Virtual implementation by bounded mechanisms: Complete information*

Ritesh Jain[†] and Michele Lombardi[‡]

April 30, 2019

Abstract

A social choice rule (SCR) F maps preference profiles to lotteries over some finite set of outcomes. F is virtually implementable in (pure and mixed) Nash equilibria provided that for all $\epsilon > 0$, there exists a mechanism such that for each preference profile θ, its set of Nash equilibrium outcomes at θ is ϵ–closed to the socially desirable set $F(\theta)$. Under a domain restriction, we obtain the following result: When there are at least three agents, any F is virtually implementable in Nash equilibrium, as well as in rationalizable strategies, by a bounded mechanism. No "tail-chasing" construction, common in the constructive proofs of the literature, is used to assure that undesired strategy combinations do not form a Nash equilibrium.

Keywords: Virtual implementation, pure and mixed Nash equilibria, rationalizability, social choice rules

JEL classification: C79, D82

*We would like to thank Arunava Sen, Debasis Mishra, Hannu Vartiainen, Kirby Neilsen, Olivier Bochet, Paul Healy, Yaron Azrieli, Ville Korpela seminar participants at the Ohio State University, Indian Statistical Institute and National Taiwan University. Ritesh Jain greatly acknowledges financial support from the Ministry of Science and Technology Grant, Taiwan under the grant: 107-2410-H-001-113-MY2. All remaining errors are ours.

[†]Institute of Economics, Academia Sinica. e-mail: ritesh@econ.sinica.edu.tw.

[‡]Adam Smith Business School, University of Glasgow. e-mail: michele.lombardi@glasgow.ac.uk.

1 Introduction

In a seminal paper, Abreu and Matsushima (1992a) (henceforth, AM) show that any single-valued SCR, or social choice function (SCF) is virtually implementable in Nash equilibrium (pure and mixed), as well as rationalizable strategies. The devised mechanism, known as the Abreu-Matsushima mechanism (henceforth, AM-mechanism), has several appealing properties. Firstly, its construction does not rely on any tail-chasing procedures to eliminate unwanted equilibria, such as integer or modulo games —indeed, it is a finite mechanism. The finiteness of the mechanism allows one to interpret its strategy spaces in natural terms. Secondly, the devised mechanism is robust to strategic uncertainty as it induces any socially optimal outcome as the unique rationalizable outcome[1]. Last but not least, it ensures continuous virtual implementation of the SCF, in the sense of Oury and Tercieux (2012).

When there are at least three agents, this paper devises a mechanism that virtually implements any SCR in Nash equilibrium strategies. The devised mechanism is an extension of AM-mechanism satisfying all the features described above. Specifically, the devised mechanism is a bounded mechanism. A mechanism is bounded if, for each agent i, whenever a strategy m_i' is not rationalizable, there exists a rationalizable strategy m_i which is weakly better than m_i' when others are playing a rationalizable strategy.

In contrast to AM, we are interested in implementing SCRs. For this reason, our construction may possess multiple Nash equilibria for certain preference profiles. This multiplicity of equilibria creates strategic uncertainty, and so agents may fail to coordinate on a specific equilibrium. An attractive feature of our bounded mechanism is that, in addition to implementation in Nash equilibria, it also achieves implementation in rationalizable strategies. Thus, it is robust to strategic uncertainty. This double implementation is surprising because implementation in Nash equilibrium strategies does not imply implementation in rationalizable strategies, and vice versa. Moreover, this result allows us to establish that our extension of AM-mechanism ensures continuous virtual implementation of F when its range is finite.

There are several reasons why focussing only on SCFs can be considered unsatisfactory (see, for instance, Thomson (1996)). Firstly, multi-valued SCRs typically represent many social decisions. Prominent examples include the Pareto, the Walrasian, the Condorcet, and the no-envy correspondences. Secondly and foremost, since F represents the social objectives that the society or its representative want to achieve, its full implementation is the correct objective of the society. It would be unacceptable to partially implement F by implementing a SCF f which systematically picks, for each preference profile θ, a socially optimal outcome

[1]In the literature on exact implementation robustness to strategic uncertainty has recently been studied in Bergemann et al. (2011), Kunimoto and Serrano (2018) and Jain (2019). The constructive proofs used in this literature rely on the integer game construction.

$f(\theta) \in F(\theta)$. The reason is that the implementation of this subselection $f \in F$ may violate some of the normative properties that led the society or its representatives to choose F. For instance, it most certainly will be unacceptable when F embodies some minimal concerns about fairness.

For the sake of concreteness, let us consider the classical divide-and-choose mechanism, which induces envy-free divisions of a vector of resources over which two agents have equal rights. The rules of the mechanism are as follows. One agent, called the divider, divides each resource into two pieces; the other agent, called the chooser, chooses one share for each type of resources. The divider obtains the remaining shares. Since this procedure always induces agents to select envy-free allocations of resources, it partially implements the no-envy correspondence—this correspondence selects, for each preference profile, the set of allocations at which no agent would like to exchange his or her consumption bundle with that of another agent. However, the no-envy allocation selected by the mechanism is always the one that the divider prefers; the chooser would benefit from acting as the divider. Thus such a partial implementation violates the fundamental fairness objective embodied in the no-envy correspondence, namely, that of equal treatment of agents (Crawford (1977)).

Another argument made in favor of a partial implementation of F is based on the interpretation that the mechanism designer views the outcomes in $F(\theta)$ as equally good (Abreu and Sen (1991); Mezzetti and Renou (2012)). Under this interpretation of planner's indifference, a simple way to implement a multi-valued SCR consists of virtually implementing a SCF which puts an equal probability on each outcome in $F(\theta)$, for each profile θ. Under this view, it is easy to see that we can use the AM-mechanism. A shortcoming of this interpretation is that in some situations we do not know whether the mechanism designer is indifferent between socially optimal outcomes or not.

The classical interpretation of implementation of F requires that each outcome in $F(\theta)$ must be supported by a distinct equilibrium (Maskin (1999); Abreu and Sen (1991); Bochet and Maniquet (2010)) and does not assume planner's indifference. This paper shows that this classical interpretation is not restrictive when the objective is to implement F virtually. This is in sharp contrast to the case of "exact" implementation where implementation in the classical sense is more restrictive than implementation under the assumption of planner's indifference (Mezzetti and Renou (2012)).

The AM-mechanism has been the focus of attention in several strands of implementation theory, such as in robust mechanism design (Bergemann and Morris (2009), Müller (2016)), in implementation in level-k (Serrano et al. (2018)), and in implementation with verification (Matsushima (2018)). Indeed, these papers provide constructive proofs that, directly or indirectly, rely on the AM-mechanism. We believe that our construction may play an important role in these strands of the literature when the objective of society is represented by a multivalued SCR F.[2]

We divide the remainder of this paper into four sections. Section 2 sets out the theoretical framework and outlines the basic model. Section 3 provides the characterization result as well as a discussion of the implementing mechanism. Section 4 provides the formal proof. Section 5 discusses continuous implementation. Section 6 concludes.

2 Model

Preliminaries

Let $N = \{1, 2, \ldots, n\}$ be a finite set of agents with $n \geq 3$. The finite set of (pure) *outcomes* is denoted by X. The set of all lotteries over X is denoted by Y. Agent i's utility function is indexed by a parameter θ_i. We refer to θ_i as agent i's type. The set of possible types for agent i is assumed to be finite, and it is denoted by Θ_i. Agent i's preferences over lotteries is described by a utility function $u_i : Y \times \Theta_i \to \mathbb{R}$, where $u_i\left(y, \theta_i\right)$ is agent i's utility of the lottery y when is of type θ_i. We assume that each agent i is an expected utility maximizer. A state or type profile is described by an n-tuple of types $\theta \in \prod_{i \in N} \Theta_i = \Theta$. For any type profile $\theta \in \Theta$, θ_{-i} denotes the $n-1$-tuple $(\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_n)$. The set of admissible type profiles is denoted by $\overline{\Theta}$, where $\overline{\Theta} = \prod_{i \in N} \overline{\Theta}_i$ is a nonempty subset of Θ. We assume that the true type profile is common knowledge among the agents. However, the planner does not know it. This is the case of complete information.

Following AM, we introduce a domain restriction that will play a key role.

Assumption. (AM). For every $i \in N$ and $\theta \in \overline{\Theta}$, there exist $\overline{a}(i, \theta) \in Y$ and $\underline{a}(i, \theta) \in Y$ such that

$$u_i(\overline{a}(i, \theta), \theta_i) > u_i(\underline{a}(i, \theta), \theta_i)$$

$$u_j(\underline{a}(i, \theta), \theta_j) \geq u_j(\overline{a}(i, \theta), \theta_j)$$

for all $j \in N \setminus \{i\}$.

