agent can induce crucially depends on what message the other agent chooses. If we are only concerned with pure Nash implementation, there is a loss of generality to use $C_{\theta'}(i) \subseteq \{c, d\}$ as the set possible outcomes agent $i \in \mathcal{G}_{\theta'}$ can induce in the mechanism. For example, even if $d$ is the best alternative in the set $C_{\theta'}(i)$ at state $\theta$, this does not necessarily imply that $d$ is a Nash equilibrium outcome at that state. Indeed, each agent can unilaterally induce outcome $a$ at any message profile that yields outcome $d$ in the above mechanism. In contrast, in our mechanism, the set $C_{\theta'}(i)$ defines all possible outcomes agent $i \in \mathcal{G}_{\theta'}$ can induce unilaterally, once the forum $\mathcal{F}_{\theta'}$ was selected. This is an important distinction between our canonical mechanism and other mechanisms such as the one above, because our mechanism enables us to restrict the set of outcomes an agent can induce when contemplating a mixed strategy, once the forum has been selected. The cost of having this feature in our mechanism is the possibility that some undesirable outcomes can easily be realized as pure Nash equilibria. This motivates us to propose Condition $COM$, which shuts down such a possibility, and nevertheless, allows us to achieve compelling implementation.

# 9 Conclusion

We present a concept of compelling implementation, which strengthens the requirement of pure-strategy Nash implementation with an additional property that every mixed strategy equilibrium is either socially desirable or "uncompelling" in the sense that its outcome is strictly Pareto dominated by the socially desirable outcome. The main contribution of this paper is to propose Condition $COM$ under which compelling implementation is possible by finite mechanisms in environments with at least three agents. We construct an example that satisfies Condition $COM$ and show that Condition $COM$ is indispensable for our result. We also propose Condition $COM_2$ to extend our compelling implementation result to the case of two agents. Our implementing mechanism has desirable properties: transfers are not needed at all; only finite mechanisms are used; integer games are not invoked; and agents' risk attitudes do not matter.

# 10 Appendix

In this appendix, we provide the proofs we omitted in the main body of the paper.

## 10.1 Proof of Lemma 2

**Proof of Step 1a**: Assume by way of contradiction that there exists an integer $x \in \{0, \ldots, k-1\}$ such that $\sigma_0(x) > 0$ and $\sigma_1(x) = 0$. Then, there are two possibilities: either there exists $x^{'} \in \{0, \ldots, k-1\} \backslash \{x\}$ such that $\sigma_1(x^{'}) > 0$ or $\sigma_1(k) = 1$.

In the first case, let $x^{'} \in \arg\max_{x'' \in \{0,\ldots,k-1\}\backslash\{x\}} \sigma_1(x^{''})$. The expected payoff for agent 0 when sending message $x$ is

$$u_0(g(x, \sigma_1); \theta^{'}) = \begin{cases} \sigma_1(x+1)u_0(b; \theta^{'}) & \text{if } x < k-1 \\ \sigma_1(0)u_0(b; \theta^{'}) & \text{if } x = k-1, \end{cases}$$

where we take into account that $u_i(d; \theta^{'}) = 0$. On the other hand, the expected payoff for agent 0 when sending message $x^{'}$ is given by

$$u_0(g(x, \sigma_1); \theta^{'}) = \begin{cases} \sigma_1(x^{'})u_0(a; \theta^{'}) + \sigma_1(x^{'}+1)u_0(b; \theta^{'}) & \text{if } x^{'} < k-1 \\ \sigma_1(x^{'})u_0(a; \theta^{'}) + \sigma_1(0)u_0(b; \theta^{'}) & \text{if } x^{'} = k-1 \end{cases}$$

As $u_0(a, \theta^{'}) > u_0(b, \theta^{'})$ and $\sigma_1(x^{'}) \geq \sigma_1(x+1)$, sending message $x^{'}$ is strictly better for agent 1 than sending $x$ against $\sigma_1$, thus contradicting the hypothesis that message $x$ is played with positive probability in the Nash equilibrium $\sigma$.

Consider the second possibility where agent 1 sends $k$ with probability 1. Then, agent 0's expected payoff of sending message $x$ is $u_0(g(x, \sigma_1); \theta^{'}) = 0$, while agent 0's expected payoff of sending message $k$ is $u_0(g(\sigma_1; k); \theta^{'}) = u_0(c, \theta^{'}) > 0$, contradicting the hypothesis that message $x$ is played with positive probability in the Nash equilibrium $\sigma$. ■

**Proof of Step 1b**: Assume by way of contradiction that there exists $x \in \{0, \ldots, k-1\}$ such that $\sigma_1(x) > 0$ and $\sigma_0(x-1) = 0$ if $x \geq 1$ and $\sigma_0(k-1) = 0$ if $x = 0$. Then we decompose our argument into the following two cases: (i) there exists $x^{'} \in \{0, \ldots, k-1\}$ such that $\sigma_0(x^{'}) > 0$ or (ii) $\sigma_0(k) = 1$.

We first consider Case (i). We assume without loss of generality that $x^{'} \in \arg\max_{x'' \in \{0,\ldots,k-1\}} \sigma_0(x^{''})$. Agent 1's expected payoff of sending message $x$ against $\sigma_0$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0, x); \theta^{'}) = \sigma_0(x)u_1(a; \theta'),$$

while agent 2's expected payoff of sending message ($x^{'} + 1$ mod $k$) against $\sigma_0$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0, x^{'} + 1 \bmod k); \theta^{'}) = \begin{cases} \sigma_1(x^{'})u_1(b; \theta^{'}) + \sigma_0(x^{'}+1)u_1(a; \theta^{'}) & \text{if } x^{'} < k-1 \\ \sigma_1(x^{'})u_1(b; \theta^{'}) + \sigma_0(0)u_1(a; \theta^{'}) & \text{if } x^{'} = k-1, \end{cases}$$

where we take into account that $u_1(d;\theta^{'}) = 0$. Since $u_1(b;\theta^{'}) > u_1(a;\theta^{'}) > 0$, due to the way $x^{'}$ is defined, we have $u_1(g(\sigma_0, x^{'}+1 \bmod k);\theta^{'}) > u_1(g(\sigma_0, x);\theta^{'})$, which contradicts the hypothesis that message $x$ is sent with positive probability in the Nash equilibrium $\sigma$.

