

CONTINUOUS IMPLEMENTATION

BY MARION OURY AND OLIVIER TERCIEUX¹

In this paper, we introduce a notion of continuous implementation and characterize when a social choice function is continuously implementable. More specifically, we say that a social choice function is continuously (partially) implementable if it is (partially) implementable for types in the model under study and it continues to be (partially) implementable for types “close” to this initial model. Our results show that this notion is tightly connected to full implementation in rationalizable strategies.

KEYWORDS: High order beliefs, robust implementation.

1. INTRODUCTION

THE NOTION OF PARTIAL—as opposed to full—implementation consists in designing games under which *some* equilibrium—but not necessarily *all*—yields the outcome desired by the social planner. Despite the fact that undesirable equilibria may potentially exist, partial implementation is widely used in both theoretical and applied works. The main contribution of this paper is to give a precise sense in which, even if the social planner is only willing to partially implement the social choice function, a lack of full implementation may make the predictions based on partial implementation very unrobust to slight perturbations of the information structure. Taking into account the doubts a social planner may have about his model, we characterize when a social choice function can be partially *continuously* implemented.²

¹We thank the editor and four anonymous referees for detailed comments that led to numerous improvements in the paper. We are also thankful to D. Bergemann, P. Jehiel, A. Kajii, T. Kunimoto, E. Maskin, N. Vieille, and seminar participants at Paris School of Economics, the EEA meeting in Budapest, Princeton University, the Institute for Advanced Study in Princeton, Rutgers University, McGill University, the World Congress of Game Theory at Northwestern, the Workshop on New Topics in Mechanism Design in Madrid, GRIPS Workshop on Global Games in Tokyo, the PSE-Northwestern joint workshop, and the Recent Developments in Mechanism Design conference held in Princeton. We are especially grateful to S. Morris for helpful remarks and discussions at various stages of this research. Joint authorship of O. Tercieux with D. Bergemann and S. Morris has greatly improved this version; in particular, the notion of interim rationalizable monotonicity was developed in discussions between D. Bergemann, S. Morris, and O. Tercieux. Finally, O. Tercieux thanks the Institute for Advanced Study at Princeton for financial support through the Deutsche Bank membership. He is also grateful to the Institute for its hospitality.

²Bergemann and Morris (2005) is one of the first attempts to relax the implicit common knowledge assumptions made in the mechanism design literature. In their setting, the modeller, when choosing a mechanism, has no information on the real situation that will finally prevail among agents. Consequently, their notion of robust implementation follows a “global approach”: a social choice function is robustly (partially) implementable if it is (partially) implementable *on all possible models*. In contrast, we assume that the planner has some specific model in mind and is quite confident about it. As a consequence, our requirement is local: the social choice function must be implemented only at types close to types in the initial model.

We consider a situation where the social planner (i) is willing to *partially* implement in strict Nash equilibrium, and (ii) wants to implement in a continuous manner. More specifically, we require that, in any model that embeds the initial model, there exists an equilibrium that (i) is a strict equilibrium on the initial model, and (ii) yields the desired outcome, not only at all types of the initial model but also at all types “close” to initial types. Of course, the notion of closeness in types is critical. To formalize this, we use the method introduced by Harsanyi (1967) and developed in Mertens and Zamir (1985). Each type in the initial model is mapped into a hierarchy of beliefs. Then, following the interim approach due to Weinstein and Yildiz (2007), we define a notion of “nearby” types. This notion, formally described by the product topology in the universal type space, captures the restrictions on the modeler’s ability to observe the players’ (high order) beliefs.

Our main results state that this continuity requirement leads to necessary conditions that are tightly linked to full implementation in rationalizable strategies. More specifically, we first show that a social choice function can be partially continuously implemented only if it is monotonic. Monotonicity was first introduced by Maskin (1999) in a complete information setting, and was then generalized to the incomplete information setting by Postlewaite and Schmeidler (1986) and further studied in Jackson (1991). While partial implementation is often a quite weak requirement, monotonicity is known to have a lot of bite.³ It is necessary for full implementation in Nash equilibrium and it is sufficient under weak additional conditions when there are more than three players. Hence, our result builds a first bridge between partial and full implementation. We then generalize this result and show that interim rationalizable monotonicity⁴—a generalization of Bayesian monotonicity—is also a necessary condition for partial continuous implementation. Since, with more than three players, some weak domain restrictions are enough for any interim rationalizable monotonic social choice function to be fully implementable in rationalizable strategies, we get a connection between partial continuous implementation and full implementation in rationalizable strategies.

Finally, we discuss how far our conditions are from sufficiency. More precisely, we drop the assumption that the planner is willing to partially implement in strict equilibrium (but keep the standard requirement of pureness) and slightly relax the assumption that messages are cheap-talk (sending a message may involve an arbitrarily small cost). We provide a full characterization

³For instance, under complete information, with more than three players, any social choice function can be partially implemented, while on the universal domain of preferences, only constant social choice functions are monotonic (see Saijo (1987)).

⁴This notion is an interim formulation of a notion introduced by Bergemann and Morris (2011) in the ex post setting. It was developed in discussions between D. Bergemann, S. Morris, and O. Tercieux and is the relevant necessary condition for interim rationalizable implementation. This can be shown using the types of arguments in Bergemann and Morris (2011) and in Bergemann, Morris, and Tercieux (2011).

of partial continuous implementation: a social choice function is continuously implementable by a finite mechanism if and only if it is fully implementable in rationalizable messages.

Since the seminal paper by Rubinstein (1989) on the e-mail game, several approaches⁵ have been followed to analyze the connection between high order beliefs and strategic behavior. These works share the common assumption that, in the perturbed models, some types may have preferences over action profiles that are radically different from those of types in the initial model.⁶ Indeed, the behavior of these specific types is used as a starting point for contagion processes that drive results in these analyses. Note that the meaning of such an assumption in the mechanism design context (where the social planner fixes the game form) is problematic. However, in the present paper, we show that the logic of implementation makes this assumption unnecessary.⁷ Indeed, in mechanism design, several different states of nature are *ex ante* possible for the social planner. Our argument in the proof uses this multiplicity and shows that this setting is then rich enough: partial implementation in the initial model is used as an (endogenous) starting point for the contagion process at equilibrium.

2. SETTING

We consider a finite set of players $\mathcal{I} = \{1, \dots, I\}$. Each agent i has a bounded utility function $u_i: A \times \Theta \rightarrow \mathbb{R}$, where Θ is the countable set of states of nature and A is the set of outcomes. A mechanism specifies a message set M_i for each agent i and a mapping g from message profiles to outcomes. More precisely, we write M as an abbreviation for $\prod_{i \in \mathcal{I}} M_i$ and, for each player i , M_{-i} for $\prod_{j \neq i} M_j$.⁸ A mechanism \mathcal{M} is a pair (M, g) where M is countable and the outcome function $g: M \rightarrow A$ assigns to each message profile m an alternative $g(m) \in A$. In general, if X is a topological space, we treat it as a measurable space with its Borel sigma field, and the space of Borel probability measures on X is denoted $\Delta(X)$. Spaces $\Delta(X)$ are endowed with the topology of weak convergence of measures. Unless otherwise specified, if X is countable, it is treated as a topological space endowed with the discrete topology. The only countable spaces that we may sometimes treat differently are type spaces; this point is detailed further in the text. The space of outcomes A is not necessarily countable and is assumed to be a Hausdorff topological space. By a slight

⁵The notion of *robustness* due to Kajii and Morris (1997), the *global games* argument due to Carlsson and Van Damme (1993), and the *interim approach* due to Weinstein and Yildiz (2007).

⁶This corresponds to the notion of “crazy types” in the robustness approach and to that of “dominance regions” in the global games or in the interim approach.

⁷In particular, our two main theorems will be derived with no assumption on the space of states of nature, which is in stark contrast with the richness assumption of Weinstein and Yildiz (2007) or its weakening proposed by Chen (2012).

⁸Similar abbreviations will be used throughout the paper for analogous objects.

abuse of notation, given a space X , we will sometimes write x for the degenerate distribution in $\Delta(X)$ assigning probability 1 to $\{x\}$; g will also be extended to measures, that is, given $\alpha \in \prod_{i \in \mathcal{I}} \Delta(M_i)$, we abuse notation and write $g(\alpha)$ for the measure $\alpha \circ g^{-1}$. Similarly, utility functions u_i 's will be extended to measures over A .

There are two main classes of situations with incomplete information. The first one consists in situations with an ex ante stage during which each player observes a private signal about the payoffs, and the joint distribution of signals and payoffs is commonly known. These situations are naturally modelled using a standard type space. The second class, on which we focus in this paper, consists in genuine situations of incomplete information, that is, situations with no ex ante stage: each player begins with a hierarchy of beliefs. We follow the standard Harsanyi (1967) approach and model these hierarchies of beliefs by introducing a hypothetical ex ante stage leading to a standard type space. This allows us to study strategic behavior of players at types that are considered to be close to a given original model.

A model \mathcal{T} is a pair (T, κ) , where $T = T_1 \times \dots \times T_I$ is a countable⁹ type space that is a complete and separable topological space, and $\kappa(t_i) \in \Delta(\Theta \times T_{-i})$ denotes the associated beliefs for each $t_i \in T_i$. We also note $\kappa(t_i)[E]$ for the probability of any measurable set $E \subset \Theta \times T_{-i}$ given by $\kappa(t_i)$. The function $\kappa(t_i)$ is assumed to be continuous with respect to t_i for each player i .¹⁰ In what follows, for two models $\mathcal{T} = (T, \kappa)$ and $\mathcal{T}' = (T', \kappa')$, we will write $\mathcal{T} \supset \mathcal{T}'$ if $T \supset T'$, the sigma field on each T'_i is the relative sigma field obtained from each T_i , and, for all i , $t_i \in T'_i$: $\kappa(t_i)[E] = \kappa'(t_i)[(\Theta \times T'_{-i}) \cap E]$ for any measurable $E \subset \Theta \times T_{-i}$. Given a mechanism \mathcal{M} and a model \mathcal{T} , we write $U(\mathcal{M}, \mathcal{T})$ for the induced incomplete information game. In this game, a (behavioral) strategy of a player i is any measurable function $\sigma_i: T_i \rightarrow \Delta(M_i)$. We will write $\sigma_i(m_i|t_i)$ for the probability that strategy σ_i assigns to message m_i when player i is of type t_i . For each $i \in \mathcal{I}$ and for each belief $\pi_i \in \Delta(\Theta \times M_{-i})$, we set

$$BR_i(\pi_i|\mathcal{M}) = \arg \max_{m_i \in M_i} \sum_{(\theta, m_{-i}) \in \Theta \times M_{-i}} \pi_i(\theta, m_{-i}) [u_i(g(m_i, m_{-i}), \theta)].$$

Since we allow for infinite mechanism, this set may be empty. Given any type t_i and any strategy profile σ_{-i} , we write $\pi_i(\cdot|t_i, \sigma_{-i}) \in \Delta(\Theta \times M_{-i})$ for the joint distribution on the underlying uncertainty and the other players' messages induced by t_i and σ_{-i} .

⁹This assumption is made to ensure existence of Bayes Nash equilibrium in finite games, which will turn out to be useful when we deal with sufficient conditions for continuous implementation. See also footnote 1 in the Supplemental Material (Oury and Tercieux (2012)).

¹⁰Given that T is countable, the discrete topology makes T complete and separable. It also makes $\kappa(t_i)$ a continuous function with respect to t_i for each player i . In general, we allow any topology on T that ensures these properties.

DEFINITION 1: A profile of strategies $\sigma = (\sigma_1, \dots, \sigma_I)$ is a Bayes Nash equilibrium in $U(\mathcal{M}, \mathcal{T})$ if, for each $i \in \mathcal{I}$ and each $t_i \in T_i$,

$$m_i \in \text{Supp}(\sigma_i(t_i)) \Rightarrow m_i \in BR_i(\pi_i(\cdot|t_i, \sigma_{-i})|\mathcal{M}).$$

Given two models \mathcal{T} and \mathcal{T}' such that $\mathcal{T} \supset \mathcal{T}'$, and given a strategy σ in $U(\mathcal{M}, \mathcal{T})$, we write $\sigma_{\mathcal{T}'}$ for the strategy σ restricted to \mathcal{T}' . It is easily checked that if σ is a Bayes Nash equilibrium in $U(\mathcal{M}, \mathcal{T})$, then $\sigma_{\mathcal{T}'}$ is also a Bayes Nash equilibrium in $U(\mathcal{M}, \mathcal{T}')$.

Given a model (T, κ) and any type t_i in type space T_i , we can compute the belief of t_i on Θ by setting

$$h_i^1(t_i) = \text{marg}_{\Theta} \kappa(t_i),$$

which is called the *first-order belief* of t_i . Similarly, we can compute the second-order belief of t_i , that is, his beliefs about $(\theta, h_1^1(t_1), \dots, h_I^1(t_I))$, by setting

$$h_i^2(t_i)[F] = \kappa(t_i)[\{(\theta, t_{-i}) | (\theta, h_1^1(t_1), \dots, h_I^1(t_I)) \in F\}]$$

for each measurable $F \subset \Theta \times \Delta(\Theta)^I$. Proceeding in this way, we can compute an entire hierarchy of beliefs. Hence, a type of a player i induces an infinite hierarchy of beliefs $(h_i^1(t_i), h_i^2(t_i), \dots, h_i^\ell(t_i), \dots)$, where $h_i^1(t_i) \in \Delta(\Theta)$ is a probability distribution on Θ , representing the beliefs of i about θ , $h_i^2(t_i) \in \Delta(\Theta \times \Delta(\Theta)^I)$ is a probability distribution representing the beliefs of i about θ and the others' first-order beliefs, and so on. Let us write $h_i(t_i)$ for the resulting hierarchy and $h_i^\ell(t_i)$ for the ℓ th-order beliefs of type t_i .

