# Online Appendix

## F. An Example Where Separation Binds

Next we present a base game where (i) the set of separated BCEs is *nowhere dense* in the set of BCEs, (ii) for every player $i$, the utility function $u_i : A \to \mathbb{R}$ is *one-to-one* (i.e., no ties in the matrix below), and (iii) no action is weakly dominated.

| | $a_2$ | $b_2$ | $c_2$ |
|---|---|---|---|
| $a_1$ | $8, 8$ | $3, 7$ | $2, 6$ |
| $b_1$ | $7, 3$ | $5, 1$ | $0, 5$ |
| $c_1$ | $6, 2$ | $1, 4$ | $4, 0$ |

The game ($\Theta$ is a sigleton) has one pure Nash equilibrium and one mixed Nash equilibrium:

$$(a_1, a_2) \quad \text{and} \quad \left(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2\right).$$

The set of BCE is the set of convex combinations of the two Nash equilibria: for $t \in [0, 1]$,

$$p^t = t(a_1, a_2) + (1 - t)\left(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2\right).$$

The game has only two separated BCE, namely, the two Nash equilibria. Indeed, for every $t \in (0, 1)$ and every player $i$, the action recommendations $a_i$ and $b_i$ (or $c_i$) induce distinct posterior beliefs about the action of the opponent:

$$p^t_{a_i}(a_j) = 1 \quad \text{while} \quad p^t_{b_i}(b_j) = p^t_{b_i}(c_j) = \frac{1}{2}.$$

Yet, $a_i$ is best response to the belief induced by $b_i$:

$$\frac{1}{2}u_i(a_i, b_j) + \frac{1}{2}u_i(a_i, c_j) = \frac{5}{2} = \frac{1}{2}u_i(b_i, b_j) + \frac{1}{2}u_i(b_i, c_j).$$

Thus, $p^t$ is not separated: player $i$ does not have an incentive to acquire information about the correlation device; they could just play $a_i$ without acquiring any information.

## G. Arbitrary Information Technologies

In this section, we characterize the predictions attainable as one ranges across *all* information technologies. In particular, we do not require the information technology to be flexible or monotone. We also show it is without loss to require the technology to be flexible, and costs to be weakly monotone. Formally, a cost function $C_i$ is **weakly monotone** if less informative experiment are weakly cheaper to acquire: if $\xi_i, \xi_i' \in \mathcal{E}_i$ are such that $\xi_i \succsim \xi_i'$, then $C_i(\xi_i) \geq C_i(\xi_i')$.

**Proposition 3.** *Fix a base game $\mathcal{G}$. An information technology $\mathcal{T}$ exists that induces the outcome-value pair $(p, v)$ in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

*(i) $p$ is a BCE, and*

*(ii) for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$.*

*In addition, for every player $i$, one can choose $\mathcal{E}_i$ flexible and $C_i$ weakly monotone.*

*Proof.* "If." Let $(p, v)$ be an outcome-value pair such that $p$ is a BCE and, for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$. Since $p$ is a BCE, by Bergemann and Morris (2016) there exist an information structure $\mathcal{S} = (Z, \zeta, (X_i, \xi_i)_{i \in I})$ and a profile of action plans $\sigma = (\sigma_i)_{i \in I}$ such that $p$ is the outcome of $(\xi, \sigma)$, and for every player $i$, $\sigma_i$ maximizes

$$\sum_{a,x,z,\theta} u_i(a, \theta) \left( \sigma_i'(a_i|x_i) \xi_i(x_i|z, \theta) \prod_{j \neq i} \sigma_j(a_j|x_j) \xi_j(x_j|z, \theta) \right) \zeta(z|\theta) \pi(\theta) \tag{52}$$

over all $\sigma_i' \in \Sigma_i$.

For every player $i$, let $\mathcal{E}_i = \{\xi_i' : \xi_i \succeq \xi_i'\}$. In addition, take $\lambda_i \in [0, 1]$ such that

$$v_i = \lambda_i \underline{v}_i(p) + (1 - \lambda_i) \bar{v}_i(p).$$

For every $\xi_i' \in \mathcal{E}_i$, set $C_i(\xi_i')$ equal to

$$\lambda_i \left[ \max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a, \theta) \left( \sigma_i'(a_i|x_i) \xi_i'(x_i|z, \theta) \prod_{j \neq i} \sigma_j(a_j|x_j) \xi_j(x_j|z, \theta) \right) \zeta(z|\theta) \pi(\theta) - \underline{v}_i(p) \right].$$

Notice that $\mathcal{E}_i$ is flexible and $C_i$ is weakly monotone.

It follows from (52) that $C_i(\xi_i) = \lambda_i \left( \bar{v}_i(p) - \underline{v}_i(p) \right)$, which in turn implies that

$$\sum_{a,x,z,\theta} u_i(a, \theta) \left( \prod_{j} \sigma_j(a_j|x_j) \xi_j(x_j|z, \theta) \right) \zeta(z|\theta) \pi(\theta) - C_i(\xi_i) = v_i.$$

We also see that for every $\xi_i' \in \mathcal{E}_i$

$$\sum_{a,x,z,\theta} u_i(a,\theta)\left(\prod_j \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right)\zeta(z|\theta)\pi(\theta) - C_i(\xi_i)$$

$$= \max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta)\left(\sigma_i'(a_i|x_i)\xi_i(x_i|z,\theta)\prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right)\zeta(z|\theta)\pi(\theta) - C_i(\xi_i)$$

$$= \lambda_i \underline{v}_i(p) + (1-\lambda_i)\max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta)\left(\sigma_i'(a_i|x_i)\xi_i(x_i|z,\theta)\prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right)\zeta(z|\theta)\pi(\theta)$$

$$\geq \lambda_i \underline{v}_i(p) + (1-\lambda_i)\max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta)\left(\sigma_i'(a_i|x_i)\xi_i'(x_i|z,\theta)\prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right)\zeta(z|\theta)\pi(\theta)$$

$$= \max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta)\left(\sigma_i'(a_i|x_i)\xi_i'(x_i|z,\theta)\prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right)\zeta(z|\theta)\pi(\theta) - C_i(\xi_i'),$$

where the first equality follows from (52) and the weak inequality from $\xi_i \succeq \xi_i'$. We conclude $(\xi, \sigma)$ is an equilibrium of $(\mathcal{G}, \mathcal{T})$ with $\mathcal{T} := (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i\in I})$; in addition, $(p, v)$ is the outcome-value pair corresponding to $(\xi, \sigma)$.

"Only if." Let $(p, v)$ be the outcome-value pair of an equilibrium $(\xi, \sigma)$ of an information acquisition game $(\mathcal{G}, \mathcal{T})$, with $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i\in I})$. Define the information structure $\mathcal{S} = (Z, \zeta, (X_i, \xi_i)_{i\in I})$. Since $(\xi, \sigma)$ is an equilibrium of $(\mathcal{G}, \mathcal{T})$, $\sigma$ is an equilibrium of $(\mathcal{G}, \mathcal{S})$. It follows from Bergemann and Morris (2016) that $p$ is a BCE.

For every player $i$, $C_i(\xi_i) \geq 0$, which implies that $v_i \leq \bar{v}_i(p)$. In addition, by hypothesis there exists an experiment $\xi_i'$ such that $C_i(\xi_i') = 0$. Thus, since $(\xi_i, \sigma_i)$ is a best response to $(\xi_{-i}, \sigma_{-i})$, we have that

$$v_i \geq \max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta)\sigma_i'(a_i|x_i)\xi_i(x_i|z,\theta)\prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\zeta(z|\theta)\pi(\theta) \geq \underline{v}_i(p).$$

We conclude that $v_i \in [\bar{v}_i(p), \underline{v}_i(p)]$.

□

## H. Strict BCE: Single-Agent Settings

A BCE $p$ is **strict** if all $i \in I$, $a_i \in \text{supp}_i(p)$, and $b_i \in A_i$ with $b_i \neq a_i$,

$$\sum_{a_{-i},\theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\, p(a_i, a_{-i}, \theta) > 0.$$

In the main text, discussing Theorem 2, we mentioned the following result:

**Proposition 4.** *Let $I = \{i\}$ be a singleton. For generic $u_i$, the set of strict BCE is dense in the BCE set.*

We expect the result to be known in the literature. However, we could not find a good reference. Thus, next we provide a self-contained proof. The proof relies on two lemmas on dominated actions. A mixed action $\alpha_i \in \Delta(A_i)$ **weakly dominates** a pure action $a_i \in A_i$ if for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\sum_{b_i} u_i(b_i, a_{-i}, \theta)\alpha_i(b_i) \geq u_i(a_i, a_{-i}, \theta).$$

**Lemma 15.** *The following statements are equivalent:*

*(i) There is no belief $\mu_{a_i} \in \Delta(A_{-i} \times \Theta)$ for which $a_i$ is the unique best response.*

*(ii) There is a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$ that weakly dominates $a_i$.*

*Proof.* Condition (i) can be rewritten as

$$\max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \min_{b_i \in A_i \setminus \{a_i\}} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta) \leq 0.$$

Equivalently,

$$\max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta)\alpha_i(b_i) \leq 0.$$

By the minimax theorem (e.g., Rockafellar, 1970, Corollary 37.3.2), the above inequality holds if and only if

$$\min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta)\alpha_i(b_i) \leq 0.$$

Equivalently,

$$\min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \max_{a_{-i}, \theta} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\alpha_i(b_i) \leq 0.$$

which is another way of expressing condition (ii). □

A mixed action $\alpha_i \in \Delta(A_i)$ **strictly dominates** a pure action $a_i \in A_i$ if for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\sum_{b_i} u_i(b_i, a_{-i}, \theta)\alpha_i(b_i) > u_i(a_i, a_{-i}, \theta).$$

**Lemma 16.** *Let $I = \{i\}$ be a singleton. For generic $u_i$, if an action $a_i$ is weakly dominated by some mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$, then it is strictly dominated by some mixed action $\beta_i \in \Delta(A_i)$.*

*Proof.* Let $a_i$ be an action that is weakly dominated by a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$. Let $A_i'$ be the support of $\alpha_i$, and let $\Theta'$ be set of states $\theta$ for which

$$u_i(a_i, \theta) = \sum_{b_i} u_i(b_i, \theta)\alpha_i(b_i). \tag{53}$$

Let $m$ be the cardinality of $A_i'$, and let $n$ be the cardinality of $\Theta'$. We consider the $m \times n$ matrix $M \in \mathbb{R}^{A_i' \times \Theta'}$ given by

$$M(b_i, \theta) = u_i(a_i, \theta) - u_i(b_i, \theta).$$

For generic $u_i$, the matrix $M$ has full rank. By (53), the rows of $M$ are linearly dependent. Thus, the rank of $M$ must be $n$, the number of columns. We obtain that the row space of $M$ has dimension $n$. Hence, we can find $\beta_i \in \mathbb{R}^{A_i'}$ such that for every $\theta \in \Theta'$

$$\sum_{b_i} (u_i(a_i, \theta) - u(b_i, \theta))\beta_i(b_i) < 0.$$

For every $t > 0$, we define $\alpha_i^t \in \mathbb{R}^{A_i'}$ by

$$\alpha_i^t(b_i) = \frac{\alpha_i(b_i) + t\beta_i(b_i)}{\sum_{c_i} \alpha_i(c_i) + t\beta_i(c_i)}.$$

For $t$ sufficiently small, $\alpha_i^t$ is a mixed action that strictly dominates $a_i$. ∎

We are now ready to prove the proposition on strict BCE.