Assumption requires that for each type profile θ and each agent i, there exist two lotteries that are strictly ranked by agent i, and for which every other agent has the (weakly) opposite

[2]More recently, there is an effort to reconcile the classical exact implementation literature, which relies on mechanisms with questionable features, with the AM-mechanism, which emphasizes on virtual implementation but relies on well-behaved mechanisms. See Chen et al. (2018b) for such an exercise.

ranking. This assumption is more likely to be satisfied in environments with transferable private goods that are positively valued by agents. In the implementation literature, this kind of assumptions is often made in studies relating to well-behaved implementing mechanisms (Jackson et al. (1994); Kartik et al. (2014)). In this regards, Kunimoto and Serrano (2011) even argue that Assumption is indispensable.

The goal of the designer is to implement a SCR F, which is a correspondence $F : \overline{\Theta} \twoheadrightarrow Y$ such that for each type profile $\theta \in \overline{\Theta}$, $F(\theta)$ is a nonempty subset of Y. The common interpretation is that F represents the social objectives that the agents or its representatives want to achieve. A SCF $f : \overline{\Theta} \to Y$ is a single-valued SCR such that $f(\theta) \in Y$ for all $\theta \in \overline{\Theta}$.

The implementation problem arises from the fact that the planner's goal depends on the true type profile and he does not know it. To elicit it, the planner designs a (stochastic) mechanism. A mechanism is a game form $\Gamma = (M_1, ...M_n, g)$, where M_i is agent i's space of pure strategies, and $g : M \to Y$ is the outcome function, where $\underset{i \in N}{\Pi} M_i = M$. We denote a pure strategy of agent i by $m_i \in M_i$ and a profile of pure strategies is denoted by $m = (m_1, ..., m_n) \in M$. A mechanism Γ with a type profile θ specifies the game (Γ, θ).

We assume that the planner wants to achieve his or her goal in (pure and mixed) Nash equilibrium. A mixed strategy for agent i is a probability distribution over M_i. Let $\Delta(M_i)$ denote the set of all probability distributions over M_i, where σ_i denotes a typical mixed strategy for agent i. We denote by σ a typical mixed strategy profile, where $\sigma = (\sigma_1, ..., \sigma_n) \in \Delta(M) = \underset{i \in N}{\Pi} \Delta(M_i)$. A strategy profile σ is a Nash equilibrium of (Γ, θ) if for all $i \in N$ and all $\sigma'_i \in \Delta(M_i)$,

$$u_i(g(\sigma_i, \sigma_{-i}), \theta_i) \geq u_i(g(\sigma'_i, \sigma_{-i}), \theta_i)$$

We denote by $NE(\Gamma, \theta)$ the set of Nash equilibria of the game (Γ, θ).

Bounded mechanisms

Jackson (1992) argues that some notion of boundedness is required to rule out "tail-chasing" constructions, which are common in the constructive proofs of the literature on Nash implementation. To introduce our notion of bounded mechanisms, we need to introduce additional notation. Given a game (Γ, θ), let $R_i(\Gamma, \theta)$ denote the set of rationalizable strategies for agent i and $R(\Gamma, \theta) = \underset{i \in N}{\Pi} R_i(\Gamma, \theta)$ denote the set of rationalizable strategy profiles. $R_i(\Gamma, \theta)$ is described as an outcome of an iterative procedure. Each round in this iterative process eliminates strategies that are never best responses. Below we provide a formal definition.

Fix any agent $i \in N$. Let $R_i^0(\Gamma, \theta) = M_i$ and $R_{-i}^0(\Gamma, \theta) = \underset{j \neq i}{\times} M_j$. For any $k \geq 1$, define

$R_i^k(\Gamma,\theta)$ as follows:

$$R_i^k(\Gamma,\theta) = \left\{ m_i \in R_i^{k-1}(\Gamma,\theta) \;\middle|\; \begin{array}{c} \text{There exists } \lambda \in \Delta(R_{-i}^{k-1}(\Gamma,\theta)) \text{ such that} \\ m_i \in \underset{m_i' \in M_i}{argmax} \sum\limits_{m_{-i} \in R_{-i}^{k-1}(\Gamma,\theta)} \lambda_i(m_{-i}) u_i(g(m_i', m_{-i}),\theta) \end{array} \right\}.$$

The set of rationalizable strategies for player i is the limit of this iterative process, i.e., $R_i(\Gamma,\theta) = \bigcap_{k\geq 1} R_i^k(\Gamma,\theta)$. In the game (Γ,θ), the set of strategy profiles $B^\Gamma(\theta) \subseteq M$ satisfies the best response property at θ if for every $i \in N$ and for every $m_i \in B_i^\Gamma(\theta)$, there exists a belief $\lambda_i(m_i) \in \Delta(B_{-i}(\theta))$ such that m_i is a best response to $\lambda_i(m_i)$. One can see that $B^\Gamma(\theta) \subseteq R(\Gamma,\theta)$, if $B^\Gamma(\theta)$ satisfies the best response property at θ.

Definition 1. A mechanism Γ is bounded relative to $\overline{\Theta}$ if for all $\theta \in \overline{\Theta}$, all $i \in N$ and all $m_i' \notin R_i(\Gamma,\theta)$, there exists an $m_i \in R_i(\Gamma,\theta)$ such that for all $m_{-i} \in R_{-i}(\Gamma,\theta)$, it holds that

$$u_i(g(m_i, m_{-i}),\theta_i) \geq u_i(g(m_i', m_{-i}),\theta_i)$$

with strict inequality for some $m_{-i} \in R_{-i}(\Gamma,\theta)$.

In words, for each strategy m_i' of agent i that is not rationalizable, Definition 1 requires that this agent has a rationalizable strategy that yields her or him at least as high a payoff as does m_i' whatever the other agents' rationalizable strategies and a higher payoff than does m_i' for some rationalizable strategy profiles of the other agents. Note that this definition implies that agents who eliminated strategies that are not rationalizable have well-defined choices. Also, note that mechanisms with finite message spaces are bounded mechanisms according to our definition.

Virtual implementation

A SCR F is virtually implementable if there exists a "nearby" nonempty correspondence H that is "exactly" implementable in a solution concept. Formally, let $d(x,y)$ be the Euclidean distance between any a pair of lotteries. The SCR $F : \overline{\Theta} \twoheadrightarrow Y$ is ϵ-close to a nonempty correspondence $H : \overline{\Theta} \twoheadrightarrow Y$ if for every $\theta \in \overline{\Theta}$, there exists a bijection $\rho : F(\theta) \mapsto H(\theta)$ such that $d(x,\rho(x)) \leq \epsilon$ for all $x \in F(\theta)$.

A mechanism $\Gamma = (M,g)$ implements the correspondence H in Nash equilibrium (*resp.*, in rationalizable strategies), if for each $\theta \in \overline{\Theta}$, it holds that: (1) for each $x \in H(\theta)$, there exists $m \in NE(\Gamma,\theta)$ (*resp.*, $m \in R(\Gamma,\theta)$) such that $g(m) = x$; and (2) for each $\sigma \in NE(\Gamma,\theta)$ (*resp.*, $\sigma \in R(\Gamma,\theta)$), $\bigcup_{m\in\, supp(\sigma)} g(m) \subseteq H(\theta)$. If such a mechanism exists, we say that H is

Nash implementable (*resp.*, implementable in rationalizable strategies). Thus:

Definition 2. The SCR $F : \overline{\Theta} \twoheadrightarrow Y$ is virtually implementable in Nash equilibrium (*resp.*, implementable in rationalizable strategies) by a bounded mechanism, if for all $\epsilon > 0$ there exists a nonempty correspondence $H : \overline{\Theta} \twoheadrightarrow Y$ which is Nash implementable (*resp.*, implementable in rationalizable strategies) by a bounded mechanism as well as ϵ-close to F.

The following lemma, due to and Matsushima (1992; p. 999), requires the existence of a set of lotteries for agent i such that each of his or her type has a distinct maximal element within the set.

Lemma 1. (Abreu and Matsushima (1992a)). Let Assumption hold. Let $i \in N$. Then, there exists a function $f_i : \overline{\Theta}_i \mapsto Y$such that for each $\theta_i \in \overline{\Theta}_i$, it holds that

$$u_i(f_i(\theta_i), \theta_i) > u_i(f_i(\theta'_i), \theta_i)$$

for each $\theta'_i \in \overline{\Theta}_i \setminus \{\theta_i\}$.

3 The characterization result

The characterization result is stated below. Its proof can be found in section 4. The devised mechanism is an extension to SCRs of the AM-mechanism, which virtually implements any SCF in iteratively undominated strategies when there are at least three agents and Assumption is satisfied.