The set of all belief hierarchies for which it is common knowledge that the beliefs are coherent (i.e., each player knows his beliefs and his beliefs at different orders are consistent with each other) is the universal type space (see Mertens and Zamir (1985) and Brandenburger and Dekel (1993)). We denote by T_i^* the set of player i 's hierarchies of beliefs in this space and write $T^* = \prod_{i \in \mathcal{I}} T_i^*$.

In our formulation, two types, t and \bar{t} , are close if there exists a sufficiently large ℓ such that, for each $\ell' \leq \ell$, the ℓ' th-order beliefs $h^{\ell'}(t)$ and $h^{\ell'}(\bar{t})$ are close in the topology of convergence of measures. To be more precise, each T_i^* is endowed with the product topology, so that a sequence of types $\{t_i[n]\}_{n=0}^\infty$ converges to a type t_i , if, for each ℓ , $h_i^\ell(t_i[n]) \rightarrow h_i^\ell(t_i)$ (i.e., $h_i^\ell(t_i[n])$ converges toward $h_i^\ell(t_i)$ in the topology of weak convergence of measures¹¹). In such a case, we write $t_i[n] \rightarrow_P t_i$.

We are now in a position to introduce the notion of continuous implementation. The social planner considers a model and wants to analyze how strategic behavior is affected under his mechanism when the assumption that his

¹¹Recall that $h_i^k(t_i[n]) \in \Delta(X_{k-1})$, where $X_0 = \Theta$ and $X_k = [\Delta(X_{k-1})]^I \times X_{k-1}$.

model is common knowledge is relaxed. In the sequel, we fix a finite model $\bar{T} = (\bar{T}, \bar{\kappa})$, which is the model the planner has in mind. A social choice function¹² (SCF, henceforth) is a mapping $f: \bar{T}_0 \rightarrow A$, where $\bar{T}_0 \subset \bar{T}$. In the sequel, \bar{T}_0 will be interpreted as the set of profiles of types that the planner cares about. As will become clear, in some natural cases (e.g., complete information settings), this may be a proper subset of \bar{T} .¹³

Again, in this paper, we focus on partial implementation. We will be considering two notions of partial implementation, depending on the solution concept used, that yield two notions of continuous implementation. Traditionally, implementation theory has focused on pure Nash equilibria, and a SCF f is said to be partially implementable by a mechanism \mathcal{M} if, in the Bayesian game $U(\mathcal{M}, \bar{T})$, there is a pure Bayes Nash equilibrium that induces f . We will also consider a stronger notion of partial implementation—namely, partial strict implementation—replacing pure Bayes Nash equilibrium in the previous sentence by strict Bayes Nash equilibrium.¹⁴ Informally, a SCF will be said to be continuously (resp. strictly) implementable if, for any model containing the original benchmark model \bar{T} , there is a Bayes Nash equilibrium that (i) is pure (resp. strict) on the original model; that is, we maintain the original requirement of partial implementation in pure (resp. strict) equilibrium. As we will explain, this ensures that continuous (resp. strict) implementation is a refinement of partial (resp. strict) implementation. In addition, the Bayes Nash equilibrium of the model containing the original benchmark model \bar{T} must be such that (ii) if a type t is close to some type \bar{t} of the original model, then the outcome provided at t is close to the desired outcome $f(\bar{t})$; that is, we impose a continuity requirement. Such an equilibrium will be said to continuously implement f . Formally, we have the following definition.

DEFINITION 2: Fix a mechanism \mathcal{M} and a model \mathcal{T} such that $\bar{T} \subset \mathcal{T}$. We say that an equilibrium σ in $U(\mathcal{M}, \mathcal{T})$ *continuously* (resp. *strictly*) *implements* $f: \bar{T}_0 \rightarrow A$ if (i) $\sigma_{\bar{T}}$ is a pure (resp. strict) Nash equilibrium in $U(\mathcal{M}, \bar{T})$,

¹²In the paper, we restrict our attention to social choice functions for simplicity. For social choice sets, there are several notions of partial implementation one may use. In any case, our results apply in an obvious way: see footnote 21.

¹³For instance, Jackson (1991) assumed that each agent has a (common support) prior over \bar{T} . \bar{T}_0 is then defined as the set of profiles of types to which agents assign strictly positive probability.

¹⁴Strict Nash equilibria resist most standard refinements in game theory. The planner may hence naturally be willing to strengthen his requirement from partial implementation in pure Nash to partial implementation in strict Nash equilibrium. We note, that in some standard contexts, like complete information settings, partial implementation in strict Nash equilibrium can be achieved for any social choice functions when there are two or more players and when transfers are allowed.

(ii) for any $\bar{i} \in \bar{T}_0$ and any sequence $t[n] \rightarrow_p \bar{i}$, where, for each n , $t[n] \in T$, we have $(g \circ \sigma)(t[n]) \rightarrow f(\bar{i})$.¹⁵

We now state a formal definition of continuous implementation.

DEFINITION 3: A social choice function $f: \bar{T}_0 \rightarrow A$ is continuously (resp. strictly) implementable if there exists a mechanism \mathcal{M} such that, for each model \mathcal{T} with $\bar{T} \subset \mathcal{T}$, there is an equilibrium σ in $U(\mathcal{M}, \mathcal{T})$ that continuously (resp. strictly) implements f .¹⁶

As already stated, the notion of continuous (resp. strict) implementation is a refinement of partial (resp. strict) implementation. Indeed, if a social choice function is continuously (resp. strictly) implementable, then (taking $\mathcal{T} = \bar{T}$ and using (i) in Definition 2), there is a pure (resp. strict) equilibrium σ in $U(\mathcal{M}, \bar{T})$ satisfying $(g \circ \sigma)(\bar{i}) = f(\bar{i})$ for any $\bar{i} \in \bar{T}_0$; that is, the mechanism \mathcal{M} must partially (resp. strictly) implement f .

On the Complete Information Benchmark

Our formalism allows \bar{T}_0 to be a proper subset of \bar{T} . In other terms, the planner may not care about achieving the right outcome at all types in \bar{T} . One natural setting in which this will be the case is the complete information setting. In this situation, given a finite set of states of nature Θ , whenever the true state is θ , it is assumed to be common belief among agents. To incorporate this in our setting, we define the complete information model $\mathcal{T}^{CI} = (T^{CI}, \kappa)$ as follows: for each player i , $T_i^{CI} = \bigcup_{\theta \in \Theta} \{t_{i,\theta}\}$ and $\kappa(t_{i,\theta}) = \delta_{(\theta, t_{-i,\theta})}$ for all θ , with $\delta_{(\theta, t_{-i,\theta})}$ the probability distribution that puts probability 1 on $\{(\theta, t_{-i,\theta})\}$. Here, at the profile of types $t_\theta = (t_{i,\theta})_{i \in \mathcal{I}}$, players believe with probability 1 that state θ is the true state. They all believe that the others believe that state θ is the true state and so on. Hence, state θ is common belief among agents. Henceforth, such profiles of types will be called complete information profiles of types. We consider that the model \bar{T} the planner has in mind is \mathcal{T}^{CI} . Note that there are profiles of types in \bar{T} that are not of complete information. Recall that, by definition, the planner is willing to achieve the right outcome only at complete information profiles of types and at profiles of types that are close to these profiles of types. In our formalism, this means that \bar{T}_0 is equal to $\bigcup_{\theta \in \Theta} \{t_\theta\}$, which is a proper subset of \bar{T} .

¹⁵The function $g \circ \sigma$ is the composition of the mapping $\sigma: T \rightarrow \prod_{i \in \mathcal{I}} \Delta(M_i)$ and of the mapping $g: \prod_{i \in \mathcal{I}} \Delta(M_i) \rightarrow \Delta(A)$.

¹⁶Our continuous implementation notion can be defined with other topologies than the product topology. Two natural candidates are the uniform topology or the strategic topology (see Dekel, Fudenberg, and Morris (2006)). With these topologies, as expected, very permissive results can be shown. In particular, partial implementation in strict equilibrium with a finite mechanism is a sufficient condition for strict continuous implementation.

On SCFs Defined on Payoff-Relevant Types

Most of the mechanism design literature assumes that the social choice function depends only on states of nature (and not directly on profiles of types of agents). Our approach in that respect is more general. Consider a social choice function f that now maps Θ into A . A simple way to incorporate this formulation into our setting is to assume the existence of a mapping $\bar{\theta}: \bar{T} \rightarrow \Theta$ that, for each profile of types, specifies the state of nature that the planner associates to it. Then define $\tilde{f}: \bar{T} \rightarrow A$ by $\tilde{f}(t) = f(\bar{\theta}(t))$ for all $t \in \bar{T}$. All our results can be applied to this standard setting.

In the next section, we assume that the planner considers partial implementation in strict Nash equilibrium in his benchmark model and so requires the SCF to be strictly continuously implemented. In Section 4, we will relax the assumption that a continuous equilibrium must be strict on the original model, and show that similar results can be obtained provided that we slightly depart from the assumption that messages are costless.

3. MAIN RESULTS

The first subsection below presents our basic argument in the complete information benchmark. This allows a presentation of the core of our argument with very few technicalities. The second subsection treats the incomplete information case and provides our main results.

3.1. A Simple Illustration of the Argument Under Complete Information

In the sequel, \bar{T}_0 is assumed to be equal to $\bigcup_{\theta \in \Theta} \{t_\theta\}$. In this section, we show that any social choice function $f: \bar{T}_0 \rightarrow A$ that is strictly continuously implementable satisfies a strengthening of the well known monotonicity condition defined in Maskin (1999). To explain in words this condition, it states that in case the desired alternative differs at profiles of types t_θ and $t_{\theta'}$, there must exist at least one agent who, if the true state were θ' and he expected other agents to claim the state is θ , could be offered a reward a that would give him a strict incentive to “report” the deviation of other agents, where the reward a would not tempt him if the true state was in fact θ ; that is, he would have a (strict) incentive to “report truthfully.” So (strict) Maskin monotonicity essentially ensures the elimination of some undesirable equilibria. Formally, we have the following definition.

DEFINITION 4: The social choice function $f: \bar{T}_0 \rightarrow A$ is strict Maskin monotonic¹⁷ if, for every pair of states θ and θ' ,

¹⁷Maskin’s original definition is weaker: the strict inequality in (\star) is replaced by a weak one.

(i) $f(t_\theta) = f(t_{\theta'})$ whenever

$$(\star) \quad u_i(f(t_\theta), \theta) > u_i(a, \theta) \implies u_i(f(t_\theta), \theta') \geq u_i(a, \theta')$$

for all i and a ;

or, equivalently,

(ii) $f(t_\theta) \neq f(t_{\theta'})$ implies

$$u_i(f(t_\theta), \theta) > u_i(a, \theta) \quad \text{and} \quad u_i(a, \theta') > u_i(f(t_\theta), \theta')$$

for some i and a .

The following result, which is a first step toward our main result, reduces the gap between strict continuous implementation and full implementation in Nash equilibrium since—as established by Maskin—the above monotonicity condition is “almost” sufficient for full implementation. The strict version of Maskin monotonicity was introduced in Bergemann, Morris, and Tercieux (2011), where it was shown to be tightly connected to full implementation in rationalizable strategies.¹⁸

THEOREM 1: *The social choice function f is strictly continuously implementable only if it is strict Maskin monotonic.*

PROOF: Assume that there exists a mechanism $\mathcal{M} = (M, g)$ that strictly continuously implements f . Pick $t_\theta, t_{\theta'} \in \bar{T}_0$ such that, for each player i and for each $a \in A$, the relation (\star) is satisfied. We want to show that $f(t_\theta) = f(t_{\theta'})$.

We show that there exists a model $\mathcal{T} = (T, \kappa) \supset \mathcal{T}^{\text{CI}}$ such that, for any equilibrium σ that strictly continuously implements f , there is a sequence of types $\{t[n]\}_{n=1}^\infty$ in T such that $t[n] \rightarrow_P t_{\theta'}$ and $(g \circ \sigma)(t[n]) \rightarrow f(t_\theta)$. By point (ii) of Definition 2, $(g \circ \sigma)(t[n]) \rightarrow f(t_{\theta'})$, which implies $f(t_\theta) = f(t_{\theta'})$.

For this purpose, we build the desired model $\mathcal{T} = (T, \kappa)$ in which, for each player i , each set T_i satisfies

$$T_i = T_i^{\text{CI}} \cup \left(\bigcup_{n=1}^\infty t_i[n] \right),$$

where $t_i[n]$ and κ are defined recursively as follows. First, $t_i[1]$ is such that

$$\text{marg}_{T_{-i}} \kappa(t_i[1])[t_{-i,\theta}] = 1, \quad \text{marg}_\theta \kappa(t_i[1])[\theta'] = \frac{1}{2},$$

¹⁸Maskin (1999) showed that, with more than three players together with the assumption that f satisfies the weak condition of no veto power, monotonicity actually implies full implementation in Nash equilibrium. In addition, with more than three players, under an economic condition and restricting attention to social choice functions that are responsive (i.e., in distinct states they select distinct outcomes), Bergemann, Morris, and Tercieux (2011) showed that full implementation in rationalizable strategies of a strict monotonic social choice function can be achieved.

and

$$\text{marg}_\theta \kappa(t_i[1])[\theta] = \frac{1}{2}.$$

In addition, for each $n \geq 2$, $t_i[n]$ is defined by

$$\text{marg}_{T_{-i}} \kappa(t_i[n])[t_{-i}[n-1]] = 1, \quad \text{marg}_\theta \kappa(t_i[n])[\theta] = 1 - \frac{1}{n+1},$$

and

$$\text{marg}_\theta \kappa(t_i[n])[\theta] = \frac{1}{n+1}.$$

Observe that each T_i is countable, and so is T .