*Proof of Proposition 4.* Let $A_i^*$ be the set of actions that are not strictly dominated. Since $u_i$ is generic, it follows from Lemma 16 that each $a_i \in A_i^*$ is not weakly dominated by a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$. By Lemma 15, there is a belief $\mu_{a_i} \in \Delta(\Theta)$ for which $a_i$ is the unique best response.

Since $\pi$ has full support, we can find $\nu \in \Delta(\Theta)$ and for every $a_i \in A_i^*$, $t_{a_i} \in (0, 1)$—with $\sum_{a_i \in A_i^*} t_{a_i} \leq 1$—such that

$$\pi = \sum_{a_i \in A_i^*} t_{a_i}\mu_{a_i} + \left(1 - \sum_{a_i \in A_i^*} t_{a_i}\right)\nu.$$

Let $a_i^*$ be a best response to $\nu$; necessarily, $a_i^* \in A_i^*$. Define the outcome $p \in \Delta_\pi(A_i \times \Theta)$

as follows:

$$p(a_i,\theta)=\begin{cases} t_{a_i}\mu_{a_i}(\theta) & \text{if } a_i\in A_i^*\setminus\{a_i^*\},\\ t_{a_i^*}\mu_{a_i^*}(\theta)+\left(1-\sum_{a_i\in A_i^*}t_{a_i}\right)\nu(\theta) & \text{if } a_i=a_i^*,\\ 0 & \text{otherwise.}\end{cases}$$

The outcome $p$ is a strict BCE. Moreover, if $q$ is a BCE, then

$$\mathrm{supp}_i(q)\subseteq A_i^*=\mathrm{supp}_i(p).$$

Thus, the set of outcomes

$$\{sq+(1-s)p : s\in(0,1) \text{ and } q\in BCE\}$$

is a subset of the set of strict BCE, and it is dense in the BCE set. We conclude that (for generic $u_i$) the set of strict BCE is dense in the BCE set. □

# I. Non-generic Environments

## I.1. Proofs of Proposition 2 and Theorem 4

In this section we prove generalizations of Proposition 2 and Theorem 4 that apply locally to a closed convex set of BCEs.

Fix a base game $\mathcal{G}$; denote by $BCE$ the set of all BCEs, and by $sBCE$ the set of all sBCEs. Let $P\subseteq BCE$ be a non-empty closed convex set. For a player $i$, an action $a_i$ $P$-**jeopardizes** an action $b_i$ if, for every $p\in P$ with $b\in\mathrm{supp}_i(p)$, $a_i\in BR(p_{b_i})$. We denote by $J_P(b_i)$ the set of actions that $P$-jeopardizes $b_i$. Just like the standard jeopardization concept, one has $b_i\in J_P(b_i)$ for all $b_i$.

An outcome $p\in P$ has $P$-**maximal support** if the support of every other $q\in P$ is contained by the support of $p$. An outcome $p\in P$ is $P$-**minimally mixed** if it has $P$-maximal support and

$$q_{a_i}\neq q_{b_i} \quad \text{implies} \quad p_{a_i}\neq p_{b_i}.$$

if for every $q\in P$, $i\in I$, and $a_i,b_i\in\mathrm{supp}_i(p)$,

We are now ready to state the local versions of Proposition 2 and Theorem 4 that we prove in this section.

**Proposition 5.** *The following statements are equivalent:*

*(i) The set $sBCE\cap P$ is dense in $P$.*

*(ii) A $P$-minimally mixed sBCE exists.*

*(iii) For all $p \in P$, $i \in I$, $a_i, b_i \in \operatorname{supp}_i(p)$,*

$$p_{a_i} \neq p_{b_i} \quad \textit{implies} \quad J_P(a_i) \cap J_P(b_i) = \varnothing.$$

**Theorem 6.** *The set $sBCE \cap P$ is either dense or nowhere dense in $P$.*

As a first step, we prove a basic lemma about best responses. In what follows, for $p \in \Delta(A \times \Theta)$, $a_i \in \operatorname{supp}_i(p)$, and $b_i \in A_i$, take $u_i(b_i, p_{a_i}) \in \mathbb{R}$ to be

$$u_i(b_i, p_{a_i}) = \sum_{a_{-i}, \theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

**Lemma 17.** *For every $t \in (0,1)$, $p, q \in BCE$, $i \in I$, and $a_i \in \operatorname{supp}_i(p)$,*

$$BR\left((tp + (1-t)q)_{a_i}\right) \subseteq BR(p_{a_i}).$$

*Proof.* Take $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$. If $a_i \notin \operatorname{supp}_i(q)$, then

$$(tp + (1-t)q)_{a_i} = p_{a_i},$$

which immediately implies the desired result.

Suppose now that $a_i \in \operatorname{supp}_i(q)$. Since $p, q \in BCE$, we have

$$u_i(a_i, p_{a_i}) \geq u_i(b_i, p_{a_i}) \quad \text{and} \quad u_i(a_i, q_{a_i}) \geq u_i(b_i, q_{a_i}).$$

Simple algebra shows that there exists $s \in (0,1)$ such that

$$(tp + (1-t)q)_{a_i} = sp_{a_i} + (1-s)q_{a_i}.$$

Since $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$, we obtain that

$$\begin{aligned} su_i(b_i, p_{a_i}) + (1-s)u_i(b_i, q_{a_i}) &= u_i(b_i, sp_{a_i} + (1-s)q_{a_i}) \\ &\geq u_i(a_i, sp_{a_i} + (1-s)q_{a_i}) \\ &= su_i(a_i, p_{a_i}) + (1-s)u_i(a_i, q_{a_i}). \end{aligned}$$

We conclude that $u_i(a_i, p_{a_i}) = u_i(b_i, p_{a_i})$ and $u_i(a_i, q_{a_i}) = u_i(b_i, q_{a_i})$. It follows from $p \in BCE$ that $b_i \in BR(p_{a_i})$. □

Next, we show that taking convex combinations of BCEs usually preserve the set of action recommendations that lead to different beliefs.

**Lemma 18.** *For every $p, q \in \Delta(A \times \Theta)$, $i \in I$, and $a_i, b_i \in \text{supp}_i(p)$ with $p_{a_i} \neq p_{b_i}$, there are at most two $t \in (0,1)$ such that*

$$(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}. \tag{54}$$

*Proof.* Note that $t \in (0,1)$ is a solution of (54) if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\begin{aligned}&\left(tp(a_i, a_{-i}, \theta) + (1-t)q(a_i, a_{-i}, \theta)\right)\left(tp(b_i) + (1-t)q(b_i)\right)\\ =&\left(tp(b_i, a_{-i}, \theta) + (1-t)q(b_i, a_{-i}, \theta)\right)\left(tp(a_i) + (1-t)q(a_i)\right).\end{aligned} \tag{55}$$

Each equation (55) is polynomial in $t$, with degree at most two. Since $p_{a_i} \neq p_{b_i}$, at least one such polynomial equation does not have degree zero and, therefore, has at most two solutions. We deduce that (54) has at most two solutions for $t \in (0,1)$. □

Our next goal is to show that $P$-minimally mixed BCEs are the norm rather than the exception. As an intermediate step, we first show the set of $P$-minimally mixed BCEs is non-empty.

**Lemma 19.** *A $P$-minimally mixed BCE exists.*

*Proof.* For every $p \in P$, define the set

$$X(p) = \bigcup_i \left\{(a_i, b_i) : a_i, b_i \in \text{supp}_i(p) \text{ and } p_{a_i} \neq p_{b_i}\right\}.$$

Note that $p \in P$ is $P$-minimally mixed if and only if it has $P$-maximal support and for every $q \in P$, $X(q) \subseteq X(p)$.

Since the set $A \times \Theta$ is finite and the set $P$ is convex, we can find a $P$-maximal support $p \in P$ such that for every $P$-maximal-support $q \in P$, the cardinality of $X(p)$ is larger than the cardinality of $X(q)$.

We now show that $p$ is $P$-minimally mixed. Fix an arbitrary $q \in P$. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$, which belongs to $P$ because $P$ is convex. Since $p$ has $P$-maximal support, the same is true for $p^t$. Thus, the cardinality of $X(p)$ is larger than the cardinality of $X(p^t)$. By Lemma 18, we can find $t \in (0,1)$ such that $X(p) \subseteq X(p^t)$ and $X(q) \subseteq X(p^t)$. This shows that $X(q) \subseteq X(p)$; otherwise, the cardinality of $X(p^t)$ would be strictly larger than the cardinality of $X(p)$. We conclude that $p$ is $P$-minimally mixed. □

We now show that the $P$-minimally mixed BCEs includes most of the BCEs in $P$ in a precise sense.

**Lemma 20.** *The set of $P$-minimally mixed BCEs is open and dense in $P$.*

*Proof.* Let $P_M$ denote the set of $P$-minimally mixed BCEs. We first argue that $P_M$ is open in $P$. Towards this goal, note the following sets are open in $P$ for every $i \in I$ and $a_i, b_i \in A_i$:

$$\{p \in P : p(a_i) > 0\}, \quad \text{and} \quad \{p \in P : p(a_i)p(b_i) > 0 \text{ and } p_{a_i} \neq p_{b_i}\}.$$

Since $A$ is finite, we obtain that $P_M$ equals the intersection of a finite number of open subsets of $P_M$. It follows $P_M$ is open in $P$.

