Online Appendix to "Subgame Perfect Implementation Under Information Perturbations"

Philippe Aghion, Drew Fudenberg, Richard Holden, Takashi Kunimoto and Olivier Tercieux

A Proof of Theorem 2

We fix an SCF f which is implemented under complete information by the MR mechanism Γ^{MR}. We let μ be a complete information prior and show that for the sequence of priors ν^ε (indexed by $\varepsilon > 0$) as specified in $(***)$ of Section III.B, there is no sequence of equilibrium strategy profiles converging to truth-telling. Let $\Gamma^{MR}(\nu^\varepsilon)$ be an incomplete information game associated with the MR mechanism and a prior ν^ε. By way of contradiction, assume that for each $\varepsilon > 0$, there exists a profile of mixed equilibrium strategies of the game $\Gamma^{MR}(\nu^\varepsilon)$ such that as ε goes to 0, the probability that both players report their signals truthfully converges to 1. Fix such a sequence of mixed equilibrium strategy profiles. We then use the following notation to describe equilibrium play in the games $\Gamma^{MR}(\nu^\varepsilon)$:

- $\sigma_{k,l}^{\varepsilon,j}$ denotes the probability that player 1 with signal $s_1^{k,l}$ announces θ_1^j at Stage 1 of Phase 1;
- $\lambda_{k,l}^{\varepsilon,j}[\hat{\theta}_1]$ denotes the probability that player 2 with signal $s_2^{k,l}$ announces θ_1^j at Stage 2 of Phase 1 given that at Stage 1 of Phase 1, player 1 has announced $\hat{\theta}_1$
- $\rho_{k,l}^{\varepsilon,j}$ denotes the probability that player 2 with signal $s_2^{k,l}$ announces θ_2^j at Stage 1 of Phase 2; and
- $\tau_{k,l}^{\varepsilon,j}[\hat{\theta}_2]$ denotes the probability that player 1 with signal $s_1^{k,l}$ announces θ_2^j at Stage 2 of Phase 2 given that at Stage 1 of Phase 2, player 2 has announced $\hat{\theta}_2$.

Using the above notation, our hypothesis to derive a contradiction is summarized as follows: for all $k \neq j$, all l and all announcements $\hat{\theta}_1$, $\sigma_{k,l}^{\varepsilon,j}$ and $\lambda_{k,l}^{\varepsilon,j}[\hat{\theta}_1]$ converge to 0 as $\varepsilon \to 0$; and for all $l \neq j$, all k and all announcement $\hat{\theta}_2$, $\rho_{k,l}^{\varepsilon,j}$ and $\tau_{k,l}^{\varepsilon,j}[\hat{\theta}_2]$ converge to 0 as $\varepsilon \to 0$.

We will use the following claim about the properties of the MR mechanism under complete information:

Claim 1. *For truthtelling to be the unique subgame-perfect equilibrium of the MR mechanism under complete information, it must be that for each $\theta = (\theta_1, \theta_2)$ and each ϕ_1,*

$$u_1(f(\theta_1, \theta_2); \theta_1) > U_1(y(\phi_1, \theta_1); \theta_1) - t_{y(\phi_1, \theta_1)} - \Delta, \tag{1}$$

and

$$u_2(f(\theta_1, \theta_2); \theta_2) > U_2(x(\theta_1, \phi_1); \theta_2) + t_{x(\theta_1, \phi_1)} - \Delta. \tag{2}$$

Proof of Claim 1. Suppose first that Inequality (1) goes the other way, that is, for some $\theta = (\theta_1, \theta_2)$ and some ϕ_1, we have

$$u_1(f(\theta_1, \theta_2); \theta_1) < U_1(y(\phi_1, \theta_1); \theta_1) - t_{y(\phi_1,\theta_1)} - \Delta.$$

Then, under complete information where the true state is θ, we claim that truthtelling is not a subgame-perfect equilibrium: player 1 has an incentive to deviate by claiming some $\phi_1 \neq \theta_1$ (and player 2 challenges player 1's report at Stage 2 under truthtelling) in order to reach Stage 3 where he would pick $\{y(\phi_1, \theta_1), t_{y(\phi_1,\theta_1)} + \Delta\}$. This contradicts the hypothesis that truthtelling is a subgame perfect equilibrium of the MR mechanism under complete information.

Now, suppose instead that for some $\theta = (\theta_1, \theta_2)$, and some $\phi_1 \neq \theta_1$, we have

$$u_1(f(\theta_1, \theta_2); \theta_1) = U_1(y(\phi_1, \theta_1); \theta_1) - t_{y(\phi_1,\theta_1)} - \Delta.$$

In this case, we claim that there is a subgame-perfect equilibrium at $\theta = (\theta_1, \theta_2)$ where player 1 does not report truthfully. To see this, we propose the following strategy profile σ^*: At Stage 1 of Phase 1, player 1 reports $\phi_1 \neq \theta_1$; player 2 reports the true state θ_1 at Stage 2 irrespective of player 1's announcement; and at Stage 3, player 1 always plays his optimal action. Note here that player 1's optimal play at Stage 3 depends on what he reported at Stage 1. In Phase 2, both players always report truthfully and player 2 plays his optimal action at Stage 3. Here again, player 2's optimal action at Stage 3 depends on what he reported at Stage 1. Given the continuation strategy profile from Stage 2 induced by σ^*, player 1 is indifferent between reporting ϕ_1 and θ_1 at Stage 1, and so (if truthtelling is a subgame perfect equilibrium) this σ^* is indeed a subgame-perfect equilibrium at $\theta = (\theta_1, \theta_2)$. This contradicts the uniqueness of truthtelling as a subgame perfect equilibrium of the MR mechanism under complete information.

Similarly, we must have that for each $\theta = (\theta_1, \theta_2)$ and each ϕ_1,

$$u_2(f(\theta_1, \theta_2); \theta_2) > U_2(x(\theta_1, \phi_1); \theta_2) + t_{x(\theta_1,\phi_1)} - \Delta.$$

By way of contradiction, we argue why this must be the case. Suppose first that for some $\theta = (\theta_1, \theta_2)$ and some ϕ_1, we have

$$u_2(f(\theta_1, \theta_2); \theta_2) < U_2(x(\theta_1, \phi_1); \theta_2) + t_{x(\theta_1,\phi_1)} - \Delta.$$

Then, under complete information where the true state is $\theta = (\theta_1, \theta_2)$, we claim that truthtelling is not an equilibrium: player 2 has an incentive to deviate by claiming some $\phi_1 \neq \theta_1$ in order to reach stage 3 where player 1 would pick $\{x(\theta_1, \phi_1), t_{x(\theta_1,\phi_1)} + \Delta\}$. This contradicts the hypothesis that truthtelling is a subgame perfect equilibrium of the MR mechanism under complete information.

Now, suppose instead that for some $\theta = (\theta_1, \theta_2)$, and some $\phi_1 \neq \theta_1$, we have

$$u_2(f(\theta_1, \theta_2); \theta_2) = U_2(x(\theta_1, \phi_1); \theta_2) + t_{x(\theta_1,\phi_1)} - \Delta.$$

In this case, we claim that there is a subgame-perfect equilibrium at $\theta = (\theta_1, \theta_2)$ where player 2 does not report truthfully. To see this, we construct the following strategy profile σ^{**}: At Stage 1 of Phase 1, player 1 always reports θ_1 truthfully; player 2 reports a

false state ϕ_1 if player 1 has claimed θ_1 and otherwise challenges with θ_1; and at Stage 3, player 1 always plays his optimal action. Note here that player 1's optimal play at Stage 3 depends on what he reported at Stage 1. In Phase 2, both players always report truthfully and player 2 plays his optimal action at stage 3. Here again, player 2's optimal action at Stage 3 depends on what he reported at Stage 1. Since player 1 would choose $\{x(\theta_1,\phi_1), t_{x(\theta_1,\phi_1)} + \Delta\}$ at Stage 3, player 2 is indifferent between reporting θ_1 and ϕ_1 at Stage 2 after player 1 reported θ_1. This shows that (if truthtelling is a subgame perfect equilibrium) σ^{**} is a subgame perfect equilibrium where player 2 does not report truthfully. However, this contradicts the uniqueness of truthtelling as a subgame perfect equilibrium of the MR mechanism under complete information. This completes the proof of the claim. □

Now, let us fix the prior ν^ε (as defined in $(***)$ of Section III.B. Consider the case where player 1 receives $s_1^{k,l}$. Clearly, $\nu^\varepsilon(\theta_1^k, \theta_2^l, s_2^{k,l} | s_1^{k,l}) \to 1$ as $\varepsilon \to 0$. Hence, at Stage 1, by continuity of expected payoffs with respect to beliefs, the expected equilibrium payoff of player 1 for announcing θ_1^k converges (as ε vanishes) to

$$u_1(f(\theta_1^k, \theta_2^l); \theta_1^k),$$

while if he lies by claiming $\phi_1 \neq \theta_1^k$ at Stage 1, his expected equilibrium payoff converges to something (weakly) smaller than

$$U_1(y(\phi_1, \theta_1^k); \theta_1^k) - t_{y(\phi_1,\theta_1^k)} - \Delta.$$

By Equation (1) and choosing $\varepsilon > 0$ small enough, there is no way that the equilibrium strategies $\{\sigma_{k,l}^{\varepsilon,j}, \lambda_{k,l}^{\varepsilon,j}[\hat{\theta}_1], \rho_{k,l}^{\varepsilon,j}, \tau_{k,l}^{\varepsilon,j}[\hat{\theta}_2]\}_{k,l,j,\hat{\theta}_1,\hat{\theta}_2}$ can make player 1's best response indifferent at Stage 1. Hence, for $\varepsilon > 0$ small enough, player 1 with signal $s_1^{k,l}$ plays pure strategies at Stage 1 of Phase 1. This reasoning holds for an arbitrary choice of $s_1^{k,l}$ so that player 1 plays in pure strategies irrespective of his signal.