Now pick any equilibrium σ of the induced game $U(\mathcal{M}, \mathcal{T})$ that strictly continuously implements f . By point (i) in Definition 2, $\sigma_{|T^c}$ is a strict Nash equilibrium in $U(\mathcal{M}, \bar{T})$ and hence, $\sigma(t_\theta)$ is a strict Nash equilibrium in the game of complete information where payoffs are given for each player i by $u_i(\cdot, \theta)$. In addition, point (ii) in Definition 2 implies that $(g \circ \sigma)(t_\theta) = f(t_\theta)$. In the sequel, we note $m_i^* := \sigma_i(t_{i,\theta})$ and $m^* := \sigma(t_\theta)$. We have, for any player i and $m'_i \neq m_i^*$,

$$u_i(g(m^*), \theta) > u_i(g(m'_i, m_{-i}^*), \theta),$$

and so,

$$u_i(f(t_\theta), \theta) > u_i(g(m'_i, m_{-i}^*), \theta).$$

By (\star) , this implies that

$$u_i(f(t_\theta), \theta') \geq u_i(g(m'_i, m_{-i}^*), \theta'),$$

which in turn yields

$$u_i(g(m^*), \theta') \geq u_i(g(m'_i, m_{-i}^*), \theta'),$$

that is, m^* is a pure Nash equilibrium in the game of complete information where payoffs are given for each player i by $u_i(\cdot, \theta')$. Now, by construction, under the equilibrium σ , type $t_i[1]$ believes with probability 1 that the other players play m_{-i}^* . He assigns strictly positive weight only to state θ , under which m_i^* is the unique best response against m_{-i}^* , and to state θ' , under which m_i^* is a best response against m_{-i}^* . Hence, we must have

$$\sigma_i(t_i[1]) = m_i^*.$$

Using a similar reasoning, it is easy to show inductively that, for all $n \geq 2$,

$$\sigma_i(t_i[n]) = m_i^*.$$

This means that, for each $n \geq 1$, $(g \circ \sigma)(t[n]) = g(m^*) = f(t_\theta)$ and so, obviously, $(g \circ \sigma)(t[n]) \rightarrow f(t_\theta)$ (as $n \rightarrow \infty$), which completes the proof since $t[n] \rightarrow_P t_\theta$ (as $n \rightarrow \infty$). *Q.E.D.*

In the next section, we generalize the above result to the Bayesian setting. The argument relies on a general necessary condition (Lemma 1 below) for strict continuous implementation, which is implicitly used in the above proof. Let us briefly explain how the argument works. In the above proof, we established that strict continuous implementation implies the existence of a strict Nash equilibrium $\sigma_{|\bar{T}}$ in $U(\mathcal{M}, \bar{T})$ such that, if m^* is a Nash equilibrium at state θ' (i.e., in the game of complete information where payoffs are given for each player i by $u_i(\cdot, \theta')$) and m^* is played by $\sigma_{|\bar{T}}$ at t_θ , we must have $g(m^*) = f(t_\theta)$. Note that by a similar argument, one can show that strict continuous implementation implies the existence of a strict Nash equilibrium $\sigma_{|\bar{T}}$ in $U(\mathcal{M}, \bar{T})$ such that, if m^* is a Nash equilibrium at state θ' and $m^* \in \sigma_{|\bar{T}}(\bar{T})$, we must have $g(m^*) = f(t_{\theta'})$.¹⁹ Since this property is true for any $\theta' \in \Theta$, strict continuous implementation implies the existence of a strict Nash equilibrium $\sigma_{|\bar{T}}$ in $U(\mathcal{M}, \bar{T})$ such that, if a pure equilibrium $\tilde{\sigma}$ in $U(\mathcal{M}, \bar{T})$ satisfies $\tilde{\sigma}(\bar{T}) \subset \sigma_{|\bar{T}}(\bar{T})$, we must have $g(\tilde{\sigma}(\bar{i})) = f(\bar{i})$ for all $\bar{i} \in \bar{T}_0$. As we will see, this condition is also necessary for strict continuous implementation in the Bayesian setting (Lemma 1) and implies the usual monotonicity condition of the social choice function.

3.2. The General Argument Under Incomplete Information

We now drop the restriction that the benchmark model \bar{T} is a complete information model. In the sequel, to make the argument simpler, we assume that the planner cares about all profiles of types in \bar{T} , that is, $\bar{T}_0 = \bar{T}$. All definitions provided below can easily be adapted to the case where \bar{T}_0 is a proper subset of \bar{T} (as, for instance, in Jackson (1991)). All our results, as well, can be extended in a straightforward way.

Our results show a tight connection between the notion of partial continuous implementation and full implementation in (interim correlated) rationalizable strategies. Let us first recall the definition of (interim correlated) rationalizability given in Dekel, Fudenberg, and Morris (2006, 2007). Fix a mechanism \mathcal{M} and a type space \mathcal{T} . Let $\Sigma_i : T_i \rightarrow 2^{M_i} \setminus \{\emptyset\}$ be a specification of possible messages for each type of player i and let $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$.

¹⁹Here, given an equilibrium σ , $\sigma(\bar{T})$ stands for the image of σ , that is, $\bigcup_{\bar{i} \in \bar{T}} \sigma(\bar{i})$.

DEFINITION 5: Given a mechanism \mathcal{M} and a type space \mathcal{T} , $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ is a best-reply set in $U(\mathcal{M}, \mathcal{T})$ if, for each t_i and $m_i \in \Sigma_i(t_i)$, there exists $\pi_i \in \Delta(\Theta \times T_{-i} \times M_{-i})$, where $\text{marg}_{\Theta \times T_{-i}} \pi_i = \kappa(t_i)$, $\pi_i(\theta, t_{-i}, m_{-i}) > 0 \Rightarrow m_{-i} \in \Sigma_{-i}(t_{-i})$, and $m_i \in BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i | \mathcal{M})$.

For each player i , the set of rationalizable messages at type \bar{t}_i , denoted $R_i(\bar{t}_i | \mathcal{M}, \bar{\mathcal{T}})$, is the set of messages m_i such that $m_i \in \Sigma_i(\bar{t}_i)$ for some best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$. We also note $R(\bar{t} | \mathcal{M}, \bar{\mathcal{T}})$ for $\prod_{i \in \mathcal{I}} R_i(\bar{t}_i | \mathcal{M}, \bar{\mathcal{T}})$. We say that a SCF f is implementable in rationalizable strategies (or rationalizable implementable) on $\bar{\mathcal{T}}$ if there is a mechanism \mathcal{M} under which, for all $\bar{t} \in \bar{\mathcal{T}}$, (i) $R(\bar{t} | \mathcal{M}, \bar{\mathcal{T}}) \neq \emptyset$, and (ii) $m \in R(\bar{t} | \mathcal{M}, \bar{\mathcal{T}}) \Rightarrow g(m) = f(\bar{t})$. In the sequel, an equilibrium σ will be considered as a single-valued best-reply set. Also, given a best-reply set Σ , we will note $\Sigma(\bar{\mathcal{T}})$ for the image of Σ , that is, $\bigcup_{\bar{t} \in \bar{\mathcal{T}}} \Sigma(\bar{t})$.

As already mentioned, our argument in the proof of Theorem 1 shows that strict continuous implementation implies the existence of a strict equilibrium $\bar{\sigma}$ on $\bar{\mathcal{T}}$, under which any equilibrium σ satisfying $\sigma(\bar{\mathcal{T}}) \subset \bar{\sigma}(\bar{\mathcal{T}})$ must also satisfy $g(\sigma(\bar{t})) = f(\bar{t})$, for all $\bar{t} \in \bar{\mathcal{T}}$. The result below shows that not only must this be true for any equilibrium σ satisfying $\sigma(\bar{\mathcal{T}}) \subset \bar{\sigma}(\bar{\mathcal{T}})$, but it is also true for best-reply sets. This lemma will generalize to the incomplete information setting the idea of the contagion argument used in the proof of Theorem 1. Our main results will be derived from this lemma.

LEMMA 1: Assume $f: \bar{\mathcal{T}} \rightarrow A$ is strictly continuously implementable by a mechanism \mathcal{M} . There exists a strict equilibrium $\bar{\sigma}$ in $U(\mathcal{M}, \bar{\mathcal{T}})$ such that, if a best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{\mathcal{T}})$ satisfies $\Sigma(\bar{\mathcal{T}}) \subset \bar{\sigma}(\bar{\mathcal{T}})$, then $g(\Sigma(\bar{t})) = \{f(\bar{t})\}$ for all $\bar{t} \in \bar{\mathcal{T}}$.

See the [Appendix](#) for the proof.

Let us relate Lemma 1 to Weinstein and Yildiz (2004, 2007). To make the comparison clearer, let us fix a (finite) game and say that a pure equilibrium $\bar{\sigma}$ in the model $\bar{\mathcal{T}}$ is continuous if, for each model $\mathcal{T} \supset \bar{\mathcal{T}}$, there is an equilibrium σ such that, for any sequence $t[n] \rightarrow_p \bar{t} \in \bar{\mathcal{T}}$ (where, for each n , $t[n] \in T$), we have $\sigma(t[n]) \rightarrow \bar{\sigma}(\bar{t})$.

Mimicking the argument in Weinstein and Yildiz (2004, 2007), it is easy to show that, in *generic games*,²⁰ if a pure equilibrium $\bar{\sigma}$ in the model $\bar{\mathcal{T}}$ has full range (i.e., for each player i and each action of this player, there is a type in $\bar{\mathcal{T}}_i$ who plays this action under $\bar{\sigma}$), and if in addition, this equilibrium is continuous, then we must have $R(\bar{t} | \bar{\mathcal{T}}) = \{\bar{\sigma}(\bar{t})\}$ for all $\bar{t} \in \bar{\mathcal{T}}$. Otherwise stated, if $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ is a best-reply set in the model $\bar{\mathcal{T}}$, then $\Sigma(\bar{t}) = \{\bar{\sigma}(\bar{t})\}$ for all $\bar{t} \in \bar{\mathcal{T}}$. One important point behind this result is that if the equilibrium $\bar{\sigma}$ has full

²⁰More precisely, when the set of strict rationalizable action profiles is equal to the set of rationalizable action profiles in $\bar{\mathcal{T}}$, which is true for a generic choice of payoffs.

range, and if there is a best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in the model \bar{T} with $\Sigma(\bar{t}) \neq \{\bar{\sigma}(\bar{t})\}$ for some $\bar{t} \in \bar{T}$, then, because the action profiles played in the best-reply set Σ (i.e., the action profiles in $\Sigma(\bar{T})$) are also played under $\bar{\sigma}$ at some profile of types (i.e., $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$), we can use a contagion argument to exhibit a model \mathcal{T} and a sequence $t[n]$ in this model that converges toward \bar{t} for which any equilibrium σ (that coincides with $\bar{\sigma}$ on \bar{T}) is such that $\sigma(t[n])$ does not converge toward $\bar{\sigma}(\bar{t})$, that is, showing that $\bar{\sigma}$ is not continuous.

Given this observation, one simple generalization of Weinstein and Yildiz (2004, 2007) would state that in *generic games*, if a pure equilibrium $\bar{\sigma}$ (that need not have full range) in the model \bar{T} is continuous, then we must have that, if $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ is a best-reply set in the model \bar{T} and $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, then $\Sigma(\bar{t}) = \{\bar{\sigma}(\bar{t})\}$ for all $\bar{t} \in \bar{T}$.

An important difference between our setup and that of Weinstein and Yildiz (2004, 2007) is that, in our mechanism design setting, the game is not part of the primitives; it is an endogenous object fixed by the social planner. Hence, we cannot make the assumption that the game is generic. Nevertheless, the above lemma shows that if f is strictly continuously implementable, then a necessary condition similar to that exhibited above for the existence of a continuous equilibrium in the model \bar{T} arises.

We now show that the necessary condition for strict continuous implementation exhibited in Lemma 1 implies Bayesian monotonicity. This notion was first introduced by Postlewaite and Schmeidler (1986), followed by Jackson (1991). We first recall its definition.

DEFINITION 6: A deception is a collection of mappings $\beta = (\beta_i)_{i=1}^I$, where each $\beta_i: \bar{T}_i \rightarrow 2^{\bar{T}_i}$. A deception β is acceptable if, for all $t \in \bar{T}$, $t' \in \beta(t) \Rightarrow f(t') = f(t)$. Deception β is unacceptable if it is not acceptable.

DEFINITION 7: The social choice function $f: \bar{T} \rightarrow A$ is Bayesian monotonic on $(\bar{T}, \bar{\kappa})$ if, for every unacceptable single-valued deception $\beta: \bar{T} \rightarrow \bar{T}$, there exist $i, t_i \in \bar{T}_i$ and $y: \bar{T} \rightarrow A$ such that

$$\begin{aligned} & \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i((y \circ \beta)(t_i, \tilde{t}_{-i}), \theta) \\ & > \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i((f \circ \beta)(t_i, \tilde{t}_{-i}), \theta), \end{aligned}$$

while

$$\sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(f(\tilde{t}_i, \tilde{t}_{-i}), \theta) \geq \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(y(\beta_i(\tilde{t}_i), \tilde{t}_{-i}), \theta)$$

for all $\tilde{t}_i \in \bar{T}_i$. The social choice function $f: \bar{T} \rightarrow A$ is strict Bayesian monotonic on $(\bar{T}, \bar{\kappa})$ if the above inequality is strict whenever $\tilde{t}_i = \beta_i(t_i)$.

