Supplementary Material: Proof of Theorem 4.1 without communication device

Actions are periodically used as messages. Because players might have as few as two actions, each such communication phase might require several periods. As the actions played during this phase affect payoffs, communication phases must be short relative to regular phases. We shall not dispense with the randomization device altogether, as this allows us to achieve *exactly* the desired continuation payoff. Details on how to eliminate the public randomization device might be omitted altogether since they are the same as in the two-player case, following ideas introduced by Sorin (1986) and Fudenberg and Maskin (1991), and we refer the reader to Hörner and Lovo (2009).

Because communication requires several periods, strategies must also specify how a player plays within a communication phase if his own previous action, or his opponent's previous action already precludes him from reporting correctly his private information. The construction must ensure that continuation strategies remain optimal for all states after such histories, and this explains why the construction that follows is more involved than one might have guessed. (In particular, it is the cause for the different kinds of communication phase described below.)

Play is divided into phases (or classes of phases): Communication phases, regular phases, penitence phases, and punishment phases.

Actions

Communication Phase

The *communication phase* replaces the communication stage. There are different versions of communication phase, denoted C, C_i, or C_i^*. (Roughly, a phase is indexed by player i if i's report during this phase is essentially ignored.[1]) A communication phase lasts c periods, where

$$c \geq 1 + \max_{i \in N} \frac{\ln |\Theta_i|}{\ln |A_i|},$$

so that $|A_i|^{c-1} \geq |\Theta_i|$, all $i \in N$. We fix two arbitrary but distinct actions for each player, denoted U and B, and a mapping

$$m_i : \Theta_i \to A_i^{c-1},$$

from his set of types into sequences of actions of length $c-1$. Player i (or his play) *reports* θ_i if his play in the communication phase is equal to $(m_i(\theta_i), B)$ (so B is the action that he takes in the last period of this phase.) For any other play, he *reports* (U, n_i^U) where n_i^U is the number of periods in the communication phase in which $a_i = U$. We also write U rather than (U, n_i^U) whenever convenient, and let

$$\overline{\theta} \in \prod\nolimits_{i \in N} \Theta_i \cup \cup_{l=0}^{c} (U, l)$$

denote a *report*, or *message* profile. For $k \in K$, let $u_i^C(k, \overline{\theta})$ denote player i's average payoff from the communication phase if the state is k and the report is $\overline{\theta}$.[2]

In a communication phase C, player j's type θ_j plays the sequence $m_i(\theta_j, B)$, as long as his previous play in the phase does not preclude him from doing so. In a communication phase C_i so does player $j \neq i$, while player i plays (U, c). If a player's past play prevents him from reporting his type θ_i, he plays U in every remaining period of the phase.

[1] It cannot be entirely ignored, since we must give i incentives that do not depend on his type.

[2] This is an abuse of terminology, as payoffs are not uniquely identified by the report profile whenever a player reports U, since there might be many sequences of actions corresponding to this report. What is meant is the payoff given the actual sequence of action profiles.

Transitions are described below.

Regular Phase

A *regular phase* is denoted $R(\overline{\theta},\varepsilon)$, where $\kappa(\overline{\theta}) \neq \varnothing$, and $\varepsilon \in [-\overline{\varepsilon},\overline{\varepsilon}]^N$, for some $\overline{\varepsilon} > 0$ to be specified.

A regular phase lasts at most n periods (to be specified), where $n > c$. We fix a (possibly correlated) mixed action profile $\mu(\overline{\theta},\varepsilon) \in \triangle A$ such that, $\forall k \in \kappa(\overline{\theta})$, $\forall i \in N$, $\forall \varepsilon, \varepsilon' \in [-\overline{\varepsilon},\overline{\varepsilon}]^N$ and $\forall \theta_i' \in \Theta_i, \theta_i' \neq \overline{\theta}_i$, such that $\kappa(\theta_i',\overline{\theta}_{-i}) \neq \varnothing$,

$$u_i^R(k,\mu(\overline{\theta},\varepsilon)) := (1-\delta^n)u_i(k,\mu(\overline{\theta},\varepsilon)) + \delta^n u_i^C(k,\overline{\theta}) = v_i^k + \varepsilon_i,$$

and

$$u_i^R(k,\mu(\overline{\theta},\varepsilon)) > u_i^R(k,\mu(\theta_i',\overline{\theta}_{-i},\varepsilon')),$$

and

$$u_i^R(k,\mu(\theta_i',\overline{\theta}_{-i},\varepsilon')) \leq v_i^k - 2\overline{\varepsilon}.$$

