

## Supplement to “Dynamic games with (almost) perfect information”

(*Theoretical Economics*, Vol. 15, No. 2, May 2020, 811–859)

WEI HE

Department of Economics, The Chinese University of Hong Kong

YENENG SUN

Risk Management Institute, Departments of Economics and Mathematics, National University of Singapore

### APPENDIX B

In [Appendix B.1](#), we present the model of measurable dynamic games with partially perfect information and show the existence of subgame-perfect equilibria in [Proposition B.1](#). It covers the results in Theorem 3 (Theorem 4) for dynamic games with almost perfect information (perfect information) and in discounted stochastic games.

In [Appendix B.2](#), we present Lemmas [B.1–B.6](#) as the mathematical preparations for proving Theorem 3. We present in [Appendix B.3](#) a new equilibrium existence result for discontinuous games with stochastic endogenous sharing rules. The proof of Theorem 3 is given in [Appendix B.4](#). The proof of [Proposition B.1](#) is provided in [Appendix B.5](#), which covers Theorem 4 as a special case. One can skip Appendices [B.2](#) and [B.3](#) first, and refer to the technical results in these two sections whenever necessary.

#### B.1 *Measurable dynamic games with partially perfect information*

In this section, we generalize the model of measurable dynamic games in three directions. The ARM condition is partially relaxed such that (a) perfect information may be allowed in some stages, (b) the state transitions could have a weakly continuous component in all other stages, and (c) the state transition in any period can depend on the action profile in the current stage as well as on the previous history. The first change allows us to combine the models of dynamic games with perfect and almost perfect information. The second generalization implies that the state transitions need not be norm continuous on the Banach space of finite measures. The last modification covers the model of stochastic games as a special case.

The changes are described below.

- The state space is a product space of two Polish spaces; that is,  $S_t = \hat{S}_t \times \tilde{S}_t$  for each  $t \geq 1$ .

---

Wei He: [hewei@cuhk.edu.hk](mailto:hewei@cuhk.edu.hk)

Yeneng Sun: [ynsun@nus.edu.sg](mailto:ynsun@nus.edu.sg)

- For each  $i \in I$ , the action correspondence  $A_{ti}$  from  $H_{t-1}$  to  $X_{ti}$  is measurable, nonempty and compact valued, and sectionally continuous on  $X^{t-1} \times \hat{S}^{t-1}$ . The additional component of Nature is given by a measurable, nonempty and closed valued correspondence  $\hat{A}_{t0}$  from  $\text{Gr}(A_t)$  to  $\hat{S}_t$ , which is sectionally continuous on  $X^t \times \hat{S}^{t-1}$ . Then  $H_t = \text{Gr}(\hat{A}_{t0}) \times \tilde{S}_t$  and  $H_\infty$  is the subset of  $X^\infty \times S^\infty$  such that  $(x, s) \in H_\infty$  if  $(x^t, s^t) \in H_t$  for any  $t \geq 0$ .
- The choice of Nature depends not only on the history  $h_{t-1}$ , but also on the action profile  $x_t$  in the current stage. The state transition  $f_{t0}(h_{t-1}, x_t) = \hat{f}_{t0}(h_{t-1}, x_t) \diamond \tilde{f}_{t0}(h_{t-1}, x_t)$ , where  $\hat{f}_{t0}$  is a transition probability from  $\text{Gr}(A_t)$  to  $\mathcal{M}(\hat{S}_t)$  such that  $\hat{f}_{t0}(\hat{A}_{t0}(h_{t-1}, x_t)|h_{t-1}, x_t) = 1$  for all  $(h_{t-1}, x_t) \in \text{Gr}(A_t)$  and  $\tilde{f}_{t0}$  is a transition probability from  $\text{Gr}(\hat{A}_{t0})$  to  $\mathcal{M}(\tilde{S}_t)$ .
- For each  $i \in I$ , the payoff function  $u_i$  is a Borel measurable mapping from  $H_\infty$  to  $\mathbb{R}_{++}$ , which is sectionally continuous on  $X^\infty \times \hat{S}^\infty$ .

As in Section 3.3, we allow the possibility for the players to have perfect information in some stages. For  $t \geq 1$ , let

$$N_t = \begin{cases} 1, & \text{if } f_{t0}(h_{t-1}, x_t) \equiv \delta_{s_t} \text{ for some } s_t \text{ and} \\ & |\{i \in I: A_{ti} \text{ is not point-valued}\}| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, if  $N_t = 1$  for some stage  $t$ , then the player who is active in the period  $t$  is the only active player and has perfect information.

We drop the ARM condition in those periods with only one active player and weaken the ARM condition in other periods.

**ASSUMPTION B.1 (ARM').** (i) For any  $t \geq 1$  with  $N_t = 1$ ,  $S_t$  is a singleton set  $\{\hat{s}_t\}$  and  $\lambda_t = \delta_{\hat{s}_t}$ .

(ii) For each  $t \geq 1$  with  $N_t = 0$ ,  $\hat{f}_{t0}$  is sectionally continuous on  $X^t \times \hat{S}^{t-1}$ , where the range space  $\mathcal{M}(\hat{S}_t)$  is endowed with topology of weak convergence of measures on  $\hat{S}_t$ . The probability measure  $\tilde{f}_{t0}(\cdot|h_{t-1}, x_t, \hat{s}_t)$  is absolutely continuous with respect to an atomless Borel probability measure  $\lambda_t$  on  $\tilde{S}_t$  for all  $(h_{t-1}, x_t, \hat{s}_t) \in \text{Gr}(\hat{A}_{t0})$  and  $\varphi_{t0}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t)$  is the corresponding density.<sup>1</sup>

(iii) The mapping  $\varphi_{t0}$  is Borel measurable and sectionally continuous on  $X^t \times \hat{S}^t$ , and integrably bounded in the sense that there is a  $\lambda_t$ -integrable function  $\phi_t: \tilde{S}_t \rightarrow \mathbb{R}_+$  such that  $\varphi_{t0}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t) \leq \phi_t(\tilde{s}_t)$  for any  $(h_{t-1}, x_t, \hat{s}_t)$ .