To explain in words Bayesian monotonicity and how it relates to full implementation in Bayes Nash equilibrium, let us briefly interpret the different conditions of the definition. As for Maskin monotonicity, Bayesian monotonicity ensures elimination of some undesirable equilibria. For instance, consider an equilibrium σ that achieves the social choice function. If agents use an unacceptable (single-valued) deception β , then the strategy profile $\sigma \circ \beta$ should not be an equilibrium (since it does not achieve the social choice function). Now, the existence of y (in the definition of Bayesian monotonicity) with the stated properties allows agent i to signal that β is played: the agent is rewarded according to y , which makes him strictly better off at type t_i . The other part of the Bayesian monotonicity condition ensures that agent i cannot gain by falsely accusing the other agents of using deception β .

Our first main result is stated as follows.

THEOREM 2: *The social choice function f is strictly continuously implementable on $(\bar{T}, \bar{\kappa})$ only if it is strict Bayesian monotonic on $(\bar{T}, \bar{\kappa})$.²¹*

PROOF: Let \mathcal{M} be a mechanism that strictly continuously implements f . We first make a simple remark: since a strict equilibrium $\bar{\sigma}$ in $U(\mathcal{M}, \bar{T})$ can be considered as a single-valued best-reply set, Lemma 1 implies that, for the strict equilibrium $\bar{\sigma}$ in $U(\mathcal{M}, \bar{T})$ defined in this lemma, we must have $(g \circ \bar{\sigma})(\bar{t}) = f(\bar{t})$ for all $\bar{t} \in \bar{T}$. Now, pick this strict equilibrium $\bar{\sigma}$. Consider a single-valued deception $\beta: \bar{T} \rightarrow \bar{T}$ such that $(f \circ \beta)(\bar{t}) \neq f(\bar{t})$ for some $\bar{t} \in \bar{T}$ (i.e., β is unacceptable). Let $\hat{\sigma}$ be defined by $\hat{\sigma} = \bar{\sigma} \circ \beta$. By definition of a deception, $\beta(\bar{T}) \subset \bar{T}$ and hence, $\hat{\sigma}(\bar{T}) \subset \bar{\sigma}(\bar{T})$. Consequently, since $\hat{\sigma}$ can be considered as a single-valued correspondence and $(g \circ \hat{\sigma})(\bar{t}) = (g \circ \bar{\sigma} \circ \beta)(\bar{t}) = (f \circ \beta)(\bar{t}) \neq f(\bar{t})$, by Lemma 1, $\hat{\sigma}$ cannot be an equilibrium on $(\bar{T}, \bar{\kappa})$. Given this, to show the existence of i , $t_i \in \bar{T}_i$ and $y: \bar{T} \rightarrow A$ satisfying the conditions in the definition of Bayesian monotonicity, we just have to mimic the usual argument: there exist i , $t_i \in \bar{T}_i$ and some message $\tilde{m}_i \neq \hat{\sigma}_i(t_i)$ such that

$$\begin{aligned} & \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i(g(\tilde{m}_i, \hat{\sigma}_{-i}(\tilde{t}_{-i})), \theta) \\ & > \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i(g(\hat{\sigma}(t_i, \tilde{t}_{-i})), \theta). \end{aligned}$$

²¹For a social choice set, one may say that it is partially implementable if it has an element that is partially implementable. Our results then extend in an obvious way: a social choice set is strictly continuously partially implementable if it has some element that satisfies Bayesian monotonicity. We note that this is a bit different from the requirement that the social choice set itself is Bayesian monotonic in the sense of Jackson (1991). One may also say that a social choice set is partially implementable if there is a mechanism \mathcal{M} under which, for any element of the set, there is an equilibrium that coincides with this selection. In that case, a necessary condition for strict continuous partial implementation of a social choice set would be Jackson's Bayesian monotonicity.

Now, let $y: \bar{T} \rightarrow A$ be such that $y(\bar{t}) = g(\bar{m}_i, \bar{\sigma}_{-i}(\bar{t}_{-i}))$ for all $\bar{t} \in \bar{T}$. We obtain

$$\begin{aligned} & \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i((y \circ \beta)(t_i, \tilde{t}_{-i}), \theta) \\ & > \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(t_i)[\theta, \tilde{t}_{-i}] u_i((f \circ \beta)(t_i, \tilde{t}_{-i}), \theta). \end{aligned}$$

Since $\bar{\sigma}$ is an equilibrium in $U(\mathcal{M}, \bar{T})$ and $(g \circ \bar{\sigma})(\bar{t}) = f(\bar{t})$ for all $\bar{t} \in \bar{T}$, it follows that, for all $\tilde{t}_i \in \bar{T}_i$

$$\begin{aligned} (1) \quad & \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(g(\bar{\sigma}(\tilde{t}_i, \tilde{t}_{-i})), \theta) \\ & \geq \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(g(\bar{m}_i, \bar{\sigma}_{-i}(\tilde{t}_{-i})), \theta) \end{aligned}$$

which yields

$$(2) \quad \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(f(\tilde{t}_i, \tilde{t}_{-i}), \theta) \geq \sum_{\theta, \tilde{t}_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, \tilde{t}_{-i}] u_i(y(\beta_i(\tilde{t}_i), \tilde{t}_{-i}), \theta)$$

for all $\tilde{t}_i \in \bar{T}_i$. Note that since $\bar{m}_i \neq \hat{\sigma}_i(t_i) = \bar{\sigma}_i(\beta_i(t_i))$ and since $\bar{\sigma}$ is a strict equilibrium, for $\tilde{t}_i = \beta_i(t_i)$, the inequality (1) is strict and hence, inequality (2) is also strict. This proves that f is strict Bayesian monotonic on $(\bar{T}, \bar{\kappa})$. *Q.E.D.*

The previous result shows that in so-called economic environments defined in Jackson (1991) and under the usual additional conditions that ensure that Bayesian monotonicity is sufficient for implementation, strict continuous implementation yields full implementation in Bayes Nash equilibrium.²² Lemma 1 actually allows us to go one step further and to prove that, under some weak economic conditions, strict continuous implementation yields full implementation in rationalizable strategies. To show this, we now move to the notion of interim rationalizable monotonicity, a Bayesian version of a notion introduced in Bergemann and Morris (2011) in the ex post environment.

²²Recall that a SCF f is fully implementable in Bayesian Nash equilibrium if there is a mechanism that partially implements f and, for every Bayes Nash equilibrium, $\sigma: \bar{T} \rightarrow M$ of the induced game $g \circ \sigma = f$. Bayesian monotonicity and incentive compatibility of a SCF f are enough to ensure full implementation of f in economic environments. This need not be true in general environments; see Jackson (1991).

DEFINITION 8: The social choice function $f: \bar{T} \rightarrow A$ is interim rationalizable monotonic on $(\bar{T}, \bar{\kappa})$ if, for every unacceptable deception β , there exist $i, t_i, t'_i \in \beta_i(t_i)$ such that, for every $\psi_i \in \Delta(\Theta \times \bar{T}_{-i} \times \bar{T}_{-i})$ that satisfies the properties

- (i) $\psi_i(\theta, t_{-i}, t'_{-i}) > 0 \Rightarrow t'_j \in \beta_j(t_j)$ for each $j \neq i$,
- (ii) $\sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) ,

there exists $y^*: \bar{T}_{-i} \rightarrow A$ such that

$$\sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(y^*(t'_{-i}), \theta) > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(f(t'_i, t'_{-i}), \theta)$$

and

$$\sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(f(\tilde{t}_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(y^*(t_{-i}), \theta)$$

for each $\tilde{t}_i \in \bar{T}_i$. The social choice function $f: \bar{T} \rightarrow A$ is strict interim rationalizable monotonic on $(\bar{T}, \bar{\kappa})$ if the above inequality is strict whenever $\tilde{t}_i = t'_i$.

Let us here again briefly explain in words interim rationalizable monotonicity and how it relates to rationalizable implementation. Its role for rationalizable strategy profiles is very close to the role played for equilibria by Bayesian monotonicity. Interim rationalizable monotonicity ensures elimination of some undesirable best-reply sets. For instance, if Σ is a best-reply set containing rationalizable strategy profiles that all achieve the social choice function, then if agents use an unacceptable deception β , then $\Sigma \circ \beta$ should not be a best-reply set (since it does not achieve the social choice function). Now, interim rationalizable monotonicity ensures that, for some type t_i , for any belief of this type that is consistent with the other agents using the unacceptable deception β , there exists a function y^* with the stated properties that allows t_i to signal that β is played; the agent is rewarded according to y^* and this makes him strictly better off. As for Bayesian monotonicity, the other part of the interim rationalizable monotonicity condition ensures that agent i cannot gain by falsely accusing the other agents of using deception β .

It is easy to check that if a SCF is interim rationalizable monotonic, then it is Bayesian monotonic. We now state our third theorem, which shows that our argument relating the necessary condition in Lemma 1 to the notion of Bayesian monotonicity actually extends to the more stringent notion of interim rationalizable monotonicity.

THEOREM 3: *If $f: \bar{T} \rightarrow A$ is strictly continuously implementable on $(\bar{T}, \bar{\kappa})$, then f is strict interim rationalizable monotonic on $(\bar{T}, \bar{\kappa})$.*

See the [Appendix](#) for the proof.

As for Bayesian monotonicity, under a weak domain restriction, a SCF that is interim rationalizable monotonic is rationalizable implementable (via a stochastic mechanism, i.e., a mechanism that associates lotteries over outcomes to each profile of messages) with more than three players. Hence, Theorem 3 makes a connection between strict continuous implementation and rationalizable implementation. Let us formally state this result.

Fix $\bar{T} = (\bar{T}, \bar{\kappa})$ and recall that $\Delta(A)$ is the set of lotteries over the set of outcomes A (which is assumed to be countable). Let us fix a social choice function f . We define the economic condition that will be needed. To do so, let Y_i be the set of mappings from types of player i 's opponents to lotteries satisfying the following property. Whatever agent i 's actual type is, he would never prefer the outcome to be selected according to a mapping in Y_i to what he would obtain under the social choice function f if other agents were reporting truthfully. Formally,

$$Y_i := \left\{ y: \bar{T}_{-i} \rightarrow \Delta(A) \mid \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(f(\tilde{t}_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(y(t_{-i}), \theta) \text{ for all } \tilde{t}_i \in \bar{T}_i \right\}.$$

ASSUMPTION 1: For all i , there exists $\bar{y}_i: \bar{T}_{-i} \rightarrow \Delta(A)$ such that the following statement holds. For any $\psi_i \in \Delta(\Theta \times \bar{T}_{-i})$, there exists $y(\cdot) \in Y_i$ satisfying

$$\sum_{\theta, t'_{-i}} \psi_i(\theta, t'_{-i}) u_i(y(t'_{-i}), \theta) > \sum_{\theta, t'_{-i}} \psi_i(\theta, t'_{-i}) u_i(\bar{y}_i(t'_{-i}), \theta).$$

The above assumption says that if we consider preferences of player i over mappings from \bar{T}_{-i} to $\Delta(A)$, then there is a mapping $\bar{y}_i: \bar{T}_{-i} \rightarrow \Delta(A)$ such that, whatever player i 's beliefs over $\Theta \times \bar{T}_{-i}$ are, some mapping in Y_i is preferred to \bar{y}_i . This assumption is satisfied in several natural environments like the one where there exist transfers or bad outcomes that are not desired by the social planner. More generally, it is also naturally satisfied by the appropriate Bayesian version of the no-worst-alternative condition as defined in Cabrales and Serrano (2011) or Bergemann, Morris, and Tercieux (2011).²³ The sufficiency result is stated as follows.

PROPOSITION 1: Assume $I \geq 3$. If $f: \bar{T} \rightarrow A$ satisfies interim rationalizable monotonicity and if Assumption 1 holds, then f is interim rationalizable implementable with a stochastic mechanism.

²³The no-worst-alternative condition requires that an agent never gets his worst outcome under the social choice function.

See the [Appendix](#) for the proof.

Proposition 1, together with Theorem 3, yields the following corollary.

COROLLARY 1: *Assume $I \geq 3$. If $f: \bar{T} \rightarrow A$ is strictly continuously implementable and if Assumption 1 holds, then f is interim rationalizable implementable.*

4. CONTINUOUS IMPLEMENTATION WITH SMALL COSTS OF MESSAGES

So far, we presented necessary conditions for strict continuous implementation that are close to rationalizable implementation. Two questions are of obvious interest. First, can we relax the requirement of strict continuous implementation to continuous implementation? Then, how far are the necessary conditions from sufficiency? Let us make two simple remarks. On the one hand, it is easily shown that continuous implementation by itself is not enough to get our previous theorems.²⁴ On the other hand, given the result by Dekel, Fudenberg, and Morris (2006) that the correspondence of (interim correlated) rationalizable strategies is upper hemicontinuous in the product topology in the universal type space, we know that if a SCF f is rationalizable implementable by a finite mechanism and the set of rationalizable strategies is a singleton, then f is actually strictly continuously implementable. In the sequel, to address the two questions mentioned above, we depart from the setting introduced so far by slightly relaxing the assumption that messages are costless. Restricting our attention to finite mechanisms, we obtain a full characterization result for continuous implementation (as opposed to strict continuous implementation). More precisely, we show that a SCF is continuously implementable by a finite mechanism if and only if it is implementable in rationalizable strategies by a finite mechanism. This will allow us to derive corollaries on virtual implementation, which uses finite mechanisms.