The strict inequalities can be satisfied for δ close enough to 1 and $\overline{\varepsilon}$ close enough to 0, since v is strictly incentive compatible.

In any period of the regular phase, players play $\mu(\overline{\theta},\varepsilon)$. The regular phase $R(\overline{\theta},\varepsilon)$ stops immediately after a unilateral deviation from $\mu(\overline{\theta},\varepsilon)$, or if not, after n periods.

Transitions are described below.

Penitence Phase

A *penitence phase* is denoted $E(\overline{\theta},\varepsilon)$, where $\varepsilon \in [-\overline{\varepsilon},\overline{\varepsilon}]^N$, $\overline{\theta} \in \Theta$, $\kappa(\overline{\theta}) = \varnothing$, and $\overline{\theta} \in D$. A penitence phase lasts at most n periods. We fix a sequence $a(\overline{\theta},\varepsilon) \in A^n$ such that $\forall (i, \theta'_i) \in \Omega_{\overline{\theta}}$, $k \in \kappa(\theta'_i, \overline{\theta}_{-i}), \varepsilon \in [-\overline{\varepsilon}, \overline{\varepsilon}]^N$,

$$u_i^E(k, a(\overline{\theta}, \varepsilon)) := \frac{1-\delta}{1-\delta^n} \sum\nolimits_{t=0}^{n-1} \delta^t u_i(k, a_t(\overline{\theta}, \varepsilon)) < v_i^k - 2\overline{\varepsilon}.$$

Such a penitence phase $E(\overline{\theta}, \varepsilon)$ stops immediately after a unilateral deviation from the sequence $a(\overline{\theta}, \varepsilon)$, or if not, after n periods. In period t of the penitence phase, players play $a_t(\overline{\theta}, \varepsilon)$.

Transitions are described below.

Punishment Phase

A *punishment phase*, indexed by i, is denoted $P_i(\overline{\theta}_{-i}, t)$, where $\overline{\theta}_{-i} \in \Theta_{-i}$ is such that $\kappa(\overline{\theta}_{-i}) \neq \varnothing$ and $t = n$ or T (to be defined) denotes the length of the punishment phase.

As before, we fix an action $\underline{a}_i \in A_i$ and let $s_i^{\underline{a}_i}$ denote the strategy of playing $\underline{a}_i$ in every period, independently of the history. In the punishment phase, player i uses $s_i^{\underline{a}_i}$, and players $-i$ use $s_{-i}^{\overline{\theta}_{-i}}$.

We pick $n, T, \overline{\delta} < 1$ and $\overline{\varepsilon}$ such that, $\forall \delta > \overline{\delta}$, $\forall k \in \kappa(\overline{\theta}_{-i})$, player i's average discounted payoff over the t periods in state k is no larger than $v_i^k - 2\varepsilon$, and that it is sufficiently larger when $t = n$ than when $t = T$, as explained below. This is possible since v satisfies individual rationality strictly.

We shall write C, R, E, P for a communication, regular, penitence and punishment phase without further argument when there is no risk of confusion.

Transitions

Given any message $\overline{\theta}$, define

- whenever $\overline{\theta} \in \Theta$, $\forall \theta \in \Theta$, $\Delta_I(\theta, \overline{\theta}) := \{i \in N | \theta_i \neq \overline{\theta}_i\}$;

- whenever $\overline{\theta} \in \Theta$, $\overline{\theta} \in D$, $\Delta_D(\overline{\theta}) := \{i \in N | (i, \theta'_i) \in \Omega_{\overline{\theta}}$ for some $\theta'_i \in \Theta_i\}$;

- whenever $\overline{\theta} \notin \Theta$, $\Delta_U(\overline{\theta}) := \{i \in N | \overline{\theta}_i \notin \Theta_i\}$.