Note now that player 1 with signal $s_1^{k,l}$ could deviate and claim that $\theta_1^{k'}$ is the true state where $k' \neq k$. In this case, because player 1 plays in pure strategies (and hence, the equilibrium is fully revealing), in the first phase, after observing $\theta_1^{k'}$, player 2 believes with probability one that player 1 has received a signal of the form $s_1^{k',l'}$ for some l'. We claim that player 2 with signal $s_2^{k,l}$ will not challenge: indeed, by construction of ν^ε, player 2 with signal $s_2^{k,l}$ believes with high probability that $\theta = (\theta_1, \theta_2)$ where $\theta_2 = \theta_2^l$ is the true state. If player 2 challenges with θ_1^k, by construction of ν^ε, he expects player 1 to choose $\{x(\theta_1^{k'}, \theta_1^k), t_{x(\theta_1^{k'},\theta_1^k)} + \Delta\}$ at Stage 3. On the other hand, if he does not challenge, his expected payoff would tend to $u_2(f(\theta_1^{k'}, \theta_2^l); \theta_2^l)$ as ε vanishes. Hence, by Equation (2), for $\varepsilon > 0$ small, player 2 will be better off by not challenging. Thus, we get that $\lambda_{k,l}^{\varepsilon,k}[\theta_1^{k'}] = 0$, which is a contradiction. This completes the proof of Theorem 2.

B Proof of Theorem 3

We first introduce some notation. Given a prior μ over $\Theta \times S$, we write $\mu(\theta)$ for $[\text{marg}_\Theta \mu](\theta)$, and given $s_{-i} \in S_{-i}$, we will write $\mu(s_{-i})$ as $[\text{marg}_{S_{-i}} \mu](s_{-i})$. Finally, given an arbitrary countable space X, δ_x will denote the probability measure that puts probability 1 on $\{x\} \subset X$.

For the sake of completeness, we reproduce the definition of *Maskin monotonitity*: A social choice correspondence (SCC) $\mathcal{F}$ on a payoff relevant state space Θ is Maskin monotonic if for all pair of states of nature θ' and θ'' if $a \in \mathcal{F}(\theta')$ and

$$\{(i,b) \,|u_i(a;\theta') \geq u_i(b;\theta')\} \subseteq \{(i,b) \,|u_i(a;\theta'') \geq u_i(b;\theta'')\} \qquad (*)$$

(i.e., no individual ranks a lower when moving from θ' to θ''), then $a \in \mathcal{F}(\theta'')$.

Let μ be any complete information prior, and assume that a mechanism Γ SPE-implements a non-Maskin monotonic SCC $\mathcal{F}$. By hypothesis $\mathcal{F}$ is not Maskin monotonic, so there are θ', θ'' and $a \in \mathcal{F}(\theta')$ satisfying $(*)$ in the definition of Maskin monotonicity while $a \notin \mathcal{F}(\theta'')$. We now fix this particular θ', θ'' and a throughout.

Since the mechanism Γ SPE-implements $\mathcal{F}$, there exists a pure strategy subgame-perfect equilibrium $m^*_{\theta'}$ in $\Gamma(\theta')$ such that $g(m^*_{\theta'}) = a$. Fix one such equilibrium. Clearly, $m^*_{\theta'}$ is a Nash equilibrium of $\Gamma(\theta')$. From $(*)$ in the definition of Maskin monotonicity, it follows that $m^*_{\theta'}$ 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 $m^*_{\theta'}$ up to date t and let $\mathcal{H}^*$ denote the set of all such histories for any t. 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 $t \geq 1$.

Consider the following family of information structures ν^ε. For each player i, let τ_i represent the profile of signals $s = (s_1, ..., s_n)$ defined by $s_i = s_i^{\theta''}$ and $s_j = s_j^{\theta'}$ for all $j \neq i$. For all i, ν^ε is given by[1]

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

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

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

Before we prove Theorem 3, we introduce some notation and the formal definition of 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 (θ, s_{-i}) is realized when agent i's signal is s_i and the observed history is h_t. We will henceforth abuse notation 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 (1991a) for its precise definition). An assessment is a pair (ϕ, σ) consisting of a profile of beliefs and a pure behavior strategy profile. We formally define sequential equilibrium.

[1]This sequence of perturbations is similar to that used by Chung and Ely (2003). However, because sequential equilibrium requires verifying sequential rationality conditions that are not imposed by undominated Nash equilibrium, the body of proof is very different from that in Chung and Ely (2003).

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

$$\sum_{(\theta, s_{-i}) \in \Theta \times S_{-i}} \phi_i[\theta, s_{-i} | s_i, h_t] \left\{ u_i(g(\sigma(s); h_t); \theta) - u_i(g((\sigma_i'(s_i), \sigma_{-i}(s_{-i})); h_t); \theta) \right\} \geq 0$$

for each σ_i'.
(C) Consistency: *there exists a sequence of totally mixed strategy profiles $(\sigma_1^k, ..., \sigma_n^k)$ converging to $(\sigma_1, ..., \sigma_n)$ with Bayes consistent beliefs ϕ^k converging to ϕ.*[2]

Now we come back to the proof and in particular, build a sequential equilibrium $(\phi^\varepsilon, \sigma^\varepsilon)$ of $\Gamma(\nu^\varepsilon)$ where $g(\sigma^\varepsilon(s^{\theta''}); \emptyset) = a$ for each $\varepsilon > 0$ small enough. This will show that there exist a sequence of priors $\{\nu^\varepsilon\}_{\varepsilon > 0}$ that converges to μ and a corresponding sequence of sequential equilibria $\{(\phi^\varepsilon, \sigma^\varepsilon)\}_{\varepsilon > 0}$ such that $g(\sigma^\varepsilon(s^{\theta''}); \emptyset) \to a \notin \mathcal{F}(\theta'')$ as ε goes to 0. This will complete the proof.

In the sequel, we will omit the dependence of σ^ε with respect to ε and simply write σ for σ^ε. In the following lines, we define a strategy profile σ and a family of systems of beliefs Φ so that $g(\sigma(s^{\theta''}); \emptyset) = a$. In addition, we will show that (ϕ, σ) is a sequential equilibrium of $\Gamma(\nu^\varepsilon)$ for some $\phi \in \Phi$. We define Φ and σ as follows:

Definition of σ:

Σ1. For any player i and any $h_t \in \mathcal{H}^*$ or $h_t \notin \mathcal{H}_{-i}^*$, $\sigma_i(h_t, s_i^{\theta''}) = m_{i,\theta'}^*(h_t)$;[3]

Σ2. For any player i, any $h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^*$, $\sigma_i(h_t, s_i^{\theta''}) = \bar{m}_i(h_t)$ where $\bar{m}_i$ satisfies for any h_t,

$$\begin{aligned} h_t \in \mathcal{H}^* \text{ or } h_t \notin \mathcal{H}_{-i}^* \quad &\Rightarrow \quad \bar{m}_i(h_t) = m_{i,\theta'}^*(h_t); \\ h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^* \quad &\Rightarrow \quad \bar{m}_i(h_t) \in \arg\max \sum_{\tilde{\theta}} \nu^\varepsilon(\tilde{\theta} | s_i^{\theta''}) u_i(g((m_i', m_{-i,\theta'}^*); h_t); \tilde{\theta}), \end{aligned}$$

where the max is taken over all pure messages $m_i' \in M_i$ that differs from $\bar{m}_i$ only at h.[4] By A1 there exists such $\bar{m}_i$**;**

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

[2]Convergence in the definition of consistency is taken uniformly over messages and histories. Given that the set of messages (and so the set of histories) 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.

[3]Note that players here send the messages that m prescribes for state θ' when their signal suggests that the state is θ''.

[4]Note that the maximization above is over all pure messages $m_i' \in M_i$ that differs from $\bar{m}_i$ only at h. Hence, since player i may be playing at several stages, it might be the case that this maximization depends on what player i is playing at further histories, and these further histories may be outside $\mathcal{H}_{-i}^* \backslash \mathcal{H}^*$ (for instance in case a player j different of i does not play according to $m_{j,\theta'}^*$ at some subsequent history). This is why we also have to define $\bar{m}_i$ outside $\mathcal{H}_{-i}^* \backslash \mathcal{H}^*$.

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 pure strategy subgame-perfect equilibrium of $\Gamma(\tilde{\theta})$. (This is well-defined since $\mathcal{F}$ is implementable in subgame-perfect equilibrium under complete information.)

Definition of Φ:

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

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

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

and

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

and for all $l \neq i$ with $h_t \in \mathcal{H}^*_{-l} \backslash \mathcal{H}^*_{-i}$

(i.e., player l has deviated from the path prescibed by $m^*_{\theta'}$)

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

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 $s_i^{\tilde{\theta}} \notin \{s_i^{\theta'}, s_i^{\theta''}\}$, $\phi_i \left[\cdot \middle| s_i^{\tilde{\theta}}, h_t \right] = \delta_{(\tilde{\theta}, s^{\tilde{\theta}}_{-i})}$ where δ_x denotes the probability measure that puts probability 1 on $\{x\}$.

Note that $h_T[\sigma(s^{\theta''}), \emptyset] = h_T[m^*_{\theta'}, \emptyset]$ and so, σ generates $g(\sigma(s^{\theta''}); \emptyset) = g(m^*_{\theta'}; \emptyset) = a$. Hence, it only remains to show that (ϕ, σ) constitutes a sequential equilibrium for some $\phi \in \Phi$. In Section B.1, we show that (ϕ, σ) satisfies sequential rationality for any $\phi \in \Phi$; and we establish that (ϕ, σ) satisfies consistency for some $\phi \in \Phi$ in Section B.2.

B.1 Sequential rationality

Fix any $\phi \in \Phi$. Sequential rationality of (ϕ, σ) will be proved by Claims 2 and 3 below.

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

$$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i}) | s_i, h_t] \left[u_i(g(\sigma(s); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s_i), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta}) \right] \geq 0$$

for each σ'_i.