Small Costs of Messages

Sending a message usually requires one to fill in a questionnaire, to write a letter, or to make public some private information (like the willingness to pay for an object). All these actions are not costless for the agents, and a recent literature in implementation takes into account potential costs of messages.²⁵

²⁴To see this, consider the simple example where, for any player i , state θ , and outcome a , $u_i(a, \theta) = 0$. Fix any model \bar{T} and any nonconstant social choice function f . Consider a (direct) mechanism where each player is asked to report his type (i.e., $M_i = \bar{T}_i$ for each player i) and the outcome provided for a profile of messages m is $f(m)$. Since players are totally indifferent, for any model, any profile of strategy is a Bayes Nash equilibrium. Hence, it is easily checked that f is continuously implemented by the above mechanism. However, since players are totally indifferent and f is nonconstant, f is not Bayesian monotonic on \bar{T} .

²⁵See, for instance, Deneckere and Severinov (2008), Matsushima (2008), and Kartik and Tercieux (2012).

In this section, the social planner does not exclude that some (arbitrarily) small costs of messages may exist even if in the benchmark model messages are costless. To formalize this idea, we proceed as follows.

Given a mechanism $\mathcal{M} = (M, g)$, for each player i , we define a cost function $c_i : M_i \times \tilde{\Theta} \rightarrow \mathbb{R}_+$, where $\tilde{\Theta}$ is the space of states of nature associated with costs of messages. We assume that $\tilde{\Theta}$ is rich enough. More precisely, it is defined by

$$\tilde{\Theta} = \bigcup_{i \in \mathcal{I}} \bigcup_{m_i \in M_i} \{\tilde{\theta}^{m_i}\} \cup \{\tilde{\theta}^0\},$$

where, for each player i and each message m_i , we have $c_i(m_i, \tilde{\theta}^0) = 0$, $c_i(m_i, \tilde{\theta}^{m_i}) = 0$, and $c_i(m'_i, \tilde{\theta}^{m_i}) = \eta$ for all $m'_i \neq m_i$, where η is a strictly positive parameter that can be chosen arbitrarily close to 0.²⁶ Note that since M is finite, $\tilde{\Theta}$ is also finite. Next, we write $\Theta^* = \Theta \times \tilde{\Theta}$ for the extended set of states of nature. For a given state of nature $\theta^* = (\theta, \tilde{\theta}) \in \Theta^*$, the utility function of player i for a given message m_i and a given outcome a is $u_i(a, \theta) - c_i(m_i, \tilde{\theta})$. Recall that, in the benchmark model of the social planner, there is no cost of messages. In other terms, the model $\bar{T} = (\bar{T}, \bar{\kappa})$ is assumed to be such that, for any $\bar{t}_i \in \bar{T}_i$, $\text{marg}_{\tilde{\theta}} \bar{\kappa}(\bar{t}_i)[\tilde{\theta}^0] = 1$. All definitions given in previous sections are naturally extended to this setup where the space of states of nature is Θ^* . In particular, the definition of continuous implementation (see Definitions 2 and 3) can be maintained: a social choice function $f : \bar{T} \rightarrow A$ is continuously implementable²⁷ if there exists a mechanism \mathcal{M} such that, for each model \mathcal{T} with $\bar{T} \subset \mathcal{T}$, there is an equilibrium σ in $U(\mathcal{M}, \mathcal{T})$ that continuously implements f .

Let us briefly discuss our assumption on cost of messages. Technically, this construction is used to break ties: if a type is indifferent between several messages, we can slightly perturb his information so that this type has a unique best reply. This assumption is reminiscent of the richness assumption used in Weinstein and Yildiz (2007). Indeed, if one fixes a mechanism, the richness assumption would state that, for any player i and any message m_i , there exists a state of nature where m_i is strictly dominant for player i , which is much stronger than our assumption. However, as we already mentioned, in Weinstein and Yildiz (2007), the game is a primitive of the model, which is not the case in our context where the mechanism is chosen by the social planner. Consequently, our construction, while logically fine, is harder to assess. Nevertheless, in the literature on partial implementation, which is the main focus of our paper, only direct mechanisms (i.e., mechanisms where the message space M_i is equal to

²⁶Abusing notation, we omit the dependence of the space $\tilde{\Theta}$ (and of the costs function $c_i(\cdot, \cdot)$) with respect to the given mechanism \mathcal{M} . We also omit the dependence with respect to the parameter η .

²⁷For simplicity, we maintain the assumption that $\bar{T}_0 = \bar{T}$.

\bar{T}_i for each player i) are usually considered. Indeed, because of the revelation principle, this is without loss of generality for the characterization of incentive compatible SCFs. Direct mechanisms can naturally be considered as part of the primitives.

Full Characterization Result

In the present setting, we obtain a full characterization result for finite mechanisms stated in Theorem 4 below. Its proof is available in the Supplemental Material. The proof of the necessity part remains valid for arbitrary (countable) mechanisms, but the sufficiency part does not extend to infinite countable mechanisms.²⁸

THEOREM 4: *A social choice function f is continuously implementable by a finite mechanism if and only if it is rationalizable implementable by a finite mechanism.*²⁹

Theorem 4 allows us to derive corollaries when we relax the requirement of exact implementation to virtual implementation. Given that the virtual implementation literature uses finite mechanisms, it is natural to see how Theorem 4 can be applied there. In the sequel, we assume that A is a metric space and note d the associated metric. Given a social choice function f , for each $\delta > 0$, we write $B_\delta(f) = \{f' : \bar{T} \rightarrow A : d(f'(\bar{t}), f(\bar{t})) < \delta \text{ for all } \bar{t} \in \bar{T}\}$. Following Abreu and Matsushima (1992a), we say that a social choice function f is partially virtually implementable by finite mechanisms if, for each $\delta > 0$, there exists a social choice function $f' \in B_\delta(f)$ that is partially implementable by a finite mechanism (that may depend on δ). In the same way, we can extend the definition of continuous implementation.

DEFINITION 9: A social choice function f is virtually continuously implementable by finite mechanisms if, for all $\delta > 0$, there exists a social choice function $f' \in B_\delta(f)$ that is continuously implementable by a finite mechanism (that may depend on δ).

We also say that a social choice function f is virtually rationalizable implementable by finite mechanisms if, for all $\delta > 0$, there exists a social choice

²⁸We built an example of a countable infinite mechanism having the best-element property (see Kunimoto and Serrano (2011)) that implements a social choice function in rationalizable strategies but does not continuously implement. The example is available upon request.

²⁹For comparison with Theorems 2 and 3, note that if a SCF f is continuously implementable (under the assumptions made in Section 4) by a finite mechanism, then by Theorem 4, there is a finite mechanism that rationalizable implements f . A straightforward argument shows that f is partially implementable in pure Nash equilibrium. Hence, f is fully implementable in pure Nash equilibrium and so is Bayesian monotonic. Similarly, it can also be shown that f is rationalizable monotonic.

function $f' \in B_\delta(f)$ that is rationalizable implementable by a finite mechanism. Proposition 2 directly follows from Theorem 4.

PROPOSITION 2: *A social choice function f is virtually continuously implementable by finite mechanisms if and only if it is virtually rationalizable implementable by finite mechanisms.*

While the formulations in Proposition 2 and Theorem 4 are similar, their implications are quite different. Indeed, Abreu and Matsushima (1992a)³⁰ showed that, under complete information, and under some mild additional conditions, every social choice function is virtually rationalizable implementable with three or more players. Since mechanisms used in their paper are finite, we know by Proposition 2 that they also ensure continuous virtual implementation. Thus, under complete information, virtual continuous implementation is a lot less demanding than (exact) continuous implementation. However, we note that whether the characterization of virtual continuous implementation is weak in general is unclear. More precisely, in the Bayesian setting studied by Abreu and Matsushima (1992b), Bayesian incentive compatibility and a measurability condition are both necessary and sufficient for virtual implementation in rationalizable strategies, and the strength of the measurability condition is controversial. In particular, Serrano and Vohra (2001) showed that it may impose demanding restrictions.

APPENDIX

A. Proof of Lemma 1

To show the lemma, we build a model \hat{T} with $\bar{T} \subset \hat{T}$ and denote by σ a strict continuous equilibrium in $U(\mathcal{M}, \hat{T})$. Let $\bar{\sigma}$ be $\sigma|_{\bar{T}}$ and recall that $\bar{\sigma}$ must be a strict equilibrium in $U(\mathcal{M}, \bar{T})$. The model \hat{T} will be built in such a way that, for any best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, any $i \in \mathcal{I}$, $\bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i)$, there will exist a sequence of types $\{\hat{t}_i[n]\}_{n=0}^\infty$ in \hat{T}_i such that (i) $\hat{t}_i[n] \rightarrow_p \bar{t}_i$ and (ii) $\sigma_i(\hat{t}_i[n]) = m_i$ for all n . This will yield $(g \circ \sigma)(\hat{t}_i[n]) = g(m)$ for all n and so, by the continuity requirement, $g(m) = f(\bar{t})$. Thus, this will show that if a best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfies $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, then $g(\Sigma(\bar{t})) = \{f(\bar{t})\}$ for all $\bar{t} \in \bar{T}$, as desired.

We define the set \mathcal{E} by

$$\mathcal{E} := \bigcup_{q \in \mathbb{N}^*} \left\{ \frac{1}{q} \right\} \cup \{0\}.$$

³⁰In their setting, the (finite) set of outcomes is extended to the set of lotteries over outcomes and the natural metric is used over this set.

In the sequel, C denotes the countable³¹ set of strict Nash equilibria in $U(\mathcal{M}, \bar{T})$. We will sometimes abuse notation and write $\Sigma \subset \bar{\sigma}$ instead of $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$. Now we build the model $\hat{T} = (\hat{T}, \hat{\kappa})$ as follows. For each $\varepsilon \in \mathcal{E}$, equilibrium $\bar{\sigma} \in C$, best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, each integer ℓ , type $\bar{t}_i \in \bar{T}_i$, and message $m_i \in \Sigma_i(\bar{t}_i)$, we build inductively $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$ and set

$$\hat{T}_i = \bigcup_{\varepsilon \in \mathcal{E}} \bigcup_{\bar{\sigma} \in C} \bigcup_{\Sigma: \Sigma \subset \bar{\sigma}} \bigcup_{\ell=1}^{\infty} \bigcup_{\bar{t}_i \in \bar{T}_i} \bigcup_{m_i \in \Sigma_i(\bar{t}_i)} \hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i] \cup \bar{T}_i.$$

Note that \hat{T}_i is countable.³²

For each $\bar{\sigma} \in C$, we know that, for each best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, player i of type $\bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i)$, there exists $\pi_i^{m_i} \in \Delta(\Theta \times \bar{T}_{-i} \times M_{-i})$ such that

$$\begin{aligned} \text{marg}_{\Theta \times \bar{T}_{-i}} \pi_i^{m_i} &= \bar{\kappa}(\bar{t}_i), \\ \text{marg}_{\bar{T}_{-i} \times M_{-i}} \pi_i^{m_i}(\bar{t}_{-i}, m_{-i}) > 0 &\Rightarrow m_{-i} \in \Sigma_{-i}(\bar{t}_{-i}), \end{aligned}$$

and

$$m_i \in BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i^{m_i} | \mathcal{M}).$$

In addition, since each $\bar{\sigma} \in C$ is a strict equilibrium in $U(\mathcal{M}, \bar{T})$, we know that, for each player i and message $m_i \in \Sigma_i(\bar{t}_i) \subset \bar{\sigma}_i(\bar{T}_i)$, there exists $\pi_*^{m_i} \in \Delta(\Theta \times \bar{T}_{-i} \times M_{-i})$ such that $\{m_i\} = BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_*^{m_i} | \mathcal{M})$. We note that, for all $\varepsilon > 0$,

$$(3) \quad \{m_i\} = BR_i((1 - \varepsilon) \text{marg}_{\Theta \times M_{-i}} \pi_i^{m_i} + \varepsilon \text{marg}_{\Theta \times M_{-i}} \pi_*^{m_i} | \mathcal{M}).$$

For each equilibrium $\bar{\sigma} \in C$, best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, player $i \in \mathcal{I}$, and message $m_i \in \Sigma_i(\bar{t}_i)$, we let $t_i[\bar{\sigma}, \Sigma, m_i] \in \bar{T}_i$ be a type such that $\bar{\sigma}_i(t_i[\bar{\sigma}, \Sigma, m_i]) = m_i$, which is well defined since $m_i \in \Sigma_i(\bar{t}_i) \subset \bar{\sigma}_i(\bar{T}_i)$.

We now define $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$ inductively on ℓ . First, we let $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]$ be such that $\hat{\kappa}(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i])$ satisfies

$$(4) \quad \hat{\kappa}(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]) = (1 - \varepsilon) \pi_i^{m_i} \circ (\tau_{-i}^{\varepsilon, 1})^{-1} + \varepsilon \pi_*^{m_i} \circ (\tau_{-i}^{\varepsilon, 1})^{-1},$$

where $(\tau_{-i}^{\varepsilon, 1})^{-1}$ stands for the preimage of the function $\tau_{-i}^{\varepsilon, 1}: \Theta \times \bar{T}_{-i} \times M_{-i} \rightarrow \Theta \times \hat{T}_{-i}$, defined by $\tau_{-i}^{\varepsilon, 1}(\theta, \bar{t}_{-i}, m_{-i}) = (\theta, t_{-i}[\bar{\sigma}, \Sigma, m_{-i}])$. Note that $(\tau_{-i}^{\varepsilon, 1})^{-1}$ is a

³¹Recall that M is countable and \bar{T} is finite.