The following result shows that the existence result is still true in this more general setting.

<sup>1</sup>In this section, a property is said to hold for  $\lambda^t$ -almost all  $h_t \in H_t$  if it is satisfied for  $\lambda^t$ -almost all  $\tilde{s}^t \in \tilde{S}^t$  and all  $(x^t, \hat{s}^t) \in H_t(\tilde{s}^t)$ .

**PROPOSITION B.1.** *If an infinite-horizon dynamic game as described above satisfies the ARM' condition and is continuous at infinity, then it possesses a subgame-perfect equilibrium  $f$ . In particular, for  $j \in I$  and  $t \geq 1$  such that  $N_t = 1$  and player  $j$  is the only active player in this period,  $f_{ij}$  can be deterministic. Furthermore, the equilibrium payoff correspondence  $E_t$  is nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^{t-1} \times \hat{S}^{t-1}$ .*

**REMARK B.1.** The result above also implies a new existence result of subgame-perfect equilibria for stochastic games. In the existence result of [Mertens and Parthasarathy \(2003\)](#), the state transitions are assumed to be norm continuous with respect to the actions in the previous stage. They did not assume the ARM condition. On the contrary, our [Proposition B.1](#) allows the state transitions to have a weakly continuous component.

## B.2 Technical preparations

The following lemma shows that the space of nonempty compact subsets of a Polish space is still Polish under the Hausdorff metric topology.

**LEMMA B.1.** *Suppose that  $X$  is a Polish space and that  $\mathcal{K}$  is the set of all nonempty compact subsets of  $X$  endowed with the Hausdorff metric topology. Then  $\mathcal{K}$  is a Polish space.*

**PROOF.** By Theorem 3.88(2) of [Aliprantis and Border \(2006\)](#),  $\mathcal{K}$  is complete. In addition, Corollary 3.90 and Theorem 3.91 of [Aliprantis and Border \(2006\)](#) imply that  $\mathcal{K}$  is separable. Thus,  $\mathcal{K}$  is a Polish space.  $\square$

The following result presents a variant of Lemma 5 in terms of transition correspondences.

**LEMMA B.2.** *Let  $X$  and  $Y$  be Polish spaces, and  $Z$  a compact subset of  $\mathbb{R}_+^l$ . Let  $G$  be a measurable, nonempty and compact valued correspondence from  $X$  to  $\mathcal{M}(Y)$ . Suppose that  $F$  is a measurable, nonempty, convex and compact valued correspondence from  $X \times Y$  to  $Z$ . Define a correspondence  $\Pi$  from  $X$  to  $Z$  as*

$$\Pi(x) = \left\{ \int_Y f(x, y) g(dy|x) : \begin{array}{l} g \text{ is a Borel measurable selection of } G, \\ f \text{ is a Borel measurable selection of } F \end{array} \right\}.$$

*If  $F$  is sectionally continuous on  $Y$ , then the following statements hold:*

- (i) *The correspondence  $\tilde{F} : X \times \mathcal{M}(Y) \rightarrow Z$  as  $\tilde{F}(x, \nu) = \int_Y F(x, y) \nu(dy)$  is sectionally continuous on  $\mathcal{M}(Y)$ .*
- (ii) *The  $\Pi$  is a measurable, nonempty and compact valued correspondence.*
- (iii) *If  $F$  and  $G$  are both continuous, then  $\Pi$  is continuous.*

PROOF. (i) For any fixed  $x \in X$ , the upper hemicontinuity of  $\tilde{F}(x, \cdot)$  follows from Lemma 7.

Next we show the lower hemicontinuity. Fix any  $x \in X$ . Suppose that  $\{\nu_j\}_{j \geq 0}$  is a sequence in  $\mathcal{M}(Y)$  such that  $\nu_j \rightarrow \nu_0$  as  $j \rightarrow \infty$ . Pick an arbitrary point  $z_0 \in \tilde{F}(x, \nu_0)$ . Then there exists a Borel measurable selection  $f$  of  $F(x, \cdot)$  such that  $z_0 = \int_Y f(y) \nu_0(dy)$ .

By Lemma 3 (Lusin's theorem), for each  $k \geq 1$ , there exists a compact subset  $D_k \subseteq Y$  such that  $f$  is continuous on  $D_k$  and  $\nu_0(Y \setminus D_k) < \frac{1}{3kM}$ , where  $M > 0$  is the bound of  $Z$ . Define a correspondence  $F_k : Y \rightarrow Z$  as

$$F_k(y) = \begin{cases} \{f(y)\}, & y \in D_k, \\ F(x, y), & y \in Y \setminus D_k. \end{cases}$$

Then  $F_k$  is nonempty, convex and compact valued, and lower hemicontinuous. By Theorem 3.22 in Aliprantis and Border (2006),  $Y$  is paracompact. Then by Lemma 3 (Michael's selection theorem),  $F_k$  has a continuous selection  $f_k$ .