Theorem 1. Suppose $n \geq 3$. Let Assumption hold. Any SCR $F : \overline{\Theta} \twoheadrightarrow Y$ is virtually implementable in Nash equilibria, as well as in rationalizable strategies, by a bounded mechanism.

Proof. See section 4. □

We provide below an intuitive discussion of the basic arguments of the proof. In the devised mechanism, each agent makes $(K + 2)$ simultaneous announcements. A typical announcement is indexed by $k \in \{-1, 0, 1, \ldots, K\}$, where K is an integer which is yet to be specified.

Each agent i reports a SCF in the $k = -1$ announcement, his or her type in the $k = 1$ announcement, and an entire type profile in each of the remaining announcements. That is, agent i's message space is:

$$M_i = \mathcal{F}_F \times \overline{\Theta}_i \times \overline{\Theta} \times ... \times \overline{\Theta} = M_i^{-1} \times M_i^0 \times M_i^1 \times ... \times M_i^K$$

where $\mathcal{F}_F = \{f : \overline{\Theta} \mapsto Y | \text{for every } \theta \in \overline{\Theta},\ f(\theta) \in F(\theta)\}$ is a collection of SCFs, each of which assign, to each type profile θ, an element $f(\theta)$ of $F(\theta)$. Notice that the set M_i is finite when $\mathcal{F}_F$ is finite.

The devised mechanism is an augmentation of the AM-mechanism with a voting scheme over elements of $\mathcal{F}_F$, which happens in stage $k = -1$. The voting scheme can be described as follows. Suppose that designer has designated $f^* \in \mathcal{F}_F$ as the default SCF to be implemented. Agents can change f^* into $f \in \mathcal{F}_F$ if there is an 'almost agreement' on this change, that is, at least $n - 1$ agents agree on f. The selected SCF is used to determine the outcome of the *decision rule* of the mechanism in each stage $k \geq 1$. The augmentation is made without losing the attractive incentive properties of the AM-mechanism.

Though the constructed mechanism is a simultaneous mechanism, it can be useful to think of it as a sequential mechanism with $K + 2$ stages, where agents make simultaneous announcements in each stage.

Suppose that the default SCF f^* is to be virtually implemented by an arbitrarily small $\epsilon > 0$. According to the AM-mechanism, a lottery from Y is selected. This lottery is a probability distribution over the following components:

- (*Dictator rule*) With probability $\frac{\epsilon}{n}$, agent i is selected as a *dictator*. Based on his or her announcement at the stage $k = 0$, his or her best outcome from a predetermined set of outcomes is selected.

- (*Audit rule*) With probability $\frac{\epsilon^2}{n}$, agent i is *audited* for consistency. To conduct this audit, the designer considers all announcements made by agents from stage $k = 1$ to stage K, and compares them with m^0—that is, with the message profile reported by agents at stage $k = 0$. Agent i is punished by selecting $\underline{a}(i, m^0)$ if he or she is the first one to announce a type profile different from m^0. Otherwise, he or she is rewarded by selecting $\overline{a}(i, m^0)$.

- (*Decision rule*) With probability $\frac{(1-\epsilon-\epsilon^2)}{K}$, at each stage $k \geq 1$, the outcome is determined as follows:

  - If all agents make exactly the same announcement, then $f^*(m^k)$ is the outcome of the mechanism.
  - If all but agent i make exactly the same announcement θ', then $f_i^*(\theta')$ is the outcome of the mechanism at this stage.
  - In all other cases, an arbitrary outcome y is selected by the mechanism.

An important feature of the AM-mechanism is that if every agent reports his or her true type θ_i in his or her $k = 0$ announcement, and everyone reports the true state θ in each stage

$k \geq 1$, then $f^*(\theta)$ is implemented with probability $(1 - \epsilon - \epsilon^2)$, where $\epsilon > 0$ is an arbitrarily small parameter chosen by the designer. Another important feature is that truthful reporting is the uniquely rationalizable strategy for each agent i. This feature is due to the following two main insights.

First, for each agent i, truthful report of his or her type θ_i in stage $k = 0$ is a *strictly dominant strategy.* The possibility for each agent to be nominated as a dictator is a key towards an understanding of this insight.

To see this, suppose that agent i plays any strategy $\hat{m}_i$ such that $\hat{m}_i^0 = \hat{\theta}_i \neq \theta_i$. By changing $\hat{m}_i$ into m_i such that $m_i^0 = \theta_i$ and $m_i^k = \hat{m}_i^k$ for each $k \geq 1$, agent i has a utility gain of $u_i(f_i(\theta_i), \theta_i) - u_i(f_i(\hat{\theta}_i), \theta_i) > 0$, by Lemma 1, when he or she is chosen as the dictator. To provide agent i with the incentives to truthfully report in stage $k = 0$, this utility gain must be greater than the maximal utility gain from lying. The gain from lying is coming only from the auditing component of the mechanism. The incentives are assured by choosing ϵ appropriately. To see why lying can be profitable, let us consider the case where everyone else is truthful in all stages. In this case, lying by agent i induces punishments on other agents in the auditing component, which may be beneficial to him or her.

The second insight is that the audit component of the mechanism, as well as the appropriate choice of K, provides agents with the incentives to be truthful in each stage $k \geq 1$.

To see this, recall that by the above discussion, everyone is truthful in stage $k = 0$. Fix $k = 1$ and any agent i. Suppose that agent i plays a strategy $\hat{m}_i$ such that $\hat{m}_i^1 \neq \theta = m^0$ and that every other agent j plays m_j.

Let us suppose that agent i is not the only agent who makes a $k = 1$ announcement which is inconsistent with $m^0 = \theta$. By changing $\hat{m}_i$ into m_i such that $m_i^1 = \theta$ and $m_i^k = \hat{m}_i^k$ for each $k \neq 1$, agent i has a utility gain of $u_i(\overline{a}(i, (\theta_i, \theta_{-i})), \theta_i) - u_i(\underline{a}(i, (\theta_i, \theta_{-i})), \theta_i) > 0$ when he or she is audited—by the domain assumption. When some other agent is audited, truthtelling by agent i does not affect of the outcome of the mechanism. However, truthful report by agent i may cause him or her a utility loss in the decision component of the mechanism when stage $k = 1$ is selected by the designer. This can happen with probability $\frac{1-\epsilon-\epsilon^2}{K}$. The designer can make this loss arbitrarily small by choosing K appropriately.

Let us suppose that agent i is the only agent who makes a $k = 1$ announcement which is inconsistent with $m^0 = \theta$. By changing $\hat{m}_i$ into m_i such that $m_i^1 = \theta$ and $m_i^k = \hat{m}_i^k$ for each $k \neq 1$, agent i does not suffer any utility loss when is audited as $u_i(\overline{a}(i, (\theta_i, \theta_{-i})), \theta_i) - u_i(\underline{a}(i, (\theta_i, \theta_{-i})), \theta_i) \geq 0$. When some other agent is audited, truthtelling by agent i can not harm him or her. The reason is that by truthtelling agent i can only harm other agents in the auditing phase—agent i may only have a utility gain by the domain assumption. When stage $k = 1$ is selected by the designer, given that all agents but agent i are making the same $k = 1$

announcement, agent i's utility is $u_i(f_i^*(\theta), \theta_i)$, which, by the domain assumption, is strictly lower than the utility he or she obtains under truthtelling, that is, $u_i(f^*(\theta), \theta_i) > u_i(f_i^*(\theta), \theta_i)$.

Since our goal is to implement F, by implementing $f^* \in \mathcal{F}_F$ we have only achieved our goal partially. To implement F, as mentioned earlier, we augment the AM-mechanism with a voting rule over $\mathcal{F}_F$, which happens in stage $k = -1$.

Recall that in our augmented mechanism, agents can coordinate on any $f \in \mathcal{F}_F$ by reaching an almost unanimous consensus on f. If they fail to do so, then f^* is implemented. Note that the 'elected' SCF is used when stage $k \geq 1$ is chosen to determine the outcome of the decision component of the mechanism.

An attractive feature of the voting game is that any unanimous agreement on a SCF forms a Nash equilibrium. This feature allows us to create multiple Nash equilibrium in the augmented mechanism. Indeed, we show that the strategy profile in which every agent i plays $m_i = (f, \theta_i, \theta, \ldots, \theta)$ forms a Nash equilibrium. Moreover, we also show that agent i's rationalizable strategies are of the form $m_i = (\cdot, \theta_i, \theta, \ldots, \theta)$. Thus, even though agents fail to coordinate on one Nash equilibrium (that is, one SCF), or even though they are playing some mixed equilibrium, the realized outcome will be ϵ-close to an $f \in \mathcal{F}_F$.