We next consider Case (ii). Agent 2's expected payoff of sending message $x$ against $\sigma_1$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0, x);\theta^{'}) = 0,$$

where we take into account that $u_1(d;\theta^{'}) = 0$. On the contrary, agent 2's expected payoff of sending message $k$ against $\sigma_0$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0, k);\theta^{'}) = u_1(c;\theta^{'}).$$

Since $u_1(c;\theta^{'}) > u_1(d;\theta^{'}) = 0$, we have $u_1(g(\sigma_0, k);\theta^{'}) > u_1(g(\sigma_0, x);\theta^{'})$, contradicting the hypothesis that message $x$ is sent with positive probability in the Nash equilibrium $\sigma$ in the game $\Gamma(\theta^{'})$. ■

**Proof of Step 1c**: Assume first that $i = 0$; that is, there exists $x^{'} \in \{0, \ldots, k-1\}$ such that $\sigma_0(x^{'}) > 0$. By Step 1a, we first have that $\sigma_1(x^{'}) > 0$. Second, by Step 1b, $\sigma_1(x^{'}) > 0$ implies $\sigma_0(x^{'} - 1) > 0$ if $x^{'} \geq 1$ and $\sigma_0(k) > 0$ if $x^{'} = 0$. Third, using Step 1a once again, we conclude that $\sigma_1(x^{'} - 1) > 0$ if $x^{'} \geq 1$ and $\sigma_1(k) > 0$ if $x^{'} = 0$. Finally, iterating this argument, we are able to conclude that $\sigma_0(x) > 0$ and $\sigma_1(x) > 0$ for all $x \in \{0, \ldots, k-1\}$.

The case where $i = 1$ is analogous to the previous one, only that we start the loop by applying Step 1b first, before Step 1a. This completes the proof of Step 1c. ■

**Proof of Step 2**: Assume by way of contradiction that there exist $i \in \{0, 1\}$ and $x, x^{'} \in \{0, \ldots, k-1\}$ such that $\sigma_i(x) > \sigma_i(x^{'}) > 0$. By Step 1c, we know that $\sigma_i(\tilde{x}) > 0$ for all $\tilde{x} \in \{0, \ldots, k-1\}$. Then, we can choose $x$ and $x^{'}$ satisfying the following property:

$$x \in \arg \max_{\tilde{x} \in \{0,\ldots,k-1\}} \sigma_i(\tilde{x}) \text{ and } x^{'} \in \arg \min_{\tilde{x} \in \{0,\ldots,k-1\}} \sigma_i(\tilde{x}).$$

By Step 1c, we also know that $\sigma_j(\tilde{x}) > 0$ for each $\tilde{x} \in \{0, \ldots, k-1\}$, where $j \in \{1, 2\} \backslash \{i\}$.

Assume that $i = 1$. The expected payoff for agent 0 of sending message $x^{'}$ against $\sigma_1$ in the game $\Gamma(\theta^{'})$ is given by

$$u_0(g(x^{'}, \sigma_1);\theta^{'}) = \begin{cases} \sigma_1(x^{'})u_0(a;\theta^{'}) + \sigma_1(x^{'}+1)u_0(b;\theta^{'}) & \text{if } x^{'} < k-1 \\ \sigma_1(x^{'})u_0(a;\theta^{'}) + \sigma_1(0)u_0(b;\theta^{'}) & \text{if } x^{'} = k-1 \end{cases}$$

On the other hand, The expected payoff for agent 0 of sending message $x$ against $\sigma_1$ in the game $\Gamma(\theta^{'})$ is given by

$$u_0(g(x,\sigma_1);\theta^{'}) = \begin{cases} \sigma_1(x)u_0(a;\theta^{'}) + \sigma_1(x+1)u_0(b;\theta^{'}) & \text{if } x < k-1 \\ \sigma_1(x)u_0(a;\theta^{'}) + \sigma_1(0)u_0(b;\theta^{'}) & \text{if } x = k-1. \end{cases}$$

We compute

$$\begin{aligned} & u_0(g(x,\sigma_1);\theta^{'}) - u_0(g(x^{'},\sigma_1);\theta^{'}) \\ = \; & [\sigma_1(x) - \sigma_1(x^{'})]u_0(a;\theta^{'}) + [\sigma_1(x+1 \bmod k) - \sigma_1(x'+1 \bmod k)]u_0(b;\theta^{'}) \\ \geq \; & [\sigma_1(x) - \sigma_1(x^{'})]u_0(a;\theta^{'}) - [\sigma_1(x) - \sigma_1(x^{'})]u_0(b;\theta^{'}) \\ & (\because [\sigma_1(x+1 \bmod k) - \sigma_1(x'+1 \bmod k)] \geq -[\sigma_1(x) - \sigma_1(x^{'})], u_0(b;\theta^{'}) > 0) \\ = \; & [\sigma_1(x) - \sigma_1(x^{'})](u_0(a;\theta^{'}) - u_0(b;\theta^{'}) \\ > \; & 0. \end{aligned}$$

This implies that message $x$ is a strictly better response for agent 0 against $\sigma_1$ than $x^{'}$ in the game $\Gamma(\theta^{'})$, contradicting the hypothesis that $\sigma_0(x^{'}) > 0$.

We next assume $i = 0$. The expected payoff for agent 1 of sending message $x^{'}+1$ against $\sigma_0$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0,x^{'}+1);\theta^{'}) = \begin{cases} \sigma_0(x^{'}+1)u_1(a;\theta^{'}) + \sigma_0(x^{'})u_1(b;\theta^{'}) & \text{if } x^{'} < k-1 \\ \sigma_0(0)u_1(a;\theta^{'}) + \sigma_0(x')u_1(b;\theta^{'}) & \text{if } x^{'} = k-1 \end{cases}$$

On the other hand, The expected payoff for agent 1 of sending message $x+1$ against $\sigma_0$ in the game $\Gamma(\theta^{'})$ is given by

$$u_1(g(\sigma_0,x+1);\theta^{'}) = \begin{cases} \sigma_0(x+1)u_1(a;\theta^{'}) + \sigma_0(x)u_1(b;\theta^{'}) & \text{if } x < k-1 \\ \sigma_0(0)u_1(a;\theta^{'}) + \sigma_0(x)u_1(b;\theta^{'}) & \text{if } x = k-1. \end{cases}$$