³²Indeed, the set of best-reply sets Σ satisfying $\Sigma \subset \bar{\sigma}$ is finite.

set-valued function and that $(\tau_{-i}^{\varepsilon,1})^{-1}(\theta, t_{-i}) = \{\emptyset\}$ whenever $t_{-i} \notin \bar{T}_{-i}$. Now, for each $\ell \geq 2$, define $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$ inductively by

$$\hat{k}(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]) = (1 - \varepsilon)\pi_i^{m_i} \circ (\tau_{-i}^{\varepsilon,\ell})^{-1} + \varepsilon\pi_i^{m_i} \circ (\tau_{-i}^{\varepsilon,\ell})^{-1},$$

where $(\tau_{-i}^{\varepsilon,\ell})^{-1}$ stands for the preimage of the function $\tau_{-i}^{\varepsilon,\ell}: \Theta \times \bar{T}_{-i} \times M_{-i} \rightarrow \Theta \times \hat{T}_{-i}$, defined by $\tau_{-i}^{\varepsilon,\ell}(\theta, \bar{t}_{-i}, m_{-i}) = (\theta, \hat{t}_{-i}[\varepsilon, \bar{\sigma}, \Sigma, \ell - 1, \bar{t}_{-i}, m_{-i}])$.

CLAIM 1: For each $\bar{\sigma} \in C$, best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, $i, \bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i): \hat{t}_i[\hat{\varepsilon}(\ell), \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i] \rightarrow_P \bar{t}_i$ as $\ell \rightarrow \infty$ for some mapping $\hat{\varepsilon}$ taking values in $\mathcal{E} \setminus \{0\}$.

To prove this claim, we will use the following well known lemma.

LEMMA 2—Mertens and Zamir (1985) and Brandenburger and Dekel (1993): Let $\mathcal{T} = (T, \kappa)$ be any model such that $\Theta \times T$ is complete and separable and $\kappa(\cdot)$ is a continuous function of t_i for each player i . Then the mapping $h: T \rightarrow T^*$ is continuous.

PROOF OF CLAIM 1: In the sequel, we will denote \bar{h} to be the mapping that projects \bar{T} into T^* and, in a similar way, denote \hat{h} to be the mapping from \hat{T} to T^* . By Lemma 2, these mappings are continuous.³³

In the sequel, we fix $\bar{\sigma}$ and a best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$. Since for all $\ell' \geq 1$ and all $\bar{t}_i, m_i: \hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i] \rightarrow \hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]$ as $\varepsilon \rightarrow 0$, by continuity, for all $\ell' \geq 1$, all $\ell \geq 1$, and all $\bar{t}_i, m_i: \hat{h}_i^\ell(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) \rightarrow \hat{h}_i^\ell(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i])$ as $\varepsilon \rightarrow 0$.

Let us now show inductively on ℓ that, for all $\ell \geq 1$ and $\ell' \geq \ell: \hat{h}_i^\ell(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) = \bar{h}_i^\ell(\bar{t}_i)$ for each i, \bar{t}_i and $m_i \in \Sigma_i(\bar{t}_i)$. First notice that the first-order beliefs are equal, that is, for all $\ell' \geq 1$,

$$\begin{aligned} \hat{h}_i^1(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) &= \text{marg}_\Theta \hat{k}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) \\ &= \text{marg}_\Theta \pi_i^{m_i} \circ (\tau_{-i}^{0,\ell'})^{-1} \\ &= \text{marg}_\Theta \pi_i^{m_i} = \text{marg}_\Theta \bar{\kappa}(\bar{t}_i) = \bar{h}_i^1(\bar{t}_i), \end{aligned}$$

³³A type in \hat{T}_i is either in \bar{T}_i —which is endowed with the discrete topology, say $\tau_{\bar{T}_i}$ —or in $\hat{T}_i \setminus \bar{T}_i$. Any point in $\hat{T}_i \setminus \bar{T}_i$ is identified with an element of the set $\mathcal{E} \times C \times \{\Sigma: \bar{T} \rightarrow 2^M \setminus \{\emptyset\}\} \Sigma(\bar{T}) \in \mathcal{S} \times \mathbb{N} \times \bar{T}_i \times M_i$, where \mathcal{S} is defined as the set of profiles of messages M' such that $M' \subset \bar{\sigma}(\bar{T})$ for some $\bar{\sigma}$ in C . The sets $C, \{\Sigma: \bar{T} \rightarrow M \mid \Sigma(\bar{T}) \in \mathcal{S}\}, \mathbb{N}, \bar{T}_i$, and M_i are all (countable and) endowed with the discrete topology, while \mathcal{E} is endowed with the usual topology on \mathbb{R} induced on \mathcal{E} . Finally, $\mathcal{E} \times C \times \{\Sigma: \bar{T} \rightarrow M \mid \Sigma(\bar{T}) \in \mathcal{S}\} \times \mathbb{N} \times \bar{T}_i \times M_i$ is endowed with the product topology; call this topology $\tau_{\hat{T}_i \setminus \bar{T}_i}$. The topology over \hat{T}_i is the coarsest topology that contains $\tau_{\bar{T}_i} \cup \tau_{\hat{T}_i \setminus \bar{T}_i}$. It can easily be checked that, under such a topology, \hat{T} satisfies the conditions of Lemma 2.

where the third and the fourth equalities are by definition of $\tau_{-i}^{0,\ell'}$ and $\pi_{\bar{t}_i}^{m_i}$, respectively. Now fix some $\ell \geq 2$ and let L be the set of all belief profiles of players other than i at order $\ell - 1$. Toward an induction, assume that, for all $\ell' \geq \ell - 1$: $\hat{h}_j^{\ell-1}(\hat{t}_j[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_j, m_j]) = \bar{h}_j^{\ell-1}(\bar{t}_j)$ for each j , $\bar{t}_j \in \bar{T}_j$ and $m_j \in \Sigma_j(\bar{t}_j)$. Then for all $\ell' \geq \ell$: $\text{proj}_{\Theta \times L} \circ (\text{id}_{\Theta} \times \hat{h}_{-i}) \circ \tau_{-i}^{0,\ell'} = \overline{\text{proj}}_{\Theta \times L} \circ (\text{id}_{\Theta} \times \bar{h}_{-i} \times \text{id}_{M_{-i}})$, where id_{Θ} (resp. $\text{id}_{M_{-i}}$) is the identity mapping from Θ to Θ (resp. from M_{-i} to M_{-i}), while $\text{proj}_{\Theta \times L}$ (resp. $\overline{\text{proj}}_{\Theta \times L}$) is the projection mapping from $\Theta \times T_{-i}^*$ to $\Theta \times L$ (resp. from $\Theta \times T_{-i}^* \times M_{-i}$ to $\Theta \times L$); hence, for all $\ell' \geq \ell$,

$$\begin{aligned} & \text{marg}_{\Theta \times L} \hat{\kappa}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) \circ (\text{id}_{\Theta} \times \hat{h}_{-i})^{-1} \\ &= \text{marg}_{\Theta \times L} \pi_{\bar{t}_i}^{m_i} \circ (\tau_{-i}^{0,\ell'})^{-1} \circ (\text{id}_{\Theta} \times \hat{h}_{-i})^{-1} \\ &= \pi_{\bar{t}_i}^{m_i} \circ (\tau_{-i}^{0,\ell'})^{-1} \circ (\text{id}_{\Theta} \times \hat{h}_{-i})^{-1} \circ (\text{proj}_{\Theta \times L})^{-1} \\ &= \pi_{\bar{t}_i}^{m_i} \circ (\text{id}_{\Theta} \times \bar{h}_{-i} \times \text{id}_{M_{-i}})^{-1} \circ (\overline{\text{proj}}_{\Theta \times L})^{-1} \\ &= \text{marg}_{\Theta \times L} \pi_{\bar{t}_i}^{m_i} \circ (\text{id}_{\Theta} \times \bar{h}_{-i} \times \text{id}_{M_{-i}})^{-1} \\ &= \text{marg}_{\Theta \times L} \bar{\kappa}(\bar{t}_i) \circ (\text{id}_{\Theta} \times \bar{h}_{-i})^{-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \hat{h}_i^{\ell}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) \\ &= \delta_{\hat{t}_i^{\ell-1}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i])} \\ & \quad \times \text{marg}_{\Theta \times L} \hat{\kappa}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) \circ (\text{id}_{\Theta} \times \hat{h}_{-i})^{-1} \\ &= \delta_{\bar{t}_i^{\ell-1}(\bar{t}_i)} \times \text{marg}_{\Theta \times L} \bar{\kappa}(\bar{t}_i) \circ (\text{id}_{\Theta} \times \bar{h}_{-i})^{-1} = \bar{h}_i^{\ell}(\bar{t}_i), \end{aligned}$$

showing that $\hat{h}_i^{\ell}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) = \bar{h}_i^{\ell}(\bar{t}_i)$. Thus, we have proved that, for all $\ell \geq 1$ and $\ell' \geq \ell$: $\hat{h}_i^{\ell}(\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]) = \bar{h}_i^{\ell}(\bar{t}_i)$ for each i , \bar{t}_i and $m_i \in \Sigma_i(\bar{t}_i)$, which means that $\hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i] \rightarrow_P \bar{t}_i$ as $\ell' \rightarrow \infty$ for each i , \bar{t}_i and $m_i \in \Sigma_i(\bar{t}_i)$. In addition, we know that, for all $\ell' \geq 1$, all i, \bar{t}_i, m_i : $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i] \rightarrow_P \hat{t}_i[0, \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i]$ as $\varepsilon \rightarrow 0$. Since T^* is a metrizable space, $\hat{t}_i[\hat{\varepsilon}(\ell'), \bar{\sigma}, \Sigma, \ell', \bar{t}_i, m_i] \rightarrow_P \bar{t}_i$ as $\ell' \rightarrow \infty$ for some function $\hat{\varepsilon}: \mathbb{N}^* \rightarrow \mathcal{E} \setminus \{0\}$ satisfying $\lim_{\ell' \rightarrow \infty} \hat{\varepsilon}(\ell') = 0$. Q.E.D.

CLAIM 2: For each $\bar{\sigma} \in C$, each best-reply set $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ in $U(\mathcal{M}, \bar{T})$ satisfying $\Sigma(\bar{T}) \subset \bar{\sigma}(\bar{T})$, $\varepsilon \in \mathcal{E} \setminus \{0\}$, $\ell, i, \bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i)$, we have $\sigma_i(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]) = m_i$ for any equilibrium σ of $U(\mathcal{M}, \hat{T})$ satisfying $\sigma_{|\bar{T}} = \bar{\sigma}$.

PROOF: Fix an equilibrium σ of $U(\mathcal{M}, \hat{T})$ satisfying $\sigma_{\bar{T}} = \bar{\sigma}$. We will show by induction on ℓ that, for all $\varepsilon > 0$ and $\ell \geq 1$: $\sigma_i(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]) = m_i$ for all player i , type $\bar{t}_i \in \bar{T}_i$, and all messages $m_i \in \Sigma_i(\bar{t}_i)$.

Recall that, by construction, for all $\bar{t}_i \in \bar{T}_i$ and $m_i \in \Sigma_i(\bar{t}_i)$: $t_i[\bar{\sigma}, \Sigma, m_i] \in \bar{T}_i$ is the type such that $\sigma_i(t_i[\bar{\sigma}, \Sigma, m_i]) = \bar{\sigma}_i(t_i[\bar{\sigma}, \Sigma, m_i]) = m_i$. First, fix $\varepsilon > 0$, $\bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i)$, and let us prove that $\sigma_i(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]) = m_i$. For each $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]$, define the belief

$$\pi_i^{\varepsilon, 1} = \hat{\kappa}(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]) \circ \gamma^{-1} \in \Delta(\Theta \times \hat{T}_{-i} \times M_{-i}),$$

where $\gamma: (\theta, t_{-i}[\bar{\sigma}, \Sigma, m_{-i}]) \mapsto (\theta, t_{-i}[\bar{\sigma}, \Sigma, m_{-i}], m_{-i})$. Note that by construction, $\pi_i^{\varepsilon, 1}$ is the belief of type $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]$ on $\Theta \times \hat{T}_{-i} \times M_{-i}$ when he believes that m_{-i} is played at each $(\theta, t_{-i}[\bar{\sigma}, \Sigma, m_{-i}])$. Hence, for each $\varepsilon \geq 0$, $\pi_i^{\varepsilon, 1}$ corresponds to beliefs of type $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]$ when the equilibrium σ is played. Now, by Equation (4), the belief $\pi_i^{\varepsilon, 1}$ of type $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]$ satisfies

$$\begin{aligned} \text{marg}_{\Theta \times M_{-i}} \pi_i^{\varepsilon, 1} &= (1 - \varepsilon) \text{marg}_{\Theta \times M_{-i}} \pi_i^{m_i} \circ (\tau_{-i}^{\varepsilon, 1})^{-1} \circ (\gamma)^{-1} \\ &\quad + \varepsilon \text{marg}_{\Theta \times M_{-i}} \pi_i^* \circ (\tau_{-i}^{\varepsilon, 1})^{-1} \circ (\gamma)^{-1} \\ &= (1 - \varepsilon) \text{marg}_{\Theta \times M_{-i}} \pi_i^{m_i} + \varepsilon \text{marg}_{\Theta \times M_{-i}} \pi_i^*. \end{aligned}$$

For $\varepsilon > 0$, by Equation (3) and the construction of $\pi_i^{\varepsilon, 1}$, $\{m_i\} = BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i^{\varepsilon, 1} | \mathcal{M})$ and $\sigma_i(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, 1, \bar{t}_i, m_i]) = m_i$.