For each  $k$ , since  $\nu_j \rightarrow \nu_0$ , and  $f_k$  is bounded and continuous,  $\int_Y f_k(y) \nu_j(dy) \rightarrow \int_Y f_k(y) \nu_0(dy)$  as  $j \rightarrow \infty$ . Thus, there exists a subsequence  $\{\nu_{j_k}\}$  such that  $\{j_k\}$  is an increasing sequence and, for each  $k \geq 1$ ,

$$\left\| \int_Y f_k(y) \nu_{j_k}(dy) - \int_Y f_k(y) \nu_0(dy) \right\| < \frac{1}{3k},$$

where  $\|\cdot\|$  is the Euclidean norm on  $\mathbb{R}^l$ .

Since  $f_k$  coincides with  $f$  on  $D_k$ ,  $\nu_0(Y \setminus D_k) < \frac{1}{3kM}$ , and  $Z$  is bounded by  $M$ ,

$$\left\| \int_Y f_k(y) \nu_0(dy) - \int_Y f(y) \nu_0(dy) \right\| < \frac{2}{3k}.$$

Thus,

$$\left\| \int_Y f_k(y) \nu_{j_k}(dy) - \int_Y f(y) \nu_0(dy) \right\| < \frac{1}{k}.$$

Let  $z_{j_k} = \int_Y f_k(y) \nu_{j_k}(dy)$  for each  $k$ . Then  $z_{j_k} \in \tilde{F}(x, \nu_{j_k})$  and  $z_{j_k} \rightarrow z_0$  as  $k \rightarrow \infty$ . By Lemma 1,  $\tilde{F}(x, \cdot)$  is lower hemicontinuous.

(ii) Since  $G$  is measurable and compact-valued, there exists a sequence of Borel measurable selections  $\{g_k\}_{k \geq 1}$  of  $G$  such that  $G(x) = \overline{\{g_1(x), g_2(x), \dots\}}$  for any  $x \in X$  by Lemma 2(v). For each  $k \geq 1$ , define a correspondence  $\Pi^k$  from  $X$  to  $Z$  by letting  $\Pi^k(x) = \tilde{F}(x, g_k(x)) = \int_Y F(x, y) g_k(dy|x)$ . Since  $F$  is convex-valued, so is  $\Pi^k$ . By Lemma 5,  $\Pi^k$  is also measurable, nonempty and compact valued.

Fix any  $x \in X$ . It is clear that  $\Pi(x) = \tilde{F}(x, G(x))$  is nonempty-valued. Since  $G(x)$  is compact, and  $\tilde{F}(x, \cdot)$  is compact-valued and continuous,  $\Pi(x)$  is compact by Lemma 2. Thus,  $\bigcup_{k \geq 1} \Pi^k(x) \subseteq \Pi(x)$ .

Fix any  $x \in X$  and  $z \in \Pi(x)$ . There exists a point  $\nu \in G(x)$  such that  $z \in \tilde{F}(x, \nu)$ . Since  $\{g_k(x)\}_{k \geq 1}$  is dense in  $G(x)$ , it has a subsequence  $\{g_{k_m}(x)\}$  such that  $g_{k_m}(x) \rightarrow \nu$ . As

$\tilde{F}(x, \cdot)$  is continuous,  $\tilde{F}(x, g_{k_m}(x)) \rightarrow \tilde{F}(x, \nu)$ ; that is,

$$z \in \overline{\bigcup_{k \geq 1} \tilde{F}(x, g_k(x))} = \overline{\bigcup_{k \geq 1} \Pi^k(x)}.$$

Therefore,  $\overline{\bigcup_{k \geq 1} \Pi^k(x)} = \Pi(x)$  for any  $x \in X$ . Lemma 2(i) and (ii) imply that  $\Pi$  is measurable.

(iii) Define a correspondence  $\hat{F}: \mathcal{M}(X \times Y) \rightarrow Z$  as

$$\hat{F}(\tau) = \left\{ \int_{X \times Y} f(x, y) \tau(d(x, y)) : f \text{ is a Borel measurable selection of } F \right\}.$$

By (i),  $\hat{F}$  is continuous. Define a correspondence  $\hat{G}: X \rightarrow \mathcal{M}(X \times Y)$  as  $\hat{G}(x) = \{\delta_x \otimes \nu : \nu \in G(x)\}$ . Since  $\hat{G}$  and  $\hat{F}$  are both nonempty-valued,  $\Pi(x) = \hat{F}(\hat{G}(x))$  is nonempty. As  $\hat{G}$  is compact-valued and  $\hat{F}$  is continuous,  $\Pi$  is compact-valued by Lemma 2. As  $\hat{G}$  and  $\hat{F}$  are both continuous,  $\Pi$  is continuous by Lemma 1(vii).  $\square$

The following lemma shows that a measurable and sectionally continuous correspondence on a product space is approximately continuous on the product space.