We compute

$$\begin{aligned} & u_1(g(\sigma_0,x+1);\theta^{'}) - u_1(g(\sigma_0,x^{'}+1);\theta^{'}) \\ = \; & [\sigma_0(x+1) - \sigma_0(x^{'}+1)]u_1(a;\theta^{'}) + [\sigma_0(x) - \sigma_0(x')]u_1(b;\theta^{'}) \\ \geq \; & [\sigma_0(x+1) - \sigma_0(x^{'}+1)]u_1(b;\theta^{'}) - [\sigma_0(x) - \sigma_0(x^{'})]u_1(a;\theta^{'}) \\ & (\because [\sigma_0(x+1 \bmod k) - \sigma_0(x'+1) \bmod k)] \geq -[\sigma_0(x) - \sigma_0(x^{'})], u_1(a;\theta^{'}) > 0) \\ = \; & [\sigma_0(x) - \sigma_0(x^{'})](u_1(b;\theta^{'}) - u_1(a;\theta^{'})) \\ > \; & 0. \end{aligned}$$

This implies that message $x+1$ is a strictly better response for agent 1 against $\sigma_0$ than $x^{'}+1$ in the game $\Gamma(\theta^{'})$, contradicting the hypothesis that $\sigma_1(x^{'}+1) > 0$. This completes the proof of Step 2. ■

## 10.2 Condition $COM$ holds in Example 1

We shall show that all properties of Condition $COM$ are satisfied.

- Property 1-i

  $f(\theta_a) = a$ is $C_{\theta_a}$-acceptable at $(\theta_a, u)$ for any $u \in \mathcal{U}^\varepsilon$ because $a$ is the best outcome within $C_{\theta_a}(i) = A$ for each $i \in \{0, 1\}$. $f(\theta_b) = b$ is $C_{\theta_b}$-acceptable at $(\theta_b, u)$ for any $u \in \mathcal{U}^\varepsilon$ because $b$ is the best outcome within $C_{\theta_b}(i) = A$ for each $i \in \{1, 2\}$. $f(\theta_c) = c$ is $C_{\theta_c}$-acceptable at $(\theta_c, u)$ for any $u \in \mathcal{U}^\varepsilon$ because $c$ is the best outcome within $C_{\theta_c}(i) = \{c, d, z\}$ for each $i \in \{1, 2\}$. $f(\theta_d) = d$ is $C_{\theta_d}$-acceptable at $(\theta_d, u)$ for any $u \in \mathcal{U}^\varepsilon$ because $d$ is the best outcome within $C_{\theta_d}(i) = \{c, d, z\}$ for each $i \in \{0, 2\}$.

- Property 1-ii

  By construction, we set $C_{\tilde{\theta}}(i) = C_{\tilde{\theta}}(j)$ for all $\tilde{\theta} \in \Theta$ and $i, j \in \mathcal{G}_{\tilde{\theta}}$. This guarantees that $f(\theta) \in C_\theta(i) \cap C_\theta(j)$ for all $\theta \in \Theta$ and $i, j \in \mathcal{G}_\theta$.

  - When $\theta = \theta_a$ and $\hat{\theta} = \theta_b$, it follows that $f(\theta_a) = a \in C_{\theta_b}(i) = A$ for all $i \in \{1, 2\}$. When $\theta = \theta_a$ and $\hat{\theta} = \theta_c$, it follows that $f(\theta_a) = a \notin C_{\theta_c}(i) = \{c, d, z\}$ for all $i \in \{1, 2\}$. When $\theta = \theta_a$ and $\hat{\theta} = \theta_d$, it follows that $f(\theta_a) = a \notin C_{\theta_d}(i) = \{c, d, z\}$ for all $i \in \{0, 2\}$.
  - When $\theta = \theta_b$ and $\hat{\theta} = \theta_a$, it follows that $f(\theta_b) = b \in C_{\theta_a}(i) = A$ for all $i \in \{0, 1\}$. When $\theta = \theta_b$ and $\hat{\theta} = \theta_c$, it follows that $f(\theta_b) = b \notin C_{\theta_c}(i) = \{c, d, z\}$ for all $i \in \{1, 2\}$. When $\theta = \theta_b$ and $\hat{\theta} = \theta_d$, it follows that $f(\theta_b) = b \notin C_{\theta_d}(i) = \{c, d, z\}$ for all $i \in \{0, 2\}$.
  - When $\theta = \theta_c$ and $\hat{\theta} = \theta_a$, it follows that $f(\theta_c) = c \in C_{\theta_a}(i) = A$ for all $i \in \{0, 1\}$. When $\theta = \theta_c$ and $\hat{\theta} = \theta_b$, it follows that $f(\theta_c) = c \in C_{\theta_b}(i) = A$ for all $i \in \{1, 2\}$. When $\theta = \theta_c$ and $\hat{\theta} = \theta_d$, it follows that $f(\theta_c) = c \in C_{\theta_d}(i) = \{c, d, z\}$ for all $i \in \{0, 2\}$.
  - When $\theta = \theta_d$ and $\hat{\theta} = \theta_a$, it follows that $f(\theta_d) = d \in C_{\theta_a}(i) = A$ for all $i \in \{0, 1\}$. When $\theta = \theta_d$ and $\hat{\theta} = \theta_b$, it follows that $f(\theta_c) = c \in C_{\theta_b}(i) = A$ for all $i \in \{1, 2\}$. When $\theta = \theta_d$ and $\hat{\theta} = \theta_c$, it follows that $f(\theta_d) = d \in C_{\theta_c}(i) = \{c, d, z\}$ for all $i \in \{1, 2\}$.

- Property 2

  When $\theta = \theta_a$, for any $\hat{\theta} \in \{\theta_a, \theta_b\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $a$ is the unique outcome in $\Delta(A)$ that is $C_{\hat{\theta}}$-acceptable at $(\theta_a, u)$. Then, we have $f(\theta_a) = a$.

  - When $\theta = \theta_a$ and $\hat{\theta} = \theta_c$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_c}$-acceptable at $(\theta_a, u)$ because $u_1(d, \theta_a) > u_1(c, \theta_a)$ and $u_2(d, \theta_a) < u_2(c, \theta_a)$ and $C_{\theta_c}(i) = \{c, d, z\}$ for each $i \in \mathcal{G}_{\theta_c}$.

– When $\theta = \theta_a$ and $\hat{\theta} = \theta_d$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_d}$-acceptable at $(\theta_a, u)$ because $u_0(d, \theta_a) > u_0(c, \theta_a)$ and $u_2(d, \theta_a) < u_2(c, \theta_a)$ and $C_{\theta_d}(i) = \{c, d, z\}$ for each $i \in \mathcal{G}_{\theta_d}$.