To see $P_M$ is dense in $P$, fix some $q \in P$. Take $p$ to be a $P$-minimally mixed BCE, which exists by Lemma 19. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$. Because $p$ has $P$-maximal support, the same is true for $p^t$ for all $t \in (0,1)$. Moreover, by Lemma 18, a finite set $T \subseteq (0,1)$ exists such that for all $t \in (0,1) \setminus T$, $i \in I$, and $a_i, b_i \in \text{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \text{implies} \quad p^t_{a_i} \neq p^t_{b_i}.$$

Thus, $p^t$ is a $P$-minimally mixed BCE for all $t \in (0,1) \setminus T$. Thus, $q$ is a limit point of $\{p^t : t \in (0,1) \setminus T\}$, which implies it is a limit point of $P_M$. □

We are now ready to prove Proposition 5 and Theorem 6.

*Proof of Proposition 5.* That (i) implies (ii) follows from Lemma 20.

We now show (ii) implies (iii). Let $q$ be a $P$-minimally mixed sBCE. Fix any $p \in P$, $i \in I$ and $a_i, b_i \in \text{supp}_i(p)$ such that $p_{a_i} \neq p_{b_i}$. Since $q$ is $P$-minimally mixed, $a_i, b_i \in \text{supp}_i(q)$ (because $q$ has $P$-maximal support) and $q_{a_i} \neq q_{b_i}$. Thus,

$$\varnothing = BR(q_{a_i}) \cap BR(q_{b_i}) \supseteq J_P(a_i) \cap J_P(b_i),$$

where first we use the separation constraint, and then the fact that $J_P(c_i) = \cap_{\tilde{p} \in P} BR(\tilde{p}_{c_i})$ for all $c_i \in A_i$. We conclude (ii) implies (iii).

Finally, we argue (iii) implies (i). Fix any $p \in P$. Because $A$ is finite and $P$ is convex, it follows from Lemma 17 that we can find $q \in P$ such that $q$ has $P$-maximal support and

$$BR(q_{a_i}) = J_P(q_{a_i}) \tag{56}$$

for all $i \in I$ and $a_i \in \text{supp}_i(q)$.

For $t \in (0,1)$, let $p^t = tp + (1-t)q$. We claim that $p^t \in sBCE \cap P$. That $p^t \in P$ follows from $P$ being convex. To see $p^t$ is a sBCE, take any $i \in I$ and $a_i, b_i \in \text{supp}_i(p^t)$ such that $p^t_{a_i} \neq p^t_{b_i}$. Since $q$ has maximal support, $a_i, b_i \in \text{supp}_i(q)$. Then,

$$BR(p^t_{a_i}) \cap BR(p^t_{b_i}) \subseteq BR(q_{a_i}) \cap BR(q_{b_i}) = J_P(a_i) \cap J_P(b_i) = \varnothing,$$

where first we use Lemma 17, then (56), and finally Proposition 5-(iii). We conclude $p^t \in sBCE \cap P$ for all $t \in (0,1)$. Proposition 5-(i) then follows from $p = \lim_{t\to 1} p^t$. $\square$

*Proof of Theorem 6.* It is enough to prove that if $sBCE \cap P$ is not nowhere dense in $P$, then it is dense in $P$. Suppose $sBCE \cap P$ is dense in some non-empty set $\tilde{P} \subseteq P$ that is open in $P$. Let $P_M$ the set of $p \in P$ that are $P$-minimally mixed. Note $\tilde{P} \cap P_M$ is open (in $P$) and non-empty by Lemma 20. Because $sBCE \cap P$ is dense in $\tilde{P}$, we obtain that $(sBCE \cap P) \cap (\tilde{P} \cap P_{MM})$ is non-empty. Thus, we have found a $P$-minimally mixed sBCE. That $sBCE \cap P$ is dense in $P$ then follows from Proposition 5. $\square$

## I.2. Checking for Equal Beliefs

To check the conditions of Proposition 2, knowing which actions induce different beliefs for some BCE is useful. In this section, we prove a result that shows how to find actions that lead to different beliefs in a closed convex set of outcomes $P \subseteq \Delta(A \times \Theta)$.[23]

For a player $i$, say an action $a_i$ is **$P$-coherent** if a $p \in P$ exists with $p(a_i) > 0$.[24] Let $\mathbf{0}$ be the all-zeros vector in $\mathbb{R}^{A_{-i}\times\Theta}$; in what follows, we use the convention that $p_{a_i} = \mathbf{0}$ for every $p \in \Delta(A \times \Theta)$ and $a_i \in A_i$ such that $p(a_i) = 0$. As in the previous section, we say that an outcome $p \in P$ has **$P$-maximal support** if the support of every other $q \in P$ is contained by the support of $p$.

**Proposition 6.** *Fix a player $i$ and two $P$-coherent actions $a_i, b_i \in A_i$. Then every $p \in P$ with $a_i, b_i \in \operatorname{supp}_i(p)$ has $p_{a_i} = p_{b_i}$ if and only if one of the following two conditions hold:*

*(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $p \in \operatorname{ext}(P)$, $\{p_{a_i}, p_{b_i}\} \subseteq \{\mu, \boldsymbol{0}\}$.*

*(ii) A constant $\lambda > 0$ exists such that for all $p \in \operatorname{ext}(P)$, $p(a_i)p_{a_i} = \lambda p(b_i)p_{b_i}$.*

Thus, to know whether a pair of actions leads to the same beliefs in all outcomes in $P$, it is enough to check the extreme points of $P$ for one of two properties. The first property states these actions induce the same beliefs in all of the set's extreme points. The second property requires the likelihood ratio for these actions to be constant across all these extreme points.

To prove the proposition, we need the following lemma.

**Lemma 21.** *Fix a player $i$ and two actions $a_i, b_i \in A_i$. Let $p, q \in \Delta(A \times \Theta)$ such that $\{a_i, b_i\} \subseteq \operatorname{supp}_i(p) \cup \operatorname{supp}_i(q)$. Suppose $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \operatorname{supp}_i(r)$. If $(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}$ for some $t \in (0,1)$, then one of the following two conditions hold:*

[23] Neither the obedience nor the separation constraint play any role in this section.

[24] Our notion of $P$-coherent is inspired by the notion of coherence in Nau and McCardle (1990).

*(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $r \in \{p, q\}$, $\{r_{a_i}, r_{b_i}\} \subseteq \{\mu, \boldsymbol{0}\}$.*

*(ii) A constant $\lambda > 0$ exists such that for all $r \in \{p, q\}$, $r(a_i) = \lambda r(b_i)$.*

*Proof.* Let $p^t := tp + (1-t)q$. We proceed by contradiction: we assume that Lemma 21-(i) and Lemma 21-(ii) both fail and show that $p^t_{a_i} \neq p^t_{b_i}$.

We begin by noting that one can rewrite the condition that $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \mathrm{supp}_i(r)$ as

$$r(a_i)r(b_i)r_{a_i} = r(a_i)r(b_i)r_{b_i} \text{ for all } r \in \{p, q\}. \tag{57}$$

Because $\mathrm{supp}_i(p^t) = \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$ and $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$, we have $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p^t)$. Thus, applying Bayes rule, we obtain that $p^t_{a_i} = p^t_{b_i}$ if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, one has

$$p^t(a_i)p^t(b_i, a_{-i}, \theta) - p^t(b_i)p^t(a_i, a_{-i}, \theta) = 0.$$

Expanding the left hand side of the above equation by substituting in the definition of $p^t$, rearranging terms as a polynomial in $t$, and using (57), delivers that the above display equation is equivalent to

$$(t - t^2)\Big[p(a_i)q(b_i, a_{-i}, \theta) + q(a_i)p(b_i, a_{-i}, \theta) - q(b_i)p(a_i, a_{-i}, \theta) - p(b_i)q(a_i, a_{-i}, \theta)\Big] = 0.$$

Since $t \in (0, 1)$, we get that $p^t_{a_i} = p^t_{b_i}$ if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, one has

$$p(a_i)q(b_i, a_{-i}, \theta) + q(a_i)p(b_i, a_{-i}, \theta) - q(b_i)p(a_i, a_{-i}, \theta) - p(b_i)q(a_i, a_{-i}, \theta) = 0.$$

Writing the above in vector notation delivers that $p^t_{a_i} = p^t_{b_i}$ is equivalent to

$$p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} = \mathbf{0}. \tag{58}$$

We now divide the proof into cases. Consider first the case in which $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p) \cap \mathrm{supp}_i(q)$. In this case, (57) implies $p_{a_i} = p_{b_i}$ and $q_{a_i} = q_{b_i}$, and so we get that

$$\begin{aligned} p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} &= \\ = (p(a_i)q(b_i) - p(b_i)q(a_i))(q_{a_i} - p_{a_i}) &\neq \mathbf{0}, \end{aligned}$$

where the inequality follows from failure of Lemma 21-(i) and Lemma 21-(ii). We conclude (58) fails.

Consider now the case in which $\{a_i, b_i\} \nsubseteq \mathrm{supp}_i(p) \cap \mathrm{supp}_i(q)$. Because Lemma 21-(ii) fails, we can assume $p(a_i) = 0 < p(b_i)$ without loss of generality. Since the lemma assume

$a_i \in \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$, it follows $q(a_i) > 0$. Therefore, we can use failure of Lemma 21-(ii) to deduce that $p_{b_i} \neq q_{a_i}$. Using these facts, we obtain that

$$p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} = q(a_i)p(b_i)(p_{b_i} - q_{a_i}) \neq \mathbf{0}.$$

It follows that (58) fails. $\square$

We are now ready to prove Proposition 6. The "if" portion is straightforward; the "only if" portion uses Lemma 21.

*Proof of Proposition 6.* We first prove the "if" portion. Let $p \in P$ and $a_i, b_i \in \mathrm{supp}_i(p)$. Let $t^1, \ldots, t^n > 0$ and $p^1, \ldots, p^n \in \mathrm{ext}(P)$ such that

$$p = \sum_{m=1}^{n} t^m p^m.$$

Simple algebra shows that for all $c_i \in \mathrm{supp}_i(p)$

$$p_{c_i} = \sum_{m=1}^{n} \frac{t^m p^m(c_i)}{\sum_{l=1}^{n} t^l p^l(c_i)} p^m_{c_i}.$$

If Proposition 6-(i) holds, then

$$\begin{aligned} p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} &= \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} \mu \\ &= \mu \\ &= \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} \mu = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}. \end{aligned}$$

Suppose now that Proposition 6-(ii) holds. For every $m$, $p^m(a_i)p^m_{a_i} = \lambda p^m(b_i)p^m_{b_i}$ implies $p^m(a_i) = \lambda p^m(b_i)$ and $p^m_{a_i} = p^m_{b_i}$. Thus,

$$p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} = \sum_{m=1}^{n} \frac{t^m \lambda p^m(b_i)}{\sum_{l=1}^{n} t^l \lambda p^l(b_i)} p^m_{b_i} = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}.$$

This concludes the proof of the proposition's "if" portion.

