Implementation with Near-Complete Information: The Case of Subgame Perfection*

Takashi Kunimoto† and Olivier Tercieux ‡

First Draft: February 2009

This Draft: April 2009

Preliminary

Abstract

While monotonicity is a necessary and almost sufficient condition for Nash implementation and often a demanding one, almost any (non-monotonic, for instance) social choice rule can be implemented using undominated Nash or subgame perfect equilibrium. By requiring solution concepts to have closed graph in the limit of complete information, Chung and Ely (2003) show that only monotonic social choice rules can be implemented in the closure of the undominated Nash equilibrium correspondence. In this paper, we show that only monotonic social choice rules can be implemented in the closure of the subgame perfect equilibrium/sequential equilibrium correspondence. Our robustness result helps understand the limits of subgame pefect implementation, which is widely used in applications. We also argue that static mechanisms might outperform sequential mechanisms when one insists on robustness.

Journal of Economic Literature Classification Numbers: C79, D82

Keywords: Monotonicity, Subgame Perfect Implementation, Robust implementation

---

*This research began when Olivier Tercieux was Member of the Institute for Advanced Study at Princeton. He gratefully acknowledge the Institute's hospitality and financial support through the Deutsche Bank membership. Takashi Kunimoto gratefully acknowledges financial supports from FQRSC and SSHRC of Canada. This paper benefits from the audiences at 2009Game Theory Workshop at Hitotsubashi, McGill, 2008NSF/NBER conference on general equilibrium and mathematical economics at Brown

†Department of Economics, McGill University and CIREQ, Montreal, Quebec, Canada; takashi.kunimoto@mcgill.ca

‡Paris School of Economics, Paris, France; tercieux@pse.ens.fr

1 Introduction

Suppose that the society has a social choice rule which associates with each environment a subset of possible outcomes. The theory of implementation is concerned with characterizing the relationship between the structure of the institution (or mechanism) through which individuals interact and the outcome of that interaction, given a social choice rule and a domain of environments.

Maskin (1999) shows a condition called *monotonicity* is necessary and almost sufficient for Nash implementation. It turns out that monotonicity is quite a demanding condition and the literature tried to obtain less restrictive characterizations using *refinments* of Nash equilibrium. Using subgame perfect equilibrium, Moore and Repullo (1988) dispense with monotonicity and provide a sufficient condition for subgame perfect implementation. Abreu and Sen (1990) further refine the analysis of Moore and Repullo (1988) and obtain a necessary and almost sufficient condition for subgame perfect implementation. Finally, Vartiainen (2007) obtains its full characterization. In fact, Miyagawa (2002) shows that while many axiomatic bargaining solutions are not monotonic, they can be implemented in subgame perfect equilibrium by a four-stage sequential mechanism. As a different refinement, Palfrey and Srivastva (1991) propose *undominated Nash equilibrium* and prove that almost any social choice rule is implementable in undominated Nash equilibrium. Therefore, allowing for the use of refinements of Nash equilibrium, one can significantly expand the class of implementable social choice rules.

Chung and Ely (2003) investigate the robustness of undominated Nash implementation to incomplete information.[1] In so doing, they require that solution concepts have closed graph in the limit of complete information. Then, Chung and Ely (2003) conclude that when preferences are strict (or more generally hedonic), only monotonic social choice rules can be implemented in the closure of the undominated Nash equilibrium correspondence. Following the approach by Chung and Ely (2003), this paper investigates the robustness of *any* subgame perfect implementing mechanism to incomplete information. We show that only monotonic social choice rules can be implemented in the closure of the subgame perfect/sequential equilibrium correspondence. Hence, our result implies that there might

[1]The type of perturbation used in Chung and Ely (2003) weakens common knowledge into common p-belief with p close to 1. Common p-belief is introduced in Monderer and Samet (1989). This is a "smaller"perturbation and less demanding than the one used for instance in Oury and Tercieux (2009). See also Kunimoto (2008) for a characterization of the perturbation used in this paper.

be little difference between sequential mechanisms and static mechanisms, once we insist on robustness. This is due to the fact that a small amount of incomplete information opens up a plethora of sequential equilibria, some of which could be "bad" equilibria and undermine the original implementing mechanism.

There is a closely related paper by Aghion, Fudenberg, and Holden (henceforth, AFH) (2007). They also consider the question of subgame perfect implementation with almost complete information. AFH (2007) focus on a particular mechanism in the spirit of the one defined in Section 5 of Moore and Repullo (1988). Under the assumption of complete information, given any social choice rule, Moore and Repullo (1988) provide a mechanism in which telling the truth is the unique subgame perfect equilibrium. AFH (2007) exhibit *one* social choice rule where telling the truth is not an (sequential) equilibrium when introducing a small amount of incomplete information. On the contrary, our paper shows that the introduction of a small amount of incomplete information may induce new "bad" equilibria, i.e. equilibria that do not implement. When considering implementation problems, we believe that this is a meaningful requirement that indeed follows previous approaches (see Chung and Ely (2003)). While the motivation in AFH (2007) is similar to the present paper in spirit, our "robustness tests" are different and the results are also very different: (1) Our result is mechanism-free: we do not consider a fixed mechanism but a very general class of mechanisms that contains the one studied by AFH (2007); (2) our non-robustness result applies to any social choice rule that is not monotonic, while AFH (2007) focus on a single non-monotonic social choice function that fails their robustness test.

We put our result in a broader perspective. Since the early works of Grossman and Hart (1986) and Hart and Moore (1988), the incomplete contracts literature often cites indescribable contingencies as a major obstacle to the creation of complete contracts. Maskin and Tirole (1999), however, argue that the literature's justification for incomplete contracts is conceptually problematic. Using the agents' minimum foresight concerning the possible payoff contingencies, they show that the inability to describe future contingencies by itself places no constraints on contracting. This is the so-called irrelevance theorem. To show this, Maskin and Tirole reduce their task to checking sufficient conditions for subgame perfect implementation. Then, our result enables us to assess the robustness of Maskin and Tirole's irrelevance theorem. In fact, we can conclude that their implementing mechanism is not robust because a small amount of incomplete information necessitates that we should focus only on monotonic social choice rules. [2]

It is also not difficult to find many other applications of subgame perfect implementation in the literature. For example, Miyagawa's (2002) mechanism to implement bargaining solutions cannot also escape from our robustness argument. In sum, we view the current paper as a first step towards understanding the robustness of sequential mechanisms.

The rest of the paper is organized as follows: In Section 2 we introduce the preliminary notation and definitions. Section 3 defines robust subgame perfect implementation. Section 4 has two subsections: in Section 4.1, we state the main theorem and illustrate the main idea of this paper through an example; and in Section 4.2, we prove the theorem. Section 5 concludes.

2 Setting

There is a finite set $N = \{1, ..., n\}$ of players, and a set A of social alternatives. There is a finite set Θ of states of nature. Associated with each state θ is a preference profile $\succeq^\theta$ which is a list $(\succeq_1^\theta, ..., \succeq_n^\theta)$. Players do not observe the state directly, but are informed of the state via signals. Player i's signal set is S_i which for simplicity we identify $\{s_i^\theta\}_{\theta\in\Theta}$ with $|S_i| = |\Theta|$ for each i. A signal profile is an element $s = (s_1, ..., s_n) \in S \equiv \times_{i\in N} S_i$. When the realized signal profile is s, each player i observes only his own signal s_i. We let μ denote the prior probability over $\Theta \times S$, and let $\mathcal{P}$ be the set of all such priors. We note $\mu(\cdot \mid s_i)$ for the probability measure over $\Theta \times S$ conditional on s_i. Let s^θ be the signal profile in which each player's signal is s_i^θ. *Complete information* refers to the environments in which $\mu(\theta, s) = 0$ whenever $s \neq s^\theta$ (μ will be then referred to as a complete information prior). Under complete information, the state, and hence the full profile of preferences is always common knowledge among agents. We will assume for each i and $\theta : \mu(s_i^\theta) \equiv [\text{marg}_{S_i}\mu](s_i^\theta) > 0$ so that Bayes rule is well-defined. Given a prior μ over $\Theta \times S$, we will sometimes abuse notations and write $\mu(\theta)$ for $[\text{marg}_\Theta\mu](\theta)$. Besides, given $s_{-i} \in S_{-i}$, we will also write $\mu(s_{-i})$ as $[\text{marg}_{S_{-i}}\mu](s_{-i})$. Finally, given some arbitrary countable space X, δ_x will denote the probability measure that puts probability 1 on $\{x\} \subset X$.

A *social choice correspondence* (SCC) is a mapping $\mathcal{F}$ which associates a subset of A

[2]In particular, the simple mechanism used in Section 4 of Maskin and Tirole (1999) is the most relevant here.

with each $\theta \in \Theta$. A single-valued social choice correspondence is a social choice *function* denoted f. Hence, any selection of SCC $\mathcal{F}$ is a social choice function. A *mechanism* is an extensive game form $\Gamma = (\mathcal{H}, M, g)$ where $\mathcal{H}$ is a set of histories h. $M = M_1 \times \cdots \times M_n$ and $M_i = \times_{h \in H} M_i(h)$ for all i. An element of $M(h) = M_1(h) \times \cdots \times M_n(h)$, say $m(h) = (m_1(h), ..., m_n(h))$ is a message profile at h while $m_i(h)$ is i's message at h. If $\#M_i(h) > 1$ and $\#M_j(h) > 1$ then agents i and j move simultaneously after history h, whereas if $\#M_i(h) > 1$ and $\#M_j(h) = 1$ for all $j \neq i$ then agent i is the only one to move. Histories and messages are tied together by the property that $M(h) = \{m : (h, m) \in \mathcal{H}\}$. An element of M_i is a pure strategy; and an element of M is a pure strategy profile. We sometimes write $m \mid_h = (m_1 \mid_h, ..., m_n \mid_h)$ for the profile of pure strategies starting from history h.

There is an initial history $\emptyset \in \mathcal{H}$, and each history h_t is represented by a sequence with finite length $t : (\emptyset, m^1, m^2, ..., m^{t-1}) = h_t$ where for each $k : m^k \in M(h_k)$.[3] If for $t' \geq t+1 : h_{t'} = (h_t, m^t, ..., m^{t'-1})$, then $h_{t'}$ *follows* history h_t. As Γ contains finitely many stages, there is a set of terminal histories[4] $H_T \subset \mathcal{H}$ such that $H_T = \{h \in \mathcal{H}$:there is no h' following $h\}$. Given any strategy profile m and any history h, there is a unique terminal history denoted $h_T[m, h]$. Formally, let $\mathcal{Z} : M \times \mathcal{H} \to \mathcal{H}$ be the mapping where

$$\mathcal{Z}[m, h] = \begin{cases} (h, m(h)) \text{ if } h \notin H_T \\ h \text{ otherwise} \end{cases}$$

is the history that immediately follows h whenever possible given that strategy profile m has been played; and so $h_T[m, h] = \lim_{k \to \infty} \mathcal{Z}^k[m, h]$ where $\mathcal{Z}^k[m, h] = \mathcal{Z}[m, \mathcal{Z}^{k-1}[m, h]]$. Finally, the *outcome function* $g : H_T \to A$ specifies an outcome for each terminal history. We will also note $g(m; h_t)$ for the outcome that obtains when agents use strategy profile m starting from history h_t i.e. $g(m; h_t) = g(h_T[m, h_t])$.

Assumption 1 *$M_i(h)$ is countable for each i and h.*

Remark: This assumption is useful when using sequential equilibrium and avoids technical complications due to the use of measures over uncountable spaces. We, however, do not believe that our results depend on the countability assumption. We refer the

[3] As Moore and Repullo (1988), we restrict ourselves to mechanisms with finitely many stages. We allow agents to move simultaneously at some nodes, so mechanisms need not be with perfect information. However, at each node, all agents are assumed to know the entire history of the play.

[4] Note that $M(h) = \{m : (h, m) \in \mathcal{H}\} = \emptyset$ for any $h \in H_T$.

reader to Duggan (1997) for the treatment of the general (uncountable) message space. In addition, in our setting where the set of states has been assumed to be finite, the famous mechanism by Moore and Repullo (Section 5) uses only a finite set of messages.

A stage mechanism Γ together with a profile $\theta \in \Theta$ defines an extensive game $\Gamma(\theta)$. A (pure strategy) Nash equilibrium for the game $\Gamma(\theta)$ is an element $m^* \in M$ such that, for each agent i, $g(m^*;\emptyset) \succeq_i^\theta g((m_i, m^*_{-i});\emptyset)$ for all $m_i \in M_i$. A (pure strategy) subgame perfect equilibrium for the game $\Gamma(\theta)$ is an element $m^* \in M$ such that, for each agent i, $g(m^*;h) \succeq_i^\theta g((m_i, m^*_{-i});h)$ for all $m_i \in M_i$ and all $h \in \mathcal{H} \backslash H_T$. Let $SPE(\Gamma(\theta))$ denote the set of subgame perfect equilibria of the game $\Gamma(\theta)$. Let also $NE(\Gamma(\theta))$ denote the set of Nash equilibria of the game $\Gamma(\theta)$.