Claim 2 states that for any player i with any signal $s_i \neq s_i^{\theta''}$, σ_i is a best response to σ_{-i} given his belief ϕ_i. This will be checked by considering three classes of histories: (1) Histories where all players have played according to the equilibrium $m^*_{\theta'}$ (i.e., in $\mathcal{H}^*$); (2) histories where player i has not played according to $m^*_{i,\theta'}$ but all other players have (i.e., in $\mathcal{H}^*_{-i}\backslash\mathcal{H}^*$); and finally (3) histories where some player other than i has not played according to $m^*_{\theta'}$ (i.e., outside $\mathcal{H}^*_{-i}$).

In particular, in the non-trivial case where $s_i = s_i^{\theta'}$, we will show that for any of these histories h_t, whenever player i follows σ_i against σ_{-i}, player i believes with probability one that the outcome will be given by $g(m^*_{\theta'}; h_t)$, while if player i deviates from $\sigma_i(s_i)$ to some m_i' , player i believes with probability one that the outcome will be given by $g(m_i', m^*_{-i,\theta'}; h_t)$. Because $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$ and player i with signal $s_i^{\theta'}$ believes with probability one that θ' is the true state, this will prove the claim.

Proof of Claim 2. Fix any player i. This claim is obvious for $s_i^{\tilde{\theta}} \neq s_i^{\theta'}$ because by $\mathbf{\Phi 3}$, $\phi_i\left[\cdot\middle|s_i^{\tilde{\theta}}, h_t\right] = \delta_{(\tilde{\theta}, s_{-i}^{\tilde{\theta}})}$ 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, $\nu^{\varepsilon}(\theta'|s_i^{\theta'}) = 1$ and so this player knows the state is θ', and he knows the profile of signals is either $s^{\theta'}$ or τ_k for some $k \neq i$. We partition the set of all histories into three classes $\mathcal{H}^*$; $\mathcal{H}^*_{-i}\backslash\mathcal{H}^*$ and $\mathcal{H}\backslash\mathcal{H}^*_{-i}$ and consider the following three cases: Case (1) $h_t \in \mathcal{H}^*$; Case (2) $h_t \in \mathcal{H}^*_{-i}\backslash\mathcal{H}^*$; and Case (3) $h_t \notin \mathcal{H}^*_{-i}$.

- Case (1): $h_t \in \mathcal{H}^*$

  In this case, each player has played according to $m^*_{\theta'}$ and if players $j \neq i$ received signals of either $s_j^{\theta'}$ or $s_j^{\theta''}$, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, this will continue to be the case as long as all players conform to σ. So when players are playing strategy σ, and the profile of signals received is $s^{\theta'}$ or τ_k**,** for $k \neq i$ any subsequent history also falls into $\mathcal{H}^*$. Thus, $g(\sigma(s^{\theta'}); h_t) = g(\sigma(\tau_k); h_t) = g(m^*_{\theta'}; h_t)$.

  Now suppose player i deviates to a strategy σ_i' so that $\sigma_i'(s_i^{\theta'}) = m_i'$. Clearly, since $m_i' \neq \sigma_i(s_i^{\theta'})$, there is a date at which player i does not play according to $m^*_{i,\theta'}$. Thus, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, when the profile of signals received is either $s^{\theta'}$ or τ_k for $k \neq i$, any subsequent history of h_t either falls in $\mathcal{H}^*$ (player i has played according to $m^*_{i,\theta'}$ so far) or does not fall in $\mathcal{H}^*_{-k}$ for each $k \neq i$ (at some point in this history, player i has not played according to $m^*_{i,\theta'}$). In each of these cases, again by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, player i's opponents are playing according to $m^*_{-i,\theta'}$. So we get [5]

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

  Here again, since $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$, we have

  $$u_i(g(m^*_{\theta'}; h_t); \theta') \geq u_i(g(m_i', m^*_{-i,\theta'}; h_t); \theta').$$

[5] We abuse notation because we should use $\sigma_{-i}(\tau_l\backslash s_i^{\theta'})$ instead of $\sigma_{-i}(\tau_l)$.

Thus, we get $u_i(g(\sigma(s^{\theta'}); h_t); \theta') \geq u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t); \theta')$ and $u_i(g(\sigma(\tau_k); h_t); \theta') \geq u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(\tau_k); h_t); \theta')$ for each $k \neq i$. Now since by $\Phi\mathbf{2}$, $\phi_i[\cdot \mid s_i^{\theta'}, h_t]$ may assign strictly positive weight only to $(\theta', s_{-i}^{\theta'})$ and (θ', τ_k) for each $k \neq i$, we can conclude

$$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i})|s_i^{\theta'}, h_t] \left[u_i(g(\sigma_i(s_i^{\theta'}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0.$$

- Case (2): $h_t \in \mathcal{H}_{-i}^* \backslash \mathcal{H}^*$

  Since $h_t \in \mathcal{H}_{-i}^*$ and $h_t \notin \mathcal{H}^*$, only player i has not played according to $m_{i,\theta'}^*$**.** Then, it is clear that h_t does not fall in $\mathcal{H}_{-k}^*$ for each $k \neq i$ (recall that $\mathcal{H}_{-k}^*$ is the set of histories under which every player j other than k has played according to $m_{j,\theta'}^*$**).** It is also clear that any subsequent history does not fall in $\mathcal{H}_{-k}^*$ for each $k \neq i$**.** By $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, we thus obtain that each player k other than i will play according to $m_{k,\theta'}^*$ at any subsequent history when receiving signal $s_k^{\theta'}$ or $s_k^{\theta''}$. Hence,

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

  Consider the case where player i deviates to a strategy σ_i' so that $\sigma_i'(s_i^{\theta'}) = m_i'$. Here, since (by a similar argument as above) any history that player i can achieve by deviating does not fall in $\mathcal{H}_{-k}^*$ for each $k \neq i$, each player k other than i will be playing according to $m_{k,\theta'}^*$ at any subsequent history whether he receive $s_k^{\theta'}$ or $s_k^{\theta''}$, which implies

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

  Since $m_{\theta'}^*$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$, we already have $u_i(g(m_{\theta'}^*; h_t); \theta') \geq u_i(g(m_i', m_{-i,\theta'}^*; h_t); \theta')$. Thus, we also get

  $$\begin{aligned} &u_i(g(\sigma(s^{\theta'}); h_t); \theta') \geq u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t); \theta') \text{ and} \\ &u_i(g(\sigma(\tau_k); h_t); \theta') \geq u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(\tau_k); h_t); \theta') \text{ for each } k \neq i. \end{aligned}$$

  Now, since by $\Phi\mathbf{2}$ we know that $\phi_i[\cdot|s_i^{\theta'}, h_t]$ assigns a strictly positive weight only to $(\theta', s_{-i}^{\theta'})$ and (θ', τ_k) for each $k \neq i$, we can conclude

  $$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i})|s_i^{\theta'}, h_t] \left[u_i(g(\sigma(s_i^{\theta'}), \sigma_{-i}(s_{-i}), h_t); \tilde{\theta}) - u_i(g(\sigma_i'(s_i^{\theta'}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0.$$

- Case (3): $h_t \notin \mathcal{H}_{-i}^*$

  In this case, at least one player $j \neq i$ has not played according to $m_{j,\theta'}^*$.

  By $\Sigma\mathbf{3}$, we know that when each player j receives signal $s_j^{\theta'}$, then these players play according to $m_{j,\theta'}^*$, so $\sigma(s^{\theta'}) = m_{\theta'}^*$. Thus, at history h_t, the outcome achieved by

playing σ when the profile of signals is $s^{\theta'}$ must be the same as the one when playing $m^*_{\theta'}$, i.e.,

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

In addition, for each $l \neq i$ with $h_t \notin \mathcal{H}^*_{-l}$, by definition, some player j other than l has not played according to $m^*_{j,\theta'}$ and obviously this will continue to be the case at any subsequent histories. Hence, any subsequent histories does not belong to $\mathcal{H}^*_{-l}$ either. At any such histories, we know by $\Sigma\mathbf{1}$, that player l will be playing according to $m^*_{l,\theta'}$ when he receives $s_l^{\theta''}$ while when players j other than l receive signal $s_j^{\theta'}$, by $\Sigma\mathbf{3}$ they will also be playing according to $m^*_{j,\theta'}$. Hence, we get that the outcome achieved from history h_t when playing σ and when the profile of signals received is τ_l is equal to the outcome achieved from history h_t when playing $m^*_{\theta'}$. Otherwise stated, for each $l \neq i$ with $h_t \notin \mathcal{H}^*_{-l}$, we have

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

Now, when player i deviates to some strategy σ'_i such that $\sigma'_i(s_i^{\theta'}) = m'_i$, using the argument above, when the other players receive signal profile $s_{-i}^{\theta'}$, we know that the outcome achieved is

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

while for each $l \neq i$ with $h_t \notin \mathcal{H}^*_{-l}$, we know that

$$g(\sigma'_i(s_i^{\theta'}), \sigma_{-i}(\tau_l); h_t) = g(m'_i, m^*_{-i,\theta'}; h_t).$$

Since $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$, we have $u_i(g(m^*_{\theta'}; h_t); \theta') \geq u_i(g(m'_i, m^*_{-i,\theta'}; h_t); \theta')$. Thus, we get

$$u_i(g(\sigma(s^{\theta'}); h_t); \theta') \geq u_i(g(\sigma'_i(s_i^{\theta'}), \sigma_{-i}(s_{-i}^{\theta'}); h_t); \theta')$$

and for each $l \neq i$ such that $h_t \notin \mathcal{H}^*_{-l}$, $u_i(g(\sigma(\tau_l); h_t); \theta') \geq u_i(g(\sigma'_i(s_i^{\theta'}), \sigma_{-i}(\tau_l); h_t); \theta')$. Because by $\Phi\mathbf{1}$, $\phi_i[\cdot|s_i^{\theta'}, h_t]$ may assign strictly positive weight only to $(\theta', s_{-i}^{\theta'})$ and (θ', τ_l) for each $l \neq i$ such that $h_t \notin \mathcal{H}^*_{-l}$, we can conclude

$$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i})|s_i^{\theta'}, h_t] \left[u_i(g(\sigma(s_i^{\theta'}, s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s_i^{\theta'}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0.$$

This completes the proof of the claim.