**LEMMA B.3.** *Let  $S$ ,  $X$ , and  $Y$  be Polish spaces endowed with the Borel  $\sigma$ -algebras, and  $\lambda$  a Borel probability measure on  $S$ . Denote  $\mathcal{B}(S)$  as the completion of the Borel  $\sigma$ -algebra  $\mathcal{B}(S)$  of  $S$  under the probability measure  $\lambda$ . Suppose that  $D$  is a  $\mathcal{B}(S) \otimes \mathcal{B}(Y)$ -measurable subset of  $S \times Y$ , where  $D(s)$  is nonempty and compact for all  $s \in S$ . Let  $A$  be a nonempty and compact valued correspondence from  $D$  to  $X$ , which is sectionally continuous on  $Y$  and has a  $\mathcal{B}(S \times Y \times X)$ -measurable graph. Then the following statements hold:*

- (i) *We have that  $\tilde{A}(s) = \text{Gr}(A(s, \cdot))$  is an  $\mathcal{S}$ -measurable mapping from  $S$  to the set of nonempty and compact subsets  $\mathcal{K}_{Y \times X}$  of  $Y \times X$ .*
- (ii) *There exist countably many disjoint compact subsets  $\{S_m\}_{m \geq 1}$  of  $S$  such that (a)  $\lambda(\bigcup_{m \geq 1} S_m) = 1$ , and (b) for each  $m \geq 1$ ,  $D_m = D \cap (S_m \times Y)$  is compact, and  $A$  is nonempty and compact valued, and continuous on each  $D_m$ .*

**PROOF.** (i) Given that  $A(s, \cdot)$  is continuous and  $D(s)$  is compact,  $\text{Gr}(A(s, \cdot)) \subseteq Y \times X$  is compact by Lemma 2. Thus,  $\tilde{A}$  is nonempty and compact valued. Since  $A$  has a measurable graph,  $\tilde{A}$  is an  $\mathcal{S}$ -measurable mapping from  $S$  to the set of nonempty and compact subsets  $\mathcal{K}_{Y \times X}$  of  $Y \times X$  by Lemma 1(iv).

(ii) Define a correspondence  $\tilde{D}$  from  $S$  to  $Y$  such that  $\tilde{D}(s) = \{y \in Y : (s, y) \in D\}$ . Then  $\tilde{D}$  is nonempty and compact valued. As in (i),  $\tilde{D}$  is  $\mathcal{S}$ -measurable. By Lemma 3 (Lusin's theorem), there exists a compact subset  $S_1 \subseteq S$  such that  $\lambda(S \setminus S_1) < \frac{1}{2}$ , and  $\tilde{D}$  and  $\tilde{A}$  are continuous functions on  $S_1$ . By Lemma 1(iii),  $\tilde{D}$  and  $\tilde{A}$  are continuous correspondences on  $S_1$ . Let  $D_1 = \{(s, y) \in D : s \in S_1, y \in \tilde{D}(s)\}$ . Since  $S_1$  is compact and  $\tilde{D}$  is continuous,  $D_1$  is compact (see Lemma 2 (6)).

Following the same procedure, for any  $m \geq 1$ , there exists a compact subset  $S_m \subseteq S$  such that (a)  $S_m \cap (\bigcup_{1 \leq k \leq m-1} S_k) = \emptyset$  and  $D_m = D \cap (S_m \times Y)$  is compact, (b)  $\lambda(S_m) > 0$

and  $\lambda(S \setminus (\bigcup_{1 \leq k \leq m} S_m)) < \frac{1}{2m}$ , and (c)  $A$  is nonempty and compact valued, and continuous on  $D_m$ . This completes the proof.  $\square$

The lemma below states an equivalence property for the weak convergence of Borel probability measures obtained from the product of transition probabilities.

**LEMMA B.4.** *Let  $S$  and  $X$  be Polish spaces, and  $\lambda$  a Borel probability measure on  $S$ . Suppose that  $\{S_k\}_{k \geq 1}$  is a sequence of disjoint compact subsets of  $S$  such that  $\lambda(\bigcup_{k \geq 1} S_k) = 1$ . For each  $k$ , define a probability measure on  $S_k$  as  $\lambda_k(D) = \frac{\lambda(D)}{\lambda(S_k)}$  for any measurable subset  $D \subseteq S_k$ . Let  $\{\nu_m\}_{m \geq 0}$  be a sequence of transition probabilities from  $S$  to  $\mathcal{M}(X)$  and  $\tau_m = \lambda \diamond \nu_m$  for any  $m \geq 0$ . Then  $\tau_m$  weakly converges to  $\tau_0$  if and only if  $\lambda_k \diamond \nu_m$  weakly converges to  $\lambda_k \diamond \nu_0$  for each  $k \geq 1$ .*

**PROOF.** First, we assume that  $\tau_m$  weakly converges to  $\tau_0$ . For any closed subset  $E \subseteq S_k \times X$ , we have  $\limsup_{m \rightarrow \infty} \tau_m(E) \leq \tau_0(E)$ ; that is,  $\limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E) \leq \lambda \diamond \nu_0(E)$ . For any  $k$ ,  $\frac{1}{\lambda(S_k)} \limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E) \leq \frac{1}{\lambda(S_k)} \lambda \diamond \nu_0(E)$ , which implies that  $\limsup_{m \rightarrow \infty} \lambda_k \diamond \nu_m(E) \leq \lambda_k \diamond \nu_0(E)$ . Thus,  $\lambda_k \diamond \nu_m$  weakly converges to  $\lambda_k \diamond \nu_0$  for each  $k \geq 1$ .