Given a prior μ, the mechanism determines a Bayesian game $\Gamma(\mu)$ in which each player's type is his signal, and after observing his signal, player i selects a strategy from the set M_i. A strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ lists a strategy for each player where $\sigma_i : S_i \to M_i$ and $\sigma_i(h_t, s_i)$ is the message in $M_i(h_t)$ given history h_t and signal s_i. Alternatively, we will sometimes let σ_i be a (mixed) behavior strategy i.e. a function that maps the set of possible histories and signals into the set of probability distributions over messages: $\sigma_i(\cdot \mid h_t, s_i) \in \Delta(M_i(h_t))$ is the probability distribution over $M_i(h_t)$ given history h_t and signal s_i.

An act is a mapping $\alpha : \Theta \times S \to A$. Let $\mathcal{A}$ be the set of acts. A belief is a probability β on $\Theta \times S$. In order to analyze incomplete information games, we must extend the original preferences to the ones under uncertainty. We assume that for each belief β each player i has a preference relation $\succeq_i^\beta$ over acts. We only make the following assumption (which is obviously satisfied by expected utility models but much weaker than that) on this order:

Assumption 2 *Let α and $\hat{\alpha}$ be two acts, and β a belief. Then*

$$[\alpha(\theta, s) \succeq_i^\theta \hat{\alpha}(\theta, s) \text{ for all } (\theta, s) \in supp(\beta)] \Rightarrow \alpha \succeq_i^\beta \hat{\alpha},$$

where $supp(\beta)$ denotes the support of β.

Let σ be a pure strategy profile. Given a profile of pure strategies $\sigma = (\sigma_1, ..., \sigma_n)$, we will note $g(\sigma; h_t)$ for the act that obtains when each agent uses strategy σ_i starting after history h_t occurred, i.e. each pair (θ, s) is mapped to $g(\sigma(s); h_t) \in A$. The act α_σ^Γ induced by σ under the mechanism Γ is defined by $\alpha_\sigma^\Gamma(\theta, s) = g(\sigma(s);\emptyset)$ for any (θ, s).

We will also assume that in the game induced by a stage mechanism, for each player best replies are always well-defined in the neighborhood of complete information when the opponents are playing according to some Nash equilibrium. In general, best-responses need not be well-defined since we allow $M_i(h)$ to be countably infinite. For instance, integer games are such an example with countably infinite message spaces in which best replies need not be well defined.[5] The next assumption ensures that in the neighborhood of complete information, against any Nash equilibrium strategy of his opponents, player i has a strategy that is optimal at histories in some given set H and equal to some fixed strategy at every other histories.

Assumption 3 *A sequential mechanism Γ has well-defined **best replies**: for any player i, any set of histories $H \subseteq \mathcal{H}$, any $\theta \in \Theta$, any $(m_i, m_{-i}) \in NE(\Gamma(\theta))$, there exists $\bar{\xi}(i, H, \theta, m_i, m_{-i}) > 0$ such that for any $\beta(\cdot|s_i^\theta) \in \Delta(\Theta \times S_{-i})$ with $\beta(\theta, s_{-i}^\theta|s_i^\theta) \geq 1 - \bar{\xi}(i, H, \theta, m_i, m_{-i})$, there exists $\sigma_i^*[i, H, \theta, m_i, m_{-i}, \beta]$, or simply σ_i^*, satisfying*

$$
\begin{aligned}
&h \notin H \Rightarrow \sigma_i^*(h; s_i^\theta) = m_i(h); \\
&h \in H \Rightarrow g((\sigma_i^*, \sigma_{-i}); h) \succeq_i^\beta g((\sigma_i', \sigma_{-i}); h)
\end{aligned}
$$

for any σ_i' that differs from σ_i^* only at h and any σ_{-i} such that $\sigma_{-i}(s_{-i}) = m_{-i}$ for any s_{-i} with $\beta(s_{-i}) > 0$.

Remark: This property is satisfied in any finite mechanism as for instance the simple mechanism in Section 5 of Moore and Repullo (1988) that uses a finite set of messages.[6] Note also that when the set of outcomes is finite, this assumption is trivially satisfied.

3 $\overline{SPE}$-implementation

When we perturb a complete information situation introducing a slight incomplete information, we must specify the equilibrium concept we use. In this paper we will focus on sequential equilibrium. Since our result provides necessary conditions, it will hold for any coarser equilibrium concept as for instance perfect Bayesian equilibrium, subgame perfect equilibrium. We now recall the definition of sequential equilibrium as defined in Kreps and Wilson (1982).

[5]If there is some player for whom there is no maximum with respect to his preference order at some state of nature, then best-replies are indeed not well-defined at this state in standard integer games.

[6]Recall that we have assumed that the set of state of nature is finite.

Sequential Equilibrium:

A *system of beliefs* of agent i is defined as a function $\phi_i : S_i \times \mathcal{H} \to \Delta(\Theta \times S_{-i})$. Let $\phi_i[(\theta, s_{-i}) \mid s_i, h_t]$ denote agent i's belief that the state (θ, s_i, s_{-i}) is realized when agent i's signal is s_i and the observed history is h_t. We will henceforth abuse notations and sometimes consider $\phi_i[(\theta, s_{-i}) \mid s_i, h_t]$ as an element of $\Delta(\Theta \times S)$. We also say a vector of beliefs $\phi = (\phi_1, \ldots, \phi_n)$ is Bayes consistent with a strategy profile σ if beliefs are updated from one stage to the next using Bayes' rule whenever possible (see Fudenberg and Tirole (1991) for its precise definition). An *assessment* is a pair (ϕ, σ) consisting of a profile of beliefs and a pure behavior strategy profile.

Definition 1 *A sequential equilibrium is an assessment (ϕ, σ) that satisfies condition (S) and (C):*
(S) Sequential rationality: *for all $i \in N$, $s_i \in S_i, h_t \in \mathcal{H}$:*

$$g(\sigma, h_t) \succeq_i^{\phi_i[\cdot|s_i,h_t]} g((\sigma_i', \sigma_{-i}), h_t)$$

for each σ_i'.
(C) Consistency: *there exists a sequence of totally mixed strategy profiles $(\sigma_1^k, ..., \sigma_n^k)$ converging uniformly*[7] *to $(\sigma_1, ..., \sigma_n)$ with Bayes consistent beliefs ϕ^k converging to ϕ.* [8]

Henceforth, we assume that A is an arbitrary topological space, and that $\mathcal{A} = A^{\Theta \times S}$ is endowed with the product topology. Given a mechanism Γ, we denote the sequential equilibrium correspondence by $\psi_\Gamma^{SE} : \mathcal{P} \to \mathcal{A}$ where each element α of $\psi_\Gamma^{SE}(\mu)$ is an act (or outcome) corresponding to some sequential equilibrium outcome of $\Gamma(\mu)$, which describes the alternative $\alpha(\theta, s)$ that will result for each (θ, s) (where SE stands for sequential equilibrium). Formally, $\psi_\Gamma^{SE}(\mu) \equiv \left\{\alpha \in \mathcal{A} : \alpha = \alpha_\sigma^\Gamma \text{ where } (\phi, \sigma) \text{ is a sequential equilibrium for some } \phi\right\}$. Let

$$\text{graph } \psi_\Gamma^{SE} \equiv \{(\mu, \alpha) : \alpha \in \psi_\Gamma^{SE}(\mu)\}.$$

The following notation will be convenient. If $\mathcal{B}$ is a set of acts such that for any $(\theta, s) \in \text{supp}(\mu)$ and any $a \in \mathcal{F}(\theta)$, there is $\alpha \in \mathcal{B}$ for which $\alpha(\theta, s) = a$, then we will write

[7] Given that the set of messages can be countably infinite, two natural convergence notions can be used: *point-wise* convergence or *uniform* convergence. The set of sequential equilibria is smaller when one assumes uniform convergence. Hence, the use of uniform convergence strengthens our main result.

[8] See also Kreps and Wilson (1982) for the detail of the definition.

$\mathcal{B} \sqsupset_\mu \mathcal{F}$. Further, if $\mathcal{B}$ is a set of acts such that $\alpha(\theta, s) \in \mathcal{F}(\theta)$ for each $\alpha \in \mathcal{B}$ and any $(\theta, s) \in \text{supp}(\mu)$, then we will write $\mathcal{B} \sqsubset_\mu \mathcal{F}$. If $\mathcal{B} \sqsubset_\mu \mathcal{F}$ and $\mathcal{B} \sqsupset_\mu \mathcal{F}$, then we write $\mathcal{B} =_\mu \mathcal{F}$.

Definition 2 *A stage mechanism Γ SE-implements an SCC $\mathcal{F} : \Theta \to A$ under μ if $\psi_\Gamma^{SE}(\mu) =_\mu \mathcal{F}$.*

When μ is a complete information prior, the above definition is equivalent to the standard definition of subgame perfect implementation. The next lemma is its formalization. We provide it with no proof.

Lemma 1 *Let μ be a complete information prior. A stage mechanism Γ SE-implements an SCC $\mathcal{F} : \Theta \to A$ under μ if and only if for each $(\theta, s^\theta) \in \Theta \times S$ with $\mu(\theta, s^\theta) > 0$, we have $g(SPE(\Gamma(\theta)); \emptyset) = \mathcal{F}(\theta)$,*

As in Chung and Ely (2003), we consider the "closure" of the solution correspondence ψ_Γ^{SE}. Define

$$\overline{\psi_\Gamma^{SE}}(\mu) = \{\alpha : (\mu, \alpha) \in \overline{\text{graph } \psi_\Gamma^{SE}}\}.$$

Note that the topology used when we consider the closure is characterized by Kunimoto (2008). Recall that $(\mu, \alpha) \in \overline{\text{graph } \psi_\Gamma^{SE}}$ if there exists a sequence $\{(\mu^k, \alpha^k)\}_{k=1}^\infty$ such that (i) $(\mu^k, \alpha^k) \in \text{graph } \psi_\Gamma^{SE}$ for each k and (ii) $(\mu^k, \alpha^k) \to (\mu, \alpha)$. The following is our definition of robust implementation, denoted $\overline{SPE}$ *implementation.*

Definition 3 *A mechanism Γ $\overline{SE}$-implements an SCC $\mathcal{F} : \Theta \to A$ under μ if $\overline{\psi_\Gamma^{SE}}(\mu) =_\mu \mathcal{F}$. When μ is a complete information prior, we say that Γ $\overline{SPE}$-implements $\mathcal{F}$ under μ. Finally we say that an SCC $\mathcal{F} : \Theta \to A$ is $\overline{SPE}$-implementable under complete information if there exists a mechanism Γ that $\overline{SE}$-implements $\mathcal{F}$ under some complete information prior μ.*

4 Monotonicity as a Necessary Condition

4.1 Theorem and Illustration

We now recall the definition of monotonicity as defined in Maskin (1999).

Definition 4 *An SCC $\mathcal{F}$ is said to be monotonic if, for any $\theta, \theta^{'} \in \Theta$ and any $a \in \mathcal{F}(\theta)$,*

$$(*) \quad \forall i \in N, \forall b \in A, \ a \succeq_i^{\theta} b \Longrightarrow a \succeq_i^{\theta^{'}} b,$$

we have $a \in \mathcal{F}(\theta^{'})$.

We are now in a position to state our main Theorem.

Theorem 1 *Suppose that Assumption 1, 2 and 3 are satisfied. If an SCC is $\overline{SPE}$-implementable under complete information, it is necessarily monotonic.*

Remark: This result seems to contradict Proposition 2 of Kreps and Wilson (1982), which shows that the sequential equilibrium correspondence is upper hemi-continuous. This apparent inconsistency comes from the very fact that the sequential equilibrium correspondence is upper hemi-continuous provided that μ has full support over $\Theta \times S$ (as is assumed in Kreps and Wilson (1982)). However – as shown in our illustration – when μ assigns probability 0 to some profile (θ, s), upper hemi-continuity may not hold.

Let us illustrate the main idea of the proof of Theorem 1 through the simple mechanism proposed in Section 5 of Moore and Repullo (1988). The set of payoff states is $\{\theta, \theta'\}$. There are two agents, called 1 and 2. For each $i = 1, 2$, agent i's preference relation in state θ is given by $\succeq_i^{\theta}$. The agents commonly observe the state, but the planer does not observe it.