We now show the proposition's "only if" portion. We proceed by contradiction: we assume that Proposition 6-(i) and Proposition 6-(ii) both fail and show that there exists $p \in P$ such that $a_i, b_i \in \mathrm{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. As we are done if $p_{a_i} \neq p_{b_i}$ for some $p \in \mathrm{ext}(P)$ with $a_i, b_i \in \mathrm{supp}_i(p)$, assume $p_{a_i} = p_{b_i}$ holds for all such $p$.

Since Proposition 6-(i) fails, and $a_i$ and $b_i$ are $P$-coherent, there exist $p, q \in \mathrm{ext}(P)$ such that $p(a_i) > 0$, $q(b_i) > 0$, and $p_{a_i} \neq q_{b_i}$. As we are done if $(0.5p + 0.5q)_{a_i} \neq (0.5p + 0.5q)_{b_i}$, assume $(0.5p+0.5q)_{a_i} = (0.5p+0.5q)_{b_i}$. Since $p_{a_i} \neq q_{b_i}$, Lemma 21-(i) fails. Thus, Lemma 21-(ii) must hold: there exist $\lambda > 0$ such that $p(a_i) = \lambda p(b_i)$ and $q(a_i) = \lambda q(b_i)$; in particular, $p(b_i) > 0$ and $q(a_i) > 0$.

Since Proposition 6-(ii) fails, there must exist $r \in \text{ext}(P)$ such that $r(a_i) \neq \lambda r(b_i)$; in particular, $r(a_i) > 0$ or $r(b_i) > 0$. Let $c_i \in \{a_i, b_i\}$ such that $r(c_i) > 0$. Since $p_{a_i} \neq q_{b_i}$, either $r_{c_i} \neq p_{a_i}$, or $r_{c_i} \neq p_{b_i}$, or both. Thus, by Lemma 21, either $(0.5p + 0.5r)_{a_i} \neq (0.5p + 0.5r)_{c_i}$, or $(0.5q + 0.5r)_{b_i} \neq (0.5q + 0.5r)_{c_i}$, or both. In any case, we have found $p \in P$ such that $a_i, b_i \in \text{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. □

# J. Vanishing Cost Equilibria

## J.1. Proof of Theorem 5

The "only if" side of the theorem follows from Theorem 1. Next we prove the "if" side.

Let $p$ be a complete-information Nash equilibrium. Following the notation in the main text, let $\alpha_{\theta,i} \in \Delta(A_i)$ be the distribution of $i$'s action given $\theta$.

Let $(p^n)_{n\in\mathbb{N}}$ be a sequence of sBCE that converges to $p$. For every player $i$ and every $n \in \mathbb{N}$, define $\mathcal{A}_i^{p^n}$ as in Section A.2. Thus, $\mathcal{A}_i^{p^n}$ is the partition of $A_i$ such that $a_i$ and $b_i$ are in the same cell if and only if either $a_i, b_i \in \text{supp}_i(p^n)$ and $p^n_{a_i} = p^n_{b_i}$, or $a_i, b_i \notin \text{supp}_i(p^n)$. Without loss of generality, we assume that $\mathcal{A}_i^{p^n} = \mathcal{A}_i^{p^m}$ for all $i \in I$ and $m, n \in \mathbb{N}$ (pass to a subsequence if necessary). To ease the exposition, we write $\mathcal{A}_i$ instead of $\mathcal{A}_i^{p^n}$; we also write

$$\mathcal{A}_{-i} = \left\{ \prod_{j\neq i} B_j : B_j \in \mathcal{A}_j \text{ for all } j \neq i \right\},$$
$$\mathcal{A} = \{B_i \times B_{-i} : B_i \in \mathcal{A}_i \text{ and } B_{-i} \in \mathcal{A}_{-i}\}.$$

Next, we construct one canonical representation $(\mathcal{S}^n, \sigma^n)$ for each $p^n$ (see Section A.2 for the general definition of canonical representation). The information structure $\mathcal{S}^n = (Z, \zeta^n, (X_i, \xi_i)_{i\in I})$ is specified as follows:

- To construct $Z \subseteq \prod_{i\in I} \Delta(\mathcal{A}_i)$, let $\delta_{B_i} \in \Delta(\mathcal{A}_i)$ be the Dirac measure concentrated on $B_i \in \mathcal{A}_i$. Moreover, let $\bar{\alpha}_{\theta,i}$ be the measure on $\mathcal{A}_i$ induced by $\alpha_{\theta,i}$: for $B_i \in A_i$,

$$\bar{\alpha}_{\theta,i}(B_i) = \sum_{a_i\in B_i} \alpha_{\theta,i}(a_i).$$

We set $Z = \prod_i Z_i$ where for every player $i$,

$$Z_i = \{\delta_{B_i} : B_i \in A_i\} \bigcup \{\bar{\alpha}_{\theta,i} : \theta \in \Theta\}.$$

- To construct $\zeta^n : \Theta \to \Delta(Z)$, first we take a sequence $(t^n)_{n\in\mathbb{N}}$ in $(0,1)$ such that (i) $t^n \to 1$, and (ii) for every $n$, $t^n p < p^n$ (such a sequence exists because $p^n \to p$). For every $n$, we define the outcome $r^n \in \Delta(A \times \Theta)$ by

$$r^n(a,\theta) = \frac{p^n(a,\theta) - t^n p(a,\theta)}{1 - t^n},$$

and the Markov kernel $\zeta^{r,n} : \Theta \to \Delta(Z)$ by

$$\zeta^{r,n}(z|\theta) = \begin{cases} \sum_{a\in B} \frac{r^n(a,\theta)}{\pi(\theta)} & \text{if } B \in \mathcal{A} \text{ and } z = (\delta_{B_i})_{i\in I}, \\ 0 & \text{otherwise.} \end{cases}$$

We also denote by $\zeta^p : \Theta \to \Delta(Z)$ the Markov kernel given by

$$\zeta^p(z|\theta) = \begin{cases} 1 & \text{if } z = (\bar{\alpha}_{\theta,i})_{i\in I}, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, we construct $\zeta^n : \Theta \to \Delta(Z)$ as follows:

$$\zeta^n(z|\theta) = (1 - t^n)\zeta^{r,n}(z|\theta) + t^n \zeta^p(z|\theta).$$

- For every player $i$, we take $X_i$ sufficiently rich so that $\mathcal{A}_i \subseteq X_i$.

- For every $i \in I$, $x_i \in X_i$, $z \in Z$, and $\theta \in \Theta$, we define the experiment $\xi_i$ by

$$\xi_i(x_i|z,\theta) = \begin{cases} z_i(x_i) & \text{if } x_i \in \mathcal{A}_i, \\ 0 & \text{otherwise.} \end{cases}$$

The profile of action plans $\sigma^n = (\sigma_i^n)_{i\in I}$ is given by, for all $i \in I$, $a_i \in A_i$, and $x_i \in \mathcal{A}_i$,

$$\sigma_i^n(a_i|x_i) = \begin{cases} \frac{p^n(a_i)}{\sum_{b_i\in x_i} p^n(b_i)} & \text{if } a_i \in x_i \text{ and } x_i \subseteq \text{supp}_i(p^n), \\ \frac{1}{|A_i| - |\text{supp}_i(p^n)|} & \text{if } a_i \in x_i \text{ and } x_i = A_i \setminus \text{supp}_i(p^n), \\ 0 & \text{otherwise.} \end{cases}$$

**Lemma 22.** *The pair $(\mathcal{S}^n, \sigma^n)$ is a canonical representation of $p^n$.*

*Proof.* We need to verify that $p^n$ is the measure induced by $(\mathcal{S}^n, \sigma^n)$ on $A \times \Theta$. We will use the following claim:

*Claim* 15. For all $B \in \mathcal{A}$ and $\theta \in \Theta$,

$$\sum_{a \in B} p^n(a,\theta) = \sum_z \left[\prod_i z_i(B_i)\right] \zeta^n(z|\theta)\pi(\theta) \tag{59}$$

*Proof of the claim.* By definition of $\zeta^n$, the right-hand side of (59) is equal to

$$(1-t^n)\sum_z \left[\prod_i z_i(B_i)\right] \zeta^{r,n}(z|\theta)\pi(\theta) + t^n \sum_z \left[\prod_i z_i(B_i)\right] \zeta^p(z|\theta)\pi(\theta). \tag{60}$$

We observe that, by definition of $\zeta^{r,n}$,

$$\sum_z \left[\prod_i z_i(B_i)\right] \zeta^{r,n}(z|\theta)\pi(\theta) = \sum_{a \in B} r^n(a,\theta). \tag{61}$$

Moreover, by definition of $\zeta^p$,

$$\sum_z \left[\prod_i z_i(B_i)\right] \zeta^p(z|\theta)\pi(\theta) = \left[\prod_i \bar{\alpha}_{\theta,i}(B_i)\right] \pi(\theta) = \sum_{a \in B} p(a,\theta). \tag{62}$$

Combining (60)-(62), we deduce that the right-hand side of (59) is equal to

$$(1-t^n)\sum_{a \in B} r^n(a,\theta) + t^n \sum_{a \in B} p(a,\theta) = \sum_{a \in B} \left[(1-t^n) r^n(a,\theta) + t^n p(a,\theta)\right].$$

Since $p^n = (1-t^n)r^n + t^n p$, we conclude that (59) holds. □

By Lemma 7, for all $a \in A$ and $\theta \in \Theta$, we can decompose $p^n(a,\theta)$ as

$$p^n(a,\theta) = \left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i))\right] \sum_{b \in \mathcal{A}(a)} p^n(b,\theta), \tag{63}$$

where $\mathcal{A}_i(a_i)$ is the cell of the partition that contains $a_i$, and $\mathcal{A}(a) = \prod_i \mathcal{A}_i(a_i)$. Putting together (59) and (63), we obtain that

$$\begin{aligned}
p^n(a,\theta) &= \left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i))\right]\sum_z\left[\prod_i z_i(\mathcal{A}_i(a_i))\right]\zeta^n(z|\theta)\pi(\theta)\\
&= \sum_z\left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i))z_i(\mathcal{A}_i(a_i))\right]\zeta^n(z|\theta)\pi(\theta)\\
&= \sum_{z,x}\left[\prod_i \sigma_i^n(a_i|x_i)z_i(x_i)\right]\zeta^n(z|\theta)\pi(\theta)\\
&= \sum_{z,x}\left[\prod_i \sigma_i^n(a_i|x_i)\xi_i(x_i|z,\theta)\right]\zeta^n(z|\theta)\pi(\theta)
\end{aligned}$$

where the first equality follows from (59) and (63), the second equality is just algebra, the third equality holds because $\sigma^n(a_i|x_i)>0$ if and only if $a_i\in x_i$, and the last equality by definition of $\xi_i$. We conclude that $p^n$ is the measure induced by $(\mathcal{S}^n,\sigma^n)$ on $A\times\Theta$. □