Given a unilateral deviation from a sequence $a(\overline{\theta}, \varepsilon)$, or from a mixed action $\mu(\overline{\theta}, \varepsilon)$, let Δ_A denote the index of the player who deviated.[3] Finally, given a set $\Delta \subset N$, let $-\Delta := N \setminus \Delta$.

From a communication phase

The transition depends on the message $\overline{\theta}$ during C, the phase $\Phi \in \{R, P, E, C\}$ immediately preceding C, and the play during Φ. Roughly speaking, if there is no unilateral deviation during Φ, and if $\overline{\theta} \in \Theta$, a regular or a penitence phase follows, while if $\overline{\theta} \notin \Theta$, either a punishment or a communication phase follows. If there is a unilateral deviation during Φ by player i, then if $\overline{\theta}_{-i} \in \Theta_{-i}$, a punishment phase follows. More precisely, if there is a unilateral deviation from $\Phi = E, R$, with $\Delta_A = \{i\}$, then the next phase is

1. if $\overline{\theta}_{-i} \in \Theta_{-i}$, $\kappa(\overline{\theta}_{-i}) \neq \varnothing$: $P_i(\overline{\theta}_{-i}, T)$;

2. otherwise, it is C.

On the other hand, if there is no unilateral deviation from Φ, or if $\Phi = P, C$, and

1. Φ equals $R(\theta, \varepsilon)$ or $E(\theta, \varepsilon)$, the next phase is:

   (a) if $\overline{\theta} \in \Theta$, $\kappa(\overline{\theta}) \neq \varnothing$: $R(\overline{\theta}, \varepsilon_{-\Delta_I(\theta,\overline{\theta})}, -\overline{\varepsilon}_{\Delta_I(\theta,\overline{\theta})})$;

   (b) if $\overline{\theta} \in \Theta$, $\overline{\theta} \in D$: $E(\overline{\theta}, \varepsilon_{-\Delta_D(\overline{\theta})}, -\overline{\varepsilon}_{\Delta_D(\overline{\theta})})$;

[3] Recall that there is a public randomization device, so that we always assume that players use a pure action profile, as a function of the realization of the public randomization device, so that the mixed action profile obtains in expectations.

(c) if $\Delta_U(\overline{\theta}) = \{i\}$, $\kappa(\overline{\theta}_{-i}) \neq \varnothing$: $P_i(\overline{\theta}_{-i}, n)$;

(d) otherwise, C;

2. Φ equals $P_i(\theta_{-i}, t)$, $t = n, T$, the next phase is:

(a) if $\overline{\theta} \in \Theta$, $\kappa(\overline{\theta}) \neq \varnothing$: $R(\overline{\theta}, \tilde{\varepsilon}(\theta, \overline{\theta}))$, where $\tilde{\varepsilon}_i(\theta, \overline{\theta}) \in [-\overline{\varepsilon}, \overline{\varepsilon}]$ is chosen so that, given $\overline{\theta}$ and $s_{-i}^{\theta_{-i}}$, using $s_i^{\underline{a}_i}$ is optimal in the punishment phase for player i; and further, if $\overline{\theta}_{-i} = \theta_{-i}$, player i's continuation payoff in the repeated game, evaluated at the beginning of the punishment phase, is equal to, for all $k \in \kappa(\overline{\theta})$,

$$(1 - \delta^t)(v_i^k - 2\overline{\varepsilon}) + \delta^t(v_i^k - \overline{\varepsilon});$$

and for $j \neq i$, $\tilde{\varepsilon}_j(\theta, \overline{\theta})$ is chosen so that, given $\overline{\theta}$, $s_{-j}^{\theta_{-j}}$ and $s_i^{\underline{a}_i}$, $s_j^{\theta_j}$ is optimal for player j in the punishment phase. Further $\tilde{\varepsilon}_j(\theta, \overline{\theta}) \in [\overline{\varepsilon}/4, 3\overline{\varepsilon}/4]$ if $\theta_j = \overline{\theta}_j$ and $\tilde{\varepsilon}_j(\theta, \overline{\theta}) \in [-3\overline{\varepsilon}/4, -\overline{\varepsilon}/4]$ otherwise;