– When $\theta = \theta_b$, for any $\hat{\theta} \in \{\theta_a, \theta_b\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $b$ is the unique outcome in $\Delta(A)$ that is $C_{\hat{\theta}}$-acceptable at $(\theta_b, u)$. Then, we have $f(\theta_b) = b$.

– When $\theta = \theta_b$ and $\hat{\theta} = \theta_c$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_c}$-acceptable at $(\theta_b, u)$ because $u_1(d, \theta_b) < u_1(c, \theta_b)$ and $u_2(d, \theta_b) > u_2(c, \theta_b)$ and $C_{\theta_c}(i) = \{c, d, z\}$ for each $i \in \mathcal{G}_{\theta_c}$.

– When $\theta = \theta_b$ and $\hat{\theta} = \theta_d$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_d}$-acceptable at $(\theta_b, u)$ because $u_0(d, \theta_a) < u_0(c, \theta_a)$ and $u_2(d, \theta_a) > u_2(c, \theta_a)$ and $C_{\theta_d}(i) = \{c, d, z\}$ for each $i \in \mathcal{G}_{\theta_d}$.

– When $\theta = \theta_c$ and $\hat{\theta} = \theta_a$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_a}$-acceptable at $(\theta_c, u)$ because $u_0(a, \theta_c) > u_0(b, \theta_c)$ and $u_1(a, \theta_c) < u_1(b, \theta_c)$ and $C_{\theta_a}(i) = A$ for each $i \in \mathcal{G}_{\theta_a}$.

– When $\theta = \theta_c$ and $\hat{\theta} = \theta_b$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_b}$-acceptable at $(\theta_c, u)$ because $u_1(a, \theta_c) < u_1(b, \theta_c)$ and $u_2(a, \theta_c) > u_2(b, \theta_c)$ and $C_{\theta_b}(i) = A$ for each $i \in \mathcal{G}_{\theta_b}$.

– When $\theta = \theta_c$ and $\hat{\theta} = \theta_d$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $c \in \Delta(A)$ is the unique outcome that is $C_{\theta_d}$-acceptable at $(\theta_c, u)$. Then, we have $f(\theta_c) = c$.

– When $\theta = \theta_d$ and $\hat{\theta} = \theta_a$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_a}$-acceptable at $(\theta_d, u)$ because $d$ is agent 0's best outcome within $C_{\theta_a}(0)$ in state $\theta_d$, while $b$ is agent 1's best outcome within $C_{\theta_a}(1)$ in state $\theta_b$ and $C_{\theta_a}(i) = A$ for each $i \in \mathcal{G}_{\theta_a}$.

– When $\theta = \theta_d$ and $\hat{\theta} = \theta_b$, for any $u \in \mathcal{U}^\varepsilon$, it follows that there is no $x \in \Delta(A)$ that is $C_{\theta_b}$-acceptable at $(\theta_d, u)$ because $b$ is agent 1's best outcome within $C_{\theta_b}(1)$ in state $\theta_d$, while $a$ is agent 2's best outcome within $C_{\theta_b}(2)$ in state $\theta_d$ and $C_{\theta_b}(i) = A$ for each $i \in \mathcal{G}_{\theta_b}$.

– When $\theta = \theta_d$ and $\hat{\theta} = \theta_c$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $d \in \Delta(A)$ is the unique outcome that is $C_{\theta_c}$-acceptable at $(\theta_c, u)$. Then, we have $f(\theta_d) = d$.

- Property 3

  This property is satisfied due to the very construction of $\mathcal{U}^\varepsilon$.

- Property 4

- When $\theta = \theta_a$ and $\hat{\theta} = \theta_b$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_a) = a$ is $C_{\theta_b}$-acceptable at $(\theta_a, u)$. Since $a$ is the best outcome for both agents 1 and 2 in state $\theta_a$ and $\mathcal{G}_{\theta_b} = \{1, 2\}$, this property holds.
- When $\theta = \theta_a$, for any $\hat{\theta} \in \{\theta_c, \theta_d\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_a)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta_a, u)$. Hence, the property holds.
- When $\theta = \theta_b$ and $\hat{\theta} = \theta_a$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_b) = b$ is not $C_{\theta_a}$-acceptable at $(\theta_b, u)$ because $u_0(a, \theta_b) > u_0(b, \theta_b)$, while $u_1(a, \theta_b) < u_1(b, \theta_b)$ and $C_{\theta_a}(i) = A$ for each $i \in \mathcal{G}_{\theta_a}$. Hence, this property holds.
- When $\theta = \theta_b$, for any $\hat{\theta} \in \{\theta_c, \theta_d\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_b)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta_b, u)$. Hence, the property holds.
- When $\theta = \theta_c$, for any $\hat{\theta} \in \{\theta_a, \theta_b\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_c) = c$ is not $C_{\hat{\theta}}$-acceptable at $(\theta_c, u)$. Hence, the property holds.
- When $\theta = \theta_c$ and $\hat{\theta} = \theta_d$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_c) = c$ is $C_{\theta_d}$-acceptable at $(\theta_c, u)$. Since $c$ is the best outcome within $\{c, d, z\}$ for both agents 0 and 2 in state $\theta_c$ and $C_{\theta_d}(i) = \{c, d, z\}$ for each $i \in \mathcal{G}_{\theta_d}$, the property holds.
- When $\theta = \theta_d$, for any $\hat{\theta} \in \{\theta_a, \theta_b\}$ and $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_d) = d$ is not $C_{\hat{\theta}}$-acceptable at $(\theta_c, u)$. Hence, the property holds.
- When $\theta = \theta_d$ and $\hat{\theta} = \theta_c$, for any $u \in \mathcal{U}^\varepsilon$, it follows that $f(\theta_d) = d$ is $C_{\theta_c}$-acceptable at $(\theta_d, u)$. For any $\tilde{a} \in \{c, z\}$, we have $u_i(\tilde{a}, \theta_d) \neq u_i(f(\theta_d), \theta_d)$ for each $i \in \{1, 2\} = \mathcal{G}_{\theta_d}$. Hence, the property holds.

Therefore, this environment satisfies Condition $COM$ with respect to the SCF $f$ and $\mathcal{U}^\varepsilon$.