□

Claim 3. *For any $i \in N$, $s_i = s_i^{\theta''}$, and $h_t \in \mathcal{H}$:*

$$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i})|s_i, h_t] \left[u_i(g(\sigma(s); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s_i), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0$$

for each σ'_i.

This claim states that for any player i with signal $s_i^{\theta''}$, σ_i is a best response to σ_{-i} given his belief ϕ_i. Here again we consider the same partition of histories as in Claim 2. When h_t is a history where each player has played according to $m^*_{\theta'}$ (i.e., $h_t \in \mathcal{H}^*$), player i assigns positive probability to both θ'' and θ'. However, we will show that here again player i believes with probability one that the other players will be playing according to $m^*_{-i,\theta'}$, whether he deviates or not. Hence, if he does not deviate and $h_t \in \mathcal{H}^*$, he gets a while if he deviates to m_i' he gets $g(m_i', m^*_{-i,\theta'}; h_t)$. Because $m^*_{\theta'}$ is a subgame-perfect equilibrium in $\Gamma(\theta')$, we know that the deviation is not profitable if θ' is the true state, and Maskin monotonicity (Condition $(*)$ of Maskin monotonicity) implies that this is also not profitable if the state is θ''. Since these are the only states to which player i assigns strictly positive probability, this will complete the argument for this class of histories.

The easy case occurs when h_t is a history where a player other than i has not played according to $m^*_{\theta'}$ (i.e., $h_t \notin \mathcal{H}^*_{-i}$). In such a case, player i believes with probability one that θ' is the true state. In addition we will check that whenever player i uses σ_i against σ_{-i}, player i believes with probability one that the outcome will be given by $g(m^*_{\theta'}; h_t)$, while if player i deviates from $\sigma_i(s_i)$ to m_i', player i believes with probability one that the outcome will be given by $g(m_i', m^*_{-i,\theta'}; h_t)$. Here again, the fact that $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game will lead to the desired result. Finally, in the last case where player i has not played according to $m^*_{\theta'}$ while all other players have (i.e., $h_t \in \mathcal{H}^*_{-i} \backslash \mathcal{H}^*$), we will also check that player i assigns probability one to his opponent playing $m^*_{-i,\theta'}$. But σ_i has been constructed (see $\Sigma\mathbf{2}$) so that playing σ_i is better than any one-shot deviation. Then the one-shot deviation principle for sequential equilibrium will complete the proof of Claim 3. Taken together, Claims 2 and 3 establish sequential rationality of (ϕ, σ).

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

- Case (1): $h_t \in \mathcal{H}^*$

  In this case, each player has played according to $m^*_{\theta'}$. Note that, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, if each player j received signals of either $s_j^{\theta'}$ or $s_j^{\theta''}$, this will continue to be the case as long as all players conform to σ. So when players are playing strategy σ, and player i's opponents received either signal profile $s_{-i}^{\theta'}$ or $s_{-i}^{\theta''}$**,** any subsequent history also falls into $\mathcal{H}^*$. Thus,

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

  Now suppose that player i deviates to a strategy σ_i' so that $\sigma_i'(s_i^{\theta''}) = m_i'$. Since $m_i' \neq \sigma_i(s_i^{\theta''})$, there must exist a date at which player i does not play according to $m^*_{i,\theta'}$. Thus, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, when player i's opponents receive signal $s_{-i}^{\theta'}$ or $s_{-i}^{\theta''}$, any subsequent history of h_t either falls in $\mathcal{H}^*$ (player i has played according to $m^*_{i,\theta'}$ so far) or does not fall in $\mathcal{H}^*_{-k}$ for each $k \neq i$ (at some point in this history, player i has not played according to $m^*_{i,\theta'}$). In each of these cases, by $\Sigma\mathbf{1}$ and $\Sigma\mathbf{3}$, player i's opponents are playing according to $m^*_{-i,\theta'}$. So we get

  $$g(\sigma_i'(s_i^{\theta''}), \sigma_{-i}(s_{-i}^{\theta'}); h_t) = g(\sigma_i'(s_i^{\theta''}), \sigma_{-i}(s_{-i}^{\theta''}); h_t) = g(m_i', m^*_{-i,\theta'}; h_t). \tag{3}$$

Here again, since $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$, we have

$$u_i(g(m^*_{\theta'}; h_t); \theta') \geq u_i(g(m'_i, m^*_{-i,\theta'}; h_t); \theta').$$

Thus, we also get

$$u_i(g(\sigma(s_i^{\theta''}, s_{-i}^{\theta'}); h_t); \theta') \geq u_i(g(\sigma'_i(s_i^{\theta''}), \sigma_{-i}(s_{-i}^{\theta'}); h_t); \theta'). \tag{4}$$

The above inequality, together with (3), also implies

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

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 $(*)$ in the definition of Maskin monotonicity, we get that

$$u_i(g(\sigma(s_i^{\theta''}, s_{-i}^{\theta''}); h_t); \theta'') \geq u_i(g(\sigma'_i(s_i^{\theta''}), \sigma_{-i}(s_{-i}^{\theta''}); h_t); \theta''). \tag{5}$$

Now, since by $\Phi\mathbf{2}$, $\phi_i[\cdot|s_i^{\theta''}, h_t]$ assigns a strictly positive weight only to $(\theta', s_{-i}^{\theta'})$ and $(\theta'', s_{-i}^{\theta''})$, we conclude (4) and (5) imply that:

$$\sum_{(\tilde{\theta}, s_{-i}))} \phi_i[(\tilde{\theta}, s_{-i})|s_i^{\theta''}, h_t] \left[u_i(g(\sigma(s_i^{\theta''}, s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s_i^{\theta''}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0.$$

- Case (2): $h_t \notin \mathcal{H}^*_{-i}$

  In this case, at least one player $j \neq i$ has not played according to $m^*_{j,\theta'}$; This is still the case for any subsequent histories, so that they all fall outside $\mathcal{H}^*_{-i}$. By $\Sigma\mathbf{1}$, if player i plays according to σ_i, from h_t, he will play according to $m^*_{i,\theta'}$. Now, by $\Sigma\mathbf{3}$, we know that when player j other than i receives signal $s_j^{\theta'}$, then he plays according to $m^*_{j,\theta'}$. Thus, the outcome achieved when the profile of signals is $(s_i^{\theta''}, s_{-i}^{\theta'})$ must be the same as the outcome achieved when $m^*_{\theta'}$ is played. That is, we obtain

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

  Suppose player i deviates to a strategy σ'_i so that $\sigma'_i(s_i^{\theta''}) = m'_i$. Since, if the other players are receiving signal profile $s_{-i}^{\theta'}$, they will all be playing according to $m^*_{-i,\theta'}$, we obtain

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

  Since $m^*_{\theta'}$ is a subgame-perfect equilibrium in the complete information game $\Gamma(\theta')$, we have $u_i(g(m^*_{\theta'}; h_t); \theta') \geq u_i(g(m'_i, m^*_{-i,\theta'}; h_t); \theta')$. Thus, we also get

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

  Because by $\Phi\mathbf{1}$, $\phi_i[(\theta', s_{-i}^{\theta'})|s_i^{\theta''}, h_t] = 1$, so we can conclude

  $$\sum_{(\tilde{\theta}, s_{-i})} \phi_i[(\tilde{\theta}, s_{-i})|s_i^{\theta''}, h_t] \left[u_i(g(\sigma(s_i^{\theta''}, s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s_i^{\theta''}), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta})\right] \geq 0.$$

- 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 player i has not played according to $m^*_{i,\theta'}$. Then h_t does not fall in $\mathcal{H}^*_{-k}$ for each $k \neq i$ (recall that $\mathcal{H}^*_{-k}$ is the set of histories under which every player j other than k has played according to $m^*_{j,\theta'}$**).** It is also clear that any subsequent history does not fall in $\mathcal{H}^*_{-k}$ for each $k \neq i$**.** By Σ**1** and Σ**3**, whether player i's opponents have received $s^{\theta'}_{-i}$ or $s^{\theta''}_{-i}$, they all play according to $m^*_{-i,\theta'}$. By Φ**2** we know that $\phi_i[\cdot|s^{\theta''}_i, h_t] = \nu^\varepsilon(\cdot|s^{\theta''}_i)$ assigns a strictly positive weight only to $(\theta', s^{\theta'}_{-i})$ and $(\theta'', s^{\theta''}_{-i})$. In addition, we have that for any $h \in \mathcal{H}^*$ or $h \notin \mathcal{H}^*_{-i}$: $\sigma_i(h, s^{\theta''}_i) = m^*_{i,\theta'}(h, s^{\theta''}_i)$. Since $h_t \in \mathcal{H}^*_{-i} \backslash \mathcal{H}^*$, we conclude with Σ**2** that:

$$\sum_{(\tilde{\theta}, s_{-i})} \nu^\varepsilon(\tilde{\theta}, s_{-i}|s^{\theta''}_i) \left[u_i(g(\sigma(s^{\theta''}_i, s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s^{\theta''}_i), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta}) \right] \geq 0$$

  for any σ'_i that differs from σ_i only at h_t. By this and case (1) and (2), we know that at any history players have no profitable one-shot deviation, by the one-shot deviation principle (see Hendon, Jacobsen, and Sloth (1996)[6]) this yields:

$$\sum_{(\tilde{\theta}, s_{-i})} \nu^\varepsilon(\tilde{\theta}, s_{-i}|s^{\theta''}_i) \left[u_i(g(\sigma(s^{\theta''}_i, s_{-i}); h_t); \tilde{\theta}) - u_i(g(\sigma'_i(s^{\theta''}_i), \sigma_{-i}(s_{-i}); h_t); \tilde{\theta}) \right] \geq 0$$

  for any σ'_i. This completes the proof.