We extend the set of outcomes A to $\tilde{A} \equiv A \times \mathbb{R}^2$ and define extended preferences over $\tilde{A}$ as follows. An element of $\tilde{A}$ is now a tuple (a, t_1, t_2) where a is an outcome while for each player $i : t_i$ denotes the transfer to player i. Preferences over A are naturally extended to preferences over $\tilde{A}$ denoted by $\trianglerighteq_i^{\tilde{\theta}}$ i.e. given any transfer t_i to agent $i : a \succeq_i^{\tilde{\theta}} b$ if and only if $(a, t_i, \cdot) \trianglerighteq_i^{\tilde{\theta}} (b, t_i, \cdot)$. To fix ideas, one instance of this extension is the setting with transfers and quasilinear preferences.

Since transfers to player $-i$ do not affect player i's ordering, throughout this example, when considering i's evaluations over outcomes, we ignore agent $j(\neq i)$'s monetary transfer from the expression, i.e. we will abuse notations and for instance, simply note $(a, t_i) \trianglerighteq_i^{\theta} (b, t'_i)$ instead of $(a, t_i, \cdot) \trianglerighteq_i^{\theta} (b, t'_i, \cdot)$.

We assume that $f(\theta) \neq f(\theta')$ and $f : \Theta \rightarrow \tilde{A}$ is "non-monotonic" and therefore not Nash implementable. With this, we must satisfy the following condition:

$$\forall i, \forall b \in \tilde{A} : f(\theta) \trianglerighteq_i^{\theta} b \Rightarrow f(\theta) \trianglerighteq_i^{\theta'} b \quad (**)$$

Following Section 5 of Moore and Repullo (1988), we argue that this non-monotonic f can be implemented as the unique subgame perfect equilibrium outcome of the following 3-stage mechanism, under some assumptions that are naturally satisfied in a setting with (large) transfers and quasi-linear preferences.

Stage 1: Agent 1 announces the state θ (resp., θ'). Then, the game moves to Stage 2.
Stage 2: If agent 2 agrees (i.e. announces the same state as agent 1), then the game ends here and $f(\theta)$ (resp., $f(\theta')$) is chosen. If agent 2 challenges by announcing θ' (resp., θ), the game moves to Stage 3.
Stage 3: Conditioning on agent 1's announcement θ (resp., θ') at Stage 1, agent 1 has to choose between $x(\theta)$ (resp., $x(\theta')$) and $y(\theta)$ (resp., $y(\theta')$) such that

$$x(\theta) \succ_1^{\theta} y(\theta), \text{ and}$$
$$(\text{resp., } x(\theta') \succ_1^{\theta'} y(\theta'), \text{ and})$$
$$y(\theta) \succ_1^{\theta'} x(\theta).$$
$$(\text{resp., } y(\theta') \succ_1^{\theta} x(\theta').)$$

Further, if agent 1 chooses $x(\theta)$ (resp., $x(\theta')$), then agent 1 receives $(x(\theta), -\Delta)$ (resp., $(x(\theta'), -\Delta)$); agent 2 receives $(x(\theta), -\Delta)$ (resp., $(x(\theta), -\Delta)$); and the planner nets 2Δ – whereas if agent 1 chooses $y(\theta)$ (resp., $y(\theta')$), then agent 1 receives $(y(\theta), -\Delta)$ (resp., $(y(\theta'), -\Delta)$); agent 2 receives $(y(\theta), +\Delta)$ (resp., $(y(\theta'), +\Delta)$); and the planner breaks even. [9] The game stops here. It is assumed that Δ is "large enough" i.e., Δ satisfies $(f(\theta'), 0) \rhd_1^{\theta'} (y(\theta), -\Delta)$; $(y(\theta), +\Delta) \rhd_2^{\theta'} (f(\theta), 0)$; and $(f(\theta'), 0) \rhd_2^{\theta'} (x(\theta'), -\Delta)$. Similarly, $(f(\theta), 0) \rhd_1^{\theta} (y(\theta'), -\Delta)$; $(y(\theta'), +\Delta) \rhd_2^{\theta} (f(\theta'), 0)$; and $(f(\theta), 0) \rhd_2^{\theta} (x(\theta), -\Delta)$. Note that this implies in particular that $(f(\theta'), 0) \rhd_1^{\theta'} (x(\theta), -\Delta)$ and $(f(\theta), 0) \rhd_1^{\theta} (x(\theta'), -\Delta)$.

Denote by $m_i^*(\tilde{\theta}; h)$ agent i's strategy in state $\tilde{\theta}$ at history h. The strategy we focus on here is given below:

- $m_1^*(\theta; \emptyset) = \theta$ and $m_1^*(\theta'; \emptyset) = \theta'$;
- $m_2^*(\theta; \theta) = \theta; m_2^*(\theta'; \theta') = \theta', m_2^*(\theta; \theta') = \theta$; and $m_2^*(\theta'; \theta) = \theta'$; and
- $m_1^*(\theta; (\theta, \theta')) = x(\theta); m_1^*(\theta; (\theta', \theta)) = y(\theta'); m_1^*(\theta'; (\theta', \theta)) = x(\theta')$; and $m_1^*(\theta'; (\theta, \theta')) = y(\theta)$.

[9] The existence of such $x(\cdot)$ and $y(\cdot)$ is guaranteed by the following weak domain restriction: for any pair (θ, θ') with $\theta \neq \theta'$, there are $a, b \in A$ and an agent i such that $a \succ_i^{\theta} b$ and $b \succ_i^{\theta'} a$ (preference reversal).

We will show that m^* constitutes the unique subgame perfect equilibrium. First, note that m^* prescribes the outcome where agent 1 will announce the true state and agent 2 will not challenge. Suppose that agent 1 announces the state θ (resp., θ'). If agent 1 lies, then agent 2 can challenge her with the truth, and at stage 3 agent 1 will choose $y(\theta)$ (resp., $y(\theta')$). This is so by construction. Given the choice of Δ, this must be worse for agent 1 than whatever the social choice function f offers. Equally, given the definition of Δ, agent 2 will be satisfied with his reward of Δ. On the other hand, if agent 1 tells the truth, then agent 2 will not (falsely) challenge, since agent 1 would now choose $x(\theta)$ (resp., $x(\theta')$) at Stage 3, which incurs a penalty of Δ for agent 2.

Suppose that the agents have a common prior that $\mu(\theta, s_1^\theta, s_2^\theta) = p$ and $\mu(\theta', s_1^{\theta'}, s_2^{\theta'}) = 1-p$, where $0 < p < 1$. [10] Now let us introduce the following perturbation of the complete information structure ν^ε.

ν^ε	s_1^θ, s_2^θ	$s_1^\theta, s_2^{\theta'}$	$s_1^{\theta'}, s_2^\theta$	$s_1^{\theta'}, s_2^{\theta'}$
θ	$p(1-\varepsilon)$	$p\varepsilon/2$	$p\varepsilon/2$	0
θ'	0	0	0	$1-p$

Observe that $\nu^\varepsilon \to \mu$ as $\varepsilon \to 0$. In this perturbation, if agent i receives s_i^θ, he knows that the state is θ but does know which signal the other agent receives. We propose the following strategy profile σ^* of the game $\Gamma(\nu^\varepsilon)$:

- $\sigma_1^*(s_1^\theta, \emptyset) = \sigma_1^*(s_1^{\theta'}, \emptyset) = \theta$;
- $\sigma_2^*(s_2, \tilde{\theta}) = \theta$ for any $\tilde{\theta} \in \{\theta, \theta'\}$ and any $s_2 \in \{s_2^\theta, s_2^{\theta'}\}$; and
- $\sigma_1^*(s_1^\theta, (\theta, \theta')) = x(\theta); \sigma_1^*(s_1^{\theta'}, (\theta, \theta')) = x(\theta); \sigma_1^*(s_1^\theta, (\theta', \theta)) = y(\theta')$; and $\sigma_1^*(s_1^{\theta'}, (\theta', \theta)) = x(\theta')$.

Note that, $\alpha_{\sigma^*}^\Gamma$, the act induced by σ^*, is such that $\alpha_{\sigma^*}^\Gamma(\theta', s_1^{\theta'}, s_2^{\theta'}) = f(\theta)$. Hence, if each player i receives a signal $s_i^{\theta'}$ and plays according to σ_i^*, the outcome provided is $f(\theta)$.

For each player i, his belief ϕ_i^* is defined as follows:

- $\phi_1^*\left[\cdot \middle| s_1, (\theta, \theta')\right] = \delta_{(\theta, s_2^\theta)}$ for each $s_1 \in \{s_1^\theta, s_1^{\theta'}\}$; and $\phi_2^*[\cdot|s_2, \theta'] = \delta_{(\theta, s_1^\theta)}$ for each $s_2 \in \{s_2^\theta, s_2^{\theta'}\}$;
- $\phi_i^*[\cdot|s_i, \emptyset] = \nu^\varepsilon(\cdot|s_i)$ for each $i = 1, 2$ and each $s_i \in \{s_i^\theta, s_i^{\theta'}\}$;

[10] The common prior assumption is completely dispensable for the rest of arguments.

- $\phi_2^*[\cdot|s_2,\theta] = \nu^\varepsilon(\cdot|s_2)$ for each $s_2 \in \{s_2^\theta, s_2^{\theta'}\}$; and

- $\phi_1^*[\cdot|s_1,(\theta',\theta)] = \nu^\varepsilon(\cdot|s_1)$ for any $s_1 \in \{s_1^\theta, s_1^{\theta'}\}$.

What we want to show is that the proposed assessment (ϕ^*, σ^*) constitutes a sequential equilibrium of the game $\Gamma(\nu^\varepsilon)$ for any $\varepsilon > 0$ small enough. In this case, since $\alpha^\Gamma_{\sigma^*}(\theta', s_1^{\theta'}, s_2^{\theta'}) = f(\theta)$ and $\nu^\varepsilon(\theta', s_1^{\theta'}, s_2^{\theta'}) = 1-p > 0$, this shows that with probability $1-p$, a bad outcome is provided (i.e. $f(\theta)$ instead of $f(\theta')$); this is indeed enough to show that the mechanism provided in this section does not $\overline{SPE}$-implements f.

First, we will check sequential rationality of (ϕ^*, σ^*). At $h_3 = (\theta, \theta')$, agent 1 has to choose between $x(\theta)$ and $y(\theta)$. Due to the construction of ϕ_1^*, regardless of the signal received, agent 1 believes with probability one that the state is θ. Then, by construction of $x(\theta)$ and $y(\theta)$, it is optimal for her to choose $x(\theta)$. Let $h_3 = (\theta', \theta)$. Suppose agent 1 received s_1^θ. In this case, by construction of ϕ_1^* and $\nu^\varepsilon(\cdot|s_1^\theta)$, agent 1 knows that the state is θ. Here, agent 1 has to choose between $x(\theta')$ and $y(\theta')$. By construction, it is optimal for her to choose $y(\theta')$, regardless of ε. Suppose that agent 1 received $s_1^{\theta'}$. Our finite setup guarantees that agent 1 always has a best reply in this perturbed environment. Due to the construction of ϕ_1^* and small enough $\varepsilon > 0$, agent 1 believes with arbitrarily high probability that the state is θ'. With an additional assumption of continuity of preferences, we proceed to argue that it is optimal for her to choose $x(\theta')$. [11]

With this in mind, we move to Stage 2. Suppose that $h_2 = \theta$. In this case, if agent 2 chooses θ', he knows that agent 1 will choose $x(\theta)$. Assume that agent 2 received s_2^θ. In this case, by construction of ϕ_2^* and $\nu^\varepsilon(\cdot|s_2^\theta)$, agent 2 knows that the state is θ. But since $(f(\theta), 0) \rhd_2^\theta (x(\theta), -\Delta)$, by Assumption 2, we can conclude that it is optimal for agent 2 to choose θ.

Assume that agent 2 received $s_2^{\theta'}$. As we argued before, agent 2 knows that agent 1 will choose $x(\theta)$. We also know that $(f(\theta), 0) \rhd_2^\theta (x(\theta), -\Delta)$. By condition $(**)$, we can obtain that $(f(\theta), 0) \trianglerighteq_2^{\theta'} (x(\theta), -\Delta)$ as well. Since $\nu^\varepsilon(\cdot|s_2^{\theta'})$ assigns strictly positive weights only to $(\theta, s_1^\theta, s_2^{\theta'})$ and $(\theta', s_1^{\theta'}, s_2^{\theta'})$, by Assumption 2, we can conclude that it is again optimal for agent 2 to choose θ.

Suppose that $h_2 = \theta'$. In this case, due to the construction of ϕ_2^*, agent 2 believes with probability one that the state is θ and agent 1 will choose $y(\theta')$ at Stage 3. But

[11] As shown in the proof of Theorem 1, the same argument can go through even if (perhaps due to the lack of continuity of preferences) $y(\theta')$ is a best reply.

we know that $(y(\theta'), +\Delta) \rhd_2^{\theta} (f(\theta'), 0)$. Since ϕ_2^* assigns a strictly positive weight only to either $(\theta, s_1^{\theta}, s_2^{\theta})$ if agent 2 received s_2^{θ} or $(\theta, s_1^{\theta}, s_2^{\theta'})$ if agent 2 received $s_2^{\theta'}$. By Assumption 2 and the construction ϕ_2^*, we can conclude that for any $s_2 \in \{s_2^{\theta}, s_2^{\theta'}\}$, it is optimal for agent 2 to choose θ.