It is worth emphasizing that our mechanism does not rely on any tail chasing construction. The reason is that the constructed mechanism is bounded with respect to rationalizability. We show that the iterative deletion of strictly never best reply does not rely on the cardinality of $\mathcal{F}_F$—see Lemma 4.

4 Proof of Theorem 1

Suppose that $n \geq 3$ and that Assumption holds. Thus, Lemma 1 holds. Let us define $\Gamma = (M, g)$ as follows.

$$M = \times_{i \in N} M_i,$$
$$M_i = M_i^{-1} \times M_i^0 \times M_i^1 \times ... \times M_i^K,$$

where the integer K is yet to be specified, where

$$M_i^{-1} = \mathcal{F}_F, \quad M_i^0 = \overline{\Theta}_i, \quad M_i^k = \overline{\Theta} \text{ for all } k \in \{1, ..., K\},$$

and where $\mathcal{F}_F = \left\{ f : \overline{\Theta} \mapsto Y | \forall \theta \in \overline{\Theta}, \, f(\theta) \in F(\theta) \right\}$.

Since Lemma (1) holds and since $\overline{\Theta}_i$ is finite, it follows that there exists a real number

$\eta > 0$ such that for each agent $i \in N$ and each of his or her type $\theta_i \in \overline{\Theta}$, it holds that

$$u_i\left(f_i\left(\theta_i\right), \theta_i\right) - u_i\left(f_i\left(\theta_i'\right), \theta_i\right) > \eta \text{ for all } \theta_i' \in \overline{\Theta}_i \setminus \{\theta_i\}.$$

Define a function $\xi : N \times M \mapsto Y$ by:

$$\xi\left(i, m\right) = \begin{cases} \underline{a}\left(i, m^0\right) & \begin{array}{c}\text{if for some } k \in \{1, ..., K\},\ m_j^h = m^0 \text{ for all} \\ h = 1, .., k-1 \text{ and all } j \in N, \text{ and } m_i^k \neq m^0;\end{array} \\ \bar{a}\left(i, m^0\right) & \text{otherwise;} \end{cases}$$

where the lotteries $\underline{a}\left(i, m^0\right)$ and $\bar{a}\left(i, m^0\right)$ are those specified by Assumption.

Given Assumption, for every $f \in \mathcal{F}_F$, there exists a nearby SCF $\hat{f} : \overline{\Theta} \mapsto Y$ and, for each $i \in N$, a nonempty single-valued funtion $\hat{f}_i : \overline{\Theta} \mapsto Y$ such that

$$u_i\left(\hat{f}\left(\theta\right), \theta_i\right) > u_i\left(\hat{f}_i\left(\theta\right), \theta_i\right) \tag{1}$$

for all $\theta \in \overline{\Theta}$. To see it, for a small number $\alpha > 0$, let us define $\hat{f}$ and $\hat{f}_i$ as follows:

$$\hat{f}\left(\theta\right) = \left(1 - n\alpha\right) f\left(\theta\right) + \alpha \textstyle\sum_{j \in N} \bar{a}\left(j, \theta\right)$$

and

$$\hat{f}_i\left(\theta\right) = \left(1 - n\alpha\right) f\left(\theta\right) + \alpha \textstyle\sum_{j \in N} \bar{a}\left(j, \theta\right) + \alpha \underline{a}\left(i, \theta\right).$$

By Assumption and by definition of $\hat{f}$ and $\hat{f}_i$, one can see that (1) holds.

For every $k \in \{1, ..., K\}$, define a function $\rho : M^k \times M^{-1} \mapsto Y$ as follows:

Rule 1: If $\left|\left\{i \in N | m_i^{-1} = f\right\}\right| \geq n - 1$ for some $f \in \mathcal{F}_F$, then:

(a) If $m_i^k = \theta$ for all $i \in N$, then $\rho\left(m^k, m^{-1}\right) = f\left(\theta\right)$.

(b) For all $i \in N$, if $m_j^k = \theta$ for all $j \in N \setminus \{i\}$ and $m_i^k \neq \theta$, then $\rho\left(m^k, m^{-1}\right) = f_i\left(\theta\right)$.

(c) Otherwise, $\rho\left(m^k, m^{-1}\right) = y$ for some $y \in Y$.

Rule 2: Otherwise, for some $f^* \in \mathcal{F}_F$:

(a) If $m_i^k = \theta$ for all $i \in N$, then $\rho\left(m^k, m^{-1}\right) = f^*\left(\theta\right)$.

(b) For all $i \in N$, if $m_j^k = \theta$ for all $j \in N \setminus \{i\}$ and $m_i^k \neq \theta$, then $\rho\left(m^k, m^{-1}\right) = f_i^*\left(\theta\right)$.

(c) Otherwise, $\rho\left(m^k, m^{-1}\right) = y$ for some $y \in Y$.

Following AM, let $\epsilon > 0$ be an arbitrary small number such that $(1-\epsilon-\epsilon^2) > 0$. The outcome function $g : M \mapsto Y$ is defined by

$$g(m) = \frac{\epsilon}{n}\sum_{i\in N} f_i\left(m_i^0\right) + \frac{\epsilon^2}{n}\sum_{i\in N}\xi(i,m) + \frac{(1-\epsilon-\epsilon^2)}{K}\sum_{k=1}^{K}\rho\left(m^k, m^{-1}\right), \tag{2}$$

for all $m \in M$.

For each $\theta_i \in \overline{\Theta}_i$, let

$$E_i(\theta_i) = \max_{m\in M}\left(\sum_{j\in N}\left|u_i(\xi(j,m),\theta_i)\right|\right).$$

In what follows, we will choose a small enough $\epsilon > 0$ such that

$$\eta > 2\epsilon E_i(\theta_i) \tag{3}$$

for all $i \in N$ and all $\theta_i \in \overline{\Theta}_i$.

For each $i \in N$ and each $\theta \in \overline{\Theta}$, define

$$B_i(\theta) = u_i(\bar{a}(i,\theta),\theta_i) - u_i(\underline{a}(i,\theta),\theta_i)$$

and

$$D_i(\theta) = \max_{\left(m^k,m^{-1}\right)\in M^k\times M^{-1}}\left[u_i\left(\rho\left(m^k,m^{-1}\right),\theta_i\right) - u_i\left(\rho\left(\left(m_{-i}^k,\bar{m}_i^k\right),m^{-1}\right),\theta_i\right)\right],$$

where $\bar{m}_i^k = \theta$.

By Assumption, it holds that $B_i(\theta) > 0$ for all $i \in N$ and all $\theta \in \overline{\Theta}$. Therefore, there exists an integer $K > 0$ such that for all $i \in N$ and all $\theta \in \overline{\Theta}$, it holds that

$$K\frac{\epsilon^2}{n}B_i(\theta) > \left(1-\epsilon-\epsilon^2\right)D_i(\theta). \tag{4}$$

Fix any SCR F. We show that F is virtually Nash implementable and that Γ is a bounded mechanism.

To this end, we prove the following lemmata for any $\theta \in \overline{\Theta}$.

Lemma 2. For all $f \in \mathcal{F}_F$ and all $m \in M$, if $m_i = (f,\theta_i,\theta,...,\theta)$ for all $i \in N$, then $m \in NE(\Gamma,\theta)$.

Proof. Take any $m \in M$ and suppose that for some $f \in \mathcal{F}_F$, $m_i = (f,\theta_i,\theta,...,\theta)$ for all $i \in N$. We will show that the strategy profile m forms a pure strategy Nash equilibrium. Let $m = (m_i, m_{-i})$ and $m' = (m'_i, m_{-i})$. Notice that by construction in each stage $k \in \{1,\ldots,K\}$

Rule 1-a will be used and hence the SCF f will be used under both m and m'. In other words any unilateral deviation from m does not change the choice of social choice function in any stage k.

$$\begin{aligned}
u_i\left(g\left(m\right),\theta_i\right)-u_i\left(g\left(m'\right),\theta_i\right) &= \frac{\epsilon}{n}\left[u_i\left(f_i\left(\theta_i\right),\theta_i\right)-u_i\left(f_i\left(m_i'^1\right),\theta_i\right)\right] \\
&\quad+\frac{\epsilon^2}{n}\textstyle\sum_{i\in N}\left[u_i\left(\xi\left(i,m\right),\theta_i\right)-u_i\left(\xi\left(i,m'\right),\theta_i\right)\right] \\
&\quad+\frac{\left(1-\epsilon-\epsilon^2\right)}{K}\textstyle\sum_{k\in K}\left[u_i\left(f\left(\theta\right),\theta_i\right)-u_i\left(f\left(i,\theta\right),\theta_i\right)\right] \\
&\geq \eta-2\epsilon E_i\left(\theta_i\right)+\frac{\left(1-\epsilon-\epsilon^2\right)}{K}\textstyle\sum_{k\in K}\left[u_i\left(f\left(\theta\right),\theta_i\right)-u_i\left(f\left(i,\theta\right),\theta_i\right)\right] \\
&\geq \frac{\left(1-\epsilon-\epsilon^2\right)}{K}\textstyle\sum_{k\in K}\left[u_i\left(f\left(\theta\right),\theta_i\right)-u_i\left(f\left(i,\theta\right),\theta_i\right)\right]\text{ (by 3)} \\
&\geq 0\text{ (by 1)}
\end{aligned}$$

Thus we have shown that there is no profitable unilateral deviation from the strategy profile m for agent i. Since the choice of agent, i is arbitrary, there is no unilateral profitable deviation for any of the agent involved, and so $m\in NE\left(\Gamma,\theta\right)$, as we claimed to prove. □

Lemma 3. For all $m\in M$, $m\in R\left(\Gamma,\theta\right)$ if and only if for all $i\in N$, $m_i^0=\theta_i$ and $m_i^k=\theta$ for all $k=1,...,K$.