□

B.2 Consistency

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

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^k_i as

$$\sigma^k_i(m^t_i|h_t, s^{\theta''}_i) = \begin{cases} 1 - \eta_k^{T \times n} \text{ if } m^t_i = \sigma_i(h_t, s^{\theta''}_i); \\ \eta_k^{T \times n} \times \xi_i(h_t, s^{\theta''}_i, m^t_i) \text{ otherwise} \end{cases}$$

where T is the (finite) length of the longest final history, and for any signal $s_i \neq s^{\theta''}_i$:

$$\sigma^k_i(m^t_i|h_t, s_i) = \begin{cases} 1 - \eta_k \text{ if } m^t_i = \sigma_i(h_t, s_i); \\ \eta_k \times \xi_i(h_t, s_i, m^t_i) \text{ otherwise} \end{cases}.$$

[6]Hendon, Jacobsen, and Sloth (1996) assume that for each i and h, $M_i(h)$ is finite, which is our A1. It is easy to check that their argument goes through in case $M_i(h)$ is countably infinite. This fact is implicitly used in Section C.

Let ϕ^k be the unique consistent belief associated with each σ^k. It is easy to check that σ^k converges to σ and also that ϕ^k converges.[7] Let $\phi \equiv \lim_{k\to\infty} \phi^k$. In what follows, 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

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

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 4. *ϕ satisfies* **Φ1**.

Claim 4 says that, for any player i who sees signal $s_i^{\theta''}$ and has an opportunity to play after some other player has not played according to $m^*_{\theta'}$ (i.e., $h_t \notin \mathcal{H}^*_{-i}$), then under $\phi \equiv \lim_{k\to\infty} \phi^k$, player i believes with probability one that the state is θ', and that the other players have received $s^{\theta'}_{-i}$. In order to show that, we observe that if every player other than i has received a signal $s_j \in \{s_j^{\theta'}, s_j^{\theta''}\}$, then at such a history some player j other than i has deviated from σ. Then, since under the sequence of totally mixed strategies built above, it is (infinitely) more likely (as η_k tends to 0) that a deviation occurred at $s_j^{\theta'}$ rather than at $s_j^{\theta''}$. In the limit, Bayes' rule will then put probability one on $s_j^{\theta'}$ and given that the prior ν^ε assigns strictly positive weight only to $(\theta'', s^{\theta''}_{-i})$ and $(\theta', s^{\theta'}_{-i})$, Bayes rule will then put probability arbitrarily close to one on $(\theta', s^{\theta'}_{-i})$. In case player i received the private signal $s_i^{\theta'}$, if h_t is a history under which all players other than l have played according to $m^*_{\theta'}$ (i.e. $h_t \in \mathcal{H}^*_{-l}$), then the deviating player is l and again using a similar argument as above, we show that player i must assign probability 0 to player l receiving $s_l^{\theta''}$ and so to τ_l**.**

Consider player i at history $h_t \notin \mathcal{H}^*_{-i}$. The proof is reduced to checking the following two cases:

Proof of Claim 4. **Case 1:** $s_i = s_i^{\theta''}$

[7] As will become clear from the proof, the sequence $\{\phi^k\}_k$ does converge. Moreover, convergence in the definition of consistency is taken uniformly over messages and histories. In the case where $M_i(h)$ is countably infinite (we will discuss this case in Section C of this online appendix), 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 result.

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

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

We now show that the ratio below converges to 0 as $k \to \infty$:

$$\prod_{j\neq i}\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta''}) \Big/ \prod_{j\neq i}\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'}) \to 0 \text{ as } k \to \infty.$$

This will show that $\phi_i^k[(\theta', s_{-i}^{\theta'})|s_i^{\theta''}, h_t] \to 1$ and $\phi_i^k[(\theta'', s_{-i}^{\theta''})|s_i^{\theta''}, h_t] \to 0$ as $k \to \infty$.

Note first that in case every player j other than i receives signal $s_j \in \{s_j^{\theta'}, s_j^{\theta''}\}$, there must exist a player $\hat{\jmath} \neq i$ and a date $\hat{t} \leq t-1$ so that $\hat{\jmath}$ has not played according to $\sigma_{\hat{\jmath}}$, i.e. $\sigma_{\hat{\jmath}}(h_{\hat{t}}, s_{\hat{\jmath}}) \neq m_{\hat{\jmath}}^{\hat{t}}$. To see this, we proceed by contradiction and assume that $\sigma_{-i}(h_{t'}, s_{-i}) = m_{-i}^{t'}$ for all $t' \leq t-1$. This implies that whenever $h_{t'-1} \in \mathcal{H}_{-i}^*$, we must have $h_{t'} \in \mathcal{H}_{-i}^*$, because $h_{t'-1} \in \mathcal{H}_{-i}^*$ 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., player i has deviated). 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 $h_{t'} \in \mathcal{H}_{-i}^*$. Since $h_1 = \emptyset \in \mathcal{H}^* \subseteq \mathcal{H}_{-i}^*$, this simple inductive argument shows that $h_t \in \mathcal{H}_{-i}^*$, a contradiction.

By construction of σ^k, this implies that for some $\hat{\jmath} \neq i$ and $\hat{t} \leq t-1$:

$$\sigma_{\hat{\jmath}}^k(m_{\hat{\jmath}}^{\hat{t}}|h_{\hat{t}}, s_{\hat{\jmath}}^{\theta''}) = \eta_k^{T\times n}\xi_{\hat{\jmath}}(h_{\hat{t}}, s_{\hat{\jmath}}^{\theta''}, m_{\hat{\jmath}}^{\hat{t}}). \tag{6}$$

Now, we have:

$$\begin{aligned}
&\frac{\prod_{j\neq i}\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta''})}{\prod_{j\neq i}\prod_{t'=1}^{t-1} \sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'})} \leq \frac{\eta_k^{T\times n} \times \xi_{\hat{\jmath}}(h_{\hat{t}}, s_{\hat{\jmath}}^{\theta''}, m_{\hat{\jmath}}^{\hat{t}}) \times 1}{\prod_{j\neq i}\prod_{t'=1}^{t-1} \eta_k \xi_j(h_{t'}, s_j^{\theta'}, m_j^{t'})} \\
&= \frac{\eta_k^{T\times n}}{\eta_k^{(t-1)(n-1)}} \times \frac{\xi_{\hat{\jmath}}(h_{\hat{t}}, s_{\hat{\jmath}}^{\theta''}, m_{\hat{\jmath}}^{\hat{t}})}{\prod_{j\neq i}\prod_{t'=1}^{t-1} \xi_j(h_{t'}, s_j^{\theta'}, m_j^{t'})} \to 0 \ \text{(as } k \to \infty).
\end{aligned}$$

Here, the inequality is assured by (6) and the construction of σ^k that, for all j and $t' \leq t-1$, $\sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'}) \geq \eta_k \times \xi_j(h_{t'}, s_j^{\theta'}, m_j^{t'})$.

Case 2: $s_i = s_i^{\theta'}$

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

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

for some positive functions $c_z(k)$. We now show that if $h_t \in \mathcal{H}_{-l}^*$, then the ratio below converges to $+\infty$ as $k \to \infty$:

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

This will show that $\phi_i^k[(\theta', \tau_l)|s_i^{\theta'}, h_t] \to 0$ for all l if $h_t \in \mathcal{H}_{-l}^*$; and hence that ϕ satisfies **Φ1**. Assume that $h_t \in H_{-l}^*$ for some l, as we already claimed, if every player j other than i has received a signal $s_j \in \{s_j^{\theta'}, s_j^{\theta''}\}$, there is a player $\hat{\jmath} \neq i$ and a date $\hat{t} \leq t-1$ so that $\hat{\jmath}$ has not played according to $\sigma_{\hat{\jmath}}$, i.e., $\sigma_{\hat{\jmath}}(h_{\hat{t}}, s_{\hat{\jmath}}) \neq m_{\hat{\jmath}}^{\hat{t}}$. Now, since $h_t \in \mathcal{H}_{-l}^*$, we claim that $\hat{\jmath} = l$. Indeed, $h_t \in \mathcal{H}_{-l}^*$ means that any player j other than l has played according to $m_{j,\theta'}^*$. So if player l had played according to σ_l (i.e., for all $t' : \sigma_l(h_{t'}, s_l) = m_l^{t'}$), repeated applications of **Σ1** and **Σ3** would yield to $h_t = h_t^* \in \mathcal{H}_{-i}^*$ which is false by assumption.

By construction of σ^k, this 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:

$$\sigma_l^k(m_l^{\hat{t}}|h_{\hat{t}}, s_l^{\theta''}) = \eta_k^{T \times n} \xi_l(h_{\hat{t}}, s_l^{\theta''}, m_l^{\hat{t}}). \tag{7}$$

Now, we have

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

Where the inequality is assured by (7) and (assuming without loss of generality 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 5. *ϕ satisfies* **Φ2.**

Claim 5 says that if player i gets signal $s_i^{\theta'}$ or $s_i^{\theta''}$ then at a history h_t under which each of his opponent has played according to $m^*_{\theta'}$, ϕ is the same as his beliefs given only by his private signal.

To prove this, we show that if every player $j \neq i$ has received a signal $s_j \in \{s_j^{\theta'}, s_j^{\theta''}\}$ then at histories where all players other than i have played according to $m^*_{\theta'}$, each player other than i has played according to σ at each previous stage. This ensures that for any $h_t \in \mathcal{H}^*_{-i}$, no player other than i has deviated from the candidate for sequential equilibrium strategy profile σ and so player i's beliefs must be given by his private signal.

Proof of Claim 5. Consider player i at history $h_t \in \mathcal{H}^*_{-i}$. Here again, 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 strictly positive weight only to $(\theta'', s_{-i}^{\theta''})$ and $(\theta', s_{-i}^{\theta'})$. Hence,

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

We now show that the ratio below converges to 1 as $k \to \infty$:

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

This will show that $\phi_i^k[(\theta'', s_{-i}^{\theta''})|s_i^{\theta''}, h_t] \to \nu^\varepsilon((\theta'', s_{-i}^{\theta''})|s_i^{\theta''})$ and $\phi_i^k[(\theta', s_{-i}^{\theta'})|s_i^{\theta''}, h_t] \to \nu^\varepsilon((\theta', s_{-i}^{\theta'})|s_i^{\theta''})$.