Since $p^n$ is a separated BCE and $(\mathcal{S}^n,\sigma^n)$ is a canonical representation of $p^n$, it follows from Lemma 8 that for every player $i$, there exists a monotone $C_i^n:\Delta(X_i)^{Z\times\Theta}\to\mathbb{R}_+$ such that $(\xi,\sigma^n)$ is an equilibrium of $(\mathcal{G},\mathcal{T}^n)$ with $\mathcal{T}^n=(Z,\zeta^n,(X_i,\Delta(X_i)^{Z\times\Theta},C_i^n)_{i\in I})$. Moreover, we can choose $C_i^n$ such that

$$\max_{\xi_i'\in\Delta(X_i)^{Z\times\Theta}} C_i^n(\xi_i')\le\frac{1}{n}+\hat{v}_i\left(\mathcal{S}^n,\sigma^n\right)-\left[\frac{n-1}{n}\bar{v}_i\left(p^n\right)+\frac{1}{n}\underline{v}_i\left(p^n\right)\right].$$

As $n\to\infty$, the upper bound on $i$'s costs converges to

$$\sum_\theta\pi(\theta)\left[\max_{a_i\in A_i}\sum_{B_{-i}\in\mathcal{A}_{-i}}\sum_{a_{-i}\in B_{-i}}u_i(a_i,a_{-i},\theta)\prod_{j\neq i}\frac{p(a_j)\bar{\alpha}_{\theta,j}(B_j)}{\sum_{b_j\in B_j}p(b_j)}\right]-\bar{v}_i\left(p\right), \tag{64}$$

where we adopt the convention that $\frac{0}{0}=0$. Hence, if we show that (64) is equal to zero, we can conclude that $p$ is a vanishing cost equilibrium, as desired.

To prove that (64) is equal to zero, we need the following intermediate result:

**Lemma 23.** *For all $\theta\in\theta$, $i\in I$, $B_i\in\mathcal{A}_i$, and $a_i\in B_i$,*

$$\frac{p(a_i)\bar{\alpha}_{\theta,i}(B_i)}{\sum_{b_i\in B_i}p(b_i)}=\alpha_{\theta,i}(a_i).$$

*Proof.* We divide the proof in three cases. Case (i): Assume $\sum_{b_i\in B_i}p(b_i)=0$. Then $\bar{\alpha}_{\theta,i}(B_i)=0$ (because $\sum_{b_i\in B_i}p(b_i)=\sum_\theta\pi(\theta)\bar{\alpha}_{\theta,i}(B_i)$) and $\alpha_{\theta,i}(a_i)=0$ (because $a_i\in B_i$).

Thus,

$$\frac{p(a_i)\bar{\alpha}_{\theta,i}(B_i)}{\sum_{b_i\in B_i} p(b_i)} = \frac{0}{0} = 0 = \alpha_{\theta,i}(a_i).$$

Case (ii): Assume $\sum_{b_i\in B_i} p(b_i) > 0$ and $p(a_i) = 0$. Then $\alpha_{\theta,i}(a_i) = 0$ (because $p(a_i) = \sum_\theta \pi(\theta)\alpha_{\theta,i}(a_i)$). Thus,

$$\frac{p(a_i)\bar{\alpha}_{\theta,i}(B_i)}{\sum_{b_i\in B_i} p(b_i)} = 0 = \alpha_{\theta,i}(a_i).$$

Case (iii): Assume $\sum_{b_i\in B_i} p(b_i) > 0$ and $p(a_i) > 0$. Take any $b_i \in B_i$ such that $p(b_i) > 0$. Since $p^n \to p$, $p^n(a_i) > 0$ and $p^n(b_i) > 0$ for all $n$ sufficiently large. Thus, by definition of $\mathcal{A}_i$, $p^n_{a_i} = p^n_{b_i}$ for all sufficiently large $n$ (in fact, for all $n$). We deduce that $p_{a_i} = p_{b_i}$: for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\frac{p(a_i, a_{-i}, \theta)}{p(a_i)} = \frac{p(b_i, a_{-i}, \theta)}{p(b_i)} \tag{65}$$

Since $p$ is a complete-information Nash equilibrium, players' actions are conditionally independent given the payoff state. Thus, (65) becomes

$$\frac{\alpha_{\theta,i}(a_i)}{p(a_i)} = \frac{\alpha_{\theta,i}(b_i)}{p(b_i)},$$

which is the same as

$$\frac{\alpha_{\theta,i}(a_i)}{p(a_i)} p(b_i) = \alpha_{\theta,i}(b_i).$$

Clearly, the equality also holds for $b_i \in B_i$ such that $p(b_i) = 0$. Thus, summing over all $b_i \in B_i$, we obtain that

$$\frac{\alpha_{\theta,i}(a_i)}{p(a_i)} \left( \sum_{b_i\in B_i} p(b_i) \right) = \overline{\alpha}_{\theta,i}(B_i).$$

Rearranging the equality, we conclude that

$$\frac{p(a_i)\bar{\alpha}_{\theta,i}(B_i)}{\sum_{b_i\in B_i} p(b_i)} = \alpha_{\theta,i}(a_i).$$

□

It follows from Lemma 23 that (64) is equal to

$$\sum_\theta \pi(\theta) \left[ \max_{a_i} \sum_{a_{-i}} u_i(a_i, a_{-i}, \theta) \prod_{j\neq i} \alpha_{\theta,j}(a_j) \right] - \bar{v}_i(p),$$

which, in turn, is equal to zero since $p$ is complete-information Nash equilibrium:

$$\begin{aligned}\bar{v}_i\,(p) &= \sum_{\theta}\pi(\theta)\left[\sum_{a}u_i(a,\theta)\alpha_{\theta,i}(a_i)\prod_{j\neq i}\alpha_{\theta,j}(a_j)\right]\\ &= \sum_{\theta}\pi(\theta)\left[\max_{a_i}\sum_{a_{-i}}u_i(a_i,a_{-i},\theta)\prod_{j\neq i}\alpha_{\theta,j}(a_j)\right],\end{aligned}$$

where the second equality holds because, given $\theta$, $\alpha_{\theta,i}$ is a best response to $(\alpha_{\theta,j})_{j\neq i}$. We conclude that $p$ is a vanishing cost equilibrium.

## J.2. Full Characterization

In this section we provide a characterization of *all* vanishing cost equilibria. When information costs are negligible, both the payoff state and the correlation state become freely learnable. Thus, every vanishing cost equilibrium must be (i) a convex combination of complete-information Nash equilibria, and (ii) the limit of a sequence of separated BCEs (from Theorem 1). It turns out that vanishing cost equilibria satisfy not only (i) and (ii), but also additional "measurability" conditions that we present next.

Fix a base game $\mathcal{G}$. For every player $i$, let $\mathcal{A}_i$ be a partition of $A_i$. We denote by $B_i$ a generic element of $\mathcal{A}_i$. Define

$$\begin{aligned}\mathcal{A}_{-i} &= \left\{\prod_{j\neq i}B_j : B_j\in\mathcal{A}_j \text{ for all } j\neq i\right\},\\ \mathcal{A} &= \{B_i\times B_{-i} : B_i\in\mathcal{A}_i \text{ and } B_{-i}\in\mathcal{A}_{-i}\}.\end{aligned}$$

We refer to $\mathcal{A}_{-i}$ and $\mathcal{A}$ as **product partitions** of $A_{-i}$ and $A$, respectively.

Given a product partition $\mathcal{A}$ of $A$, we say an outcome $p$ is **$\mathcal{A}$-measurable** if for every $i\in I$, $B_i\in\mathcal{A}_i$, and $a_i,b_i\in B_i\cap\operatorname{supp}_i(p)$,

$$p_{a_i}=p_{b_i}.$$

For an intuition, take the perspective of a mediator who wants to implement an outcome $p$. If $p$ is $\mathcal{A}$-measurable, then the mediator can implement it as follows: draw a payoff state $\theta$ and an element $B$ of $\mathcal{A}$ with probability $\sum_{a\in B}p(\theta,a)$; communicate to each player $i$ the realized $B_i$, and let them privately draw an action $a_i\in B_i$ with probability $p(a_i)/\sum_{b_i\in B_i}p(b_i)$.

As in Section A.2, let $\mathcal{A}^p$ be the product partition of $A$ such that, for every player $i$, actions $a_i$ and $b_i$ are in the same cell if and only if either $a_i,b_i\in\operatorname{supp}_i(p)$ and $p_{a_i}=p_{b_i}$, or $a_i,b_i\notin\operatorname{supp}_i(p)$. Note that $p$ is measurable with respect to $\mathcal{A}_p$; except for zero-probability actions, $\mathcal{A}_p$ is the coarsest product partition for which $p$ is measurable.

Given an outcome $p$ and a product partition $\mathcal{A}$ of $A$, we say an outcome $q$ is **$(\mathcal{A}, p)$-decomposable** if $q$ is $\mathcal{A}$-measurable, and for every $i \in I$, $B_i \in \mathcal{A}_i$, and $a_i, b_i \in B_i$,

$$q(a_i)p(b_i) = p(a_i)q(b_i). \tag{66}$$

For an intuition, take the perspective of a mediator who wants to implement an outcome $q$. If $q$ is $(\mathcal{A}, p)$-decomposable, then the mediator can implement it as follows: draw a payoff state $\theta$ and an element $B$ of $\mathcal{A}$ with probability $\sum_{a \in B} q(\theta, a)$; communicate to each player $i$ the realized $B_i$, and let them privately draw an action $a_i \in B_i$ with probability $p(a_i)/\sum_{b_i \in B_i} p(b_i) = q(a_i)/\sum_{b_i \in B_i} q(b_i)$.