Finally, we move to Stage 1. If agent 1 chooses θ, she knows that agent 2 will choose θ so that $f(\theta)$ is chosen. On the other hand, suppose agent 1 chooses θ'. Assume also that she received s_1^{θ}. Then, she knows that the state is θ and that agent 2 will choose θ at Stage 2. We know that $(f(\theta), 0) \rhd_1^{\theta} (x(\theta'), -\Delta)$ and $(f(\theta), 0) \rhd_1^{\theta} (y(\theta'), -\Delta)$. Since $\phi_1^*[\cdot | s_1^{\theta}, \emptyset]$ assigns strictly positive weights only to $(\theta, s_1^{\theta}, s_2^{\theta})$ and $(\theta, s_1^{\theta}, s_2^{\theta'})$, by Assumption 2, we can conclude that it is optimal for her to choose θ.

Assume, on the contrary, that agent 1 received $s_1^{\theta'}$. If agent 1 deviates to θ', she knows that agent 2 will choose θ at Stage 2 and she herself will choose either $x(\theta')$ or $y(\theta')$ at Stage 3. As we argued above, we have chosen $\Delta > 0$ so that $(f(\theta), 0) \rhd_1^{\theta} (x(\theta'), -\Delta)$ and $(f(\theta), 0) \rhd_1^{\theta} (y(\theta'), -\Delta)$. By condition $(**)$, we also obtain $(f(\theta), 0) \trianglerighteq_1^{\theta'} (x(\theta'), -\Delta)$ and $(f(\theta), 0) \trianglerighteq_1^{\theta'} (y(\theta'), -\Delta)$. Since $\phi_1^*[\cdot | s_1^{\theta'}, \emptyset]$ assigns strictly positive weights only to $(\theta, s_1^{\theta'}, s_2^{\theta})$ and $(\theta', s_1^{\theta'}, s_2^{\theta'})$, by Assumption 2, we can conclude that it is optimal for agent 1 to choose θ at Stage 1.

We conclude that (ϕ^*, σ^*) so constructed satisfies sequential rationality.

Next we will check consistency of (ϕ^*, σ^*). Let $\{\eta_k\}_{k=1}^{\infty}$ be a sequence such that $\eta_k > 0$ for each k and $\eta_k \to 0$ as $k \to \infty$. Let a sequence of totally mixed strategy profiles $\{\sigma^k\}_{k=1}^{\infty}$ be defined as follows:

$$\sigma_1^k(s_1^{\theta}, \emptyset) = \begin{cases} \theta & \text{w.p. } 1 - \eta_k \\ \theta' & \text{w.p. } \eta_k \end{cases}$$

$$\sigma_1^k(s_1^{\theta'}, \emptyset) = \begin{cases} \theta & \text{w.p. } 1 - \eta_k^2 \\ \theta' & \text{w.p. } \eta_k^2 \end{cases}$$

$$\sigma_2^k(s_2^{\theta}, \theta) = \sigma_2^k(s_2^{\theta}, \theta') = \begin{cases} \theta & \text{w.p. } 1 - \eta_k \\ \theta' & \text{w.p. } \eta_k \end{cases}$$

$$\sigma_2^k(s_2^{\theta'}, \theta) = \sigma_2^k(s_2^{\theta'}, \theta') = \begin{cases} \theta & \text{w.p. } 1 - \eta_k^2 \\ \theta' & \text{w.p. } \eta_k^2 \end{cases}$$

$$\sigma_1^k(s_1^\theta,(\theta,\theta')) = \begin{cases} x(\theta) & \text{w.p. } 1-\eta_k \\ y(\theta) & \text{w.p. } \eta_k \end{cases}$$

$$\sigma_1^k(s_1^\theta,(\theta',\theta)) = \begin{cases} x(\theta') & \text{w.p. } \eta_k \\ y(\theta') & \text{w.p. } 1-\eta_k \end{cases}$$

$$\sigma_1^k(s_1^{\theta'},(\theta,\theta'))\ (resp., \sigma_1^k(s_1^{\theta'},(\theta',\theta)) = \begin{cases} x(\theta)\ (\text{resp., } x(\theta')) & \text{w.p. } 1-\eta_k^2 \\ y(\theta)\ (\text{resp., } y(\theta')) & \text{w.p. } \eta_k^2 \end{cases}$$

Note that $\sigma^k \to \sigma^*$ by construction. We can define a belief profile ϕ^k associated with σ^k. We claim that $\phi^k \to \phi^*$ as $k \to \infty$. For simplicity, we only pay attention to checking off the equilibrium beliefs. This can be done by explicitly computing the following:

$$\begin{aligned} & \phi_1^k[(\theta, s_2^{\theta'})|s_1^\theta,(\theta,\theta')] \\ = \quad & \frac{\nu^\varepsilon(\theta, s_1^\theta, s_2^{\theta'}) \times \sigma_1^k(\theta \mid \emptyset, s_1^\theta) \times \sigma_2^k(\theta' \mid \theta, s_2^{\theta'})}{\nu^\varepsilon(\theta, s_1^\theta, s_2^{\theta'}) \times \sigma_1^k(\theta \mid \emptyset, s_1^\theta) \times \sigma_2^k(\theta' \mid \theta, s_2^{\theta'}) + \nu^\varepsilon(\theta, s_1^\theta, s_2^\theta) \times \sigma_1^k(\theta \mid \emptyset, s_1^\theta) \times \sigma_2^k(\theta' \mid \theta, s_2^\theta)} \\ = \quad & \frac{(p\varepsilon/2)(1-\eta_k)\eta_k^2}{(p\varepsilon/2)(1-\eta_k)\eta_k^2 + p(1-\varepsilon)(1-\eta_k)\eta_k} = \frac{p\varepsilon\eta_k/2}{p\varepsilon\eta_k/2 + p(1-\varepsilon)} \to 0 \ \ (\text{as } k \to \infty) \end{aligned}$$

$$\phi_1^k[(\theta', s_2^{\theta'})|s_1^{\theta'},(\theta,\theta')] = \frac{(1-p)(1-\eta_k^2)\eta_k^2}{(1-p)(1-\eta_k^2)\eta_k^2 + (p\varepsilon/2)(1-\eta_k^2)(\eta_k)} = \frac{(1-p)\eta_k}{(1-p)\eta_k + p\varepsilon/2} \to 0 \ \ (\text{as } k \to \infty)$$

$$\phi_2^k[(\theta, s_1^{\theta'})|s_2^\theta,\theta'] = \frac{(p\varepsilon/2)\eta_k^2}{(p\varepsilon/2)\eta_k^2 + p(1-\varepsilon)\eta_k} = \frac{p\varepsilon\eta_k/2}{p\varepsilon\eta_k/2 + p(1-\varepsilon)} \to 0 \ \ (\text{as } k \to \infty)$$

$$\phi_2^k[(\theta', s_1^{\theta'})|s_2^{\theta'},\theta'] = \frac{(1-p)\eta_k^2}{(1-p)\eta_k^2 + (p\varepsilon/2)\eta_k} = \frac{(1-p)\eta_k}{(1-p)\eta_k + p\varepsilon/2} \to 0 \ \ (\text{as } k \to \infty)$$

4.2 Proof of Theorem 1

Let μ be a complete information prior, and let $\mathcal{F}$ be a $\overline{SPE}$-implementable SCC with implementing mechanism Γ. Fix any $\theta, \theta' \in \Theta$ and any $a \in \mathcal{F}(\theta)$. Suppose θ and θ' are two possible states satisfying $(*)$. We will show that $a \in \mathcal{F}(\theta')$.

Since Γ $\overline{SPE}$-implements $\mathcal{F}$, it must also SPE-implements $\mathcal{F}$. Thus, by Lemma 1, there exists a subgame perfect equilibrium m_θ^* in $\Gamma(\theta)$ such that $g(m_\theta^*) = a$. Clearly, m_θ^* is actually a Nash equilibrium of $\Gamma(\theta)$. From $(*)$, it follows that m_θ^* is also a Nash equilibrium of $\Gamma(\theta')$. Recall that $\mathcal{H}$ denotes the set of all possible histories. For each $t \geq 0$, let h_t^* be the history induced by m_θ^* up to date t and denote $\mathcal{H}^*$ for the set of all such histories. In addition, for each player i, let $\mathcal{H}^*_{-i}$ be the set of histories h along which every player $j \neq i$ has chosen the message $m^*_{\theta,j}(h')$; formally, $\mathcal{H}^*_{-i} \equiv \{h \in \mathcal{H} : h = (\emptyset, m^1, m^2, ..., m^{t-1})$ for some t and $m^{t'}_j = m^{*,t'}_{j,\theta}$ for all $t' \leq t-1$ and all $j \neq i\}$. Note that $h^*_t \in \mathcal{H}^*_{-i}$ for each $t \geq 1$.

Fix $\varepsilon > 0$ to be sufficiently small so that for each $\tilde{\theta}$, we have $\nu^{\varepsilon}((\tilde{\theta}, s^{\tilde{\theta}}_{-i}) \mid s^{\tilde{\theta}}_i) \geq 1 - \bar{\xi}(i, \mathcal{H}^*_{-i} \backslash \mathcal{H}^*, \tilde{\theta}, m^*_{-i,\theta})$ where $\bar{\xi}(i, \mathcal{H}^*_{-i} \backslash \mathcal{H}^*, \tilde{\theta}, m^*_{-i,\theta})$ is in Assumption 3. Consider the following family of information structure ν^{ε}. For each player i, let τ_i represent the profile of signals $s = (s_1, ..., s_n)$ defined by $s_i = s^{\theta'}_i$ and $s_j = s^{\theta}_j$ for all $j \neq i$. For all i, ν^{ε} describes

$$\begin{aligned}
&\nu^{\varepsilon}(\theta, \tau_i) = \frac{\varepsilon}{n}\mu(\theta, s^{\theta}); \\
&\nu^{\varepsilon}(\theta, s^{\theta}) = (1-\varepsilon)\mu(\theta, s^{\theta}); \text{ and} \\
&\nu^{\varepsilon}(\tilde{\theta}, s^{\tilde{\theta}}) = \mu(\tilde{\theta}, s^{\tilde{\theta}}) \;\; \forall \tilde{\theta} \neq \theta.
\end{aligned}$$

In this information structure when the state is anything other than θ or θ', the state is common knowledge. Furthermore, when a player observes θ, he knows that the state is θ. Obviously, $\nu^{\varepsilon} \to \mu$ as $\varepsilon \to 0$. [12] The support of ν^{ε} is denoted

$$\mathrm{supp}(\nu^{\varepsilon}) = \{(\tilde{\theta}, s^{\tilde{\theta}}) : \tilde{\theta} \in \Theta\} \cup \{(\theta, \tau_i) : i \in N\}.$$

We build a sequential equilibrium (ϕ, σ) of $\Gamma(\nu^{\varepsilon})$ for any $\varepsilon > 0$ small enough where σ induces an act α^{Γ}_{σ} for which $\alpha^{\Gamma}_{\sigma}(\theta', s^{\theta'}) = a$. Hence, this will show that $(\nu^{\varepsilon}, \alpha^{\Gamma}_{\sigma}) \in$ graph ψ^{SE}_{Γ} for all $\varepsilon > 0$ small enough. Note that although σ depends on ε, the induced act α^{Γ}_{σ} does not. Hence, $(\nu^{\varepsilon}, \alpha^{\Gamma}_{\sigma}) \to (\mu, \alpha^{\Gamma}_{\sigma}) \in \overline{\text{graph } \psi^{SE}_{\Gamma}}$ as $\varepsilon \to 0$. Thus since Γ $\overline{\psi^{SE}_{\Gamma}}$-implements $\mathcal{F}$ under μ, we must have $a = \alpha^{\Gamma}_{\sigma}(\theta', s^{\theta'}) \in \mathcal{F}(\theta')$, which will complete the proof.

In the following lines, we define a strategy σ and a family of system of beliefs Φ so that σ induces an act α^{Γ}_{σ} for which $\alpha^{\Gamma}_{\sigma}(\theta', s^{\theta'}) = a$. In addition, we will show that (ϕ, σ) is a sequential equilibrium of $\Gamma(\nu^{\varepsilon})$ for some $\phi \in \Phi$. Φ and σ are defined as follows:

Definition of Φ:

$\phi \in \Phi$ if and only ϕ satisfies the following three properties.