## 10.3 Proof of Theorem 2

**Proof of Theorem 2**: Lemma 4 allows us to make use of many of the arguments established in the proof of Theorem 1 to prove Theorem 2. The main difference between the two proofs lies on the fact that the outcome function $\tilde{g}^k$ is now slightly different from the outcome function $g^k$ used for Theorem 1. In particular, we need to modify the arguments in Steps 1 and 2 for Theorem 1 because establishing two-person pure Nash implementation requires a separate proof. Therefore, we first need to establish the two-person counterparts of these two steps. Once that is done, by Lemma 4, we will show that the rest of the proof follows from the same argument in Theorem 1.

Let $f$ be an SCF and $I = \{0, 1\}$. Suppose that the finite environment $\mathcal{E}$ satisfies Condition $COM_2$ with respect to $f$ and $\hat{\mathcal{U}}$. The rest of the proof is completed by a series of steps in a very similar way Theorem 1 is proved.

**Step 1:** For each $k \geq 2$, $f$ is pure Nash implementable by the mechanism $\tilde{\Gamma}^k$.

**Proof of Step 1:** Let $\theta \in \Theta$ be a true state and $u \in \hat{\mathcal{U}}$ be an arbitrary cardinal utility function. Fix $k \geq 2$. By Property 1, we know that $f(\theta)$ is $C_\theta$-acceptable at $(\theta, u)$. Thus, $f(\theta) \in C_\theta(i)$ for some $i \in I$. We assume, without loss of generality, that $i = 0$. Let $m$ be a message profile with the following characteristics:

- $m_j^1 = \theta$ for all $j \in I$;
- $m_j^2[\theta] = 0$ for all $j \in I$;
- $m_0^3[\theta] = f(\theta)$.

By construction of $m$, we have $\tilde{g}^k(m) = f(\theta)$. We first establish $m \in pureNE(\tilde{\Gamma}^k(\theta))$. Notice that any deviation by either agents 0 or 1 will yield a lottery that is in either $C_\theta(0)$ (if agent 0 is deviating) or $C_\theta(1)$ (if agent 1 is deviating). This is immediate for the case in which the deviating strategies preserve $m_i^1 = \theta$, but even a deviation involving a different choice for either $\theta_0$ or $\theta_1$ will result in $z(\theta_0, \theta)$ or $z(\theta, \theta_1)$, respectively, both of which are in the sets $C_\theta(0)$ and $C_\theta(1)$. Then, Property 1 of Condition $COM_2$ ensures that no lottery in $\{C_\theta(i)\}_{i \in \{0,1\}}$, which includes all possible outcomes which are induced by Rules 1 or 2, can be strictly preferred to $f(\theta)$. Thus, we show that $m \in pureNE(\tilde{\Gamma}^k(\theta))$.

We next show that $m \in pureNE(\tilde{\Gamma}^k(\theta))$ implies $\tilde{g}^k(m) = f(\theta)$. We assume, by way of contradiction, that there exists $m \in pureNE(\tilde{\Gamma}^k(\theta))$ such that $\tilde{g}^k(m) \neq f(\theta)$. The, there are two cases we need to consider. The first case occurs when there exists $\hat{\theta} \in \Theta$ such that $\tilde{g}^k(m) \in C_{\hat{\theta}}(0) \cup C_{\hat{\theta}}(1)$. The second case occurs when no such $\hat{\theta}$ exists, which implies that $\tilde{g}^k(m) = z(\theta_0, \theta_1)$, where $\theta_0 \equiv m_0^1$ and $\theta_1 \equiv m_1^1$.

Suppose that the first case applies. It follows from Property 2 of Condition $COM_2$ that, for any $\hat{\theta} \in \Theta$, $\tilde{g}^k(m)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$. Therefore, there exist agent $i \in \{0, 1\}$ and $x \in C_{\hat{\theta}}(i)$ such that $u_i(x, \theta) > u_i(\tilde{g}^k(m), \theta)$. We define $\hat{m}_i^3[\hat{\theta}] = x$ and $\hat{m}_i^2[\hat{\theta}] \in \{0, \ldots, k-1\}$ such that $m_i^2[\hat{\theta}] + m_j^2[\hat{\theta}] \pmod{k} = i$. We define $\hat{m}_i$ to be identical to $m_i$ except that $m_i^3[\hat{\theta}]$ is replaced by $\hat{m}_i^3[\hat{\theta}]$ and $m_i^2[\hat{\theta}]$ is replaced by $\hat{m}_i^2[\hat{\theta}]$. Then, $\hat{m}_i$ is agent $i$'s profitable deviation such that agent $i$ is the winner of the modulo game. Thus, $m \notin pureNE(\tilde{\Gamma}^k(\theta))$, a desired contradiction.

Next suppose that the second case applies. Then, by Property 5 of Condition $COM_2$, there exists either $a_{(\theta_0,\theta_1)} \in C_{\theta_1}(0)$ such that $u_0(a_{(\theta_0,\theta_1)}, \theta) > u_0(z(\theta_0, \theta_1), \theta)$ or $b_{(\theta_0,\theta_1)} \in C_{\theta_0}(1)$ such that $u_1(b_{(\theta_0,\theta_1)}, \theta) > u_1(z(\theta_0, \theta_1), \theta)$. Assume the first case applies. We define $\hat{m}_0^1 = \theta_1$, $\hat{m}_0^3[\theta_1] = a_{(\theta_0,\theta_1)}$, and $\hat{m}_0^2[\theta_1]$ such that $\hat{m}_0^2[\theta_1] + m_1^2[\theta_1] \pmod{k} = 0$. Then, we define $\hat{m}_0$ to be identical to $m_0$ except that $m_0^1$ is replaced by $\hat{m}_0^1$, $m_0^3[\theta_1]$ is replaced by $\hat{m}_0^3[\theta_1]$, and $m_0^2[\theta_1]$ is replaced by $\hat{m}_0^2[\theta_1]$. This implies that $\hat{m}_0$, which ensures that agent 0 is the winner of the modulo game, is a profitable deviation from $m$ so that $m \notin pureNE(\tilde{\Gamma}^k(\theta))$, a desired contradiction.