Proof. The proof of this statement is based on the proof of Abreu and Matsushima (1992a). We report it for the sake of completeness. Suppose that $m\in R\left(\Gamma,\theta\right)$. Let us first show that for all $i\in N$, $m_i^0=\theta_i$. Assume, to the contrary, that $m_i^0\neq\theta_i$ for some $i\in N$. Let $\bar{m}_i\in M_i$ be such that $\bar{m}_i^0=\theta_i\neq m_i^0$, $\bar{m}_i^{-1}=m_i^{-1}$ and $\bar{m}_i^k=m_i^k$ for each $k=1,...,K$. Let us show that m_i is strictly dominated by $\bar{m}_i$ - that is, $u_i\left(g\left(\bar{m}_i,m_{-i}\right),\theta_i\right)>u_i\left(g\left(m_i,m_{-i}\right),\theta_i\right)$ for all $m_{-i}\in M_{-i}$. Fix any $m_{-i}\in M_{-i}$. To save space, let $\left(\bar{m}_i,m_{-i}\right)=\bar{m}$ and $\left(m_i,m_{-i}\right)=m$. Note that, by construction, m and $\bar{m}$ fall into the same rule. By definition of g given in (2), Lemma 1 and the fact that $\bar{m}_i^{-1}=m_i^{-1}$, it follows that

$$\begin{aligned}
u_i\left(g\left(\bar{m}\right),\theta_i\right)-u_i\left(g\left(m\right),\theta_i\right) &= \frac{\epsilon}{n}\left[u_i\left(f_i\left(\bar{m}_i\right),\theta_i\right)-u_i\left(f_i\left(m_i\right),\theta_i\right)\right] \\
&\quad+\frac{\epsilon^2}{n}\textstyle\sum_{i\in N}\left[u_i\left(\xi\left(i,\bar{m}\right),\theta_i\right)-u_i\left(\xi\left(i,m\right),\theta_i\right)\right] \\
&> \eta-2\epsilon E_i\left(\theta_i\right) \\
&> 0\text{ (by 3)},
\end{aligned}$$

Since the choice of $m_{-i}\in M_{-i}$ is arbitrary, it follows that m_i is strictly dominated by

$\bar{m}_i$, which contradicts the initial supposition that $m_i \in R_i(\Gamma, \theta)$.

Let $P(h)$ be the statement "if $m \in R^h(\Gamma, \theta)$, then for all $i \in N$,

$$m_i^0 = \theta_i \text{ and } m_i^\ell = \theta \text{ for each } \ell = 1, ..., h\text{".}$$

We have already shown that $P(0)$ holds, that is, $m_i^0 = \theta_i$ for each $i \in N$. Assume that the statement $P(h-1)$ holds for each $i \in N$, where $0 \leq h < K$. Then, for each $i \in N$,

$$m_i^0 = \theta_i \text{ and } m_i^\ell = \theta \text{ for each } \ell = 1, ..., h-1\text{".}$$

We show that $P(h)$ holds, that is, $m_i^h = \theta$ for each $i \in N$. Assume, to the contrary, that $m_i^h \neq \theta$ for some $i \in N$. Let $\bar{m}_i \in M_i$ be such that $\bar{m}_i^q = m_i^q$ for each $q = 0, ..., h-1$, $\bar{m}_i^{-1} = m_i^{-1}$ and $\bar{m}_i^h = \theta \neq m_i^h$.

Take any $m_{-i} \in R_{-i}^h(\Gamma, \theta)$. To save space, let $(\bar{m}_i, m_{-i}) = \bar{m}$ and $(m_i, m_{-i}) = m$. Note that, by construction, m and $\bar{m}$ fall into the same rule. We proceed according to the following two cases.

Case 1: There exists an agent $j \neq i$ such that $m_j^h \neq m^0$

By definition of g given in (2), Assumption and by the fact that $\bar{m}_i^{-1} = m_i^{-1}$, it follows that

$$\begin{aligned}
u_i(g(\bar{m}), \theta_i) - u_i(g(m), \theta_i) &= \frac{\epsilon^2}{n}\left[u_i\left(\bar{a}\left(i, \bar{m}^0\right), \theta_i\right) - u_i\left(\underline{a}\left(i, m^0\right), \theta_i\right)\right] \\
&\quad + \frac{(1-\epsilon-\epsilon^2)}{K}\left[u_i\left(\rho\left(\bar{m}^h, m^{-1}\right), \theta_i\right) - u_i\left(\rho\left(m^h, m^{-1}\right), \theta_i\right)\right] \\
&= \frac{\epsilon^2}{n}B_i(\theta) - \frac{(1-\epsilon-\epsilon^2)}{K}\left[u_i\left(\rho\left(m^h, m^{-1}\right), \theta_i\right) - u_i\left(\rho\left(\bar{m}^h, m^{-1}\right), \theta_i\right)\right] \\
&\geq \frac{\epsilon^2}{n}B_i(\theta) - \frac{(1-\epsilon-\epsilon^2)}{K}D_i(\theta) \\
&> 0 \text{ (by 4).}
\end{aligned}$$

Since the choice of $m_{-i} \in R_{-i}^h(\Gamma, \theta)$ is arbitrary, this contradicts our initial supposition that $m_i \in R_i(\Gamma, \theta)$.

Case 2: For all $j \neq i$, $m_j^\ell = m^0$

We distinguish whether $\xi(i, \bar{m}) = \xi(i, m)$ or not. Suppose that $\xi(i, \bar{m}) = \xi(i, m)$. It simplifies the argument, and causes no loss of generality, to assume that $m_j^{-1} = f$ for all

$j \neq i$. Then, m and $\bar{m}$ fall into Rule 1. Then, by definition of g given in (2), it follows that

$$\begin{aligned} u_i\left(g\left(\bar{m}\right),\theta_i\right)-u_i\left(g\left(m\right),\theta_i\right) &= \frac{\left(1-\epsilon-\epsilon^2\right)}{K}\left[u_i\left(\rho\left(\bar{m}^h,m^{-1}\right),\theta_i\right)-u_i\left(\rho\left(m^h,m^{-1}\right),\theta_i\right)\right] \\ &= \frac{\left(1-\epsilon-\epsilon^2\right)}{K}\left[u_i\left(f\left(\theta\right),\theta_i\right)-u_i\left(f\left(i,\theta\right),\theta_i\right)\right] \\ &> 0 \text{ (by 1)}. \end{aligned}$$

Otherwise, suppose that $\xi\left(i,\bar{m}\right) \neq \xi\left(i,m\right)$. Then, by applying the same reasoning used in Case 1, one can see that $u_i\left(g\left(\bar{m}\right),\theta_i\right)-u_i\left(g\left(m\right),\theta_i\right) > 0$. In either case, since the choice of $m_{-i} \in R^h_{-i}\left(\Gamma,\theta\right)$ is arbitrary, this contradicts our initial supposition that $m_i \in R_i\left(\Gamma,\theta\right)$.

By the principle of mathematical induction, it follows that if $m \in R\left(\Gamma,\theta\right) = R^K\left(\Gamma,\theta\right)$, then for all $i \in N$, $m_i^0 = \theta_i$ and $m_i^\ell = \theta$ for each $\ell = 1,...,K$, as we aimed to achieve.