Φ1. Fix any $i \in N$, any $h_t \notin \mathcal{H}^*_{-i}$,

$$\phi_i\left[\cdot | s^{\theta'}_i, h_t\right] = \delta_{(\theta, s^{\theta}_{-i})}$$

[12] We use exactly the same information structures as in Chung and Ely (2003).

also

$$\operatorname{supp}\left(\phi_i\left[\cdot|s_i^{\theta},h_t\right]\right)\subseteq\operatorname{supp}\left(\nu^{\varepsilon}\left[\cdot|s_i^{\theta}\right]\right)$$

and for all l with $h_t\in\mathcal{H}^*_{-l}$: (i.e., l has deviated)

$$\phi_i[(\theta,\tau_l)\mid s_i^{\theta},h_t]=0.$$

Φ2. For any $i\in N$, any $h_t\in\mathcal{H}^*_{-i}$, any $s_i\in\{s_i^{\theta},s_i^{\theta'}\}$:

$$\phi_i[\cdot|s_i,h_t]=\nu^{\varepsilon}(\cdot|s_i).$$

Φ3. For any $i\in N$, any $h_t\in\mathcal{H}$ and any $s_i^{\tilde{\theta}}\notin\{s_i^{\theta},s_i^{\theta'}\}$, we just assume that $\phi_i\left[\cdot\mid s_i^{\tilde{\theta}},h_t\right]=\delta_{(\tilde{\theta},s_{-i}^{\tilde{\theta}})}$ where δ_x denotes the probability measure that puts probability 1 on $\{x\}$.

Definition of σ:

Σ1. For any player i and any $h_t\in\mathcal{H}^*$ or $h_t\notin\mathcal{H}^*_{-i}:\sigma_i(h_t,s_i^{\theta'})=m^*_{\theta,i}(h_t)$;

Σ2. For any player i and any $h_t\in\mathcal{H}^*_{-i}\backslash\mathcal{H}^*,\ \sigma_i(h_t,s_i^{\theta'})=\sigma_i^*(h_t,s_i^{\theta'})$ where $\sigma_i^*=\sigma_i^*[i,\mathcal{H}^*_{-i}\backslash\mathcal{H}^*,\theta',m^*_{i,\theta},m^*_{-i,\theta},\nu^{\varepsilon}]$ as defined in Assumption 3 and so satisfies:

$$h\in\mathcal{H}^*\text{ or }h\notin\mathcal{H}^*_{-i}\Rightarrow\sigma_i^*(h,s_i^{\theta'})=m^*_{i,\theta}(h);$$

$$h\in\mathcal{H}^*_{-i}\backslash\mathcal{H}^*\Rightarrow g((\sigma_i^*,\hat{\sigma}_{-i});h)\succeq_i^{\nu^{\varepsilon}(\cdot|s_i^{\theta'})}g((\sigma_i',\hat{\sigma}_{-i});h)$$

for any $\sigma_i^{'}$ that differs from σ_i^* only at h (one-shot deviation) and any $\hat{\sigma}_{-i}$ satisfying $\hat{\sigma}_{-i}(s_{-i})=m^*_{\theta,-i}$ for any s_{-i} with $\nu^{\varepsilon}(s_{-i}|s_i^{\theta'})>0$. This is well-defined by Assumption 3 because ε is small enough so that $\nu^{\varepsilon}(\theta^{'},s_{-i}^{\theta^{'}}|s_i^{\theta^{'}})\geq 1-\bar{\xi}(i,\mathcal{H}^*_{-i}\backslash\mathcal{H}^*,\theta',m^*_{i,\theta},m^*_{-i,\theta},\nu^{\varepsilon})$;

Σ3. For any player i and any $h_t\in\mathcal{H}:\sigma_i(h_t,s_i^{\theta})=m^*_{\theta,i}(h_t)$;

Σ4. And for any $h_t\in\mathcal{H}$, $\sigma_i(h_t,s_i^{\tilde{\theta}})=m^*_{\tilde{\theta},i}(h_t)$ for $\tilde{\theta}\neq\theta,\theta'$ where $m^*_{\tilde{\theta}}$ is an arbitrary subgame perfect equilibrium of $\Gamma(\tilde{\theta})$. This is well-defined since $\mathcal{F}$ is implementable in subgame perfect equilibrium under complete information.

Note that $h_T[\sigma(s^{\theta'}),\emptyset]=h_T[m^*_{\theta},\emptyset]$ and so, σ generates an act α_{σ}^{Γ} for which $\alpha_{\sigma}^{\Gamma}(\theta',s^{\theta'})=g(\sigma(s^{\theta'});\emptyset)=g(m^*_{\theta};\emptyset)=a$. Hence, it only remains to show that (ϕ,σ) constitutes a sequential equilibrium for some $\phi\in\Phi$. In Section 4.2.1, we will show that (ϕ,σ) satisfies sequential rationality for any $\phi\in\Phi$; and we will also establish that (ϕ,σ) satisfies consistency for some $\phi\in\Phi$ in Section 4.2.2.

4.2.1 Sequential rationality

Fix any $\phi \in \Phi$. Sequential rationality of (ϕ, σ) will be proved by Claims 1 and 2 below.

Claim 1 *For any $i \in N$, $s_i \neq s_i^{\theta'}, h_t \in \mathcal{H}$:*

$$g(\sigma; h_t) \succeq_i^{\phi_i[\cdot|s_i, h_t]} g((\sigma_i', \sigma_{-i}); h_t)$$

for each σ_i'.

Proof of Claim 1: Fix any player i. It is obvious for $s_i^{\tilde{\theta}} \neq s_i^{\theta}$ because by $\Phi\mathbf{3}$, $\phi_i \left[\cdot \mid s_i^{\tilde{\theta}}, h_t\right] = \delta_{(\tilde{\theta}, s_{-i}^{\tilde{\theta}})}$ and so state $\tilde{\theta}$ is common knowledge. By $\Sigma\mathbf{4}$, we can further conclude that $\sigma(s^{\tilde{\theta}}) = m_{\tilde{\theta}}^*$ is a subgame perfect equilibrium in the complete information game $\Gamma(\tilde{\theta})$. Hence, we focus on the case where $s_i = s_i^{\theta}$. By construction, $\nu^{\varepsilon}(\theta \mid s_i^{\theta}) = 1$ and so this player knows that his preference is given by $\succeq_i^{\theta}$. The uncertainty he faces is rather on the signals of his opponents, i.e. whether the profile of signals is s^{θ} or τ_k for some $k \neq i$.

Let $h_t \notin \mathcal{H}_{-i}^*$. By $\Sigma\mathbf{3}$ we know that $\sigma(s^{\theta}) = m_{\theta}^*$. Hence, $h_T[\sigma(s^{\theta}), h_t] = h_T[m_{\theta}^*, h_t]$ and so

$$g(\sigma(s^{\theta}); h_t) = g(m_{\theta}^*; h_t).$$

In addition, for each $l \neq i$ with $h_t \notin \mathcal{H}_{-l}^*$, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ we know that $\sigma_{-i}(\tau_l, h_t) = m_{-i,\theta}^*(h_t)$. [13] For any history $h_{t'}$ that follows h_t, we must have $h_{t'} \notin \mathcal{H}_{-l}^*$. By applying again $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ we get that $\sigma_{-i}(\tau_l) \mid_{h_t} = m_{-i,\theta}^* \mid_{h_t}$. Hence, we obtain $h_T[\sigma(\tau_l), h_t] = h_T[m_{\theta}^*, h_t]$ and so for each $l \neq i$ with $h_t \notin \mathcal{H}_{-l}^*$, we have

$$g(\sigma(\tau_l); h_t) = g(m_{\theta}^*; h_t).$$

In case player i deviates to σ_i', he can induce the following terminal histories: $h_T[\sigma_i'(s_i^{\theta}), \sigma_{-i}(s_{-i}^{\theta}), h_t] = h_T[m_i', m_{-i,\theta}^*, h_t]$ for some strategy m_i' and so

$$g(\sigma_i'(s_i^{\theta}), \sigma_{-i}(s_{-i}^{\theta}); h_t) = g(m_i', m_{-i,\theta}^*; h_t).$$

In addition, for each $l \neq i$ with $h_t \notin \mathcal{H}_{-l}^*$, we know that $\sigma_{-i}(\tau_l) \mid_{h_t} = m_{-i,\theta}^* \mid_{h_t}$. Hence, $h_T[\sigma_i'(s_i^{\theta}), \sigma_{-i}(\tau_l), h_t] = h_T[m_i', m_{-i,\theta}^*, h_t]$ and so for each $l \neq i$ with $h_t \notin \mathcal{H}_{-l}^*$, we have

$$g(\sigma_i'(s_i^{\theta}), \sigma_{-i}(\tau_l); h_t) = g(m_i', m_{-i,\theta}^*; h_t).$$

[13] We abuse the notation because we should use $\sigma_{-i}(\tau_l \backslash s_i^{\theta}, h_t)$ instead of $\sigma_{-i}(\tau_l, h_t)$. This abuse will be used everywhere.

Since m^*_θ is a subgame perfect equilibrium in the complete information game $\Gamma(\theta)$, we have $g(m^*_\theta; h_t) \succeq^\theta_i g(m'_i, m^*_{-i,\theta}; h_t)$. Thus, we get $g(\sigma(s^\theta); h_t) \succeq^\theta_i g(\sigma'_i(s^\theta_i), \sigma_{-i}(s^\theta_{-i}); h_t)$ and for each $l \neq i$ such that $h_t \notin \mathcal{H}^*_{-l}$: $g(\sigma(\tau_l); h_t) \succeq^\theta_i g(\sigma'_i(s^\theta_i), \sigma_{-i}(\tau_l); h_t)$. Because by **Φ1**, we have $\phi_i[\cdot \mid s^\theta_i, h_t]$ assigns a strictly positive weight only to (θ, s^θ_{-i}) and (θ, τ_l) for each $l \neq i$ such that $h_t \notin \mathcal{H}^*_{-l}$, we can conclude with Assumption 2

$$g(\sigma; h_t) \succeq_i^{\phi_i[\cdot|s^\theta_i, h_t]} g((\sigma'_i, \sigma_{-i}); h_t).$$

Let $h_t \in \mathcal{H}^*_{-i}$. Let us distinguish two cases. First, assume that $h_t \in \mathcal{H}^*_{-i} \backslash \mathcal{H}^*$. Since $h_t \in \mathcal{H}^*_{-i}$ and $h_t \notin \mathcal{H}^*$, there must exist $t' < t$ such that $\sigma_i(h_{t'}, s^\theta_i) \neq m^*_{i,\theta}(h_{t'})$ where $h_{t'}$ is a truncation of history h_t. Then, for any history $h_{t''}$ following $h_{t'}$ (and so in particular, following h_t), we have $h_{t''} \notin \mathcal{H}^*_{-k}$ for each $k \neq i$. By **Σ1** and **Σ3**, we thus obtain $\sigma(h_{t''}, s^\theta) = \sigma(h_{t''}, \tau_k) = m^*_\theta(h_{t''})$ for each $k \neq i$. Hence, for each $k \neq i$ we have $h_T[\sigma(s^\theta), h_t] = h_T[\sigma(\tau_k), h_t] = h_T[m^*_\theta, h_t]$, which further implies

$$g(\sigma(s^\theta); h_t) = g(\sigma(\tau_k); h_t) = g(m^*_\theta; h_t).$$

Consider the case where player i deviates to σ'_i. Here, **Σ1** and **Σ3** allow us to conclude that for each $k \neq i$, player i can induce the following terminal histories: $h_T[\sigma'_i(s^\theta_i), \sigma_{-i}(s^\theta_{-i}), h_t] = h_T[\sigma'_i(s^\theta_i), \sigma_{-i}(\tau_k), h_t] = h_T[m'_i, m^*_{-i,\theta}, h_t]$ for some strategy m'_i, which implies

$$g(\sigma'_i(s^\theta_i), \sigma_{-i}(s^\theta_{-i}); h_t) = g(\sigma'_i(s^\theta_i), \sigma_{-i}(\tau_k); h_t) = g(m'_i, m^*_{-i,\theta}; h_t).$$

Since m^*_θ is a subgame perfect equilibrium in the complete information game $\Gamma(\theta)$, we already have $g(m^*_\theta; h_t) \succeq^\theta_i g(m'_i, m^*_{-i,\theta}; h_t)$. Thus, we also get $g(\sigma(s^\theta); h_t) \succeq^\theta_i g(\sigma'_i(s^\theta_i), \sigma_{-i}(s^\theta_{-i}); h_t)$ and $g(\sigma(\tau_k); h_t) \succeq^\theta_i g(\sigma'_i(s^\theta_i), \sigma_{-i}(\tau_k); h_t)$ for each $k \neq i$. Now, since by **Φ2** we know that $\phi_i[\cdot \mid s^\theta_i, h_t]$ assigns a strictly positive weight only to (θ, s^θ_{-i}) and (θ, τ_k) for each $k \neq i$, we can conclude with Assumption 2

$$g(\sigma, h_t) \succeq_i^{\phi_i[\cdot|s^\theta_i; h_t]} g((\sigma'_i, \sigma_{-i}); h_t).$$