Assume the second case applies. We define $\hat{m}_1^1 = \theta_0$, $\hat{m}_1^3[\theta_0] = b_{(\theta_0,\theta_1)}$, and $\hat{m}_1^2[\theta_0]$ such that $\hat{m}_1^2[\theta_0] + m_0^2[\theta_0] \pmod{k} = 1$. Then, we define $\hat{m}_1$ to be identical to $m_1$ except that $m_1^1$ is replaced by $\hat{m}_1^1$, $m_1^3[\theta_0]$ is replaced by $\hat{m}_1^3[\theta_0]$, and $m_1^2[\theta_0]$ is replaced by $\hat{m}_1^2[\theta_0]$. This implies that $\hat{m}_1$, which ensures that agent 1 is the winner of the modulo game, is a profitable deviation from $m$ so that $m \notin pureNE(\tilde{\Gamma}^k(\theta))$, a desired contradiction. This completes the proof. ■

Throughout the proof, we denote by $\theta$ the true state and by $\hat{\theta}$ the state determined by the agents' announcement in the mechanism. Let $\tilde{\Gamma}^k = (M^k, \tilde{g}^k)$ be our canonical mechanism where $k \geq 3$. We define $C_{\hat{\theta}} \equiv \bigcup_{i \in I} C_{\hat{\theta}}(i)$ for each $\hat{\theta} \in \Theta$, and $C \equiv \bigcup_{\hat{\theta} \in \Theta} C_{\hat{\theta}}$. Note that $C_{\hat{\theta}}$ and $C$ are both finite. In the rest of the proof, we use the same notation used in the proof of Theorem 1. The difference centers around the fact that we use a different outcome function $\tilde{g}^k$ rather than $g^k$. Therefore, to stress this difference, we add "tilde" to the notation used in the proof of Theorem 1. Specifically, we use $\tilde{M}^*(\hat{\theta}, q, x)$, $\tilde{P}^\sigma(q|\hat{\theta})$, $\tilde{\ell}^k(\hat{\theta}, \sigma)$, and so on.

We also need to introduce a new set of notation for probabilities and lotteries that have no counterparts in Theorem 1. This is because the two agents announce different states in the first component of their message.

$$\tilde{M}^*(\theta_0, \theta_1) \equiv \left\{ m \in M^k \middle| \; m_0^1 = \theta_0 \text{ and } m_1^1 = \theta_1 \right\};$$

$$\tilde{P}^\sigma(\theta_0, \theta_1) \equiv \sum_{m \in \tilde{M}^*(\theta_0, \theta_1)} \sigma(m);$$

$$\tilde{\ell}_z^k(\sigma) = \sum_{\theta_0 \neq \theta_1} \tilde{P}^\sigma(\theta_0, \theta_1) z(\theta_0, \theta_1) \left( 1 - \sum_{\theta \in \Theta} P^\sigma(\theta) \right)^{-1}.$$

Using this notation, we establish the two-person counterpart of Step 2 in the proof of Theorem 1:

**Step 2:** For any mixed strategy profile $\sigma$ in the mechanism $\tilde{\Gamma}^k = (M^k, \tilde{g}^k)$,

$$\tilde{g}^k(\sigma) = \sum_{\hat{\theta} \in \Theta} \tilde{P}^\sigma(\hat{\theta}) \tilde{\ell}^k(\hat{\theta}, \sigma) + \left( 1 - \sum_{\hat{\theta} \in \Theta} \tilde{P}^\sigma(\hat{\theta}) \right) \tilde{\ell}_z^k(\sigma),$$

where, for each $\hat{\theta} \in \Theta$,

$$\tilde{\ell}^k(\hat{\theta}, \sigma) = \tilde{P}^\sigma(q = 0|\hat{\theta})(\tilde{\ell}_0^k(\hat{\theta}, \sigma) - z(\hat{\theta}, \hat{\theta})) + \tilde{P}^\sigma(q = 1|\hat{\theta})(\tilde{\ell}_1^k(\hat{\theta}, \sigma) - z(\hat{\theta}, \hat{\theta})) + z(\hat{\theta}, \hat{\theta}).$$

**Proof of Step 2:** This comes from how our mechanism and the lotteries are constructed. Fix $\sigma$. For each $\hat{\theta} \in \Theta$, the event that $m_0^1 = m_1^1 = \hat{\theta}$ occurs with probability $\tilde{P}^\sigma(\hat{\theta})$. Following the argument in Step 2 in the proof of Theorem 1, we can show that the outcome must be given by $\tilde{\ell}^k(\hat{\theta}, \sigma)$. On the contrary, the event that $m_0^1 = \theta_0 \neq \theta_1 = m_1^1$ occurs with the remaining probability, $1 - \sum_{\hat{\theta} \in \Theta} \tilde{P}^\sigma(\hat{\theta})$. In this case, the mechanism triggers Rule 2, which will yield $z(\theta_0, \theta_1)$, which occurs with probability $\tilde{P}^\sigma(\theta_0, \theta_1)$. This is represented by the lottery $\tilde{\ell}_z^k(\sigma)$. ■

We will now show how the rest of the argument closely follows Steps 3 through 8 developed in the proof of Theorem 1. We note that Property 3 of Condition $COM_2$ implies that $u_i(\tilde{\ell}_z^k(\sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in \{0, 1\}$, as $\tilde{\ell}_z^k(\sigma)$ is nothing but a weighted average over different punishment outcomes. In turn, this structure allows us to focus on the set of lotteries $\{\ell^k(\hat{\theta}, \sigma)\}_{\hat{\theta} \in \Theta}$ because it follows from Step 2 that if each $\tilde{\ell}^k(\hat{\theta}, \sigma)$ is strictly Pareto dominated by $f(\theta)$, then $\tilde{g}^k(\sigma)$ will also be strictly Pareto dominated by $f(\theta)$. For all $\hat{\theta} \in \Theta$, we set $\mathcal{G}_{\hat{\theta}} = \{0, 1\}$, $w_{\hat{\theta}}(q) = q$ for all $q \in \{0, 1\}$, and $z_{\hat{\theta}} = z(\hat{\theta}, \hat{\theta})$. Then, we can replicate all of the arguments in the remaining steps in the proof of Theorem 1. In what follows, we will briefly go over each of them.

**Step 3**: Let $\sigma \in NE(\tilde{\Gamma}^k(\theta, u))$ for some $u \in \hat{\mathcal{U}}$. Then, for any $m \in \text{supp}(\sigma)$, $q \in \{0, 1\}$, and $\hat{\theta} \in \Theta$, if $m \in \tilde{M}^*(\hat{\theta}, q)$, then $m_i^3[\hat{\theta}] \in C_{\hat{\theta}}^*(i, \theta)$, where $i = w_{\hat{\theta}}(q)$.