Finally, suppose that for all $i \in N$, m_i is such that $m_i^0 = \theta_i$ and $m_i^k = \theta$ for all $k = 1,...,K$. We show that $m \in R\left(\Gamma,\theta\right)$. To this end, we need to show that $m_i \in R_i(\Gamma,\theta)$ for each i. Fix any i. Suppose that $m_i^{-1} = f$ for some f in M_i^{-1}.

For each $j \in N\backslash\{i\}$, let $m'_j = (f,\theta_j,\theta,...,\theta) \in M_j$. Lemma 2 implies that $\left(m_i,m'_{-i}\right) \in NE\left(\Gamma,\theta\right)$. It follows that $m_i \in R_i\left(\Gamma,\theta\right)$. Since the choice of agent i is arbitrary, it follows that $m \in R\left(\Gamma,\theta\right)$, as we wanted to prove. □

From the above lemmata, it follows that agent $i \in N$'s set of rationalizable strategies of (Γ,θ) is given by

$$R_i\left(\Gamma,\theta\right) = \left\{m_i \in M_i | m_i^0 = \theta_i \text{ and } m_i^k = \theta \text{ for all } k = 1,...,K\right\}. \tag{5}$$

We now show that Γ satisfies Definition 1.

Lemma 4. Γ is a bounded mechanism.

Proof. Let us show that Γ satisfies Definition 1. Suppose that $m_i \notin R_i(\Gamma,\theta)$ for some $i \in N$. We show that $m'_i = (m_i^{-1},\theta_i,\theta,\theta,\ldots,\theta) \in R_i(\Gamma,\theta)$ dominates m_i for every $m_{-i} \in R_{-i}(\Gamma,\theta)$, that is,

$$u_i(g(m'_i,m_{-i}),\theta_i) \geq u_i(g(m_i,m_{-i}),\theta_i)$$

with strict inequality for some $m_{-i} \in R_{-i}(\Gamma,\theta)$.

Since $m_i \notin R_i(\Gamma,\theta)$, it follow from (5) that $m_i^0 \neq \theta_i$ or $m_i^h \neq \theta$ for some $h \in \{1,\ldots,K\}$. We proceed according to whether $m_i^0 \neq \theta_i$ or not. We only prove the case $m_i^0 \neq \theta_i$ given the other case can be proved similarly.

Let $n_i^0 \in M_i$ be such that $n_i^{0,-1} = m_i^{-1}$, $n_i^{0,0} = \theta_i \neq m_i^{-1}$ and $n_i^{0,k} = m_i^k$ for all $k \in \{1, ..., K\}$. By the same arguments of the proof of Lemma 3, it follows that

$$u_i\left(g\left(n_i^0, m_{-i}\right), \theta_i\right) > u_i\left(g\left(m_i, m_{-i}\right), \theta_i\right)$$

for all $m_{-i} \in R_{-i}(\Gamma, \theta)$. If $n_i^0 \in R_i(\Gamma, \theta)$, then Definition 1 is satisfied. Otherwise, let h^1 be the lowest integer in $\{1, ..., K\}$ such that $n_i^{0,0} = \theta_i$, $n_i^{0,k} = \theta$ for all $1 \leq k < h^1$, and $n_i^{0,h^1} \neq \theta$. Note that h^1 exists by (5).

Let $n_i^1 \in M_i$ be such that $n_i^{1,h^1} = \theta$ and $n_i^{1,k} = n_i^{0,k}$ for all $k \in \{-1, 0, ..., K\}$ such that $k \neq h^1$. By the same arguments of the proof of Lemma 3, it follows that

$$u_i\left(g\left(n_i^1, m_{-i}\right), \theta_i\right) > u_i\left(g\left(n_i^0, m_{-i}\right), \theta_i\right)$$

for all $m_{-i} \in R_{-i}(\Gamma, \theta)$. If $n_i^1 \in R_i(\Gamma, \theta)$, then Definition 1 is satisfied. Otherwise, let h^2 be the lowest integer in $\{h^1 + 1, ..., K\}$ such that $n_i^{1,0} = \theta_i$, $n_i^{1,k} = n^{0,k} = \theta$ for all $1 \leq k < h^2$, and $n_i^{1,h^2} \neq \theta$. Note that h^2 exists by (5). By repeating the same reasoning, we have that there exists $n_i^2 \in M_i$ such that

$$u_i\left(g\left(n_i^2, m_{-i}\right), \theta_i\right) > u_i\left(g\left(n_i^1, m_{-i}\right), \theta_i\right)$$

for all $m_{-i} \in R_{-i}(\Gamma, \theta)$. And so on. After a finite number $J \leq K$ of iterations, a sequence of messages $n_i^0, ..., n_i^J \in M_i$ can be derived, where n_i^J is such that $n_i^{J,0} = \theta_i$, $n_i^{J,k} = \theta$ for all $k \in \{1, ..., K\}$ and $n_i^{J,-1} = m_i^{-1}$, such that

$$u_i\left(g\left(n_i^J, m_{-i}\right), \theta_i\right) > ... > u_i\left(g\left(n_i^0, m_{-i}\right), \theta_i\right) > u_i\left(g\left(m_i, m_{-i}\right), \theta_i\right)$$

for all $m_{-i} \in R_{-i}(\Gamma, \theta)$. It follows that

$$u_i\left(g\left(n_i^J, m_{-i}\right), \theta_i\right) > u_i\left(g\left(m_i, m_{-i}\right), \theta_i\right)$$

for all $m_{-i} \in R_{-i}(\Gamma, \theta)$, and so Definition 1 is satisfied.

$\square$

To complete the proof, let us show that Γ virtually implements F in Nash equilibria and rationalizable strategies. To this end, we first need to define a correspondence $H : \overline{\Theta} \twoheadrightarrow Y$ that is ϵ-close to F.

For each $x \in F(\theta)$, define $\gamma(x)$ as follows:

$$\gamma(x) = \frac{\epsilon}{n}\sum_{i\in N} f_i(\theta_i) + \frac{\epsilon^2}{n}\sum_{i\in N}\overline{a}(i,\theta) + \frac{1-\epsilon-\epsilon^2}{K}x$$

Note that, by definition, it holds that $d(x, \gamma(x)) \leq \epsilon$. Moreover, let us define the correspondence $H(\theta)$ by $H(\theta) = \{\gamma(x) | x \in F(\theta)\}$. One can easily check that γ is a bijection from $F(\theta)$ to $H(\theta)$. It follows that H is ϵ-close to F. We need the following useful result.

Lemma 5. *For all $m \in R(\Gamma, \theta)$, $g(m) = \gamma(x)$ for some $\gamma(x) \in H(\theta)$.*

Proof. Take any $m \in R(\Gamma, \theta)$. Then, $m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$, by Lemma 3. This means that in each stage $k \in \{1, \ldots, K\}$ under m either Rule 1-a or Rule 2-a applies . Suppose Rule 1-a applies in each stage $k \in \{1, \ldots, K\}$, that SCF f is selected and $F(\theta) = x$ then $g(m) = \gamma(x)$ where $\gamma(x) \in H(\theta)$ since $x \in F(\theta)$. Suppose Rule 2-a applies in each stage $k \in \{1, \ldots, K\}$, that SCF f^* is selected and $f^*(t) = x^*$ then $g(m) = \gamma(x^*)$ where $\gamma(x^*) \in H(\theta)$ since $x^* \in F(\theta)$. □

Next, we show that Γ implement H in Nash equilibria and rationalizable strategies. In particular, we show that (1) for each $\gamma(x) \in H(\theta)$, there exists $m \in NE(\Gamma, \theta)$ (*resp.*, $m \in R(\Gamma, \theta)$) such that $g(m) = \gamma(x)$; and (2) for each $\sigma \in NE(\Gamma, \theta)$ (*resp.*, $\sigma \in R(\Gamma, \theta)$), if $m' \in supp(\sigma)$, then $g(m') = \gamma(x)$ for some $\gamma(x) \in H(\theta)$.

To show part (1), let us consider the strategy profile m such that $m_i = (f, \theta_i, \theta, \ldots, \theta)$ for each i such that $F(\theta) = x$. Note that $g(m) = \gamma(x)$, by definition of the mechanism. Since Lemma 2 holds, $m \in NE(\Gamma, \theta)$, and so $m \in R(\Gamma, \theta)$. Since the choice of $\gamma(x)$ is arbitrary, it follows that part (1) is satisfied.

To show part (2), by Lemma 5 we know that for every $m' \in R(\Gamma, \theta)$, it holds that $g(m') = \gamma(x)$ for some $\gamma(x) \in H(\theta)$. Then, H is implementable in rationalizable strategies.