Now, for each $\ell \geq 2$, proceed by induction and assume that $\sigma_{-i}(\hat{t}_{-i}[\varepsilon, \bar{\sigma}, \Sigma, \ell - 1, \bar{t}_{-i}, m_{-i}]) = m_{-i}$ for any $\bar{t}_{-i} \in \bar{T}_{-i}$, $m_{-i} \in \Sigma_{-i}(\bar{t}_{-i})$, and $\varepsilon > 0$. Fix $\varepsilon > 0$, i , $\bar{t}_i \in \bar{T}_i$, and $m_i \in \Sigma_i(\bar{t}_i)$. For each $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$, define the belief

$$\pi_i^{\varepsilon, \ell} = \hat{\kappa}(\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]) \circ \gamma_\ell^{-1} \in \Delta(\Theta \times \hat{T}_{-i} \times M_{-i}),$$

where $\gamma_\ell: (\theta, \hat{t}_{-i}[\varepsilon, \bar{\sigma}, \Sigma, \ell - 1, \bar{t}_{-i}, m_{-i}]) \mapsto (\theta, \hat{t}_{-i}[\varepsilon, \bar{\sigma}, \Sigma, \ell - 1, \bar{t}_{-i}, m_{-i}], m_{-i})$.

Note that, by construction, $\pi_i^{\varepsilon, \ell}$ is the belief of type $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$ on $\Theta \times \hat{T}_{-i} \times M_{-i}$ when he believes that m_{-i} is played at each $(\theta, \hat{t}_{-i}[\varepsilon, \bar{\sigma}, \Sigma, \ell - 1, \bar{t}_{-i}, m_{-i}])$. Hence, by the induction hypothesis, for each $\varepsilon \geq 0$, $\pi_i^{\varepsilon, \ell}$ corresponds to beliefs of type $\hat{t}_i[\varepsilon, \bar{\sigma}, \Sigma, \ell, \bar{t}_i, m_i]$ when the equilibrium σ is played. The end of the proof mimics the case $\ell = 1$. Q.E.D.

B. Proof of Theorem 3

Suppose f is strictly continuously implementable on $(\bar{T}, \bar{\kappa})$ and fix a mechanism \mathcal{M} that strictly continuously implements f . Take the strict equilibrium

$\bar{\sigma}: \bar{T} \rightarrow M$ obtained in Lemma 1 and pick any unacceptable deception $\beta: \bar{T} \rightarrow \bar{T}$. Note that there must exist some \bar{t} and \bar{t}' in \bar{T} such that $\bar{t}' \in \beta(\bar{t})$ with $f(\bar{t}') \neq f(\bar{t})$. Given that $g \circ \bar{\sigma} = f$, this yields $(g \circ \bar{\sigma} \circ \beta)(\bar{t}) \neq \{f(\bar{t})\}$. Thus, Lemma 1, together with the fact that the correspondence $\bar{\sigma} \circ \beta = \times_{i \in \mathcal{I}} (\bar{\sigma}_i \circ \beta_i)$ satisfies $(\bar{\sigma} \circ \beta)(\bar{T}) \subset \bar{\sigma}(\bar{T})$, implies that $\bar{\sigma} \circ \beta$ cannot be a best-reply set. In other words, there exist $i, t_i, t'_i \in \beta_i(t_i)$ and $m_i^* = \bar{\sigma}_i(t'_i) \in \bar{\sigma}_i \circ \beta_i(t_i)$ such that, for all $v_i \in \Delta(\Theta \times \bar{T}_{-i} \times M_{-i})$ satisfying (i) $v_i(\theta, t_{-i}, m_{-i}) > 0 \Rightarrow m_j \in (\bar{\sigma}_j \circ \beta_j)(t_j)$ for all $j \neq i$ and (ii) $\sum_{m_{-i}} v_i(\theta, t_{-i}, m_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) , there exists $\tilde{m}_i \neq m_i^*$ such that

$$(5) \quad \sum_{\theta, m_{-i}} \text{marg}_{\Theta \times M_{-i}} v_i(\theta, m_{-i}) u_i(g(\tilde{m}_i, m_{-i}), \theta) > \sum_{\theta, m_{-i}} \text{marg}_{\Theta \times M_{-i}} v_i(\theta, m_{-i}) u_i(g(m_i^*, m_{-i}), \theta).$$

Now, let $\psi_i \in \Delta(\Theta \times \bar{T}_{-i} \times \bar{T}_{-i})$ be such that

- (a) $\psi_i(\theta, t_{-i}, t'_{-i}) > 0 \Rightarrow t'_j \in \beta_j(t_j)$ for each $j \neq i$, and,
- (b) $\sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) .

Set $v_i(\theta, t_{-i}, m_{-i}) := \sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) \bar{\sigma}_{-i}(m_{-i} | t'_{-i})$. Note that v_i satisfies (i) and (ii) above. Thus, there exists $\tilde{m}_i \neq m_i^*$ such that Equation (5) is satisfied. Equivalently,

$$\sum_{\theta, m_{-i}} \sum_{t_{-i}} \sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) \bar{\sigma}_{-i}(m_{-i} | t'_{-i}) u_i(g(\tilde{m}_i, m_{-i}), \theta) > \sum_{\theta, m_{-i}} \sum_{t_{-i}} \sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) \bar{\sigma}_{-i}(m_{-i} | t'_{-i}) u_i(g(m_i^*, m_{-i}), \theta),$$

and so

$$\sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(g(\tilde{m}_i, \bar{\sigma}_{-i}(t'_{-i})), \theta) > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(g(m_i^*, \bar{\sigma}_{-i}(t'_{-i})), \theta),$$

which yields

$$\sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(y^*(t'_{-i}), \theta) > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(f(t'_i, t'_{-i}), \theta),$$

where $y^*(\cdot) := g(\tilde{m}_i, \bar{\sigma}_{-i}(\cdot))$ and using the fact that $m_i^* = \bar{\sigma}_i(t'_i)$ and $g \circ \bar{\sigma} = f$. Now to complete the proof, we have to show that, for each $\tilde{t}_i \in \tilde{T}_i$,

$$\sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(f(\tilde{t}_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(y^*(t_{-i}), \theta).$$

This is true because, for each $\tilde{t}_i \in \tilde{T}_i$,

$$\begin{aligned} & \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(y^*(t_{-i}), \theta) \\ &= \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(g(\tilde{m}_i, \bar{\sigma}_{-i}(t_{-i})), \theta) \\ &\leq \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(g(\bar{\sigma}_i(\tilde{t}_i), \bar{\sigma}_{-i}(t_{-i})), \theta) \\ &= \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(f(\tilde{t}_i, t_{-i}), \theta), \end{aligned}$$

where the inequality above comes from the fact that $\bar{\sigma}$ is an equilibrium in $U(\mathcal{M}, \tilde{T})$, while the last equality is obtained using the fact that $g \circ \bar{\sigma} = f$. Note that since $\bar{\sigma}$ is a strict equilibrium and $\tilde{m}_i \neq m_i^* = \bar{\sigma}_i(t'_i)$, the above inequality is strict when $\tilde{t}_i = t'_i$. This proves that f is strict interim rationalizable monotonic.

C. Proof of Proposition 1

Before proving this proposition, let us make the following simple remark.

REMARK 1: It is easily checked that Assumption 1 yields that, for all i , there exists $\hat{y}_i \in \Delta(A)$ such that the following holds. For any $\phi_i \in \Delta(\Theta)$, there exists $y \in \Delta(A)$ such that

$$\sum_{\theta} \phi_i(\theta) u_i(y, \theta) > \sum_{\theta} \phi_i(\theta) u_i(\hat{y}_i, \theta).$$

Note that here \hat{y}_i and y are lotteries (and not mappings).

Before proving the proposition, we will need the following lemma, which is an extension to the Bayesian setting of Lemma 1 in Bergemann and Morris (2011).³⁴

³⁴It is easily checked that as in Bergemann and Morris (2011), we actually prove here that f is semi-strict incentive compatible, that is, a strengthening of incentive compatibility.

LEMMA 3: *If f is interim rationalizable monotonic, then f is incentive compatible; that is, for all i and $t_i \in \bar{T}_i$,*

$$\sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t'_i, t_{-i}), \theta)$$

for all t'_i .

PROOF: Assume f is interim rationalizable monotonic. Fix any player i and t_i, t'_i . If $f(t_i, t_{-i}) = f(t'_i, t_{-i})$ for all t_{-i} , we trivially have

$$\sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t'_i, t_{-i}), \theta).$$

Now assume that $f(t_i, \hat{t}_{-i}) \neq f(t'_i, \hat{t}_{-i})$ for some \hat{t}_{-i} . To complete the proof, we will show that

$$\sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t_i, t_{-i}), \theta) > \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t'_i, t_{-i}), \theta).$$

Build the deception β as follows: for each $j \neq i$, $\beta_j(t_j) = \{t_j\}$ and

$$\beta_i(\tilde{t}_i) = \begin{cases} \{t_i, t'_i\}, & \text{if } \tilde{t}_i = t_i, \\ \{\tilde{t}_i\}, & \text{otherwise.} \end{cases}$$

Note that β is an unacceptable deception; indeed, $(t'_i, \hat{t}_{-i}) \in \beta(t_i, \hat{t}_{-i})$ while $f(t'_i, \hat{t}_{-i}) \neq f(t_i, \hat{t}_{-i})$. It is easily checked that, by interim rationalizable monotonicity, the following condition holds for t_i and $t'_i \in \beta_i(t_i)$. For every $\psi_i \in \Delta(\Theta \times \bar{T}_{-i} \times \bar{T}_{-i})$ such that

- (i) $\psi_i(\theta, t_{-i}, t'_{-i}) > 0 \Rightarrow t'_j \in \beta(t_j)$ for each $j \neq i$,
- (ii) $\sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) ,

there exists $y^* : \bar{T}_{-i} \rightarrow A$ such that

$$\sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(y^*(t'_{-i}), \theta) > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(f(t'_i, t'_{-i}), \theta)$$

and

$$\sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(f(\tilde{t}_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(\tilde{t}_i)[\theta, t_{-i}] u_i(y^*(t_{-i}), \theta)$$

for each $\tilde{t}_i \in \bar{T}_i$.

Now to complete the proof, build $\psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]\delta_{\{t_{-i}\}}(t'_{-i})$, where $\delta_{\{t_{-i}\}}$ is the Dirac measure on $\{t_{-i}\}$; it is clear that (i) $\psi_i(\theta, t_{-i}, t'_{-i}) > 0 \Rightarrow t'_j \in \beta_j(t_j)$ for each $j \neq i$, and that (ii) $\sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) . Hence, there must exist $y^* : \bar{T}_{-i} \rightarrow \Delta(A)$ such that

$$(6) \quad \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(y^*(t_{-i}), \theta) > \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(f(t'_{-i}, t_{-i}), \theta)$$

and

$$(7) \quad \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(f(t_i, t_{-i}), \theta) \geq \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(y^*(t_{-i}), \theta).$$

Equations (6) and (7) yield

$$\sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(f(t_i, t_{-i}), \theta) > \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}]u_i(f(t'_{-i}, t_{-i}), \theta),$$

the desired result. Q.E.D.

We are now in a position to prove Proposition 1.

PROOF OF PROPOSITION 1: The proof is by construction of a canonical mechanism \mathcal{M} that implements, in rationalizable strategies, f when $I \geq 3$. Each agent i sends a message $m_i = (m_i^1, m_i^2, m_i^3, m_i^4)$, where $m_i^1 \in \bar{T}_i$, $m_i^2 \in \mathbb{N}$, $m_i^3(\cdot) \in Y_i$, and $m_i^4 \in \Delta(A)$. The outcome $g(m)$ is determined by the following rules:

Rule 1: If $m_i^2 = 1$ for all i , pick $f(m^1)$.

Rule 2: If there exists j such that $m_i^2 = 1$ for all $i \neq j$ and $m_j^2 > 1$, then pick $m_j^3(m_{-j}^1)$ with probability $1 - \frac{1}{m_j^2+1}$ and $\bar{y}_j(m_{-j}^1)$ with probability $\frac{1}{m_j^2+1}$, where $\bar{y}_j(\cdot)$ is as defined in Assumption 1.

Rule 3: In all other cases, for each i , with probability $\frac{1}{I}(1 - \frac{1}{m_i^2+1})$ pick m_i^4 and with probability $\frac{1}{I}(\frac{1}{m_i^2+1})$ pick \hat{y}_i , where \hat{y}_i is defined in Remark 1.

The following three claims complete the proof.

CLAIM 3: For each player i and type t_i , $R_i(t_i|\mathcal{M}, \bar{T})$ is nonempty.