Next we use $\mathcal{A}$-measurability and $(\mathcal{A}, p)$-decomposability to characterize vanishing cost equilibrium:

**Theorem 7.** *An outcome $p$ is a vanishing cost equilibrium if and only if there exists a product partition $\mathcal{A}$ of $A$ such that*

*(i) $p$ is a convex combination of finitely many $(\mathcal{A}, p)$-decomposable complete-information Nash equilibria, and*

*(ii) $p$ is the limit of a sequence $(p^n)_{n=1}^{\infty}$ of separated BCEs, with $\mathcal{A}^{p^n} = \mathcal{A}$ for all $n$.*

Theorem 7 generalizes Theorem 5. To see the relationship between the two results, let $p$ be a complete-information Nash equilibrium that is the limit of a sequence $(p^n)_{n=1}^{\infty}$ of separated BCEs, as in Theorem 5. Passing to a subsequence, we can assume that $\mathcal{A}^{p^n} = \mathcal{A}^{p^1}$ for all $n$. Since $p^n \to p$, $p$ is $\mathcal{A}^1$-measurable. Because (66) trivially holds for $q = p$, $p$ is $(\mathcal{A}^1, p)$-decomposable. Thus, $p$ satisfies the hypotheses of Theorem 7 with $\mathcal{A} := \mathcal{A}^1$. In particular, one can obtain Theorem 5 as a corollary of Theorem 7.

We divide the proof of the theorem 7 in two parts.

### J.3. Proof of the "if" side of Theorem 7

Let $\mathcal{A}$ be a product partition of $A$, and let $\{q^1, \ldots, q^L\}$ be a finite set of complete-information Nash equilibria. For every $l = 1, \ldots, L$, we denote by $\alpha^l_{\theta,i} \in \Delta(A_i)$ the conditional distribution of $i$'s action given $\theta$. Let $p$ be an outcome in the convex hull of $\{q^1, \ldots, q^L\}$:

$$p = \sum_{l=1}^{L} s^l q^l.$$

Without loss of generality, suppose that $s^l > 0$ for all $l = 1, \ldots, L$.

Assume that for all $l = 1, \dots, L$, $q^l$ is $(\mathcal{A}, p)$-decomposable. Furthermore, assume that $p$ is the limit of a sequence $(p^n)_{n=1}^{\infty}$ of separated BCEs such that $\mathcal{A}^{p^n} = \mathcal{A}$ for all $n$. We want to prove that $p$ is a vanishing cost equilibrium.

We begin with constructing one canonical representation $(\mathcal{S}^n, \sigma^n)$ for each $p^n$ (see Section A.2 for the general definition of canonical representation). The information structure $\mathcal{S}^n = (Z, \zeta^n, (X_i, \xi_i)_{i\in I})$ is specified as follows:

- To construct $Z \subseteq \prod_{i\in I} \Delta(\mathcal{A}_i)$, let $\delta_{B_i} \in \Delta(\mathcal{A}_i)$ be the Dirac measure concentrated on $B_i \in \mathcal{A}_i$. Moreover, let $\bar{\alpha}^l_{\theta,i}$ be the measure on $\mathcal{A}_i$ induced by $\alpha^l_{\theta,i}$: for $B_i \in A_i$,

$$\bar{\alpha}^l_{\theta,i}(B_i) = \sum_{a_i \in B_i} \alpha^l_{\theta,i}(a_i).$$

  We set $Z = \prod_i Z_i$ where for every player $i$,

$$Z_i = \{\delta_{B_i} : B_i \in A_i\} \bigcup \left\{\bar{\alpha}^l_{\theta,i} : \theta \in \Theta\right\}.$$

- To construct $\zeta^n : \Theta \to \Delta(Z)$, first we take a sequence $(t^n)_{n=1}^{\infty}$ in $(0, 1)$ such that
  - $t^n \to 1$, and
  - $t^n p < p^n$ for all $n$.

  Such a sequence exists because $p^n \to p$. For each $n$, we define the outcome $r^n \in \Delta(A \times \Theta)$ by

$$r^n(a, \theta) = \frac{p^n(a, \theta) - t^n p(a, \theta)}{1 - t^n},$$

  and the Markov kernel $\zeta^{r,n} : \Theta \to \Delta(Z)$ by

$$\zeta^{r,n}(z|\theta) = \begin{cases} \sum_{a\in B} \frac{r^n(a,\theta)}{\pi(\theta)} & \text{if } B \in \mathcal{A} \text{ and } z = (\delta_{B_i})_{i\in I}, \\ 0 & \text{otherwise.} \end{cases}$$

  We also denote by $\zeta^p : \Theta \to \Delta(Z)$ the Markov kernel given by

$$\zeta^p(z|\theta) = \begin{cases} s^l & \text{if } l = 1, \dots, L \text{ and } z = \left(\bar{\alpha}^l_{\theta,i}\right)_{i\in I}, \\ 0 & \text{otherwise.} \end{cases}$$

  Finally, we construct $\zeta^n : \Theta \to \Delta(Z)$ as follows:

$$\zeta^n(z|\theta) = (1 - t^n)\zeta^{r,n}(z|\theta) + t^n \zeta^p(z|\theta).$$

- For every player $i$, we take $X_i$ sufficiently rich so that $\mathcal{A}_i \subseteq X_i$.

- For every $i \in I$, $x_i \in X_i$, $z \in Z$, and $\theta \in \Theta$, we define the experiment $\xi_i$ by

$$\xi_i(x_i|z,\theta) = \begin{cases} z_i(x_i) & \text{if } x_i \in \mathcal{A}_i, \\ 0 & \text{otherwise.} \end{cases}$$

The profile of action plans $\sigma^n = (\sigma_i^n)_{i \in I}$ is given by, for all $i \in I$, $a_i \in A_i$, and $x_i \in \mathcal{A}_i$,

$$\sigma_i^n(a_i|x_i) = \begin{cases} \frac{p^n(a_i)}{\sum_{b_i \in x_i} p^n(b_i)} & \text{if } a_i \in x_i \text{ and } x_i \subseteq \operatorname{supp}_i(p^n), \\ \frac{1}{|A_i| - |\operatorname{supp}_i(p^n)|} & \text{if } a_i \in x_i \text{ and } x_i = A_i \setminus \operatorname{supp}_i(p^n), \\ 0 & \text{otherwise.} \end{cases}$$

**Lemma 24.** *The pair $(\mathcal{S}^n, \sigma^n)$ is a canonical representation of $p^n$.*

*Proof.* We need to verify that $p^n$ is the measure induced by $(\mathcal{S}^n, \sigma^n)$ on $A \times \Theta$. We will use the following claim:

*Claim* 16. For all $B \in \mathcal{A}$ and $\theta \in \Theta$,

$$\sum_{a \in B} p^n(a,\theta) = \sum_z \left[\prod_i z_i(B_i)\right] \zeta^n(z|\theta)\pi(\theta) \tag{67}$$

*Proof of the claim.* By definition of $\zeta^n$, the right-hand side of (67) is equal to

$$(1-t^n) \sum_z \left[\prod_i z_i(B_i)\right] \zeta^{r,n}(z|\theta)\pi(\theta) + t^n \sum_z \left[\prod_i z_i(B_i)\right] \zeta^p(z|\theta)\pi(\theta). \tag{68}$$

We observe that, by definition of $\zeta^{r,n}$,

$$\sum_z \left[\prod_i z_i(B_i)\right] \zeta^{r,n}(z|\theta)\pi(\theta) = \sum_{a \in B} r^n(a,\theta). \tag{69}$$

Moreover, by definition of $\zeta^p$,

$$\begin{aligned} \sum_z \left[\prod_i z_i(B_i)\right] \zeta^p(z|\theta)\pi(\theta) &= \sum_l s^l \left[\prod_i \bar{\alpha}^l_{\theta,i}(B_i)\right] \pi(\theta) \\ &= \sum_l s^l \left[\sum_{a \in B} \frac{q^l(a,\theta)}{\pi(\theta)}\right] \pi(\theta) = \sum_{a \in B} p(a,\theta). \end{aligned} \tag{70}$$

Combining (68)-(70), we deduce that the right-hand side of (67) is equal to

$$(1-t^n) \sum_{a \in B} r^n(a,\theta) + t^n \sum_{a \in B} p(a,\theta) = \sum_{a \in B} \left[(1-t^n) r^n(a,\theta) + t^n p(a,\theta)\right].$$

Since $p^n = (1-t^n)r^n + t^n p$, we conclude that (67) holds. $\square$

By Lemma 7, for all $a \in A$ and $\theta \in \Theta$, we can decompose $p^n(a,\theta)$ as

$$p^n(a,\theta) = \left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i))\right] \sum_{b\in\mathcal{A}(a)} p^n(b,\theta), \tag{71}$$

where $\mathcal{A}_i(a_i)$ is the cell of the partition that contains $a_i$, and $\mathcal{A}(a) = \prod_i \mathcal{A}_i(a_i)$. Putting together (67) and (71), we obtain that

$$\begin{aligned}
p^n(a,\theta) &= \left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i))\right] \sum_z \left[\prod_i z_i(\mathcal{A}_i(a_i))\right] \zeta^n(z|\theta)\pi(\theta) \\
&= \sum_z \left[\prod_i \sigma_i^n(a_i|\mathcal{A}_i(a_i)) z_i(\mathcal{A}_i(a_i))\right] \zeta^n(z|\theta)\pi(\theta) \\
&= \sum_{z,x} \left[\prod_i \sigma_i^n(a_i|x_i) z_i(x_i)\right] \zeta^n(z|\theta)\pi(\theta) \\
&= \sum_{z,x} \left[\prod_i \sigma_i^n(a_i|x_i) \xi_i(x_i|z,\theta)\right] \zeta^n(z|\theta)\pi(\theta)
\end{aligned}$$

where the first equality follows from (67) and (71), the second equality is just algebra, the third equality holds because $\sigma^n(a_i|x_i) > 0$ if and only if $a_i \in x_i$, and the last equality by definition of $\xi_i$. We conclude that $p^n$ is the measure induced by $(\mathcal{S}^n, \sigma^n)$ on $A \times \Theta$. $\square$

Since $p^n$ is a separated BCE and $(\mathcal{S}^n, \sigma^n)$ is a canonical representation of $p^n$, it follows from Lemma 8 that for every player $i$, there exists a monotone $C_i^n : \Delta(X_i)^{Z\times\Theta} \to \mathbb{R}_+$ such that $(\xi, \sigma^n)$ is an equilibrium of $(\mathcal{G}, \mathcal{T}^n)$ with $\mathcal{T}^n = (Z, \zeta^n, (X_i, \Delta(X_i)^{Z\times\Theta}, C_i^n)_{i\in I})$. Moreover, we can choose $C_i^n$ such that

$$\max_{\xi_i' \in \Delta(X_i)^{Z\times\Theta}} C_i^n(\xi_i') \leq \frac{1}{n} + \hat{v}_i(\mathcal{S}^n, \sigma^n) - \left[\frac{n-1}{n}\bar{v}_i(p^n) + \frac{1}{n}\underline{v}_i(p^n)\right].$$

As $n \to \infty$, the upper bound on $i$'s costs converges to

$$\sum_\theta \pi(\theta) \sum_l s^l \left[\max_{a_i\in A_i} \sum_{B_{-i}\in\mathcal{A}_{-i}} \sum_{a_{-i}\in B_{-i}} u_i(a_i, a_{-i}, \theta) \prod_{j\neq i} \frac{p(a_j)\bar{\alpha}^l_{\theta,j}(B_j)}{\sum_{b_j\in B_j} p(b_j)}\right] - \bar{v}_i(p), \tag{72}$$

where we adopt the convention that $\frac{0}{0} = 0$. Hence, if we show that (72) is equal to zero, we can conclude that $p$ is a vanishing cost equilibrium, as desired.