Note now that if players $j \neq i$ receive signal $s_j \in \{s_j^{\theta'}, s_j^{\theta''}\}$, then for all $t' \leq t-1$, $\sigma_j(h_{t'}, s_j) = m_j^{t'}$. To see this, note that for any $t' \leq t-1 : h_{t'} \in \mathcal{H}^*_{-i}$, thus, either every player has played according to $m^*_{\theta'}$ (i.e., $h_{t'} \in \mathcal{H}^*$) or player i has not played according to $m^*_{i,\theta'}$ (i.e., $h_{t'} \notin \mathcal{H}^*_{-j}$ for all $j \neq i$). In each of these cases we know, by Σ**1** and Σ**3**, that σ_j presribes to play according to $m^*_{j,\theta'}$. Since $h_{t'} \in \mathcal{H}^*_{-i}$ this implies that $\sigma_j(h_{t'}, s_j) = m^*_{j,\theta'}(h_{t'}) = m_j^{t'}$.

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

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

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 strictly positive weight only to $(\theta', s_{-i}^{\theta'})$ and (θ', τ_l) for $l \neq i$. Hence,

$$\phi_i^k[(\theta', s_{-i}^{\theta'})|s_i^{\theta'}, h_t]$$

$$= \frac{\nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'}) \times \prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'})}{\nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'}) \times \prod_{j\neq i}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'}) + \sum_{l\neq i}\nu^\varepsilon(\theta', \tau_l) \times \prod_{j\neq i,l}\prod_{t'=1}^{t-1}\sigma_j^k(m_j^{t'}|h_{t'}, s_j^{\theta'}) \times \prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'}|h_{t'}, s_l^{\theta''})}$$

$$= \frac{\nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'})}{\nu^\varepsilon(\theta', s_{-i}^{\theta'}, s_i^{\theta'}) + \sum_{l\neq i}\nu^\varepsilon(\theta', \tau_l) \times \frac{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'} \mid h_{t'}, s_l^{\theta''})}{\prod_{t'=1}^{t-1}\sigma_l^k(m_l^{t'} \mid h_{t'}, s_l^{\theta'})}}$$

We now show that for each $l \neq i$, the ratio below converges to 1 as $k \to \infty$:

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

This will show that $\phi_i^k[(\theta', s_{-i}^{\theta'})|s_i^{\theta'}, h_t] \to \nu^\varepsilon((\theta', s_{-i}^{\theta'})|s_i^{\theta'})$ and similar reasoning shows that for each $l \neq i : \phi_i^k[(\theta', \tau_l)|s_i^{\theta'}, h_t] \to \nu^\varepsilon((\theta', \tau_l)|s_i^{\theta'})$, and hence, ϕ satisfies $\Phi\mathbf{2}$. □

Now, by similar reasoning as in the case above, we get that for all $l \neq i$ and $t' \leq t-1$:

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

Thus,

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

Finally, observing that for $s_i^{\tilde{\theta}} \notin \{s_i^{\theta'}, s_i^{\theta''}\}$, $\nu^\varepsilon(\cdot, s_i^{\tilde{\theta}})$ assigns a weight one to $(\tilde{\theta}, s_{-i}^{\tilde{\theta}})$, we have established the following claim, which completes the proof of Theorem 3.

Claim 6. *ϕ satisfies $\Phi\mathbf{3}$.*

C Theorem 3 extends to countable messages

Here we extend Theorem 3 to mechanisms that have countably infinite message spaces. This extension is important because some of the literature on implementation theory uses "integer games" where each player has to announce an integer and becomes the dictator when his integer is the largest one, as in Maskin (1999) and in Moore and Repullo (1988).

Assumption A2. *$M_i(h)$ is countable for each i and h.*

The next assumption says that against any profile of strategy in the complete information game, in the neighborhood of complete information, each player i has a non-empty set of best responses. This condition is vacuously satisfied under A1, so Theorems 3 and 4 show that if a mechanism can implement a non-Maskin monotonic social choice correspondence (SCC) both under complete information and under small information perturbations, then under this mechanism players must not have well-defined best responses. In addition, we show in Section C.2 that when the state space is finite (this is our case), Moore and Repullo's general mechanism has well-defined best-responses (under weak assumptions) and so our argument also applies there.

Assumption A3. *The sequential mechanism Γ has well-defined best replies: for any player i, any $\theta \in \Theta$, any $m_{-i} \in M_{-i}$, there exists $\bar{\xi}(i,\theta,m_{-i}) > 0$ such that for any $\beta \in \Delta(\Theta)$ with $\beta(\theta) \geq 1 - \bar{\xi}(i,\theta,m_{-i})$, for any $m_i \in M_i$ we have for all $h \in \mathcal{H}$:*

$$\arg\max \sum_{\tilde{\theta}} \beta(\tilde{\theta}) u_i(g\left(\left(m_i', m_{-i}\right)\right); h); \tilde{\theta}) \neq \emptyset$$

where the max *is taken over all pure messages $m_i' \in M_i$ that differ from m_i only at h.*

Remark C.1. *If the mechanism is not finite but the set of outcomes is, A3 is also vacuously satisfied. We also note that A3 is not needed for sequential mechanisms in which each player moves only once.*[8]

Theorem C.1. *Assume A2 and A3. Suppose that a mechanism Γ SPE-implements a non-Maskin monotonic SCC $\mathcal{F}$. Fix any complete information prior μ. There exist a sequence of priors $\{\nu^\varepsilon\}_{\varepsilon>0}$ that converges to μ and a corresponding sequence of sequential equilibria $\{(\phi^\varepsilon, \sigma^\varepsilon)\}_{\varepsilon>0}$ such that as ε tends to 0, $g(\sigma^\varepsilon(s^\theta); \emptyset) \to a \notin \mathcal{F}(\theta)$ for some $\theta \in \Theta$ and some $a \in A$.*

Proof. The proof is essentially the same as the proof of Theorem 3 where we only consider finite mechanisms. So, we claim that there are essentially only two changes we need to extend the proof of Theorem 3 to the case of countably infinite message spaces. First, in the beginning of the proof of Theorem 3, we have to choose $\varepsilon > 0$ small enough to apply A3. Second, we will show that A3 guarantees that $\Sigma\mathbf{2}$ (which is introduced in the proof of Theorem 3) is well defined. This will be proved in the next subsection. □

[8] One can directly check this in the definition of strategy σ ($\Sigma\mathbf{2}$) used in the proof of Theorem 3. More specifically, it can be checked there that for each player, A3 is only used at histories where this player has to choose a message and at which he has previously deviated from the equilibrium. By construction, there is no such a history.

C.1 A3 guarantees that Σ2 is well-defined

Fix $\varepsilon > 0$ small enough so that $\nu^{\varepsilon}(\theta'|s_i^{\theta'}) \geq 1 - \bar{\xi}(i, \theta, m^*_{-i,\theta})$. We shall claim that A3 guarantees that one can construct $\bar{m}_i$ needed for $\Sigma\mathbf{2}$. First, for any $h_t \in \mathcal{H}^*$ or $h_t \notin \mathcal{H}^*_{-i}$, we set $\bar{m}_i(h_t) = m^*_{i,\theta}(h_t)$. Second, we define $\bar{m}_i$ by induction on the set of histories in $\mathcal{H}^*_{-i}\backslash\mathcal{H}^*$. Take any history $h_t \in \mathcal{H}^*_{-i}\backslash\mathcal{H}^*$ so that there is no subsequent history that falls into $\mathcal{H}^*_{-i}\backslash\mathcal{H}^*$. Since we already defined $\bar{m}_i(h_t) = m^*_{i,\theta}(h_t)$ for any $h_t \notin \mathcal{H}^*_{-i}\backslash\mathcal{H}^*$, $\bar{m}_i$ has been defined for any subsequent histories. By A3 we obtain

$$\arg\max \sum_{\tilde{\theta}} \nu^{\varepsilon}(\tilde{\theta}|s_i^{\theta'}) u_i(g\left(\left(m'_i, m^*_{-i,\theta}\right)\right); h_t); \tilde{\theta}) \neq \emptyset$$

where the max is taken over all pure messages $m'_i \in M_i$ that differ from $\bar{m}_i$ only at h_t and are identical at any subsequent histories (what happens before h_t is obviously irrelevant).

Now set

$$\bar{m}_i(h_t) \in \arg\max \sum_{\tilde{\theta}} \nu^{\varepsilon}(\tilde{\theta}|s_i^{\theta'}) u_i(g\left(\left(m'_i, m_{-i}\right)\right); h_t); \tilde{\theta}).$$

This establishes that one can inductively construct $\bar{m}_i$ so that $\bar{m}_i$ satisfies the properties needed for $\Sigma\mathbf{2}$.

C.2 A3 is satisfied in the Moore-Repullo canonical mechanism

We will review some of the main results of Moore and Repullo (1988) here.

Definition C.1 (Moore and Repullo (1988)). *A social choice correspondence $\mathcal{F}$ satisfies Condition C if, for every pair of profiles $\theta, \phi \in \Theta$ with $a \in \mathcal{F}(\theta)\backslash\mathcal{F}(\phi)$, there exists a finite sequence*

$$\sigma(\theta, \phi; a) \equiv \{a_0 = a, a_1, \ldots, a_k, \ldots, a_l, a_{l+1}\} \subset A,$$

with $l = l(\theta, \phi; a) \geq 1$, such that:

1. *for each $k = 0, \ldots, l-1$, there is some particular player $j(k) = j(k|\theta, \phi; a)$, for whom*

$$u_{j(k)}(a_k; \theta) \geq u_{j(k)}(a_{k+1}; \theta);$$

2. *there is some player $j(l) = j(l|\theta, \phi; a)$ for whom*

$$u_{j(l)}(a_l; \theta) \geq u_{j(l)}(a_{l+1}; \theta) \text{ and } u_{j(l)}(a_{l+1}; \phi) > u_{j(l)}(a_l; \phi).$$

Further, $l(\theta, \phi; a)$ is uniformly bounded by some $\bar{l} < \infty$.