Second, we consider the case that  $\lambda_k \diamond \nu_m$  weakly converges to  $\lambda_k \diamond \nu_0$  for each  $k \geq 1$ . For any closed subset  $E \subseteq S \times X$ , let  $E_k = E \cap (S_k \times X)$  for each  $k \geq 1$ . Then  $\{E_k\}$  are disjoint closed subsets and  $\limsup_{m \rightarrow \infty} \lambda_k \diamond \nu_m(E_k) \leq \lambda_k \diamond \nu_0(E_k)$ . Since  $\lambda_k \diamond \nu_m(E') = \frac{1}{\lambda(S_k)} \lambda \diamond \nu_m(E')$  for any  $k, m$ , and measurable subset  $E' \subseteq S_k \times X$ , we have that  $\limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E_k) \leq \lambda \diamond \nu_0(E_k)$ . Thus,

$$\sum_{k \geq 1} \limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E_k) \leq \sum_{k \geq 1} \lambda \diamond \nu_0(E_k) = \lambda \diamond \nu_0(E).$$

Since the limit superior is subadditive, we have

$$\sum_{k \geq 1} \limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E_k) \geq \limsup_{m \rightarrow \infty} \sum_{k \geq 1} \lambda \diamond \nu_m(E_k) = \limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E).$$

Therefore,  $\limsup_{m \rightarrow \infty} \lambda \diamond \nu_m(E) \leq \lambda \diamond \nu_0(E)$ , which implies that  $\tau_m$  weakly converges to  $\tau_0$ .  $\square$

The following lemma is a key that allows one to drop the continuity condition on the state variables through a reference measure in Theorem 3.

**LEMMA B.5.** *Suppose that  $X, Y$ , and  $S$  are Polish spaces, and that  $Z$  is a compact metric space. Let  $\lambda$  be a Borel probability measure on  $S$ , and  $A$  a nonempty and compact valued correspondence from  $Z \times S$  to  $X$  that is sectionally upper hemicontinuous on  $Z$  and has a  $\mathcal{B}(Z \times S \times X)$ -measurable graph. Let  $G$  be a nonempty and compact valued, and continuous correspondence from  $Z$  to  $\mathcal{M}(X \times S)$ . We assume that for any  $z \in Z$  and  $\tau \in G(z)$ , the marginal of  $\tau$  on  $S$  is  $\lambda$  and  $\tau(\text{Gr}(A(z, \cdot))) = 1$ . Let  $F$  be a measurable, nonempty, convex and compact valued correspondence from  $\text{Gr}(A) \rightarrow \mathcal{M}(Y)$  such that*

$F$  is sectionally continuous on  $Z \times X$ . Define a correspondence  $\Pi$  from  $Z$  to  $\mathcal{M}(X \times S \times Y)$  by letting

$$\Pi(z) = \{g(z) \diamond f(z, \cdot) : g \text{ is a Borel measurable selection of } G, \\ f \text{ is a Borel measurable selection of } F\}.$$

Then the correspondence  $\Pi$  is nonempty and compact valued, and continuous.

PROOF. Let  $S$  be the completion of  $\mathcal{B}(S)$  under the probability measure  $\lambda$ . By Lemma B.3,  $\tilde{A}(s) = \text{Gr}(A(s, \cdot))$  can be viewed as an  $\mathcal{S}$ -measurable mapping from  $S$  to the set of nonempty and compact subsets  $\mathcal{K}_{Z \times X}$  of  $Z \times X$ . For any  $s \in S$ , the correspondence  $F_s = F(\cdot, s)$  is continuous on  $\tilde{A}(s)$ . By Lemma 3, there exists a measurable, nonempty and compact valued correspondence  $\tilde{F}$  from  $Z \times X \times S$  to  $\mathcal{M}(Y)$  and a Borel measurable subset  $S' \subseteq S$  with  $\lambda(S') = 1$  such that for each  $s \in S'$ ,  $\tilde{F}_s$  is continuous on  $Z \times X$ , and the restriction of  $\tilde{F}_s$  to  $\tilde{A}(s)$  is  $F_s$ .

By Lemma 3 (Lusin's theorem), there exists a compact subset  $S_1 \subseteq S'$  such that  $\tilde{A}$  is continuous on  $S_1$  and  $\lambda(S_1) > \frac{1}{2}$ . Let  $K_1 = \tilde{A}(S_1)$ . Then  $K_1 \subseteq Z \times X$  is compact.

Let  $C(K_1, \mathcal{K}_{\mathcal{M}(Y)})$  be the space of continuous functions from  $K_1$  to  $\mathcal{K}_{\mathcal{M}(Y)}$ , where  $\mathcal{K}_{\mathcal{M}(Y)}$  is the set of nonempty and compact subsets of  $\mathcal{M}(Y)$ . Suppose that the restriction of  $\mathcal{S}$  on  $S_1$  is  $S_1$ . Let  $\tilde{F}_1$  be the restriction of  $\tilde{F}$  to  $K_1 \times S_1$ . Then  $\tilde{F}_1$  can be viewed as an  $S_1$ -measurable function from  $S_1$  to  $C(K_1, \mathcal{K}_{\mathcal{M}(Y)})$  (see Theorem 4.55 in Aliprantis and Border (2006)). Again by Lemma 3 (Lusin's theorem), there exists a compact subset of  $S_1$ , say itself, such that  $\lambda(S_1) > \frac{1}{2}$  and  $\tilde{F}_1$  is continuous on  $S_1$ . As a result,  $\tilde{F}_1$  is a continuous correspondence on  $\text{Gr}(A) \cap (S_1 \times Z \times X)$  and so is  $F$ . Let  $\lambda_1$  be a probability measure on  $S_1$  such that  $\lambda_1(D) = \frac{\lambda(D)}{\lambda(S_1)}$  for any measurable subset  $D \subseteq S_1$ .