Consider now the second case where $h_t \in \mathcal{H}^*$. Note that $h_{t+1} = (h_t, \sigma(h_t, s^\theta)) = (h_t, \sigma(h_t, \tau_k)) = (h_t, m^*_\theta(h_t)) = h^*_{t+1} \in \mathcal{H}^*$ where the second and third equalities are assured by **Σ1** and **Σ3** and we use the fact that $h_t \in \mathcal{H}^*$. Similar argument can be made inductively so that any subsequent history also falls into $\mathcal{H}^*$. Because $h_T[\sigma(s^\theta), h_t] = h_T[\sigma(\tau_k), h_t] = h_T[m^*_\theta, h_t]$, we obtain

$$g(\sigma(s^\theta); h_t) = g(\sigma(\tau_k); h_t) = g(m^*_\theta; h_t).$$

Now consider that player i deviates to σ_i'. Let $\hat{t} \geq t$ be the first date at which $\sigma_i'(h_{\hat{t}}, s_i^\theta) \neq \sigma_i(h_{\hat{t}}, s_i^\theta)$; or equivalently, $\sigma_i'(h_{\hat{t}}, s_i^\theta) \neq m_{i,\theta}^*(h_{\hat{t}})$. As above, one can inductively show that as long as $t' < \hat{t}$, we obtain $h_{t'+1} = (h_{t'}, \sigma_i'(h_{t'}, s_i^\theta), \sigma_{-i}(h_{t'}, s_{-i}^\theta)) = (h_{t'}, \sigma_i'(h_{t'}, s_i^\theta), \sigma_{-i}(h_{t'}, \tau_k)) = (h_{t'}, m_{i,\theta}^*(h_{t'}), m_{-i,\theta}^*(h_{t'})) \in \mathcal{H}^*$ for each $k \neq i$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_{t'} \in \mathcal{H}^*$. In addition, $h_{t'+1} \notin \mathcal{H}_{-k}^*$ for each $k \neq i$ and $t' \geq \hat{t}$. Hence, $t' \geq \hat{t}$ for $h_{t'+1} = (h_{t'}, \sigma_i'(h_{t'}, s_i^\theta), \sigma_{-i}(h_{t'}, s_{-i}^\theta)) = (h_{t'}, \sigma_i'(h_{t'}, s_i^\theta), \sigma_{-i}(h_{t'}, \tau_k)) = (h_{t'}, \sigma_i'(h_{t'}, s_i^\theta), m_{-i,\theta}^*(h_{t'}))$ for each $k \neq i$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_{t'} \notin \mathcal{H}_{-k}^*$ for each $k \neq i$. So we get $h_T[\sigma_i'(s_i^\theta), \sigma_{-i}(s_{-i}^\theta), h_t] = h_T[\sigma_i'(s_i^\theta), \sigma_{-i}(\tau_k), h_t] = h_T[m_i', m_{-i,\theta}^*, h_t]$ for some strategy m_i', which implies

$$g(\sigma_i'(s_i^\theta), \sigma_{-i}(s_{-i}^\theta); h_t) = g(\sigma_i'(s_i^\theta), \sigma_{-i}(\tau_k); h_t) = g(m_i', m_{-i,\theta}^*; h_t).$$

Here again, since m_θ^* is a subgame perfect equilibrium in the complete information game $\Gamma(\theta)$, we have $g(m_\theta^*; h_t) \succeq_i^\theta g(m_i', m_{-i,\theta}^*; h_t)$. Thus, we get $g(\sigma(s^\theta); h_t) \succeq_i^\theta g(\sigma_i'(s_i^\theta), \sigma_{-i}(s_{-i}^\theta); h_t)$ and $g(\sigma(\tau_k); h_t) \succeq_i^\theta g(\sigma_i'(s_i^\theta), \sigma_{-i}(\tau_k); h_t)$ for each $k \neq i$. Now since by $\Phi\mathbf{2}$, $\phi_i[\cdot \mid s_i^\theta, h_t]$ assigns a strictly positive weight only to (θ, s_{-i}^θ) and (θ, τ_k) for each $k \neq i$, we can conclude with Assumption 2

$$g(\sigma; h_t) \succeq_i^{\phi_i[\cdot|s_i^\theta, h_t]} g((\sigma_i', \sigma_{-i}); h_t).$$

This completes the proof. ■

Claim 2 *For any $i \in N$, $s_i = s_i^{\theta'}$, and $h_t \in \mathcal{H}$:*

$$g(\sigma, h_t) \succeq_i^{\phi_i[\cdot|s_i, h_t]} g((\sigma_i', \sigma_{-i}), h_t)$$

for each σ_i'.

Proof of Claim 2: This claim will be proved by studying three different cases depending on the type of history we consider: (1) $h_t \notin \mathcal{H}_{-i}^*$; (2) $h_t \in \mathcal{H}^*$; and (3) $h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^*$.

Let us first consider the case (1) $h_t \notin \mathcal{H}_{-i}^*$. By $\Sigma\mathbf{3}$ we know that $\sigma_{-i}(s_{-i}^\theta) = m_{-i,\theta}^*$. In addition, for any history $h_{t'}$ following h_t, we have $h_{t'} \notin \mathcal{H}_{-i}^*$. Thus, by $\Sigma\mathbf{1}$, we obtain $\sigma_i(h_{t'}, s_i^{\theta'}) = m_{i,\theta}^*(h_{t'})$ for any subsequent history $h_{t'}$. This further implies that $h_T[\sigma(s_i^{\theta'}, s_{-i}^\theta), h_t] = h_T[m_\theta^*, h_t]$ and so we obtain

$$g(\sigma(s_i^{\theta'}, s_{-i}^\theta); h_t) = g(m_\theta^*; h_t).$$

Consider that player i deviates to σ_i'. Then, we have $h_T[\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}), h_t] = h_T[m_i', m_{-i,\theta}^*, h_t]$ for some strategy m_i'. Hence, we obtain

$$g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}); h_t) = g(m_i', m_{-i,\theta}^*; h_t).$$

Since m_θ^* is a subgame perfect equilibrium in the complete information game $\Gamma(\theta)$, we have $g(m_\theta^*; h_t) \succeq_i^\theta g(m_i', m_{-i,\theta}^*; h_t)$. Thus, we also get $g(\sigma(s_i^{\theta'}, s_{-i}^{\theta}); h_t) \succeq_i^\theta g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}); h_t)$. Because by $\mathbf{\Phi 1}$, $\phi_i[(\theta, s_{-i}^{\theta}) \mid s_i^{\theta'}, h_t] = 1$, we can conclude with Assumption 2

$$g(\sigma; h_t) \succeq_i^{\phi_i[\cdot|s_i^{\theta'}, h_t]} g((\sigma_i', \sigma_{-i}); h_t).$$

Consider now the case (2) $h_t \in \mathcal{H}^*$. Note that $h_{t+1} = (h_t, \sigma(h_t, s_i^{\theta'}, s_{-i}^{\theta'})) = (h_t, \sigma(h_t, s_i^{\theta'}, s_{-i}^{\theta})) = (h_t, m_\theta^*(h_t)) = h_{t+1}^* \in \mathcal{H}^*$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_t \in \mathcal{H}^*$. Similar argument can be made inductively so that any subsequent history also falls into $\mathcal{H}^*$. Hence we have $h_T[\sigma(s_i^{\theta'}, s_{-i}^{\theta'}), h_t] = h_T[\sigma(s_i^{\theta'}, s_{-i}^{\theta}), h_t] = h_T[m_\theta^*, h_t]$, which implies

$$g(\sigma(s_i^{\theta'}, s_{-i}^{\theta'}); h_t) = g(\sigma(s_i^{\theta'}, s_{-i}^{\theta}); h_t) = g(m_\theta^*; h_t).$$

Now consider that player i deviates to σ_i'. Let $\hat{t} \geq t$ be the first date at which $\sigma_i'(h_{\hat{t}}, s_i^{\theta'}) \neq \sigma_i(h_{\hat{t}}, s_i^{\theta'})$; or equivalently, $\sigma_i'(h_{\hat{t}}, s_i^{\theta'}) \neq m_{i,\theta}^*(h_{\hat{t}})$. As above, similar argument would show that as long as $t' < \hat{t}$, we have $h_{t'+1} = (h_{t'}, \sigma_i'(h_{t'}, s_i^{\theta'}), \sigma_{-i}(h_{t'}, s_{-i}^{\theta'})) = (h_{t'}, \sigma_i'(h_{t'}, s_i^{\theta'}), \sigma_{-i}(h_{t'}, s_{-i}^{\theta})) = (h_{t'}, m_{i,\theta}^*(h_{t'}), m_{-i,\theta}^*(h_{t'})) \in \mathcal{H}^*$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_{t'} \in \mathcal{H}^*$. In addition, $h_{\hat{t}+1} = (h_{\hat{t}}, \sigma_i'(h_{\hat{t}}, s_i^{\theta'}), \sigma_{-i}(h_{\hat{t}}, s_{-i}^{\theta'})) = (h_{\hat{t}}, \sigma_i'(h_{\hat{t}}, s_i^{\theta'}), \sigma_{-i}(h_{\hat{t}}, s_{-i}^{\theta})) = (h_{\hat{t}}, \sigma_i'(h_{\hat{t}}, s_i^{\theta'}), m_{-i,\theta}^*(h_{\hat{t}}))$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_{\hat{t}} \in \mathcal{H}^*$. Note that $h_{t'} \notin \mathcal{H}_{-k}^*$ for each $k \neq i$ and for $t' \geq \hat{t}+1$. Therefore, using an inductive argument, one can show that $h_{t'+1} = (h_{t'}, \sigma_i'(h_{t'}, s_i^{\theta'}), \sigma_{-i}(h_{t'}, s_{-i}^{\theta'})) = (h_{t'}, \sigma_i'(h_{t'}, s_i^{\theta'}), \sigma_{-i}(h_{t'}, s_{-i}^{\theta})) = (h_{t'}, \sigma_i'(h_{t'}, s_i^{\theta'}), m_{-i,\theta}^*(h_{t'}))$ where the second and third equalities are assured by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ and we use the fact that $h_{t'} \notin \mathcal{H}_{-k}^*$ for each $k \neq i$. So we get $h_T[\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}), h_t] = h_T(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}), h_t] = h_T[m_i', m_{-i,\theta}^*, h_t]$ for some strategy m_i', which implies

$$g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}); h_t) = g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t) = g(m_i', m_{-i,\theta}^*; h_t). \tag{1}$$

Here again, since m_θ^* is a subgame perfect equilibrium in the complete information game $\Gamma(\theta)$, we have $g(m_\theta^*; h_t) \succeq_i^\theta g(m_i', m_{-i,\theta}^*; h_t)$. Thus, we also get

$$g(\sigma(s_i^{\theta'}, s_{-i}^{\theta}); h_t) \succeq_i^\theta g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta}); h_t). \tag{2}$$

The above preference relation together with (1) also implies

$$g(\sigma(s_i^{\theta'}, s_{-i}^{\theta'}); h_t) \succeq_i^{\theta} g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t).$$

Since $g(\sigma(s_i^{\theta'}, s_{-i}^{\theta'}); h_t) = g(m_\theta^*; h_t^*) = a$ and we have assumed that θ and θ' are two states satisfying $(*)$, we get that

$$g(\sigma(s_i^{\theta'}, s_{-i}^{\theta'}); h_t) \succeq_i^{\theta'} g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t). \tag{3}$$

Now since by $\Phi\mathbf{2}$, $\phi_i[\cdot \mid s_i^{\theta'}, h_t]$ assigns a strictly positive weight only to $(\theta, s_{-i}^{\theta})$ and $(\theta', s_{-i}^{\theta'})$, Assumption 2 together with (2) and (3) yields:

$$g(\sigma, h_t) \succeq_i^{\phi_i[\cdot|s_i^{\theta'}, h_t]} g((\sigma_i', \sigma_{-i}), h_t).$$

Finally consider the case (3) $h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^*$. Since $h_t \in \mathcal{H}_{-i}^*$ and $h_t \notin \mathcal{H}^*$ (only i has deviated up to t), there must exist $t' < t$ such that $\sigma_i(h_{t'}, s_i^{\theta'}) \neq m_{i,\theta}^*(h_{t'})$ where $h_{t'}$ is a truncation of history h_t. Then, for any history $h_{t''}$ following $h_{t'}$ (and so, in particular, following h_t), we have $h_{t''} \notin \mathcal{H}_{-k}^*$ for each $k \neq i$. Moreover, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$ we have $\sigma_{-i}(h_{t''}, s_{-i}^{\theta}) = \sigma_{-i}(h_{t''}, s_{-i}^{\theta'}) = m_{-i,\theta}^*(h_{t''})$. Otherwise stated, we have $\sigma_{-i}(s_{-i}^{\theta}) \mid_{h_t} = \sigma_{-i}(s_{-i}^{\theta'}) \mid_{h_t} = m_{-i,\theta}^* \mid_{h_t}$. By $\Phi\mathbf{2}$ we know that $\phi_i[\cdot \mid s_i^{\theta'}, h_t] = \nu^{\varepsilon}(\cdot \mid s_i^{\theta'})$ assigns a strictly positive weight only to $(\theta, s_{-i}^{\theta})$ and $(\theta', s_{-i}^{\theta'})$. In addition, we have just shown that for any $h \in \mathcal{H}^*$ or $h \notin \mathcal{H}_{-i}^* : \sigma_i(h, s_i^{\theta'}) = m_{i,\theta}^*(h, s_i^{\theta'})$. Since $h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^*$, we conclude with $\Sigma\mathbf{2}$

$$g((\sigma_i, \sigma_{-i}); h_t) \succeq_i^{\nu^{\varepsilon}(\cdot|s_i^{\theta'})} g((\sigma_i', \sigma_{-i}); h_t)$$

for any σ_i' that differs from σ_i only at h_t. By the one-shot deviation principle for sequential equilibria[14], the above is equivalent to

$$g((\sigma_i, \sigma_{-i}); h_t) \succeq_i^{\nu^{\varepsilon}(\cdot|s_i^{\theta'})} g((\sigma_i', \sigma_{-i}); h_t)$$

for any σ_i'. This completes the proof. ■

4.2.2 Consistency

In this section, we show that for some $\phi \in \Phi$, (ϕ, σ) satisfies consistency.