Assuming Condition C holds, let $\mathcal{Q}(\mathcal{F})$ be a class of subsets Q of A. A typical Q is defined as follows:

> For each pair of profiles θ and ϕ in Θ, and for each $a \in \mathcal{F}(\theta)\backslash\mathcal{F}(\phi)$, select *one* sequence $\sigma(\theta, \phi; a)$ satisfying (1) and (2) in Condition C. Then let Q be the union of the elements in these sequences.

$\mathcal{Q}(\mathcal{F})$ comprises the Q's constructed from all possible selections.

Definition C.2. *A social choice correspondence $\mathcal{F}$ satisfies Condition C^+ if it satisfies Condition C and the following condition as well: there exists a particular $Q^+ \in \mathcal{Q}(\mathcal{F})$, and a particular set $B \subseteq A$ containing Q^+, such that the following is true for each $\theta \in \Theta$:*

- *Each player i has nonempty maximal set $B_i^*(\theta) \subseteq B$ under θ, i.e., $B_i^*(\theta) = \arg\max_{a \in B} u_i(a; \theta)$.*
- *$B_i^*(\theta) \cap B_j^*(\theta) = \emptyset$ for each $\theta \in \Theta$ and each $i, j \in N$ with $i \neq j$*
- *$B_i^*(\theta) \cap Q^+ = \emptyset$ for each i and each θ.*

Let the selected sequences $\sigma(\theta, \phi; a) \in Q^+$ be labelled $\sigma^+(\theta, \phi; a)$. Define the Moore-Repullo canonical mechanism $\Gamma^{MR} = (M, g)$ as follows.

Stage 0: each player i announces some triplet $m_{i,0} = (\theta^i, a^i, n_0^i)$, where $\theta^i \in \Theta, a^i \in \mathcal{F}(\theta^i)$, and n_0^i is a nonnegative integer. There are three possibilities to consider:

1. all n players agree on a common profile θ and outcome $a \in \mathcal{F}(\theta)$, then outcome a is chosen. STOP
2. If only $n-1$ players agree on a common profile θ and outcome $a \in \mathcal{F}(\theta)$, and if the remaining player i announces a profile ϕ, and
   (a) if $a \in \mathcal{F}(\phi)$, then outcome a is implemented; STOP
   (b) if $a \notin \mathcal{F}(\phi)$ but i is not the agent $j(0)$ prescribed in $\sigma^+(\theta, \phi; a)$, then outcome a is implemented; STOP
   (c) if $a \notin \mathcal{F}(\phi)$ and $i = j(0)$, then go to Stage 1.
3. If neither (1) nor (2) apply, then the player with the highest integer n_0^i is allowed to choose an outcome from B. Ties are broken by selecting from the players who announced the highest number according to who has the smallest i. STOP

Stage $k = 1, \ldots, l$: each player i can either raise a "flag," or announce a nonnegative integer $n_k^i \in \mathbb{N}$, i.e., $m_{i,k} \in M_{i,k} \in \{\text{flag}\} \cup \mathbb{N}$. Again there are three possibilities to consider:

1. If $n-1$ or more flags are raised, then the agent $j(k-1)$ prescribed in $\sigma^+(\theta, \phi; a)$ is allowed to choose an outcome from B. STOP
2. If $n-1$ or more players announce zero, and
   (a) if the player $j(k)$ prescribed in $\sigma^+(\theta, \phi; a)$ is one of those who announce zero, then implement outcome a_k from sequence $\sigma^+(\theta, \phi; a)$; STOP
   (b) if $j(k)$ does not announce zero, then
      i. if $k < l$, go to Stage $k+1$;
      ii. if $k = l$, implement outcome a_{l+1} from sequence $\sigma^+(\theta, \phi; a)$. STOP

(c) If neither (1) nor (2) apply, then the player who announces the highest integer n_k^i is allowed to choose an outcome from B. STOP

Theorem C.2 (Moore and Repullo (1988)). *If a social choice correspondence $\mathcal{F}$ satisfies Condition C^+, and $n \geq 3$, then $\mathcal{F}$ can be implemented in subgame-perfect equilibrium.*

Moore and Repullo (1988) show the above theorem by using the mechanism described above. We note that this mechanism satisfies A3 if the set of outcomes A is finite or when each player's preferences over A are strict and utilities are bounded. Furthermore, the above mechanism satisfies A3 whenever (i) the set B given in Condition C^+ is a compact set of outcomes; (ii) $u_i : A \times \Theta \to \mathbb{R}$ is continuous in a.[9,10] It is worth noting that many researchers assume (i) and (ii) after appealing to Moore and Repullo's (1988) result. This is the case for instance in Moore and Repullo (1988)'s examples of risk-sharing (Section 6.1) or the production contract example (Section 6.2). More importantly, it is also the case in Maskin and Tirole (1999a)'s proof of the irrelevance theorems. Hence our non-robusntess result (i.e., our Theorem C.1) also apply to Maskin and Tirole's irrelevance theorems.

D Sufficiency for Robust Implementation: The Case of Social Choice Correspondences (SCCs)

In Remark 3 of Section IV, we argue that Maskin monotonic social choice functions are robustly implementable. Here we extend this argument to the case of social choice correspondences.

We need to strengthen Maskin monotonicity to the following:

Definition D.1. *An SCC $\mathcal{F}$ satisfies **strong Maskin Monotonicity** if for every SCF f selected from $\mathcal{F}$ and every pair of states θ' and θ'' such that*

$$\{(i,b) \,|\, u_i(f(\theta');\theta') > u_i(b;\theta')\} \subseteq \{(i,b) \,|\, u_i(f(\theta');\theta'') \geq u_i(b;\theta'')\}$$

then $f(\theta') \in \mathcal{F}(\theta'')$.

[9] Then, for any $\beta \in \Delta(\Theta)$,

$$\arg\max_{a \in B} \sum_{\tilde{\theta}} \beta(\tilde{\theta}) u_i(a;\tilde{\theta}) \neq \emptyset.$$

We note that a one-shot deviation of player i at stage k in Γ^{MR} allows player i possibly to fall into an integer game at stage k where he can get any outcome in B; if he cannot fall into this integer game, he can only induce a finite number of outcomes, say B_k, by deviating. In any case, he has a most preferred deviation, i.e.,

$$\arg\max_{a \in B} \sum_{\tilde{\theta}} \beta(\tilde{\theta}) u_i(a;\tilde{\theta}) \neq \emptyset; \quad \arg\max_{a \in B \cup B_k} \sum_{\tilde{\theta}} \beta(\tilde{\theta}) u_i(a;\tilde{\theta}) \neq \emptyset; \quad \text{and} \quad \arg\max_{a \in B_k} \sum_{\tilde{\theta}} \beta(\tilde{\theta}) u_i(a;\tilde{\theta}) \neq \emptyset.$$

Then A3 is satisfied whenever (i) and (ii) hold.

[10] Note that A2 need not be satisfied for these mechanisms since B need not be countable. A2 was introduced only to define sequential equilibrium in a simple manner. If one uses perfect Bayesian equilibrium instead, we believe that A2 is not required.

Strong Maskin monotonicity is equivalent to Maskin monotonicity in many economic environments.[11] For example, consider environments in which there is a private good that is both desirable and continuously transferable. Another example is an environment in which agents have strict preferences. The next definition is the no-veto-power condition, which is widely used in the literature.

Definition D.2. *An SCC $\mathcal{F}$ satisfies* ***no-veto-power*** *if whenever there is an alternative $c \in A$ such that for at least $n-1$ players i, $u_i(c;\theta) \geq u_i(b;\theta)$ for every $b \in A$, we have $c \in \mathcal{F}(\theta)$.*

We need one extra condition together with strong Maskin monotonicity and no-veto power. This is the no-worst-alternative condition as defined by Cabrales and Serrano (2011):

Definition D.3. *An SCC $\mathcal{F}$ satisfies the* ***no-worst-alternative*** *(NWA) condition if for each agent $i \in N$, $\theta \in \Theta$ and f selected from $\mathcal{F}$, there exists $z(i,\theta,f) \in A$ such that $u_i(f(\theta);\theta) > u_i(z(i,\theta,f);\theta)$.*

Let $\mathcal{P}$ denote the set of priors over $\Theta \times S$ with the following metric $d : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_+$: for any $\mu, \mu^{'} \in \mathcal{P}$,

$$d(\mu, \mu^{'}) = \max_{(\theta,s)\in\Theta\times S} |\mu(\theta,s) - \mu^{'}(\theta,s)|.$$

So, when we say $\nu^k \to \mu$, we mean that $d(\nu^k, \mu) \to 0$ as $k \to \infty$. When $\Theta \times S$ is a finite state space, Theorem 14.5 of Fudenberg and Tirole (1991a) shows that when $\nu^k \to \mu$ as $k \to \infty$, there exists $\{p^k\}_{k=1}^{\infty}$ such that (1) $p^k \to 1$ as $k \to \infty$; (2) $\nu^k(\{(\theta, s^\theta)\}_{\theta\in\Theta}) \geq p^k$ for each k; and (3) for each k, it is common p^k-belief at any profile of signals s^θ that θ has realized.[12]

We propose the following definition of robust implementation:

Definition D.4. *An SCC $\mathcal{F}$ is* ***robustly*** *implementable under the complete information prior μ if there exists a mechanism $\Gamma = (M, g)$ satisfying the following two properties: for any SCF f selected from $\mathcal{F}$ and any sequence of priors $\{\nu^\varepsilon\}_{\varepsilon>0}$ converging to μ, (1) there is a sequence of sequential equilibria $\{\sigma^\varepsilon\}_{\varepsilon>0}$ in $\{\Gamma(\nu^\varepsilon)\}_{\varepsilon>0}$ satisfying $\lim_{\varepsilon\to 0} g(\sigma^\varepsilon(s^\theta);\emptyset) = f(\theta)$ for every $\theta \in \Theta$; and (2) for any sequence of sequential equilibria $\{\sigma^\varepsilon\}_{\varepsilon>0}$ in $\{\Gamma(\nu^\varepsilon)\}_{\varepsilon>0}$, we have $\lim_{\varepsilon\to 0} g(\sigma^\varepsilon(s^\theta);\emptyset) \in \mathcal{F}(\theta)$ for every $\theta \in \Theta$.*

Remark D.1: The first requirement of robust implementation says that for any SCF f selected from a given SCC $\mathcal{F}$ and any environment near μ, there is an equilibrium whose outcome is close to that given by f whenever a signal profile s has strictly positive probability under μ (i.e., $s = s^\theta$ for some θ). The second requirement says that for any environment near μ, whenever a signal profile s has strictly positive probability under μ, equilibrium outcomes are close to that of $\mathcal{F}$. Both requirements are robust analogs of the two standard requirements of implementation.[13] Roughly speaking, the first requirement

[11] What we mean by "strong"is that we replace the first weak inequality of $(*)$ in the definition of Maskin monotonicity with a strict one. This notion also appears in Chung and Ely (2003).