For any  $z \in Z$  and  $\tau \in G(z)$ , the definition of  $G$  implies that there exists a transition probability  $\nu$  from  $S$  to  $X$  such that  $\lambda \diamond \nu = \tau$ . Define a correspondence  $G_1$  from  $Z$  to  $\mathcal{M}(X \times S)$  as follows: for any  $z \in Z$ ,  $G_1(z)$  is the set of all  $\tau_1 = \lambda_1 \diamond \nu$  such that  $\tau = \lambda \diamond \nu \in G(z)$ . It can be easily checked that  $G_1$  is also a nonempty and compact valued, and continuous correspondence. Let

$$\Pi_1(z) = \{\tau_1 \diamond f(z, \cdot) : \tau_1 = \lambda_1 \diamond \nu \in G_1(z), \\ f \text{ is a Borel measurable selection of } \tilde{F}\}.$$

By Lemma 9,  $\Pi_1$  is nonempty and compact valued, and continuous. Furthermore, it is easy to see that for any  $z$ ,  $\Pi_1(z)$  coincides with the set

$$\{(\lambda_1 \diamond \nu) \diamond f(z, \cdot) : \lambda \diamond \nu \in G(z), f \text{ is a Borel measurable selection of } F\}.$$

Repeating this procedure, one can find a sequence of compact subsets  $\{S_t\}$  such that (a) for any  $t \geq 1$ ,  $S_t \subseteq S'$ ,  $S_t \cap (S_1 \cup \dots \cup S_{t-1}) = \emptyset$  and  $\lambda(S_1 \cup \dots \cup S_t) \geq \frac{t}{t+1}$ , (b)  $F$  is continuous on  $\text{Gr}(A) \cap (S_t \times Z \times X)$ ,  $\lambda_t$  is a probability measure on  $S_t$

such that  $\lambda_t(D) = \frac{\lambda(D)}{\lambda(S_t)}$  for any measurable subset  $D \subseteq S_t$ , and (c) the correspondence

$$\Pi_t(z) = \{(\lambda_t \diamond \nu) \diamond f(z, \cdot) : \lambda \diamond \nu \in G(z), \\ f \text{ is a Borel measurable selection of } F\}.$$

is nonempty and compact valued, and continuous.

Pick a sequence  $\{z_k\}$ ,  $\{\nu_k\}$ , and  $\{f_k\}$  such that  $(\lambda \diamond \nu_k) \diamond f_k(z_k, \cdot) \in \Pi(z_k)$ ,  $z_k \rightarrow z_0$ , and  $(\lambda \diamond \nu_k) \diamond f_k(z_k, \cdot)$  weakly converges to some  $\kappa$ . It is easy to see that  $(\lambda_t \diamond \nu_k) \diamond f_k(z_k, \cdot) \in \Pi_t(z_k)$  for each  $t$ . As  $\Pi_1$  is compact-valued and continuous, it has a subsequence, say itself, such that  $z_k$  converges to some  $z_0 \in Z$  and  $(\lambda_1 \diamond \nu_k) \diamond f_k(z_k, \cdot)$  weakly converges to some  $(\lambda_1 \diamond \mu^1) \diamond f^1(z_0, \cdot) \in \Pi_1(z_0)$ . Repeating this procedure, one can get a sequence of  $\{\mu^m\}$  and  $f^m$ . Let  $\mu(s) = \mu^m(s)$  and  $f(z_0, s, x) = f^m(z_0, s, x)$  for any  $x \in A(z_0, s)$  when  $s \in S_m$ . By Lemma B.4,  $(\lambda \diamond \mu) \diamond f(z_0, \cdot) = \kappa$ , which implies that  $\Pi$  is upper hemicontinuous.

Similarly, the compactness and lower hemicontinuity of  $\Pi$  follow from the compactness and lower hemicontinuity of  $\Pi_t$  for each  $t$ .  $\square$

The next lemma presents the convergence property for the integrals of a sequence of functions and probability measures.

**LEMMA B.6.** *Let  $S$  and  $X$  be Polish spaces, and let  $A$  be a measurable, nonempty and compact valued correspondence from  $S$  to  $X$ . Suppose that  $\lambda$  is a Borel probability measure on  $S$  and that  $\{\nu_n\}_{1 \leq n \leq \infty}$  is a sequence of transition probabilities from  $S$  to  $\mathcal{M}(X)$  such that  $\nu_n(A(s)|s) = 1$  for each  $s$  and  $n$ . For each  $n \geq 1$ , let  $\tau_n = \lambda \diamond \nu_n$ . Assume that the sequence  $\{\tau_n\}$  of Borel probability measures on  $S \times X$  converges weakly to a Borel probability measure  $\tau_\infty$  on  $S \times X$ . Let  $\{g_n\}_{1 \leq n \leq \infty}$  be a sequence of functions that satisfy the following three properties.*

- (i) *For each  $n$  between 1 and  $\infty$ ,  $g_n : S \times X \rightarrow \mathbb{R}_+$  is measurable and sectionally continuous on  $X$ .*
- (ii) *For any  $s \in S$  and any sequence  $x_n \rightarrow x_\infty$  in  $X$ ,  $g_n(s, x_n) \rightarrow g_\infty(s, x_\infty)$  as  $n \rightarrow \infty$ .*
- (iii) *The sequence  $\{g_n\}_{1 \leq n \leq \infty}$  is integrably bounded in the sense that there exists a  $\lambda$ -integrable function  $\psi : S \rightarrow \mathbb{R}_+$  such that for any  $n$ ,  $s$ , and  $x$ ,  $g_n(s, x) \leq \psi(s)$ .*

Then we have

$$\int_{S \times X} g_n(s, x) \tau_n(d(s, x)) \rightarrow \int_{S \times X} g_\infty(s, x) \tau_\infty(d(s, x)).$$