To prove that (72) is equal to zero, we need the following intermediate results:

**Lemma 25.** *For all $\theta \in \theta$, $l \in \{1, \ldots, L\}$, $i \in I$, $B_i \in \mathcal{A}_i$, and $a_i \in B_i$,*

$$\frac{p(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} p(b_i)} = \frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)}$$

*where on both sides of the equation we adopt the convention that $\frac{0}{0} = 0$.*

*Proof.* If $\bar{\alpha}^l_{\theta,i}(B_i) = 0$, then (trivially)

$$\frac{p(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} p(b_i)} = 0 = \frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)}$$

Suppose now that $\bar{\alpha}^l_{\theta,i}(B_i) > 0$. Then, we have

$$\sum_{b_i \in B_i} q^l(b_i) = \sum_{\theta} \pi(\theta)\bar{\alpha}^l_{\theta,i}(B_i) > 0,$$

which in turn implies that $\sum_{b_i \in B_i} p(b_i) > 0$ (since $q^l$ is absolutely continuous with respect to $p$). Since $q^l$ is $(\mathcal{A}, p)$-decomposable, It follows from (66) that

$$\frac{p(a_i)}{\sum_{b_i \in B_i} p(b_i)} = \frac{q^l(a_i)}{\sum_{b_i \in B_i} q^l(b_i)}.$$

Multiplying both sides of the equation by $\bar{\alpha}^l_{\theta,i}(B_i)$, we obtain the desired result. □

**Lemma 26.** *For all $\theta \in \theta$, $l \in \{1, \ldots, L\}$, $i \in I$, $B_i \in \mathcal{A}_i$, and $a_i \in B_i$,*

$$\frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)} = \alpha^l_{\theta,i}(a_i).$$

*Proof.* We divide the proof in three cases. Case (i): Assume $\sum_{b_i \in B_i} q^l(b_i) = 0$. Then $\bar{\alpha}^l_{\theta,i}(B_i) = 0$ and $\alpha^l_{\theta,i}(a_i) = 0$. Thus,

$$\frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)} = \frac{0}{0} = 0 = \alpha^l_{\theta,i}(a_i).$$

Case (ii): Assume $\sum_{b_i \in B_i} q^l(b_i) > 0$ and $q^l(a_i) = 0$. Then $\alpha^l_{\theta,i}(a_i) = 0$. Thus,

$$\frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)} = 0 = \alpha^l_{\theta,i}(a_i).$$

Case (iii): Assume $\sum_{b_i \in B_i} q^l(b_i) > 0$ and $q^l(a_i) > 0$. Take any $b_i \in B_i$ such that

$q^l(b_i) > 0$. Since $q^l$ is $\mathcal{A}$-measurable, $q^l_{a_i} = q^l_{b_i}$: for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\frac{q^l(a_i, a_{-i}, \theta)}{q^l(a_i)} = \frac{q^l(b_i, a_{-i}, \theta)}{q^l(b_i)} \tag{73}$$

Since $q^l$ is a complete-information Nash equilibrium, players' actions are conditionally independent given the payoff state. Thus, (73) becomes

$$\frac{\alpha^l_{\theta,i}(a_i)}{q^l(a_i)} = \frac{\alpha^l_{\theta,i}(b_i)}{q^l(b_i)},$$

which is the same as

$$\frac{\alpha^l_{\theta,i}(a_i)}{q^l(a_i)} q^l(b_i) = \alpha^l_{\theta,i}(b_i).$$

Clearly, the equality also holds for $b_i \in B_i$ such that $q^l(b_i) = 0$. Thus, summing over all $b_i \in B_i$, we obtain that

$$\frac{\alpha^l_{\theta,i}(a_i)}{q^l(a_i)} \left( \sum_{b_i \in B_i} q^l(b_i) \right) = \overline{\alpha}^l_{\theta,i}(B_i).$$

Rearranging the equality, we conclude that

$$\frac{q^l(a_i)\bar{\alpha}^l_{\theta,i}(B_i)}{\sum_{b_i \in B_i} q^l(b_i)} = \alpha^l_{\theta,i}(a_i).$$

$\square$

It follows from Lemmas 25 and 26 that (72) is equal to

$$\sum_\theta \pi(\theta) \sum_l s^l \left[ \max_{a_i} \sum_{a_{-i}} u_i(a_i, a_{-i}, \theta) \prod_{j \neq i} \alpha^l_{\theta,j}(a_j) \right] - \bar{v}_i(p),$$

which, in turn, is equal to zero since each $q^l$ is complete-information Nash equilibrium:

$$\begin{aligned} \bar{v}_i(p) &= \sum_\theta \pi(\theta) \sum_l s^l \left[ \sum_a u_i(a, \theta) \alpha^l_{\theta,i}(a_i) \prod_{j \neq i} \alpha^l_{\theta,j}(a_j) \right] \\ &= \sum_\theta \pi(\theta) \sum_l s^l \left[ \max_{a_i} \sum_{a_{-i}} u_i(a_i, a_{-i}, \theta) \prod_{j \neq i} \alpha^l_{\theta,j}(a_j) \right], \end{aligned}$$

where the second equality holds because, given $\theta$, $\alpha^l_{\theta,i}$ is a best response to $(\alpha^l_{\theta,j})_{j \neq i}$. We conclude that $p$ is a vanishing cost equilibrium.

## J.4. Proof of the "only if" side of Theorem 7

Let $p$ be a vanishing cost equilibrium: we want to show that there exists a product partition $\mathcal{A}$ of $A$ such that Theorem 7-(i) and Theorem 7-(ii) hold.

By the definition of vanishing cost equilibrium, for every $n \in \{1, 2, \ldots\}$, we can find an unconstrained rational-inattention technology $\mathcal{T}^n = (Z^n, \zeta^n, (X_i^n, \mathcal{E}_i^n, C_i^n)_{i \in I})$ and an equilibrium $(\xi^n, \sigma^n)$ of $(\mathcal{G}, \mathcal{T}^n)$ such that, denoting by $p^n$ the outcome of $(\xi^n, \sigma^n)$,

$$\max_{\xi_i' \in \Delta(X_i^n)^{Z^n \times \Theta}} C_i^n(\xi_i') \leq \frac{1}{n} \quad \text{and} \quad |p(a, \theta) - p^n(a, \theta)| \leq \frac{1}{n}$$

for all $i \in I$, $a \in A$, and $\theta \in \Theta$. Possibly passing to a subsequence, we can assume that

$$\mathcal{A}^{p^n} = \mathcal{A}^{p^{n+1}}$$

for all $n$. Let $\mathcal{A} := \mathcal{A}^{p^n}$ be this fixed product partition. Since each $p^n$ is a sBCE by Theorem 1, we obtain that Theorem 7-(ii) holds.

In the rest of the proof, we show that Theorem 7-(i) is satisfied. We begin with introducing some notation. Let $\mathcal{M}_i$ be the set of player $i$'s mixed actions: $\mathcal{M}_i = \Delta(A_i)$; we define the Cartesian products $\mathcal{M}_{-i} = \prod_{j \neq i} \mathcal{M}_j$ and $\mathcal{M} = \mathcal{M}_{-i} \times \mathcal{M}_i$. For $n \in \mathbb{N}$, $z \in Z^n$, and $\theta \in \Theta$, we denote by $\alpha^{n,z,\theta} \in \mathcal{M}$ the induced mixed-action profile: for all $i$ and $a_i$,

$$\alpha_i^{n,z,\theta}(a_i) = \sum_{x_i} \sigma_i^n(a_i | x_i) \xi_i^n(x_i | z, \theta).$$

In addition, we define the transition kernel $\Theta \ni \theta \mapsto \chi_\theta^n \in \Delta(\mathcal{M})$ by

$$\chi_\theta^n(M) = \zeta^n\left(\left\{z \in Z^n : \alpha^{n,z,\theta} \in M\right\} | \theta\right)$$

for all Borel sets $M \subseteq \mathcal{M}$. Direct computation shows that

$$p^n(a, \theta) = \pi(\theta) \int_{\mathcal{M}} \prod_i \alpha_i(a_i) \, \mathrm{d}\chi_\theta^n(\alpha) \tag{74}$$

for all $a$ and $\theta$. By weak* compactness of $\Delta(\mathcal{M})$ (Aliprantis and Border, 2006, Theorem 15.11), it is without loss (potentially by passing to a subsequence) to assume that, for every $\theta$, $\chi_\theta^n$ converges in the weak* topology to some limit $\chi_\theta$. Since $p^n \to p$, we obtain that

$$p(a, \theta) = \pi(\theta) \int_{\mathcal{M}} \prod_i \alpha_i(a_i) \, \mathrm{d}\chi_\theta(\alpha) \tag{75}$$

for all $a$ and $\theta$.

The next lemma relates the support of $\chi^n_\theta$ (which is finite) to $p^n$:

**Lemma 27.** *For all $n \in \mathbb{N}$, $\theta \in \Theta$, $\alpha \in \operatorname{supp}(\chi^n_\theta)$, $i \in I$, $B_i \in \mathcal{A}_i$, and $a_i, b_i \in B_i$,*

$$\alpha_i(a_i)p^n(b_i) = \alpha_i(b_i)p^n(a_i). \tag{76}$$

*Proof.* Since $\alpha \in \operatorname{supp}(\chi^n_\theta)$, $\alpha$ is absolutely continuous with respect to $p^n$—see (74). Thus, (76) trivially holds if $p^n(a_i) = 0$ or $p^n(b_i) = 0$.