[12] See Monderer and Samet (1989) and/or Fudenberg and Tirole (1991a) for the precise definition of common p-belief.

[13] See, for instance, Maskin (1999) for the definition of Nash implementation.

embodies a version of lower hemi-continuity of the equilibrium correspondence and the second embodies a version of upper hemi-continuity.[14] As is clear from the proof of Theorem 3, to show that Maskin monotonicity is necessary for SCCs to be robustly implemented, we only used the second property of robust implementation and do not exploit the full strength of robust implementation. Finally, the subsequent argument provides sufficient conditions under which a static mechanism yields robust implementation. Hence, the result would hold if we were to replace sequential equilibrium by Nash equilibrium in the above statement.

We are now ready to state the result:

Theorem D.1. *Suppose there are at least three players, i.e., $|N| = n \geq 3$. If an SCC $\mathcal{F}$ satisfies strong Maskin monotonicity, no-veto-power and the NWA condition, then $\mathcal{F}$ is robustly implementable.*

Proof. We construct an implementing mechanism $\Gamma = (M, g)$.[15] For each $i \in N$, let $M_i = (\Theta \times \mathcal{F}) \cup (\mathbb{Z}_+ \times A)$ where $\mathbb{Z}_+$ is the set of nonnegative integers. That is, each agent is asked to report *either* a state and a social choice function or an integer and an alternative. Let $m^{\theta,f}$ denote the message profile $((\theta, f), (\theta, f), \ldots, (\theta, f))$, and $m^{\theta,f}\backslash m_i$ the profile obtained from $m^{\theta,f}$ by substituting m_i for agent i. We set $g(m^{\theta,f}) = f(\theta)$. If $m_i = (\theta', f')$, and if there exists an alternative $c \in A$ such that $u_i(c; \theta') > u_i(f(\theta); \theta')$ but $u_i(f(\theta); \theta) > u_i(c; \theta)$, then we set $g(m^{\theta,f}\backslash m_i) = c$. (If there is more than one such c, select one arbitrarily). For all other cases, we set $g(m^{\theta,f}\backslash m_i) = z(i, \theta, f(\theta))$ as defined for the NWA condition.

Consider any other profile of messages m. If each m_i consists of a state and a social choice function, then choose $g(m)$ to be an arbitrary element of $\mathcal{F}(\Theta)$. If at least one agent has announced an integer and an alternative, set $g(m)$ to be the alternative named by the agent whose named integer is the greatest (breaking ties by choosing the lowest index among those who announced the greatest integer).

The rest of the proof can be completed by the following three steps: in Step 1, we show that for any SCF f selected from $\mathcal{F}$, there exists a good equilibrium whose outcome coincides with that of f for any nearby environment. In Step 2, we show that any Nash equilibrium outcome is socially desirable. In Step 3, we show that this continues to be the case in nearby environments.

For any complete information prior μ, let $U(\mu)$ denote a neighborhood around μ with respect to metric d.

Step 1: Let μ be a complete information prior. For each SCF f selected from $\mathcal{F}$, there exists a neighborhood $U(\mu)$ for which there exists a strict Bayesian Nash equilibrium σ of the game $\Gamma(\nu)$ for each $\nu \in U(\mu)$ such that $g(\sigma(s^\theta)) = f(\theta)$ for every $\theta \in \Theta$.

For each SCF f selected from $\mathcal{F}$ and $\theta \in \Theta$, consider the truthful strategy of agent i as $m_i^{\theta,f} = (\theta, f)$. This yields $g(m^{\theta,f}) = f(\theta)$. By construction, if in state θ, agent i sends message $m_i \neq m_i^{\theta,f}$,

$$u_i(g(m^{\theta,f}); \theta) > u_i(g(m^{\theta,f}\backslash m_i); \theta).$$

[14] Property (2) in our definition says that the correspondence from priors to equilibrium outcomes has a closed graph. In general, this is not equivalent to upper hemi-continuity. However, the closed graph property of the equilibrium outcomes correspondence implies upper hemi-continuity if the range of the correspondence is compact (see Aliprantis and Border (1999)).

[15] The proof here is a modification of that of Theorem 2 of Chung and Ely (2003).

Hence, $m^{\theta,f}$ is a *strict* Nash equilibrium of the game $\Gamma(\theta)$. Define $\sigma_i(s_i^\theta) = (\theta, f)$ for each $s_i^\theta \in S_i$ as agent i's strategy of the game $\Gamma(\mu)$. Then σ is a strict Nash equilibrium of the game $\Gamma(\mu)$. Define

$$A[\sigma_{-i}] = \left\{ a \in A \middle| \ \exists s_{-i} \in S_{-i},\ \exists \sigma_i^{'} \text{ such that } g(\sigma_i^{'}(s_i), \sigma_{-i}(s_{-i})) = a \right\}$$

as the set of possible outcomes that can be induced by agent i's strategy $\sigma_i^{'}$ against σ_{-i}. By construction of Γ and the finiteness of S, $A[\sigma_{-i}]$ is finite. It is important to note that each agent can only induce a finite number of outcomes, while the set of strategies may be infinite. By the continuity of expected utility and the finiteness of S, N, and $A[\sigma_{-i}]$, there is a neighborhood $U(\mu)$ such that σ continues to be a strict Bayesian Nash equilibrium of the game $\Gamma(\nu)$ for every $\nu \in U(\mu)$.

Step 2: Let μ be a complete information prior and σ be a Nash equilibrium of the game $\Gamma(\mu)$. Then, $g(\sigma(s^\theta)) \in \mathcal{F}(\theta)$ for every $\theta \in \Theta$.

Suppose σ is a Nash equilibrium of $\Gamma(\mu)$. Assume further that in $\sigma(s^\theta)$, each player announces the same state and SCF $(\theta^{'}, f^{'})$. Then, $g(\sigma(s^\theta)) = f^{'}(\theta^{'})$. In this case, we claim that $f^{'}(\theta^{'}) \in \mathcal{F}(\theta)$. If this is not the case, by strong Maskin monotonicity, there exist a player i and an alternative a such that $u_i(a;\theta) > u_i(f^{'}(\theta^{'});\theta)$ but $u_i(f^{'}(\theta^{'});\theta^{'}) > u_i(a;\theta^{'})$. By construction of Γ, we can conclude that $g(\sigma(s^\theta)\backslash(\theta, f')) = a$. Thus, $\sigma(s^\theta)$ would not be a Nash equilibrium of $\Gamma(\theta)$. For any other profile $\sigma(s^\theta)$, there must be at least $n-1$ agents who can deviate from $\sigma(\theta)$ and bring about a profile in which there are at least 3 distinct messages. Thus, by construction of Γ, each of these agents could dictatorially choose his most preferred alternative from A in state θ. But since $\sigma(s^\theta)$ is a Nash equilibrium of $\Gamma(\theta)$, it must be that for each of these players i, $u_i(g(\sigma(s^\theta));\theta) \geq u_i(a;\theta)$ for every $a \in A$. Since $\mathcal{F}$ satisfies no-veto-power, $g(\sigma(s^\theta)) \in \mathcal{F}(\theta)$.

Step 3: Let μ be be a complete information. Suppose that σ is a strategy profile such that $g(\sigma(s^\theta)) \notin \mathcal{F}(\theta)$ for some $\theta \in \Theta$. It is enough for our purpose to show that there must exist a neighborhood $U(\mu)$ such that σ is not a Bayesian Nash equilibrium of the game $\Gamma(\nu)$ for every $\nu \in U(\mu)$.

Suppose σ is given such that $g(\sigma(s^\theta)) \notin \mathcal{F}(\theta)$ for some $\theta \in \Theta$. This implies that σ is not a Nash equilibrium of $\Gamma(\theta)$. Hence, there exists an agent i and a strategy $\sigma_i^{'}$ such that

$$u_i(g((\sigma_i^{'}, \sigma_{-i})(s^\theta));\theta) > u_i(g(\sigma(s^\theta));\theta).$$

By the continuity of expected utility and the finiteness of N, S, and $A[\sigma_{-i}]$, there exists a neighborhood $U(\mu)$ such that for any $\nu \in U(\mu)$,

$$\sum_{\tilde{\theta}\in\Theta} \sum_{s_{-i}\in S_{-i}} \nu(\tilde{\theta}, s_{-i}|s_i^\theta) \left[u_i(g(\sigma_i^{'}(s_i^\theta), \sigma_{-i}(s_{-i})); \tilde{\theta}) - u_i(g(\sigma_i(s_i^\theta), \sigma_{-i}(s_{-i})); \tilde{\theta}) \right] > 0.$$

This implies that σ is not a Bayesian Nash equilibrium of $\Gamma(\nu)$ for every $\nu \in U(\mu)$. □

References

[1] Aliprantis, C., and K. Border (1999), *Infinite Dimensional Analysis*, 2nd. ed., Springer Verlag, Berlin.

[2] 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.

[3] Chung, K., and J. Ely (2003), "Implementation with Near-Complete Information", *Econometrica* 71, 857-871.

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

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

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

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

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

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