[14] See for instance, Hendon, Jacobsen and Sloth (1996).

To show this part, we first fix σ as defined above and consider the following sequence $\{(\phi^k, \sigma^k)\}_{k=0}^{\infty}$ of assessments. Let $\eta_k > 0$ for each k and $\eta_k \to 0$ as $k \to \infty$. For each player i, $h_t \in \mathcal{H}$, and signal s_i, let $\xi_i(h_t, s_i, \cdot)$ be any strictly positive prior over $M_i(h_t)\backslash\{\sigma_i(s_i, h_t)\}$ and define σ_i^k as

$$\sigma_i^k(m_i^t \mid h_t, s_i^{\theta'}) = \begin{cases} 1 - \eta_k^{t \times n} \text{ if } m_i^t = \sigma_i(h_t, s_i^{\theta'}); \\ \eta_k^{t \times n} \times \xi_i(h_t, s_i^{\theta'}, m_i^t) \text{ otherwise} \end{cases}$$

and for any signal $s_i \neq s_i^{\theta'}$:

$$\sigma_i^k(m_i^t \mid h_t, s_i) = \begin{cases} 1 - \eta_k \text{ if } m_i^t = \sigma_i(h_t, s_i); \\ \eta_k \times \xi_i(h_t, s_i, m_i^t) \text{ otherwise} \end{cases}.$$

Let ϕ^k be the unique Bayes consistent belief associated with each σ^k. It is easy to check that σ^k converges uniformly to σ and we also have that ϕ^k converges[15]. Let $\phi \equiv \lim_{k\to\infty} \phi^k$. In the sequel, we show that ϕ satisfies **Φ1, Φ2** and **Φ3**. This will show that (ϕ, σ) satisfies consistency, and $\phi \in \Phi$ as claimed.

To do so, we explicitly compute each ϕ^k and study its limit as k tends to infinity. In general for each $(\tilde{\theta}, \tilde{s}_{-i}) \in \Theta \times S_{-i}$, each $h_t = (m^1, ..., m^{t-1}) \in \mathcal{H}$, and each $\tilde{s}_i \in S_i$, we have

$$\phi_i^k[(\tilde{\theta}, \tilde{s}_{-i}) \mid \tilde{s}_i, h_t] = \frac{\nu^{\varepsilon}(\tilde{\theta}, \tilde{s}_{-i}, \tilde{s}_i) \times \prod_{t'=1}^{t-1} \left[\sigma^k(m^{t'} \mid h_{t'}, \tilde{s})\right]}{\sum\limits_{(\theta', s'_{-i})} \nu^{\varepsilon}(\theta', s'_{-i}, \tilde{s}_i) \times \prod_{t'=1}^{t-1} \left[\sigma^k(m^{t'} \mid h_{t'}, s'_{-i}, \tilde{s}_i)\right]}.$$

In the above formula for each $t' \leq t$, $h_{t'}$ stands for the truncation of h_t to the first t' elements i.e., $h_{t'} = (m^1, ..., m^{t'-1})$.

Claim 3 *ϕ satisfies* **Φ1.**

Proof of Claim 3: Consider player i, $h_t \notin \mathcal{H}_{-i}^*$. First, we will establish the following lemma.

Lemma 2 *Fix player i and assume that $h_t = (\emptyset, m^1, ..., m^{t-1}) \notin \mathcal{H}_{-i}^*$. For all $j \neq i$, let $s_j \in \{s_j^{\theta}, s_j^{\theta'}\}$.*

(1) There exists $\hat{j} \neq i$ and $\hat{t} \leq t-1$ such that $\sigma_{\hat{j}}(h_{\hat{t}}, s_{\hat{j}}) \neq m_{\hat{j}}^{\hat{t}}$;

(2) If $h_t \in \mathcal{H}_{-l}^*$ for some $l \neq i$, then there exists $\hat{t} \leq t-1$ such that $\sigma_l(h_{\hat{t}}, s_l) \neq m_l^{\hat{t}}$.

[15]As will become clear from the proof, the sequence $\{\phi^k\}_k$ does converge.

Proof of Lemma 2: (1) Assume, on the contrary, that $\sigma_{-i}(h_{t'}, s_{-i}) = m_{-i}^{t'}$ for all $t' \leq t-1$. We then show by induction that for all $t' \leq t$, $h_{t'} \in \mathcal{H}_{-i}^*$, which yields a contradiction. Let $t' = 1$; in this case, $h_1 = \emptyset \in \mathcal{H}^* \subseteq \mathcal{H}_{-i}^*$. Now, toward an induction, assume that $h_{t'-1} \in \mathcal{H}_{-i}^*$ and let us show that $h_{t'} \in \mathcal{H}_{-i}^*$. It is easy to show that $h_{t'-1} \in \mathcal{H}_{-i}^*$ implies that either $h_{t'-1} \in \mathcal{H}^*$ (i.e., no player has deviated) or $h_{t'-1} \notin \mathcal{H}_{-j}^*$ for all $j \neq i$ (i.e., only i has deviated). However, in either case, $\sigma_{-i}(h_{t'-1}, s_{-i}) = m_{-i,\theta}^*(h_{t'-1})$ is obtained by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$. Since we have assumed that $\sigma_{-i}(h_{t'-1}, s_{-i}) = m_{-i}^{t'-1}$, we get $m_{-i}^{t'-1} = m_{-i,\theta}^*(h_{t'-1})$, which proves that $h_{t'} = (h_{t'-1}, (\tilde{m}_i(h_{t'-1}), m_{-i,\theta}^*(h_{t'-1}))$ for some strategy $\tilde{m}_i$ and so $h_{t'} \in \mathcal{H}_{-i}^*$. This is a contradiction as desired. (2) Since $h_t \in \mathcal{H}_{-l}^*$, we have that, for all $j \neq l$ and all $t' \leq t-1$, $m_j^{t'} = m_{j,\theta}^*(h_{t'})$. Since $h_t \notin \mathcal{H}_{-i}^*$, we must have that $h_t \in \mathcal{H}_{-l}^* \backslash \mathcal{H}^*$. Let $\tilde{t} \leq t-1$ be the first date at which $m_l^{\tilde{t}} \neq m_{l,\theta}^*(h_{\tilde{t}})$. By construction, we have that for all $t' \leq \tilde{t}$, $h_{t'} \in \mathcal{H}^*$ while for all $t' > \tilde{t}$, $h_{t'} \notin \mathcal{H}_{-j}^*$ for all $j \neq l$. This implies that for all $j \neq l$ and $t' \leq t-1$, we have $\sigma_j(h_{t'}, s_j) = m_{j,\theta}^*(h_{t'})$ by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$. This further implies that for all $j \neq l$ and $t' \leq t-1$, $\sigma_j(h_{t'}, s_j) = m_j^{t'}$. As we already proved in (1), we must have the existence of $\hat{t} \leq t-1$ such that $\sigma_l(h_{\hat{t}}, s_l) \neq m_l^{\hat{t}}$, as claimed. ■

The rest of the proof is reduced to checking the following two cases:

Case 1: $s_i = s_i^{\theta'}$. Recall that $\nu^\varepsilon(\cdot, s_i^{\theta'})$ assigns a weight strictly positive only to $(\theta', s_{-i}^{\theta'})$ and $(\theta, s_{-i}^{\theta})$. Hence,

$$\begin{aligned}
&\phi_i^k[(\theta, s_{-i}^\theta) \mid s_i^{\theta'}, h_t] \\
&= \frac{\nu^\varepsilon(\theta, s_{-i}^\theta, s_i^{\theta'}) \times \prod_{j\neq i}\left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^\theta)\right]}{\nu^\varepsilon(\theta, s_{-i}^\theta, s_i^{\theta'}) \times \prod_{j\neq i}\left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^\theta)\right] + \nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'}) \times \prod_{j\neq i}\left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}, h_{t'}, s_j^{\theta'})\right]} \\
&= \frac{\nu^\varepsilon(\theta, s_{-i}^\theta, s_i^{\theta'})}{\nu^\varepsilon(\theta, s_{-i}^\theta, s_i^{\theta'}) + \nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'}) \times \dfrac{\prod_{j\neq i}[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} | h_{t'}, s_j^{\theta'})]}{\prod_{j\neq i}[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} | h_{t'}, s_j^{\theta})]}}
\end{aligned}$$

We now show that the ratio $\prod_{j\neq i}[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta'})] \Big/ \prod_{j\neq i}[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta})]$ tends to 0 as k tends to infinity. This will show that $\phi_i^k[(\theta, s_{-i}^\theta) \mid s_i^{\theta'}, h_t] \to 1$ and $\phi_i^k[(\theta', s_{-i}^{\theta'}) \mid s_i^{\theta'}, h_t] \to 0$.

By construction of σ^k, Lemma 2 (1) implies that for some $\hat{j} \neq i$ and $\hat{t} \leq t-1$:

$$\sigma_{\hat{j}}^k(m_{\hat{j}}^{\hat{t}} \mid h_{\hat{t}}, s_{\hat{j}}^{\theta'}) = \eta_k^{t \times n} \xi_{\hat{j}}(h_{\hat{t}}, s_{\hat{j}}^{\theta'}, m_{\hat{j}}^{\hat{t}}). \tag{4}$$

Now, we have:

$$\frac{\prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta'}) \right]}{\prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta}) \right]} \leq \frac{\eta_k^{t \times n} \times \xi_{\hat{j}}(h_{\hat{t}}, s_{\hat{j}}^{\theta'}, m_{\hat{j}}^{\hat{t}}) \times 1}{\prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \eta_k \xi_j(h_{t'}, s_j^{\theta}, m_j^{t'}) \right]}$$

$$\leq \frac{\eta_k^{t \times n}}{\eta_k^{(t-1)(n-1)}} \times \frac{\xi_{\hat{j}}(h_{\hat{t}}, s_{\hat{j}}^{\theta'}, m_{\hat{j}}^{\hat{t}})}{\prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \xi_j(h_{t'}, s_j^{\theta}, m_j^{t'}) \right]} = \eta_k^{t+n-1} \times \frac{\xi_{\hat{j}}(h_{\hat{t}}, s_{\hat{j}}^{\theta'}, m_{\hat{j}}^{\hat{t}})}{\prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \xi_j(h_{t'}, s_j^{\theta}, m_j^{t'}) \right]} \to 0 \ \text{(as } k \to \infty\text{)}.$$

Where the first inequality is assured by (4) and (assuming wlog that η_k is small) we use the very construction that, for all j and $t' \leq t-1$, $\sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta}) \geq \eta_k \times \xi_j(h_{t'}, s_j^{\theta}, m_j^{t'})$.