Let us show that part (2) holds for implementation in Nash equilibria. Take any $\sigma \in NE(\Gamma, \theta)$ and any $m' \in supp(\sigma)$. By definition of $R(\Gamma, \theta)$, it follows that $supp(\sigma) \subseteq R(\Gamma, \theta)$, and so $m' \in R(\Gamma, \theta)$. Lemma 5 implies that $g(m') = \gamma(x)$ for some $\gamma(x) \in H(\theta)$. Since the choice of $\sigma \in NE(\Gamma, \theta)$ and $m' \in supp(\sigma)$ are arbitrary, it follows that part (2) is satisfied. This completes the proof.

5 Continuous implementation

In this section, we show that the constructed mechanism continuously virtually implements $\mathcal{F}$ in the spirit of Oury and Tercieux (2012). To this end, we introduce the model of Oury and Tercieux (2012).

Let $\prod_{i \in N} T_i$ be a countable type space with generic element $t = (t_1, ..., t_n) \in T$. A model $\mathcal{T}$ is a pair (T, κ) where for each $i \in N$ and $t_i \in T_i$, $\kappa(t_i)$ is the associated beliefs of agent i of type t_i on the state and the respective types of agents other than i. Formally, for each $i \in N$ and $t_i \in T_i$, $\kappa(t_i) \in \Delta(\Theta \times T_{-i})$. Given any measurable subset $A \subset \Delta(\Theta \times T_{-i})$, we associate a probability measure $\kappa[A]$. We assume that $\kappa(\cdot)$ is continuous. Given a mechanism Γ and a model $\mathcal{T}$, let $\mathcal{U}(\Gamma, \mathcal{T})$ be the induced game of incomplete information. For each $i \in N$, a (mixed) strategy for agent i is denoted $\sigma_i : T_i \to \Delta(M_i)$. The probability that σ_i assigns to message m_i when agent i is of type t_i is denoted $\sigma_i(m_i|t_i)$. Let σ denote a strategy profile. Given some belief $\pi_i \in \Delta(\Theta \times M_{-i})$ for agent i, his set of best responses is denoted by

$$BR_i(\pi_i) = \underset{m_i \in M_i}{argmax} \sum_{(\theta, m_{-i}) \in (\Theta \times M_{-i})} \pi_i(\theta, m_{-i}) u_i(g(m), \theta)$$

A strategy profile σ is a Bayes Nash equilibrium (BNE, henceforth) of $\mathcal{U}(\Gamma, \mathcal{T})$ if for each $i \in N$ and each $t_i \in T_i$,

$$m_i \in supp(\sigma_i(t_i)) \Rightarrow m_i \in BR_i(\pi_i(\cdot|t_i, \sigma_{-i}), \tag{6}$$

where $\pi(\cdot|t_i, m_{-i}) \in \Delta(\Theta \times M_{-i})$ denotes the joint distribution of states and messages given agent i's type and the strategy of agents other than i.

Because we are interested in types which are close to those in the original model, we introduce the notion of nested models. Let $\overline{\mathcal{T}}$ be our initial model, which is a model of complete information, i.e., $\overline{T_i} = \{t_i^\theta : \theta \in \Theta\}$ and $\overline{\kappa}_{t_i^\theta}[(\theta, t_{-1}^\theta)]$ for each $\theta \in \Theta$. For any $\hat{\mathcal{T}} = (\hat{T}, \hat{\kappa})$, we write $\hat{\mathcal{T}} \supset \overline{\mathcal{T}}$ if $\overline{T} \subset \hat{T}$ and for each $i \in N$ and $t_i \in \hat{T}_i$, and any measurable $E \subset \Theta \times \overline{T}_{-i}$, $\kappa(t_i)[E] = \hat{\kappa}(t_i)[(\Theta \times \hat{T}_{-i}) \cap E]$. Given a strategy profile σ in $\mathcal{U}(\Gamma, \hat{\mathcal{T}})$, we write $\sigma_{\overline{T}}$ for the strategy profile restricted to $\overline{T}$.

Given a type t_i in a model (T, κ), we can compute the hierarchies of beliefs for agent i. First, his first-order belief $h_i^1(t_i)$ (i.e., his belief about Θ) is given by the marginal distribution of the beliefs κ on Θ, $h_i^1(t_i) = marg_\Theta\, \kappa(t_i)$. Likewise, we can compute agent i's second order beliefs at t_i (i.e., his belief about the state and about the first-order beliefs) by setting

$$h_i^2(t_i)[E] = \kappa(t_i)[\{(\theta, t_{-i}) : (\theta, (h_j(t_j))_{j \in N}) \in E\}],\ \forall E \subset \Theta \times (\Delta(\Theta))^n$$

For each agent i and each type t_i, we can compute the entire hierarchy of beliefs $(h_i^1(t_i), ..., h_i^\ell(t_i), ...)$ by proceeding in this way. Note that $h_i^1(t_i) \in \Delta(\Theta)$, $h_i^2(t_i) \in \Delta(\Theta \times \Delta(\Theta)^n)$, and so on.

Following Chen et al. (2018a), let $Z^0 = \Theta$ and for each $k \geq 1 : Z^k = z^{k-1} \times \Delta[(Z^{k-1}]^n$. Note that for each $i, t_i \in T_i$, and each $k \geq 1$, $h_i^k(t_i) \in \Delta(Z^{k-1})$. Let d^0 be the discrete metric on Θ and d^1 be the Prokhorov distance on the first-order beliefs ($\Delta(\Theta)$).[3]

For any integer $k \geq 2$, let us endow the set $\Delta(Z^{k-1})$ with the Prokhorov distance d^k, where Z^{k-1} is endowed with the sup-metric induced by $d^0, d^1, ..., d^{k-1}$. The set of all beliefs hierarchies for which it is common knowledge that the beliefs are coherent is the universal type space $T^* = \underset{i \in N_i}{T}^*$ introduced by Mertens and Zamir (1985), where for each $i \in N$, $T_i^* \subset \times_{k=0}^{\infty} \Delta X^k$ is the set of agent i's hierarchies of beliefs in this space. Note that $h_i(t_i) \in T_i^*$ if there exists some type t_i' in some model such that t_i and t_i' have the same hierarchies of beliefs. Each T_i^* is endowed with the product topology. For a given $i \in N$, we say that two types t_i and $\tilde{t}_i$ are close if there exists a sufficiently large ℓ such that for all $k \leq \ell$ the kth order beliefs $h^k(t_i)$ and $h^k(\tilde{t}_i)$ are close in the topology of convergence of measures. In other words, pick a sequence of types $\{t_i[n]\}_{n=0}^{\infty}$. We say that it converges to a type t_i if for each ℓ, $h_i^{\ell}(t_i[n]) \to h_i^{\ell}(t_i)$, i.e.,

$$d_i^P(t_i[n], t_i) \equiv \sum_{k=1}^{\infty} 2^{-k} d_i^k(h_i^k(t_i[n]), h_i^k(t_i)) \to \ 0$$

In the sequel, we simply write $t_i[n] \to_P \ t_i$ for convergence of the sequence $\{t_i[n]\}_{n=0}^{\infty}$ to t_i, and $t[n] \to_P \ t$ if $t_i[n] \to_P \ t_i$ for each $i \in N$.

In this section, we study the following notion of continuous implementation.

Definition 3. A SCR F is continuously implementable by a mechanism Γ if for every model $\mathcal{T}$ such that $\overline{\mathcal{T}} \subseteq \mathcal{T}$, there exists a Bayes nash equilibrium σ of $\mathcal{U}(\Gamma, \mathcal{T})$ such that:

1. $\sigma_{\overline{\mathcal{T}}}$ is a Nash equilibrium in $\mathcal{U}(\Gamma, \overline{\mathcal{T}})$
2. for any $\bar{t} \in \overline{T}$ and any sequence $t[n] \mapsto_p \bar{t}$, where $t[n]$ is a sequence in T, there exists $\overline{n}$ such that for all $n \geq \overline{n}$, it holds that $\bigcup_{m \in supp(\sigma(t[n])} g \circ m \subseteq F(\bar{t})$.

The next result shows that F is continuously implementable when it is implementable in rationalizable strategies by a finite mechanism.

Lemma 6. *If F is implementable in rationalizable strategies by a finite mechanism, then F is continuously implementable by a finite mechanism.*

[3]The Prokhorov distance is a metric on the collection of probability measures on a given metric space. Formally, for any two $z, z' \in \Delta(Z)$ for some metric space (Z, ρ), the Prokhorov distance is given by

$$inf\{\gamma > 0 : z'(E) \leq z(E^{\gamma}) + \gamma \,\text{for every borel set}\, E \subseteq \Delta(Z)\},$$

where $E^{\gamma} = \{x \in Z : \inf_{y \in E} \rho(x, y) < \gamma\}$. Alternatively, $E^{\gamma} = \underset{e \in E}{\cup} B_{\gamma}(e)$ where $B_{\gamma}(e)$ is the open ball of radius γ and centered at $e \in E$.