**Proof of Step 3:** This proof is identical to the proof of Step 3 in the proof of Theorem 1 if we replace $M^*(\hat{\theta}, q)$ with $\tilde{M}^*(\hat{\theta}, q)$. Since Step 3 does not require any of the properties in Condition $COM$, it relies only on the fact that whenever an agent can dictate the outcome with positive probability, he must choose pick one of his best alternatives in his choice set. Therefore, the same logic applies to the modified mechanism of the two-agent case. ■

**Step 4**: Let $\sigma$ be a mixed strategy profile in the mechanism $\tilde{\Gamma}^K$, where we later choose $K \geq 3$ large enough, and fix $\hat{\theta} \in \Theta$ such that $\tilde{P}^\sigma(\hat{\theta}) > 0$. Assume that $\tilde{P}^\sigma(q = 0|\hat{\theta}) + \tilde{P}^\sigma(q = 1|\hat{\theta}) \leq 2/K$ . Then, there exists $K \in \mathbb{N}$ large enough so that $u_i(\tilde{\ell}^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$ and $u \in \hat{\mathcal{U}}$.

**Proof of Step 4:** This proof is identical to the proof of Step 4 in the proof of Theorem 1, once we replace all the conditional probabilities and lotteries with their respective counterparts in Theorem 2. In particular, if lottery $\tilde{\ell}^K(\hat{\theta}, \sigma)$ assigns a sufficiently high probability to the event in which the punishment $z_{\hat{\theta}}$ is realized, we can use Property 3 of Condition $COM$ to conclude that $\tilde{\ell}^K(\hat{\theta}, \sigma)$ is strictly Pareto dominated by $f(\theta)$. We complete the proof of Step 4, as we acknowledge that the lottery $\tilde{\ell}^K(\hat{\theta}, \sigma)$ corresponds to $\ell^K(\hat{\theta}, \sigma)$ used for the proof of Theorem 1, and Property 3 of Condition $COM_2$ has the same implications of Property 3 of Condition $COM$. ■

**Step 5**: Let $\sigma \in NE(\tilde{\Gamma}^K(\theta, u))$, where we later choose $K$ large enough. For any $\hat{\theta} \in \Theta$ with $\tilde{P}^\sigma(\hat{\theta}) > 0$, we assume that there exists $q \in \{0, 1\}$ such that $\tilde{P}^\sigma(q|\hat{\theta}) = 0$. Then, there exists $K \in \mathbb{N}$ large enough so that, for any $\hat{\theta} \in \Theta$ with $\tilde{P}^\sigma(\hat{\theta}) > 0$, if $\tilde{\ell}^K(\hat{\theta}, \sigma) \neq f(\theta)$, then $u_i(\tilde{\ell}^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

**Proof of Step 5:** It follows the same argument in Step 5 in the proof of Theorem 1. We postpone the discussion relevant here to that in Step 6. ■

**Step 6**: Let $\sigma \in NE(\tilde{\Gamma}^K(\theta, u))$, where we later choose $K$ large enough. Assume that there exists $\hat{\theta} \in \Theta$ with $\tilde{P}^\sigma(\hat{\theta}) > 0$ such that $\tilde{P}^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. If $f(\theta)$ is not $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then there exists $K \in \mathbb{N}$ large enough so that $u_i(\tilde{\ell}^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

**Proof of Step 6**: It follows the same argument as in Step 6 in the proof of Theorem 1. Steps 5 and 6 both aim to establish the following result: there exists $K \in \mathbb{N}$ large enough so that, for any $\sigma \in NE(\tilde{\Gamma}^K(\theta, u))$, if $\tilde{\ell}^K(\hat{\theta}, \sigma) \neq f(\theta)$, then $\tilde{\ell}^K(\hat{\theta}, \sigma)$ is strictly Pareto dominated by $f(\theta)$. To show this, we rely on the proof by contradiction, i.e., we construct an alternative message for an agent that constitutes a profitable deviation from $\sigma$, which contradicts the hypothesis that $\sigma$ is an equilibrium. We can replicate the same argument even in the case of two agents because these two mechanisms $\Gamma^K$ and $\tilde{\Gamma}^K$ work in the same way as far as we are concerned with how each deviation induces the changes in the second and third components of their message $m_j^2[\hat{\theta}]$ and $m_j^3[\hat{\theta}]$. In particular, $\tilde{P}^\sigma(q|\hat{\theta})$ and $\tilde{P}_i^\sigma(q|\hat{\theta})$ both work in the same way as $P^\sigma(q|\hat{\theta})$ and $P_i^\sigma(q|\hat{\theta})$ do, as both mechanisms share the same modulo game to determine the outcome after the agents agree on the state, i.e., $\theta^m = \hat{\theta}$. By Lemma 4, we know that Properties 2, 3 and 4 of Condition $COM_2$ imply Properties 2, 3, and 4 of Condition $COM$. Therefore, we can replicate the same argument as in Steps 5 and 6 in the proof of Theorem 1. ■

**Step 7**: Let $\sigma \in NE(\tilde{\Gamma}^K(\theta, u))$, where we later choose $K$ large enough. Assume that there exists $\hat{\theta} \in \Theta$ such that $\tilde{P}^\sigma(\hat{\theta}) > 0$ and $\tilde{P}^\sigma(q|\hat{\theta}) > 0$ for all $q \in \{0, 1\}$. If $\tilde{\ell}^K(\hat{\theta}, \sigma) \neq f(\theta)$ and $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then there exists $K \in \mathbb{N}$ large enough so that $u_i(\tilde{\ell}^K(\hat{\theta}, \sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

**Proof of Step 7**: As explained in the proof of Step 7 in the proof of Theorem

1, if $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then, for every $x \in C^*_{\hat{\theta}}(0, \theta)$ we have either $x = f(\theta)$ or both $u_0(x, \theta) < u_0(f(\theta), \theta)$ and $u_1(x, \theta) < u_1(f(\theta), \theta)$. Similarly, if $f(\theta)$ is $C_{\hat{\theta}}$-acceptable at $(\theta, u)$, then, for every $x \in C^*_{\hat{\theta}}(1, \theta)$ we have either $x = f(\theta)$ or both $u_0(x, \theta) < u_0(f(\theta), \theta)$ and $u_1(x, \theta) < u_1(f(\theta), \theta)$. It follows from Step 3 that both $\tilde{\ell}^K_0(\hat{\theta}, \sigma)$ and $\tilde{\ell}^K_1(\hat{\theta}, \sigma)$ are either equal to $f(\theta)$ or strictly Pareto dominated by $f(\theta)$. Furthermore, we can use Step 2 and Property 3 of Condition $COM_2$ in the same way as in Step 7 in the proof of Theorem 1 to establish that $\tilde{\ell}^k(\hat{\theta}, \sigma)$ is strictly Pareto dominated by $f(\theta)$. ■

This allows us to conclude with Step 8 much in the same way as in the proof of Theorem 1.