**PROOF.** By Theorem 2.1.3 in Castaing et al. (2004), for any integrably bounded function  $g : S \times X \rightarrow \mathbb{R}_+$  that is sectionally continuous on  $X$ , we have

$$\int_{S \times X} g(s, x) \tau_n(d(s, x)) \rightarrow \int_{S \times X} g(s, x) \tau_\infty(d(s, x)). \quad (\text{B1})$$

Let  $\{y_n\}_{1 \leq n \leq \infty}$  be a sequence such that  $y_n = \frac{1}{n}$  and  $y_\infty = 0$ . Then  $y_n \rightarrow y_\infty$ . Define a mapping  $\tilde{g}$  from  $S \times X \times \{y_1, \dots, y_\infty\}$  such that  $\tilde{g}(s, x, y_n) = g_n(s, x)$ . Then  $\tilde{g}$  is measurable on  $S$  and continuous on  $X \times \{y_1, \dots, y_\infty\}$ . Define a correspondence  $G$  from  $S$  to  $X \times \{y_1, \dots, y_\infty\} \times \mathbb{R}_+$  such that

$$G(s) = \{(x, y_n, c) : c \in \tilde{g}(s, x, y_n), x \in A(s), 1 \leq n \leq \infty\}.$$

For any  $s$ ,  $A(s) \times \{y_1, \dots, y_\infty\}$  is compact and  $\tilde{g}(s, \cdot, \cdot)$  is continuous. By Lemma 2(vi),  $G(s)$  is compact. By Lemma 1(ii),  $G$  can be viewed as a measurable mapping from  $S$  to the space of nonempty compact subsets of  $X \times \{y_1, \dots, y_\infty\} \times \mathbb{R}_+$ . Similarly,  $A$  can be viewed as a measurable mapping from  $S$  to the space of nonempty compact subsets of  $X$ .

Fix an arbitrary  $\epsilon > 0$ . By Lemma 3 (Lusin's theorem), there exists a compact subset  $S_1 \subseteq S$  such that  $A$  and  $G$  are continuous on  $S_1$  and  $\lambda(S \setminus S_1) < \epsilon$ . Without loss of generality, we can assume that  $\lambda(S \setminus S_1)$  is sufficiently small such that  $\int_{S \setminus S_1} \psi(s) \lambda(ds) < \frac{\epsilon}{6}$ . Thus, for any  $n$ ,

$$\int_{(S \setminus S_1) \times X} \psi(s) \tau_n(d(s, x)) = \int_{(S \setminus S_1)} \psi(s) \nu_n(X) \lambda(ds) < \frac{\epsilon}{6}.$$

By Lemma 2(vi), the set  $E = \{(s, x) : s \in S_1, x \in A(s)\}$  is compact. Since  $G$  is continuous on  $S_1$ ,  $\tilde{g}$  is continuous on  $E \times \{y_1, \dots, y_\infty\}$ . Since  $E \times \{y_1, \dots, y_\infty\}$  is compact,  $\tilde{g}$  is uniformly continuous on  $E \times \{y_1, \dots, y_\infty\}$ . Thus, there exists a positive integer  $N_1 > 0$  such that for any  $n \geq N_1$ ,  $|g_n(s, x) - g_\infty(s, x)| < \frac{\epsilon}{3}$  for any  $(s, x) \in E$ .

By (B1), there exists a positive integer  $N_2$  such that for any  $n \geq N_2$ ,

$$\left| \int_{S \times X} g_\infty(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(d(s, x)) \right| < \frac{\epsilon}{3}.$$

Let  $N_0 = \max\{N_1, N_2\}$ . For any  $n \geq N_0$ ,

$$\begin{aligned} & \left| \int_{S \times X} g_n(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(d(s, x)) \right| \\ & \leq \left| \int_{S \times X} g_n(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_n(d(s, x)) \right| \\ & \quad + \left| \int_{S \times X} g_\infty(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(d(s, x)) \right| \\ & \leq \left| \int_{S_1 \times X} g_n(s, x) \tau_n(d(s, x)) - \int_{S_1 \times X} g_\infty(s, x) \tau_n(d(s, x)) \right| \\ & \quad + \left| \int_{(S \setminus S_1) \times X} g_n(s, x) \tau_n(d(s, x)) - \int_{(S \setminus S_1) \times X} g_\infty(s, x) \tau_n(d(s, x)) \right| \\ & \quad + \left| \int_{S \times X} g_\infty(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(d(s, x)) \right| \\ & \leq \int_E |g_n(s, x) - g_\infty(s, x)| \tau_n(d(s, x)) + 2 \cdot \int_{(S \setminus S_1) \times X} \psi(s) \tau_n(d(s, x)) \end{aligned}$$

$$\begin{aligned}
& + \left| \int_{S \times X} g_{\infty}(s, x) \tau_n(d(s, x)) - \int_{S \times X} g_{\infty}(s, x) \tau_{\infty}(d(s, x)) \right| \\
& < \frac{\epsilon}{3} + 2 \cdot \frac{\epsilon}{6} + \frac{\epsilon}{3} \\
& = \epsilon.
\end{aligned}$$

This completes the proof.  $\square$

### B.3 Discontinuous games with endogenous stochastic sharing rules

It was proved in [Simon and Zame \(1990\)](#) that a Nash equilibrium exists in discontinuous games with endogenous sharing rules. In particular, they considered a static game with a payoff correspondence  $P$  that is bounded, nonempty, convex and compact valued, and upper hemicontinuous. They showed that there exists a Borel measurable selection  $p$  of the payoff correspondence, namely the endogenous sharing rule, and a mixed-strategy profile  $\alpha$  such that  $\alpha$  is a Nash equilibrium when players take  $p$  as the payoff function (see Lemma 10).