Proof. Suppose F is implementable in rationalizable strategies by a finite mechanism Γ. Let $\overline{T} \subset T$. Since Γ is a finite mechanism and T is countable, there exists a BNE σ of $\mathcal{U}(\Gamma, \mathcal{T})$.

First, we show that $\sigma_{\overline{T}}$ is a BNE of $\mathcal{U}(\Gamma, \overline{\mathcal{T}})$. To see this, note that $\forall \bar{t}_i \in \overline{T}, \kappa(\bar{t}_i)$ takes its support on $\Theta \times \overline{T}_{-i}$. Therefore, for all $\bar{t}_i \in \overline{T}$, it follows that:

$$m_i \in supp(\sigma_i(\bar{t}_i)) \Rightarrow m_i \in BR_i(\pi_i(\cdot|\bar{t}_i, \sigma_{-i}|\Gamma), \tag{7}$$

and so $\sigma_{\overline{T}}$ is a BNE in $\mathcal{U}(\Gamma, \overline{\mathcal{T}})$. This establishes part (1) of Definition 3.

As far as part (2) is concerned, take any sequence $t[n] \in T$ such that $t[n] \mapsto_p \bar{t}$. By the definition of BNE and by the definition of interim correlated rationalizability, provided in Dekel et al. (2007), we know that for all $t[n] \in T$,

$$supp(\sigma(t[n])) \subseteq R(t[n]|\Gamma, \mathcal{T}) \tag{8}$$

where $R(t[n]|\Gamma, \mathcal{T})$ denotes the set of interim correlated rationalizable strategies of type $t[n]$ in the game $\mathcal{U}(\Gamma, \mathcal{T})$. By Lemma 1 of Dekel et al. (2006), we know that for there exists $\overline{n}$ such that for all $n \geq \overline{n}$, it holds that

$$R(t[n]|\mathcal{M}, \mathcal{T}) \subseteq R(\bar{t}|\Gamma, \overline{\mathcal{T}}). \tag{9}$$

By (8) and (9), it follows that there exists $\overline{n}$ such that for all $n \geq \overline{n}$,

$$supp(\sigma(t[n])) \subseteq R(\bar{t}|\Gamma, \overline{\mathcal{T}}). \tag{10}$$

Fix any $n \geq \overline{n}$ and any $m \in supp(\sigma(t[n]))$. By (10), it follows that $m \in R(\bar{t}|\Gamma, \overline{\mathcal{T}})$. Since Γ implements F in rationalizable strategies, $g(m) \in F(\bar{t})$, as we sought. □

In a recent paper, Weinstein and Yildiz (2017) extend the finite game result of Dekel et al. (2006) to infinite games. By using this result, the above lemma can be extended to mechanisms with compact strategy spaces.

The above lemma implies that F is virtually continuously implementable when it is virtually implementable in rationalizable strategies by a finite mechanism.

Lemma 7. *If F is virtually implementable in rationalizable strategies by a finite mechanism, then it is virtually continuously implementable by a finite mechanism.*

From Theorem 1 and Lemma 7, one can easily see that our augmentation of the AM-mechanism virtually continuously implements any SCR F when $\mathcal{F}_F$ is finite.

Theorem 2. *Let $n \geq 3$. Any F is virtually continuously implementable when $\mathcal{F}_F$ is finite.*

Using the arguments of Weinstein and Yildiz (2017), the above theorem can be extended to the case when $F(\theta)$ is compact for every $\theta \in \Theta$.

Though Theorem 2 important, it is not entirely satisfactory. The reason is that our notion of continuous implementation does not require that every equilibrium in the true model is robust to the higher order uncertainty. To achieve implementation according to this stronger notion, one has to ensure that the perturbed model has multiple BNEs—at least one for each outcome in the range of F. We leave this challenging task for future research.

6 Concluding remarks

When there are at least three agents, this paper devises a bounded mechanism that virtually implements any SCR in Nash and rationalizable strategies. The devised mechanism is robust to strategic uncertainty and does not rely on any tail-chasing procedures to eliminate unwanted equilibria. Finally, its strategy spaces are easy to interpret in natural terms. The devised mechanism can be viewed as a natural extension of the AM-mechanism as it embeds the AM-mechanism in a voting scheme, in which an almost unanimous decision on $f \in \mathcal{F}_F$ must be taken by agents in order to virtually implement f; when there is a larger disagreement, a default SCF is virtually implemented.

The main result we obtain should be considered as providing a theoretical benchmark for SCRs. Indeed, we obtain our result under a stringent informational assumption of complete information (among agents). This assumption may not be satisfied in certain situations. Abreu and Matsushima (1992b) have generalized the mechanism constructed in Abreu and Matsushima (1992b) to Bayesian environments. They show that any SCF f which can be virtually Nash implemented in these environments must satisfy a measurability condition, a.k.a. AM-measurability. To characterize the class of social choice sets which are virtually implementable in both Bayesian Nash and interim correlated rationalizable strategies, we believe that a construction like the one presented in this paper would be useful. We believe that such a construction will hinge on the identification of the appropriate variant of the AM-measurability condition. We leave this subject for future research.

References

Abreu, D. and Matsushima, H. (1992a). Virtual implementation in iteratively undominated strategies: complete information. *Econometrica: Journal of the Econometric Society*, pages 993–1008.

Abreu, D. and Matsushima, H. (1992b). Virtual implementation in iteratively undominated strategies II: Incomplete information. Princeton University working paper.

Abreu, D. and Sen, A. (1991). Virtual implementation in nash equilibrium. *Econometrica: Journal of the Econometric Society*, pages 997–1021.

Bergemann, D. and Morris, S. (2009). Robust virtual implementation. *Theoretical Economics*, 4(1):45–88.

Bergemann, D., Morris, S., and Tercieux, O. (2011). Rationalizable implementation. *Journal of Economic Theory*, 146(3):1253–1274.

Bochet, O. and Maniquet, F. (2010). Virtual nash implementation with admissible support. *Journal of mathematical economics*, 46(1):99–108.

Chen, Y.-C., Frank, M. M., and Pai, M. M. (2018a). Truthful continuous implementation. *Working Paper*.

Chen, Y. C., Kunimoto, T., Sun, Y., and Xiong, S. (2018b). Maskin meets abreu and matsushima. *Working Paper*.

Crawford, V. P. (1977). A game of fair division. *The Review of Economic Studies*, 44(2):235–247.

Dekel, E., Fudenberg, D., and Morris, S. (2006). Topologies on types. *Theoretical Economics*, 1:275–309.

Dekel, E., Fudenberg, D., and Morris, S. (2007). Interim correlated rationalizability. *Theoretical Economics*, 2:15–40.

Jackson, M. O. (1992). Implementation in undominated strategies: A look at bounded mechanisms. *The Review of Economic Studies*, 59(4):757–775.

Jackson, M. O., Palfrey, T. R., and Srivastava, S. (1994). Undominated nash implementation in bounded mechanisms. *Games and Economic Behavior*, 6(3):474–501.

Jain, R. (2019). Rationalizable implementation of social choice correspondences.

Kartik, N., Tercieux, O., and Holden, R. (2014). Simple mechanisms and preferences for honesty. *Games and Economic Behavior*, 83:284–290.

Kunimoto, T. and Serrano, R. (2011). A new necessary condition for implementation in iteratively undominated strategies. *Journal of Economic Theory*, 146(6):2583–2595.

Kunimoto, T. and Serrano, R. (2018). Rationalizable implementation of correspondences. *Mathematics of Operations Research, forthcoming*.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *The Review of Economic Studies*, 66(1):23–38.

Matsushima, H. (2018). Implementation without expected utility: Ex-post verifiability. Technical report, Center for Advanced Research in Finance, Faculty of Economics,.

Mezzetti, C. and Renou, L. (2012). Implementation in mixed nash equilibrium. *Journal of Economic Theory*, 147(6):2357–2375.

Müller, C. (2016). Robust virtual implementation under common strong belief in rationality. *Journal of Economic Theory*, 162:407–450.

Oury, M. and Tercieux, O. (2012). Continuous implementation. *Econometrica*, 80(4):1605–1637.

Serrano, R., de Clippel, G., and Saran, R. (2018). Level-k Mechanism Design. *The Review of Economic Studies*, 86(3):1207–1227.

Thomson, W. (1996). Concepts of implementation. *The Japanese Economic Review*, 47(2):133–143.

Weinstein, J. and Yildiz, M. (2017). Interim correlated rationalizability in infinite games. *Journal of Mathematical Economics*, 72:82–87.