**Step 8**: There exists $K \in \mathbb{N}$ large enough such that, for any $u \in \hat{\mathcal{U}}$ and $\sigma \in NE(\tilde{\Gamma}^K(\theta, u))$, it follows that either $\tilde{g}^K(\sigma) = f(\theta)$ or $u_i(\tilde{g}^K(\sigma), \theta) < u_i(f(\theta), \theta)$ for all $i \in I$.

Combining Steps 1 and 8, we conclude that there exists $K \in \mathbb{N}$ large enough such that the SCF $f$ is compellingly implementable with respect to $\hat{\mathcal{U}}$ by the mechanism $\tilde{\Gamma}^K$. This completes the proof of the theorem. ■

# References


[1] Abreu, D. and H. Matsushima, (1992), "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica*, vol. 60, 993-1008.

[2] Abreu, D. and H. Matsushima, (1994), "Exact Implementation," *Journal of Economic Theory*, vol. 64, 1-19.

[3] Aghion, P., D. Fudenberg, R. Holden, T. Kunimoto, and O. Tercieux, (2012), "Subgame-Perfect Implementation under Information Perturbations," *Quarterly Journal of Economics*, vol. 127(4), 1843-1881

[4] Bergemann, D., S. Morris, and O. Tercieux, (2011), "Rationalizable Implementation," *Journal of Economic Theory*, vol. 146, 1253-1274.

[5] Chen, Y-C., T. Kunimoto, Y. Sun, and S. Xiong, (2022), "Maskin Meets Abreu and Matsushima," *Theoretical Economics*, vol. 17, 1683-1717..

[6] Chen, Y-C., T. Kunimoto, Y. Sun, and S. Xiong, (2021), "Rationalizable Implementation in Finite Mechanisms," *Games and Economic Behavior*, vol. 129, 181-197.

[7] Chatterji, S., T. Kunimoto, and P. Ramos, (2022), "Compellingness in Nash Implementation," Working Paper, Singapore Management University.

[8] Chung, K-S and J. Ely, (2003), "Implementation with Near-Complete Information," *Econometrica*, vol. 71, 857-87.

[9] Dutta, B. and A. Sen, (1991), "A Necessary and Sufficient Condition for Two-Person Nash Implementation," *Review of Economic Studies*, vol. 58, 121-128.

[10] Goltsman, M., (2011), "Nash Implementation using Simple Mechanisms without Undesirable Mixed-Strategy Equilibria," *Working Paper*.

[11] Jackson, M-O., (1991), "Bayesian Implementation," *Econometrica*, vol. 59, 461-477.

[12] Jackson, M-O., (1992), "Implementation in Undominated Strategies: A Look at Bounded Mechanisms," *Review of Economic Studies*, vol. 59, (1992), 757-775.

[13] Jackson, M-O, (2001), "A Crash Course in Implementation Theory," *Social Choice and Welfare*, vol. 39, 655-708.

[14] Jackson, M-O, T-R Palfrey, and S. Srivastava, (1994), "Undominated Nash Implementation in Bounded Mechanisms," *Games and Economic Behavior*, vol. 6, 474-501.

[15] Jain, R., (2021), "Rationalizable Implementation of Social Choice Correspondences," *Games and Economic Behavior*, vol. 127, 47-66.

[16] Kartik, N. and O. Tercieux, (2012), "A Note on Mixed-Nash Implementation," *Working Paper*.

[17] Korplela, V., (2016), "Nash Implementation with Finite Mechanisms," *Working Paper*.

[18] Kunimoto, T., (2019), "Mixed Bayesian Implementation in General Environments," *Journal of Mathematical Economics*, vol. 82, 247-263.

[19] Kunimoto, T. and R. Serrano, (2019), "Rationalizable Implementation of Correspondences," *Mathematics of Operations Research*, vol. 44, 1326-1344.

[20] Lee, J. and H. Sabourian, (2015), "Complexity and Repeated Implementation," *Journal of Economic Theory*, vol. 158, 259-292.

[21] Maskin, E., (1999), "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, vol. 66, 23-38.

[22] Maskin, E. and T. Sjöström, (2002), "Implementation Theory," in *Handbook of Social Choice and Welfare*, vol 1, edited by K.J. Arrow, A.K. Sen, and K. Suzumura, (2002), 237-288.

[23] McKelvey, R.D., (1989), "Game Forms for Nash Implementation of General Social Choice Correspondences," *Social Choice and Welfare*, vol. 6, 139-156.

[24] Mezzetti, C. and L. Renou, (2012a), "Implementation in Mixed Nash Equilibrium," *Journal of Economic Theory*, vol. 147, 2357-2375.

[25] Mezzetti, C. and L. Renou, (2012b), "Mixed Nash Implementation with Finite Mechanisms," *Working Paper*.

[26] Moore, J. and R. Repullo, (1988), "Subgame Perfect Implementation," *Econometrica*, vol. 56, 1191-1220.

[27] Moore, J. and R. Repullo, (1990), "Nash Implementation: A Full Characterization," *Econometrica*, vol. 58, 1083-1099.

[28] Saijo, T., (1988), "Strategy Space Reduction in Maskin's Theorem," *Econometrica*, vol. 56, 693-700.

[29] Serrano, R., (2004), "The Theory of Implementation of Social Choice Rules," *SIAM Review*, vol. 46, 377-414.

[30] Serrano, R. and R. Vohra, (2010), "Multiplicity of Mixed Equilibria in Mechanisms: A Unified Approach to Exact and Approximate Implementation," *Journal of Mathematical Economics*, vol. 46, 775-785.

[31] Xiong, S., (2023), "Rationalizable Implementation of Social Choice Functions: Complete Characterization," *Theoretical Economics*, vol. 18, 197-230.