Suppose now that $p^n(a_i) > 0$ and $p^n(b_i) > 0$. Paralleling the notation in Section A.2, we denote by $\nu^n \in \Delta(A \times X^n \times Z^n \times \Theta)$ the probability measure over actions, signals, and states induced by the information structure $(Z^n, \zeta^n, (X^n_i, \xi^n_i)_{i\in I})$ and the profile of action plans $\sigma^n$. Let $X^n_{a_i}$ and $X^n_{b_i}$ be the set of positive-probability signals that make player $i$ take actions $a_i$ and $b_i$:

$$\begin{aligned}
X^n_{a_i} &= \{x_i : \nu^n(x_i) > 0 \text{ and } \sigma^n_i(a_i|x_i) > 0\},\\
X^n_{b_i} &= \{x_i : \nu^n(x_i) > 0 \text{ and } \sigma^n_i(b_i|x_i) > 0\}.
\end{aligned}$$

By Lemma 9, for all $x_i \in X^n_{a_i}$ and $x'_i \in X^n_{b_i}$,

$$BR(p_{a_i}) \subseteq BR(\nu^n_{x_i}) \quad \text{and} \quad BR(p_{b_i}) \subseteq BR(\nu^n_{x'_i}),$$

where $\nu^n_{x_i} \in \Delta(A_{-i} \times X^n_{-i} \times Z^n \times \Theta)$ and $\nu^n_{x'_i} \in \Delta(A_{-i} \times X^n_{-i} \times Z^n \times \Theta)$ are the posterior beliefs generated by $x_i$ and $x'_i$. Since $a_i, b_i \in B_i$, we have $p_{a_i} = p_{b_i}$, which in turn implies $BR(p_{a_i}) = BR(p_{b_i})$. Thus,

$$BR(\nu^n_{x_i}) \cap BR(\nu^n_{x'_i}) \neq \varnothing.$$

It follows from Lemma 5-(ii) that $\nu^n_{x_i} = \nu^n_{x'_i}$.

Since $\alpha_i \in \operatorname{supp}(\chi^n_\theta)$, there must be $z \in Z^n$ and $\theta \in \Theta$ such that $\zeta^n(z|\theta) > 0$ and $\alpha_i = \alpha^{n,z,\theta}_i$. Then, given a fixed $x'_i \in X^n_{b_i}$,

$$\begin{aligned}
\frac{\alpha_i(a_i)}{p^n(a_i)} &= \sum_{x_i \in X^n_i} \frac{\sigma^n_i(a_i|x_i)\xi^n_i(x_i|z,\theta)}{p^n(a_i)} = \sum_{x_i \in X^n_{a_i}} \frac{\sigma^n_i(a_i|x_i)\xi^n_i(x_i|z,\theta)}{p^n(a_i)}\\
&= \sum_{x_i \in X^n_{a_i}} \frac{\sigma^n_i(a_i|x_i)\xi^n_i(x_i|z,\theta)\nu^n(x_i)}{p^n(a_i)\nu^n(x_i)} = \sum_{x_i \in X^n_{a_i}} \frac{\sigma^n_i(a_i|x_i)\xi^n_i(x'_i|z,\theta)\nu^n(x_i)}{p^n(a_i)\nu^n(x'_i)} = \frac{\xi^n_i(x'_i|z,\theta)}{\nu^n(x'_i)},
\end{aligned}$$

where the fourth inequality follows from $\nu^n_{x_i} = \nu^n_{x'_i}$ for all $x_i \in X^n_{a_i}$. A similar argument delivers that, given a fixed $x_i \in X^n_{a_i}$,

$$\frac{\alpha_i(b_i)}{p^n(b_i)} = \frac{\xi^n_i(x_i|z,\theta)}{\nu^n(x_i)}.$$

We conclude that
$$\frac{\alpha_i(a_i)}{p^n(a_i)} = \frac{\xi_i^n(x_i'|z,\theta)}{\nu^n(x_i')} = \frac{\xi_i^n(x_i|z,\theta)}{\nu^n(x_i)} = \frac{\alpha_i(b_i)}{p^n(b_i)},$$
where the third equality again follows from $\nu^n_{x_i'} = \nu^n_{x_i}$. □

The next lemma relates the support of $\chi_\theta$ (which may not be finite) to $p$:

**Lemma 28.** *Let $\theta \in \Theta$ and $\alpha \in \operatorname{supp}(\chi_\theta)$. Then, for all $i \in I$, $B_i \in \mathcal{A}_i$, and $a_i, b_i \in B_i$,*
$$\alpha_i(a_i)p(b_i) = \alpha_i(b_i)p(a_i). \tag{77}$$

*Proof.* Since the support correspondence is lower hemicontinuous (Aliprantis and Border, 2006, Theorem 17.14), there exists a sequence $(\alpha_i^n)_{n=1}^\infty$ such that $\alpha_i^n \to \alpha_i$, and $\alpha_i^n \in \operatorname{supp}(\chi_\theta^n)$ for all $n$. By Lemma 27,
$$\alpha_i^n(a_i)p^n(b_i) = \alpha_i^n(b_i)p^n(a_i)$$
for all $n$. Taking the limit as $n \to \infty$, we obtain (77). □

For every state $\theta$, let $\mathrm{NE}_\theta$ be the set of Nash equilibria of the complete-information game corresponding to $\theta$:
$$\mathrm{NE}_\theta = \bigcap_{i \in I} \left\{ \alpha \in \mathcal{M} : \alpha_i \in \arg\max_{\beta_i \in \mathcal{M}_i} \sum_a u_i(a,\theta)\beta_i(a_i) \prod_{j \neq i} \alpha_j(a_j) \right\}.$$
Note that the set $\mathrm{NE}_\theta$ is closed.

**Lemma 29.** *For all $\theta \in \Theta$, $\operatorname{supp}(\chi_\theta) \subseteq NE_\theta$.*

*Proof.* For $\alpha \in \mathcal{M}$, let $u_i(\alpha,\theta)$ be $i$'s expected utility in state $\theta$:
$$u_i(\alpha,\theta) = \sum_{a \in A} u_i(a,\theta) \prod_{j \in I} \alpha_j(a_j).$$
For $\alpha_{-i} \in \mathcal{M}_{-i}$, let $u_i^*(\alpha_{-i},\theta)$ be $i$'s expected utility by best responding to $\alpha_{-i}$ in state $\theta$:
$$u_i^*(\alpha_{-i},\theta) = \max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}, \theta) \prod_{j \neq i} \alpha_j(a_j).$$
Note that $\alpha \in \mathrm{NE}_\theta$ if and only if $u_i^*(\alpha_{-i},\theta) = u_i(\alpha,\theta)$ for all $i \in I$.

Consider now the information acquisition game $(\mathcal{G}, \mathcal{T}^n)$, and take player $i$'s perspective. Let $\bar{\xi}_i^n$ be a fully-revealing experiment that tells the exact value of $(z^n, \theta)$. Since information acquisition is unconstrained, the experiment $\bar{\xi}_i^n$ is feasible. Since $(\xi^n, \sigma^n)$ is an equilibrium, player $i$'s payoff from playing her equilibrium strategy should be larger than the payoff she obtains from deviating to $\bar{\xi}_i^n$ and taking the optimal action given each state. Tedious but straightforward computation shows this inequality is equivalent to

$$\sum_\theta \pi(\theta) \int_{\mathcal{M}} [u_i^*(\alpha_{-i}, \theta) - u_i(\alpha, \theta)] \, \mathrm{d}\chi_\theta^n(\alpha) \leq C_i^n(\bar{\xi}_i^n) - C_i^n(\xi_i^n).$$

Since information costs are bounded by $1/n$, we have $C_i^n(\bar{\xi}_i^n) - C_i^n(\xi_i^n) \leq 1/n$, and so we obtain that

$$\begin{aligned}
&\sum_\theta \pi(\theta) \int_{\mathcal{M}} [u_i^*(\alpha_{-i}, \theta) - u_i(\alpha, \theta)] \, \mathrm{d}\chi_\theta(\alpha) \\
&= \lim_{n \to \infty} \sum_\theta \pi(\theta) \int_{\mathcal{M}} [u_i^*(\alpha_{-i}, \theta) - u_i(\alpha, \theta)] \, \mathrm{d}\chi_\theta^n(\alpha) \leq 0.
\end{aligned}$$

Because $u_i^*(\alpha_{-i}, \theta) \geq u_i(\alpha, \theta)$ for all $\alpha \in \mathcal{M}$, we obtain that $u_i(\alpha, \theta) = u_i^*(\alpha, \theta)$ for $\chi_\theta$-almost all $\alpha \in \mathcal{M}$. Since this is true for every player $i$, we obtain that $\chi_\theta(\mathrm{NE}_\theta) = 1$. Since $\mathrm{NE}_\theta$ is closed, we conclude that $\mathrm{supp}(\chi_\theta) \subseteq \mathrm{NE}_\theta$. □

We are ready to show that Theorem 7-(ii) holds. Fix a state $\theta$, and denote by $p_\theta \in \Delta(A)$ the conditional distribution of $a$:

$$p_\theta(a) = \frac{p(a, \theta)}{\pi(\theta)}.$$

Let $M_\theta$ be the set of all $\alpha \in \mathrm{NE}_\theta$ that satisfy (77). Note that the set $M_\theta$ is compact. Denote by $\mathrm{co}(M_\theta) \subseteq \Delta(A)$ the convex hull of $M_\theta$, that is, the set of all convex combinations of (finitely many) elements of $M_\theta$ (here we identify $\alpha$ with the product measure induced on $A$). Since $M_\theta$ is compact, $\mathrm{co}(M_\theta)$ is compact. By Lemmas 28 and 29, the probability measure $\chi_\theta$ puts probability one on $M_\theta$. Since $p_\theta$ is the barycenter of $\chi_\theta$—see (75)—we obtain that $p_\theta \in \mathrm{co}(M_\theta)$ (Phelps, 2001, Proposition 1.2).

Overall, for every state $\theta$, there are $\alpha_\theta^l \in M_\theta$, with $l = 1, \ldots, L_\theta$, such that

$$p_\theta(a) = \sum_{l=1}^{L_\theta} s_\theta^l \left[ \prod_{i \in I} \alpha_{\theta,i}^l(a_i) \right]$$

with $s_\theta^l \geq 0$ for all $l \in L_\theta$, and $\sum_{l=1}^{L_\theta} s_\theta^l = 1$. It follows from a standard cake-cutting argument that, without loss of generality, we can assume that $s_\theta^l$ and $L_\theta$ are independent of $\theta$. Given $s^l := s_\theta^l$ and $L := L_\theta$, we define $q_l \in \Delta(A)$ by

$$q_l(a, \theta) = \pi(\theta) \prod_{i \in I} \alpha_{\theta,i}^l(a_i),$$

and we notice that

$$p(a,\theta)\sum_{l=1}^{L} s^l q_l(a,\theta).$$

Each $q_l$ is a $(\mathcal{A}, p)$-decomposable complete-information Nash equilibrium. We conclude that Theorem 7-(ii) holds.