Case 2: $s_i = s_i^{\theta}$. Recall that $\nu^{\varepsilon}(\cdot, s_i^{\theta})$ assigns a weight strictly positive only to $(\theta, s_{-i}^{\theta})$ and (θ, τ_l) for each $l \neq i$. Hence,

$$\begin{aligned}
&\phi_i^k[(\theta, \tau_l) \mid s_i^{\theta}, h_t] \\
&= \frac{\nu^{\varepsilon}(\theta, \tau_l) \times \prod_{j \neq l, i} \left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta}) \right] \times \left[\prod_{t'=1}^{t-1} \sigma_l^k(m_l^{t'} \mid h_{t'}, s_l^{\theta'}) \right]}{\left\{ \begin{array}{c} \sum_{z \neq i} \nu^{\varepsilon}(\theta, \tau_z) \times \prod_{j \neq z, i} \left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'} \mid h_{t'}, s_j^{\theta}) \right] \times \left[\prod_{t'=1}^{t-1} \sigma_z^k(m_z^{t'} \mid h_{t'}, s_z^{\theta'}) \right] \\ + \nu^{\varepsilon}(\theta, s_{-i}^{\theta}, s_i^{\theta}) \times \prod_{j \neq i} \left[\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}, h_{t'}, s_j^{\theta}) \right] \end{array} \right\}} \\
&= \frac{\nu^{\varepsilon}(\theta, \tau_l)}{\sum_{z \neq i} \nu^{\varepsilon}(\theta, \tau_z) \times c_z + \nu^{\varepsilon}(\theta, s_{-i}^{\theta}, s_i^{\theta}) \times \frac{\prod_{t'=1}^{t-1} \sigma_l^k(m_l^{t'}, h_{t'}, s_l^{\theta})}{\prod_{t'=1}^{t-1} \sigma_l^k(m_l^{t'} \mid h_{t'}, s_l^{\theta'})}}
\end{aligned}$$

for some positive numbers c_z. We now show that if $h_t \in \mathcal{H}_{-l}^*$, then the ratio $\prod_{t'=1}^{t-1} \sigma_l^k(m_l^{t'}, h_{t'}, s_l^{\theta}) / \prod_{t'=1}^{t-1} \sigma_l^k(m_l^{t'} \mid h_{t'}, s_l^{\theta'})$ tends to ∞ as k tends to infinity. This will show that $\phi_i^k[(\theta, \tau_l) \mid s_i^{\theta}, h_t] \to 0$ for all l such that $h_t \in \mathcal{H}_{-l}^*$; and hence that ϕ satisfies **Φ1.** Assume that $h_t \in \mathcal{H}_{-l}^*$ for some l, by construction of σ^k, Lemma 2 (2) implies that there exists $\hat{t} \leq t-1$ such that $\sigma_l(h_{\hat{t}}, s_l) \neq m_l^{\hat{t}}$ and so:

$$\sigma_l^k(m_l^{\hat{t}} \mid h_{\hat{t}}, s_l^{\theta'}) = \eta_k^{\hat{t} \times n} \xi_l(h_{\hat{t}}, s_l^{\theta'}, m_l^{\hat{t}}). \tag{5}$$

Now, we have

$$\frac{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'},h_{t'},s_l^\theta)}{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})} \geq \frac{\eta_k^{t-1}\prod_{t'=1}^{t-1}\times\xi_l(h_{t'},s_l^\theta,m_l^{t'})}{\eta_k^{t\times n}\xi_l(h_{\hat{t}},s_l^{\theta'},m_l^{\hat{t}})\times 1}$$

$$= \frac{1}{\eta_k^{t(n-1)+1}}\times\frac{\prod_{t'=1}^{t-1}\xi_l(h_{t'},s_l^\theta,m_l^{t'})}{\xi_l(h_{\hat{t}},s_l^{\theta'},m_l^{\hat{t}})}\to\infty \text{ (as } k\to\infty).$$

Where the first inequality is assured by (5) and (assuming wlog that η_k is small) we use the fact that by construction: for all $t'\leq t-1 : \sigma_l^k(m_j^{t'},h_{t'},s_l^\theta)\geq\eta_k\times\xi_l(h_{t'},s_l^\theta,m_l^{t'})$. ■

Claim 4 *ϕ satisfies* Φ**2.**

Proof of Claim 4: Consider player i, $h_t\in\mathcal{H}_{-i}^*$. The following lemma will be useful.

Lemma 3 *Fix player i and assume that $h_t=(\emptyset,m^1,...,m^{t-1})\in\mathcal{H}_{-i}^*$. For all $j\neq i$, let $s_j\in\{s_j^\theta,s_j^{\theta'}\}$. For all $j\neq i$ and $t'\leq t-1 : \sigma_j(h_{t'},s_j)=m_j^{t'}$.*

Proof of Lemma 3: Pick any $t'\leq t-1$ and note that $h_{t'}\in\mathcal{H}_{-i}^*$. Hence, it must be that either $h_{t'}\in\mathcal{H}^*$ or $h_{t'}\notin\mathcal{H}_{-j}^*$ for all $j\neq i$. In each of these cases, by Σ**1** and Σ**3**, we have for all $j\neq i : \sigma_j(h_{t'},s_j)=m_{j,\theta}^*(h_{t'})$. Since $h_{t'}\in\mathcal{H}_{-i}^*$, we have that, for all $j\neq i$, $m_j^{t'}=m_{j,\theta}^*(h_{t'})$, which completes the proof. ■

Here again, the rest of the proof is reduced to checking the following two cases.

Case 1: $s_i=s_i^{\theta'}$. Recall that $\nu^\varepsilon(\cdot,s_i^{\theta'})$ assigns a weight strictly positive only to $(\theta',s_{-i}^{\theta'})$ and (θ,s_{-i}^θ). Hence,

$$\begin{aligned}
&\phi_i^k[(\theta',s_{-i}^{\theta'})\mid s_i^{\theta'},h_t]\\
&= \frac{\nu^\varepsilon(\theta',s_{-i}^{\theta'},s_i^{\theta'})\times\prod_{j\neq i}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta'})\right]}{\nu^\varepsilon(\theta',s^{\theta'})\times\prod_{j\neq i}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta'})\right]+\nu^\varepsilon(\theta,s_{-i}^\theta,s_i^{\theta'})\times\prod_{j\neq i}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})\right]}\\
&= \frac{\nu^\varepsilon(\theta',s_{-i}^{\theta'},s_i^{\theta'})}{\nu^\varepsilon(\theta',s^{\theta'})+\nu^\varepsilon(\theta,s_{-i}^\theta,s_i^{\theta'})\times\frac{\prod_{j\neq i}[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^\theta)]}{\prod_{j\neq i}[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta'})]}}
\end{aligned}$$

We now show that the ratio $\prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})/\prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta'})$ tends to 1 as k tends to infinity. This will show that $\phi_i^k[(\theta',s_{-i}^{\theta'})\mid s_i^{\theta'},h_t]\to\nu^{\varepsilon}((\theta',s_{-i}^{\theta'})\mid s_i^{\theta'})$ and $\phi_i^k[(\theta,s_{-i}^{\theta})\mid s_i^{\theta'},h_t]\to\nu^{\varepsilon}((\theta,s_{-i}^{\theta})\mid s_i^{\theta'})$.

By construction of σ^k, Lemma 3 implies that for all $j\neq i$ and $t'\leq t-1$:

$$\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})=1-\eta_k \text{ and } \sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta'})=1-\eta_k^{t\times n}$$

Thus,

$$\prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'},h_{t'},s_j^{\theta'})\Big/\prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'},h_{t'},s_j^{\theta})\to 1 \text{ as } k\to\infty$$

Case 2: $s_i=s_i^{\theta}$. Recall that $\nu^{\varepsilon}(\cdot,s_i^{\theta})$ assigns a weight strictly positive only to (θ,s_{-i}^{θ}) and (θ,τ_l) for $l\neq i$. Hence,

$$\begin{aligned}
&\phi_i^k[(\theta,s_{-i}^{\theta})\mid s_i^{\theta},h_t]\\
=&\frac{\nu^{\varepsilon}(\theta,s_{-i}^{\theta},s_i^{\theta})\times\prod_{j\neq i}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})\right]}{\left\{\begin{array}{c}\nu^{\varepsilon}(\theta,s_{-i}^{\theta},s_i^{\theta})\times\prod_{j\neq i}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})\right]\\ +\sum_{l\neq i}\nu^{\varepsilon}(\theta,\tau_l)\times\prod_{j\neq i,l}\left[\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}\mid h_{t'},s_j^{\theta})\right]\times\left[\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})\right]\end{array}\right\}}\\
=&\frac{\nu^{\varepsilon}(\theta,s_{-i}^{\theta},s_i^{\theta})}{\nu^{\varepsilon}(\theta,s_{-i}^{\theta},s_i^{\theta})+\sum_{l\neq i}\nu^{\varepsilon}(\theta,\tau_l)\times\frac{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})}{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta})}}
\end{aligned}$$

We now show that for each $l\neq i$, the ratio $\prod_{t'=1}^{t-1}[\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})]\Big/\prod_{t'=1}^{t-1}[\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta})]$ tends to 1 as k tends to infinity. This will show that $\phi_i^k[(\theta,s_{-i}^{\theta})\mid s_i^{\theta},h_t]\to\nu^{\varepsilon}((\theta,s_{-i}^{\theta})\mid s_i^{\theta})$ and similar reasoning shows that for each $l\neq i$: $\phi_i^k[(\theta,\tau_l)\mid s_i^{\theta},h_t]\to\nu^{\varepsilon}((\theta,\tau_l)\mid s_i^{\theta})$; and hence, ϕ satisfies Φ**2.**

By construction of σ^k, Lemma 3 implies that for all $l\neq i$ and $t'\leq t-1$:

$$\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta})=1-\eta_k \text{ and } \sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})=1-\eta_k^{t\times n}$$

Thus,

$$\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta'})\Big/\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}\mid h_{t'},s_l^{\theta})\to 1 \text{ as } k\to\infty$$

$\blacksquare$

Finally, observing that for $s_i^{\tilde{\theta}} \notin \{s_i^{\theta}, s_i^{\theta'}\}$, $\nu^{\varepsilon}(\cdot, s_i)$ assigns a weight one to $(\tilde{\theta}, s_{-i}^{\tilde{\theta}})$, we establish the claim below:

Claim 5 *ϕ satisfies* **Φ3.**

5 Concluding Remarks

In this paper, we prove a necessary condition result analogous to Chung and Ely (2003) focusing on subgame perfect implementation while Chung and Ely (2003) focused on undominated Nash implementation. It is natural to check what strengthening of Maskin's monotonicity would ensure $\overline{SPE}$-implementation. Given that we will have to assume monotonicity, there is probably very little gain to build a sequential mechanism, a static one would most likely be enough. We conjecture that a simple condition of continuity on preferences as well as the usual no-veto-power condition would ensure $\overline{SPE}$-implementation of any monotonic social choice function with more than two players (in one stage). More generally, we believe that full-implementation in (strict) Nash equilibrium together with some continuity requirement on preferences is enough for $\overline{SPE}$-implementation (in one stage).

References

[1] D. Abreu and A. Sen (1990) "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition", *Journal of Economic Theory*, 50, 285-299

[2] Aghion P., D. Fudenberg, R. and Holden (2007) "Subgame Perfect Implementation With Almost Perfect Information," working paper, available at http://www.economics.harvard.edu/faculty/aghion/files/Subgame_Perfect.pdf

[3] Bergemann, D. and S. Morris, (2005) "Robust Mechanism Design," *Econometrica*, 73, 1521-1534

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

[5] J. Duggan (1997) "Virtual Bayesian Implementation," *Econometrica*, 65, 1175-1199

[6] D. Fudenberg and J. Tirole (1991) "*Game Theory*," MIT Press.

[7] S.J. Grossman and O.D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691-719.

[8] O. Hart and J. Moore (1990) "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119-1158.

[9] Hendon, E., H.J. Jacobsen and B. Sloth, (1996) "The One-Shot Deviation Principle for Sequential Rationality," *Games and Economic Behavior*, 12, 274-282

[10] Kreps, D.M. and R. Wilson (1982) "Sequential Equilibria," *Econometrica*, 50, 863-894

[11] Kunimoto, T (2008) "How Robust is Undominated Nash Implementation?," *mimeo* available at http://people.mcgill.ca/files/takashi.kunimoto/robustness-Dec-9-2008.pdf.

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

[13] Maskin, E. and J. Tirole (1999) "Unforseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66, 83-114.

[14] E. Miyagawa (2002) "Subgame-Perfect Implementation of Bargining Solutions," *Games and Economic Behavior*, 41, 292-302

[15] Monderer, D. and D. Samet (1989) "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 170-190.

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

[17] Oury, M. and O. Tercieux (2009), "Continuous Implementation," working paper, available at http://www.pse.ens.fr/tercieux/ContUTP.pdf

[18] Palfrey, T.R., and S. Srivastava (1991), "Nash Implementation Using Undominated Strategies," *Econometrica*, 59, 479-501.

[19] H. Vartiainen (2007), "Subgame Perfect Implementation: A Full Characterization," *Journal of Economic Theory*, 133, 111-126.