PROOF: We will check that, for any m_i^3 and m_i^4 , $(t_i, 1, m_i^3, m_i^4) \in R_i(t_i|\mathcal{M}, \bar{T})$. To be more precise, for each t_i , we define $\Sigma_i(t_i) = \{m_i \in M_i | m_i^1 = t_i \text{ and } m_i^2 = 1\}$ and check that $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ is a best-reply set in $U(\mathcal{M}, \bar{T})$. Fix i , t_i , and any selection $m_{-i}(\cdot)$ of $\Sigma_{-i}(\cdot)$. For $\pi_i \in \Delta(\Theta \times \bar{T}_{-i} \times M_{-i})$ defined by $\pi_i(\theta, t_{-i}, m_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]\delta_{m_{-i}(t_{-i})}(m_{-i})$, we clearly have $\text{marg}_{\Theta \times \bar{T}_{-i}} \pi_i = \bar{\kappa}(t_i)$

and $\pi_i(\theta, t_{-i}, m_{-i}) > 0 \implies m_{-i} \in \Sigma_{-i}(t_{-i})$. In addition, it is easily checked that, for all $m_i \in \Sigma_i(t_i)$,

$$\begin{aligned} & \sum_{\theta, m_{-i}} \text{marg}_{\Theta \times M_{-i}} \pi_i(\theta, m_{-i}) u_i(g(m_i, m_{-i}), \theta) \\ &= \sum_{\theta, t_{-i}} \bar{\kappa}(t_i)[\theta, t_{-i}] u_i(f(t_i, t_{-i}), \theta). \end{aligned}$$

Since, by Lemma 3, f is incentive compatible and since $m_i^3(\cdot) \in Y_i$, by construction of the mechanism, we have $m_i \in BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i | \mathcal{M})$ for all $m_i \in \Sigma_i(t_i)$, proving that $\Sigma = (\Sigma_i)_{i \in \mathcal{I}}$ is indeed a best-reply set in $U(\mathcal{M}, \bar{T})$. *Q.E.D.*

So far we have proved that the set of rationalizable strategy profiles in $U(\mathcal{M}, \bar{T})$ is nonempty. It remains to show that all rationalizable strategy profiles yield the desired outcome at any profile of types. This is established via the following two claims.

CLAIM 4: *For any player i and type t_i , it is never a best reply to send a message with $m_i^2 > 1$.*

PROOF: Fix any player i of type t_i and pick any message $m_i = (m_i^1, m_i^2, m_i^3, m_i^4)$ such that $m_i^2 > 1$. Suppose that t_i has belief $\lambda_i \in \Delta(\Theta \times T_{-i} \times M_{-i})$. Let us consider $\pi_i := \text{marg}_{\Theta \times M_{-i}} \lambda_i$. We can partition the messages of the other agents as follows. For each t_{-i} ,

$$M_{-i}^*(t_{-i}) = \{m_{-i} | m_j^2 = 1 \text{ for all } j \neq i \text{ and } m_{-i}^1 = t_{-i}\}$$

and

$$\hat{M}_{-i} = \{m_{-i} | m_j^2 > 1 \text{ for some } j \neq i\}.$$

Clearly, for $m_{-i} \in M_{-i}^*(t_{-i})$ for some t_{-i} , (m_i, m_{-i}) falls into Rule 2 of the mechanism, while for $m_{-i} \in \hat{M}_{-i}$, (m_i, m_{-i}) falls into Rule 3. Now, if for some t'_{-i} , we have $\sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) > 0$, set $\psi_i \in \Delta(\Theta \times \bar{T}_{-i})$ such that, for all (θ, t'_{-i}) , $\psi_i(\theta, t'_{-i}) := \alpha \sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i})$, where the constant α is chosen so that $\sum_{\theta, t'_{-i}} \psi_i(\theta, t'_{-i}) = 1$. By Assumption 1, there exists $y(\cdot) \in Y_i$ such that

$$\sum_{\theta, t'_{-i}} \psi_i(\theta, t'_{-i}) u_i(y(t'_{-i}), \theta) > \sum_{\theta, t'_{-i}} \psi_i(\theta, t'_{-i}) u_i(\bar{y}_i(t'_{-i}), \theta).$$

Hence, if m_i is a best response to π_i (and $\sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) > 0$ for some t'_{-i}), by construction of the mechanism, we must have

$$(8) \quad \sum_{t'_{-i}} \sum_{\theta} \sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) u_i(m_i^3(t'_{-i}), \theta) > \sum_{t'_{-i}} \sum_{\theta} \sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) u_i(\hat{y}_i(t'_{-i}), \theta).$$

Now, if $\sum_{m_{-i} \in \hat{M}_{-i}} \pi_i(\theta, m_{-i}) > 0$, set $\phi_i(\theta) := \alpha \sum_{m_{-i} \in \hat{M}_{-i}} \pi_i(\theta, m_{-i})$, where here again, the constant α is chosen so that $\sum_{\theta} \phi_i(\theta) = 1$. By Remark 1, there exists $y \in \Delta(A)$ such that

$$\sum_{\theta} \phi_i(\theta) u_i(y, \theta) > \sum_{\theta} \phi_i(\theta) u_i(\hat{y}_i, \theta).$$

Hence, if m_i is a best response to π_i (and $\sum_{m_{-i} \in \hat{M}_{-i}} \pi_i(\theta, m_{-i}) > 0$), by construction of the mechanism, we must have

$$(9) \quad \sum_{\theta} \sum_{m_{-i} \in \hat{M}_{-i}} \pi_i(\theta, m_{-i}) u_i(m_i^4, \theta) > \sum_{\theta} \sum_{m_{-i} \in \hat{M}_{-i}} \pi_i(\theta, m_{-i}) u_i(\hat{y}_i, \theta).$$

Now, Equations (8) and (9) together show that, if $m_i = (m_i^1, m_i^2, m_i^3, m_i^4)$ were a best response, then since with probability 1, either Rule 2 or 3 is sparked (i.e., either $\sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) > 0$ for some t'_{-i} or $\sum_{m_{-i} \in M_{-i}^*(t'_{-i})} \pi_i(\theta, m_{-i}) > 0$), $(m_i^1, m_i^2 + 1, m_i^3, m_i^4)$ would be an even better response, a contradiction. *Q.E.D.*

To complete the proof, it is enough to prove the following claim.

CLAIM 5: Let $\beta_i(t_i) := \{t'_i : (t'_i, 1, m_i^3, m_i^4) \in R_i(t_i | \mathcal{M}, \bar{T}) \text{ for some } (m_i^3, m_i^4)\}$. The deception $\beta := \times_i \beta_i$ is acceptable.

PROOF: By contradiction. Assume that β is an unacceptable deception. By interim rationalizable monotonicity, we know that there exist $i, t_i, t'_i \in \beta_i(t_i)$ such that, for every $\psi_i \in \Delta(\Theta \times \bar{T}_{-i} \times \bar{T}_{-i})$ satisfying

- (a) $\psi_i(\theta, t_{-i}, t'_{-i}) > 0 \Rightarrow t'_j \in \beta_j(t_j)$ for each $j \neq i$,
- (b) $\sum_{t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) ,

there exists $y^* \in Y_i$ such that

$$\sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(y^*(t'_{-i}), \theta) > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(f(t'_i, t'_{-i}), \theta).$$

Observe first that by construction of β_i , there exists $m_i \in R_i(t_i | \mathcal{M}, \bar{T})$ such that $m_i = (t'_i, 1, m_i^3, m_i^4)$ for some (m_i^3, m_i^4) . Now pick $\pi_i \in \Delta(\Theta \times \bar{T}_{-i} \times M_{-i})$

such that (i) $\pi_i(\theta, t_{-i}, m_{-i}) > 0 \Rightarrow m_j \in R_j(t_j|\mathcal{M}, \bar{T})$ for each $j \neq i$, (ii) $\sum_{m_{-i}} \pi_i(\theta, t_{-i}, m_{-i}) = \bar{\kappa}(t_i)[\theta, t_{-i}]$ for each (θ, t_{-i}) , and (iii) $m_i \in BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i|\mathcal{M})$. In addition, for all θ, t_{-i} and t'_{-i} , set

$$\psi_i(\theta, t_{-i}, t'_{-i}) := \sum_{m_{-i} \text{ s.t. } m_{-i}^1 = t'_{-i}} \pi_i(\theta, t_{-i}, m_{-i})$$

and observe that ψ_i satisfies (a) and (b) above. Hence, there exists $y^* \in Y_i$ such that

$$\begin{aligned} & \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(y^*(t'_{-i}), \theta) \\ & > \sum_{\theta, t_{-i}, t'_{-i}} \psi_i(\theta, t_{-i}, t'_{-i}) u_i(f(t'_{-i}, t'_{-i}), \theta), \end{aligned}$$

which yields

$$\begin{aligned} & \sum_{\theta, m_{-i}} \text{marg}_{\Theta \times M_{-i}} \pi_i(\theta, m_{-i}) u_i(y^*(m_{-i}^1), \theta) \\ & > \sum_{\theta, m_{-i}} \text{marg}_{\Theta \times M_{-i}} \pi_i(\theta, m_{-i}) u_i(f(t'_i, m_{-i}^1), \theta). \end{aligned}$$

Thus, by construction of the mechanism, against π_i , message $(t'_i, \ell, \tilde{m}_i^3, m_i^4)$, where $\tilde{m}_i^3(\cdot) = y^*(\cdot)$ is strictly better than $m_i = (t'_i, 1, m_i^3, m_i^4)$ for ℓ large enough, which contradicts $m_i \in BR_i(\text{marg}_{\Theta \times M_{-i}} \pi_i|\mathcal{M})$. Q.E.D.

Q.E.D.

REFERENCES

ABREU, D., AND H. MATSUSHIMA (1992a): “Virtual Implementation in Iteratively Undominated Strategies: Complete Information,” *Econometrica*, 60, 993–1008. [1624,1625]
 ——— (1992b): “Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information,” Unpublished Manuscript, Princeton University. Available at http://www.princeton.edu/~dabreu/index_files/virtual%20implementation-incomplete.pdf. [1625]
 BERGEMANN, D., AND S. MORRIS (2005): “Robust Mechanism Design,” *Econometrica*, 73, 1771–1813. [1605]
 ——— (2011): “Robust Implementation in General Mechanisms,” *Games and Economic Behavior*, 71, 261–281. [1606,1619,1631]
 BERGEMANN, D., S. MORRIS, AND O. TERCIEUX (2011): “Rationalizable Implementation,” *Journal of Economic Theory*, 146, 1253–1274. [1606,1613,1621]
 BRANDENBURGER, A., AND E. DEKEL (1993): “Hierarchies of Beliefs and Common Knowledge,” *Journal of Economic Theory*, 59, 189–198. [1609,1627]
 CABRALES, A., AND R. SERRANO (2011): “Implementation in Adaptive Better-Response Dynamics: Towards a General Theory of Bounded Rationality in Mechanisms,” *Games and Economic Behavior*, 73, 360–374. [1621]

- CARLSSON, H., AND E. VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018. [1607]
- CHEN, Y. (2012): "A Structure Theorem for Rationalizability in the Normal Form of Dynamic Games," *Games and Economic Behavior*, 75, 587–597. [1607]
- DEKEL, E., D. FUDENBERG, AND S. MORRIS (2006): "Topologies on Types," *Theoretical Economics*, 1, 275–309. [1611,1615,1622]
- (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15–40. [1615]
- DENECKERE, R., AND S. SEVERINOV (2008): "Mechanism Design With Partial State Verifiability," *Games and Economic Behavior*, 64, 487–513. [1622]
- HARSANYI, J. (1967): "Games With Incomplete Information Played by 'Bayesian' Players. Part I. The Basic Model," *Management Science*, 14, 159–182. [1606,1608]
- JACKSON, M. O. (1991): "Bayesian Implementation," *Econometrica*, 59, 461–477. [1606,1610, 1615,1617-1619]
- KAJII, A., AND S. MORRIS (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65, 1283–1309. [1607]
- KARTIK, N., AND O. TERCIEUX (2012): "Implementation With Evidence," *Theoretical Economics*, 7, 323–355. [1622]
- KUNIMOTO, T., AND R. SERRANO (2011): "A New Necessary Condition for Implementation in Iteratively Undominated Strategies," *Journal of Economic Theory*, 146, 2583–2595. [1624]
- MASKIN, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38. [1606,1612,1613]
- MATSUSHIMA, H. (2008): "Role of Honesty in Full Implementation," *Journal of Economic Theory*, 139, 353–359. [1622]
- MERTENS, J.-F., AND S. ZAMIR (1985): "Formulation of Bayesian Analysis for Games of Incomplete Information," *International Journal of Game Theory*, 14, 1–29. [1606,1609,1627]
- OURY, M., AND O. TERCIEUX (2012): "Supplement to 'Continuous Implementation,'" *Econometrica Supplemental Material*, 80, http://www.econometricsociety.org/ecta/Supmat/8577_proofs.pdf. [1608]
- POSTLEWAITE, A., AND D. SCHMEIDLER (1986): "Implementation in Differential Information Economics," *Journal of Economic Theory*, 39, 14–33. [1606,1617]
- RUBINSTEIN, A. (1989): "The Electronic Mail Game: Strategic Behavior Under 'Almost Common Knowledge,'" *American Economic Review*, 79, 385–391. [1607]
- SAIJO, T. (1987): "On Constant Maskin Monotonic Social Choice Functions," *Journal of Economic Theory*, 42, 382–386. [1606]
- SERRANO, R., AND R. VOHRA (2001): "Some Limitations of Virtual Bayesian Implementation," *Econometrica*, 69, 785–792. [1625]
- WEINSTEIN, J., AND M. YILDIZ (2004): "Finite-Order Implications of Any Equilibrium," Working Paper 03-14, MIT Department of Economics. [1616,1617]
- (2007): "A Structure Theorem for Rationalizability With Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400. [1606,1607,1616,1617,1623]

Théorie Économique, Modélisation, Application, Université de Cergy-Pontoise, Cergy-Pontoise, France; marionoury@gmail.com

and

Paris School of Economics, 48 boulevard Jourdan, 75 014 Paris, France; tercieux@pse.ens.fr.

Manuscript received May, 2009; final revision received October, 2011.