(b) if $\overline{\theta} \in \Theta$, $\overline{\theta} \in D$: $E(\overline{\theta}, 0_{-\Delta_D(\overline{\theta})}, -\overline{\varepsilon}_{\Delta_D(\overline{\theta})})$;

(c) otherwise, C.

3. Φ equals C, or C_i and θ is the report during Φ, the next phase is:

(a) if $\overline{\theta} \in \Theta$, $\kappa(\overline{\theta}) \neq \varnothing$: $R(\overline{\theta}, \hat{\varepsilon}(\theta, \overline{\theta}))$, where, if $\Phi = C$, or $j \neq i$,

$$\hat{\varepsilon}_j(\theta, \overline{\theta}) = \begin{cases} 0 & : \quad \theta_j = \overline{\theta}_j, \\ -\overline{\varepsilon}/4 + \rho n_U & : \quad \theta_j = (U, n_U), \\ -\overline{\varepsilon} & : \quad \text{otherwise}, \end{cases}$$

and if $\Phi = C_i$,

$$\hat{\varepsilon}_i(\theta, \overline{\theta}) = \begin{cases} -\overline{\varepsilon} + \rho n_U & : \quad \theta_i = (U, n_U), \\ -\overline{\varepsilon} & : \quad \text{otherwise,} \end{cases}$$

for some $\rho > 0$ to be defined;

(b) if $\overline{\theta} \in \Theta$, $\overline{\theta} \in D$: $E(\overline{\theta}, 0_{-\Delta_D(\overline{\theta})}, -\overline{\varepsilon}_{\Delta_D(\overline{\theta})})$;

(c) if $\Delta_U(\overline{\theta}) = \{i\}$, $\kappa(\overline{\theta}_{-i}) \neq \varnothing$: $P_i(\overline{\theta}_{-i}, n)$;

(d) otherwise, C.

From any other phase

Any other phase is followed by a communication phase. If there is a unilateral deviation from a phase $\Phi = R, E$, with $\Delta_A = \{i\}$, it is a communication phase C_i; otherwise, it is a communication phase C.

Initial phase

The game starts with a communication phase, at the end of which transitions occur as if the previous phase had been C, with $\theta = \overline{\theta}$, and $\varepsilon \in [-\overline{\varepsilon}, \overline{\varepsilon}]$ is such that the payoff (inclusive of the initial communication phase) is equal to v.

Verification of optimality

Consider first the incentives of player i to deviate during a regular phase. If he does so, a punishment phase P_i will start after the communication phase. Player i expects the type profile θ_{-i} reported by the other players after the deviation and before the punishment phase to be correct; since his payoff at the beginning of the punishment phase is

$$(1-\delta^T)(v_i^k - 2\overline{\varepsilon}) + \delta^T(v_i^k - \overline{\varepsilon}),$$

then he has no incentive to deviate in this case, as whether or not his own report was correct, his payoff from following the equilibrium strategies is higher.[4]

Consider next a punishment phase P_i. The definition of $\tilde{\varepsilon}_i$ guarantees that $s_i^{a_i}$ is optimal for player i. Similarly, the definition of $\tilde{\varepsilon}_j$ ensures that player $j \neq i$ has no incentive to deviate. This is true whether the punishment phase lasts n or T periods.

Consider next a possible deviation during the penitence phase. While the average payoff from the penitence phase is low, observe that it lasts only n periods (and, given the equilibrium strategies, the ensuing communication phase will be followed by a regular phase if the player refrains from deviating, independently of the history up to the contemplated deviation), while the punishment phase that the deviation would trigger lasts T periods. We pick T and n so as to ensure that no such deviation is profitable.

Consider finally a possible deviation during a communication phase. Start with a communication phase C.

1. Assume first that the history in the communication phase is consistent with (possibly, among others) some type profile $\theta \in \Theta$ (i.e., the history in the communication phase is an initial segment of $(m_1(\theta_1), \ldots, m_N(\theta_N)))$, and θ_i is indeed player i's type. If the true state is θ, then by reporting U, a punishment phase P_i of length n will be entered, the expected payoff of which ensures that it is better not to do so. If the true state is not θ, then according to the equilibrium strategies, some player $j \neq i$ will report U in this

[4]Note that the situation where the reported type profile by $-i$ is incorrect is not relevant for verifying that player i does not deviate during the regular phase. This is because, at the time of the deviation, he expects the other player to report correctly their type during the communication phase.

communication phase. If player i reports U, a communication phase C will be entered, at the end of which a regular phase will be started, for which $\varepsilon_i < 0$ (pick ρ such that $-\overline{\varepsilon}/4 + \rho c < 0$); by sticking to the report of θ_i, either a communication phase C will start (in case θ_j and $\theta_{j'}$ differ from the true state for two players j, j'), in which case, in the ensuing regular phase, player i's ε_i is zero, or a punishment phase of length n will start, at the end of which, in the ensuing regular phase, player i's ε_i is at least $\overline{\varepsilon}/4$; of course, i's payoff during the n periods can be very low, but we can deter such deviations by picking ρ sufficiently small (but not too small, see below).

2. Assume next that the history in the communication phase is consistent with some type profile $\theta \in \Theta$, but θ_i is not player i's type. Thus, the equilibrium strategy calls for player i to report U (if there is at least one period; otherwise, there is nothing to show). Suppose first that the other players' type profile is indeed θ_{-i}. By reporting U, player i triggers a punishment phase P_i of length n, but by failing to do so, he triggers the play of a regular phase for which the play does not correspond to the true type profile. We can pick n small enough to guarantee that, since the payoff during such a regular phase is less that $v_i - \overline{\varepsilon}$, player i prefers not to deviate. Suppose next that there exists exactly one other player j for which θ_j is not the true type. By reporting U, a second communication phase starts, but player i is guaranteed at least a value of $\varepsilon_i \geq -\overline{\varepsilon}/4$ in the regular phase at the end of it; if player i persists in reporting the incorrect type, a punishment phase P_j of length n follows, at the end of which player i's ε is strictly less than $-\overline{\varepsilon}/4$; finally, if there are two or more other players for which θ_j is incorrect, and if player i reports U, he also guarantees that, in the regular phase that will follow the second communication phase, $\varepsilon_i \geq -\overline{\varepsilon}/4$; if he reports differently, in the regular phase that will follow the second communication phase, $\varepsilon_i = -\overline{\varepsilon}$.

3. Assume finally that the history in the communication phase is not consistent with some type profile $\theta \in \Theta$, i.e. some player reports U already. The same arguments as before apply almost *verbatim*, since in the previous arguments, if θ_j was not the true type for one or more players, those players j were about to report U anyway. Note that postponing a report of U by one or more periods within a communication phase is suboptimal, since the argument ε_i from the relevant ensuing regular phase is increasing in the number of times player i choose U. (This is where we need that ρ be not too small, more precisely, it must be at least $(1-\delta)M$).

These arguments are readily adapted to the case in which the communication phase is C_i. Consider first the case in which the previous phase was E or R (i.e., player i deviated in actions). Suppose first that the other players' type profile θ_{-i} is consistent with the history in the communication phase. Since the equilibrium calls for a punishment phase to follow, the specification of $\tilde{\varepsilon}_j$, $\tilde{\varepsilon}_i$ ensures that no player gains from deviating: i.e., player i benefits from playing U as often as possible, and other players gain by reporting their type truthfully. Suppose now that the history in the communication phase is not consistent with some type profile $\theta_i \in \Theta_i$, then some player $-i$ will play U and a new communication phase C will follow. Also in this case player i benefits from playing U since $\varepsilon_i = -\overline{\varepsilon} + \rho n_U$ in the regular phase that will follow the new communication C .