In this section, we consider discontinuous games with endogenous stochastic sharing rules. That is, we allow the payoff correspondence to depend on some state variable in a measurable way as follows:

- Let  $S$  be a Borel subset of a Polish space,  $Y$  a Polish space, and  $\lambda$  a Borel probability measure on  $S$ .
- We have that  $D$  is a  $\mathcal{B}(S) \otimes \mathcal{B}(Y)$ -measurable subset of  $S \times Y$ , where  $D(s)$  is compact for all  $s \in S$  and  $\lambda(\{s \in S : D(s) \neq \emptyset\}) > 0$ .
- We have  $X = \prod_{1 \leq i \leq n} X_i$ , where each  $X_i$  is a Polish space.
- For each  $i$ ,  $A_i$  is a measurable, nonempty and compact valued correspondence from  $D$  to  $X_i$ , which is sectionally continuous on  $Y$ .
- We have  $A = \prod_{1 \leq i \leq n} A_i$  and  $E = \text{Gr}(A)$ .
- We have that  $P$  is a bounded, measurable, nonempty, convex and compact valued correspondence from  $E$  to  $\mathbb{R}^n$  that is essentially sectionally upper hemicontinuous on  $Y \times X$ .

A stochastic sharing rule at  $(s, y) \in D$  is a Borel measurable selection of the correspondence  $P(s, y, \cdot)$ ; i.e., a Borel measurable function  $p : A(s, y) \rightarrow \mathbb{R}^n$  such that  $p(x) \in P(s, y, x)$  for all  $x \in A(s, y)$ . Given  $(s, y) \in D$ ,  $P(s, y, \cdot)$  represents the set of all possible payoff profiles, and a sharing rule  $p$  is a particular choice of the payoff profile.

Now we prove the following proposition.

**PROPOSITION B.2.** *There exists a  $\mathcal{B}(D)$ -measurable, nonempty and compact valued correspondence  $\Phi$  from  $D$  to  $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$  such that  $\Phi$  is essentially sectionally upper hemicontinuous on  $Y$ , and for  $\lambda$ -almost all  $s \in S$  with  $D(s) \neq \emptyset$  and  $y \in D(s)$ ,  $\Phi(s, y)$  is the set of points  $(v, \alpha, \mu)$  that the following statements hold:*

- (i) We have  $v = \int_X p(s, y, x) \alpha(dx)$  such that  $p(s, y, \cdot)$  is a Borel measurable selection of  $P(s, y, \cdot)$ .<sup>2</sup>
- (ii) We have that  $\alpha \in \bigotimes_{i \in I} \mathcal{M}(A_i(s, y))$  is a Nash equilibrium in the subgame  $(s, y)$  with payoff profile  $p(s, y, \cdot)$  and action space  $A_i(s, y)$  for each player  $i$ .
- (iii) We have  $\mu = p(s, y, \cdot) \circ \alpha$ .<sup>3</sup>

In addition, denote the restriction of  $\Phi$  on the first component  $\mathbb{R}^n$  as  $\Phi|_{\mathbb{R}^n}$ , which is a correspondence from  $D$  to  $\mathbb{R}^n$ . Then  $\Phi|_{\mathbb{R}^n}$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $Y$ .

PROOF. There exists a Borel subset  $\hat{S} \subseteq S$  with  $\lambda(\hat{S}) = 1$  such that  $D(s) \neq \emptyset$  for each  $s \in \hat{S}$  and  $P$  is sectionally upper hemicontinuous on  $Y$  when it is restricted on  $D \cap (\hat{S} \times Y)$ . Without loss of generality, we assume that  $\hat{S} = S$ .

Suppose that  $\mathcal{S}$  is the completion of  $\mathcal{B}(S)$  under the probability measure  $\lambda$ . Let  $\mathcal{D}$  and  $\mathcal{E}$  be the restrictions of  $\mathcal{S} \otimes \mathcal{B}(Y)$  and  $\mathcal{S} \otimes \mathcal{B}(Y) \otimes \mathcal{B}(X)$  on  $D$  and  $E$ , respectively.

Define a correspondence  $\tilde{D}$  from  $S$  to  $Y$  such that  $\tilde{D}(s) = \{y \in Y : (s, y) \in D\}$ . Then  $\tilde{D}$  is nonempty and compact valued. By Lemma 1(iv),  $\tilde{D}$  is  $\mathcal{S}$ -measurable.

Since  $\tilde{D}(s)$  is compact and  $A(s, \cdot)$  is upper hemicontinuous for any  $s \in S$ ,  $E(s)$  is compact by Lemma 2(vi). Define a correspondence  $\Gamma$  from  $S$  to  $Y \times X \times \mathbb{R}^n$  as  $\Gamma(s) = \text{Gr}(P(s, \cdot, \cdot))$ . For all  $s$ ,  $P(s, \cdot, \cdot)$  is bounded, upper hemicontinuous, and compact-valued on  $E(s)$ ; hence, it has a compact graph. As a result,  $\Gamma$  is compact-valued. By Lemma 1(i),  $P$  has an  $\mathcal{S} \otimes \mathcal{B}(Y \times X \times \mathbb{R}^n)$ -measurable graph. Since  $\text{Gr}(\Gamma) = \text{Gr}(P)$ ,  $\text{Gr}(\Gamma)$  is  $\mathcal{S} \otimes \mathcal{B}(Y \times X \times \mathbb{R}^n)$ -measurable. Due to Lemma 1(iv), the correspondence  $\Gamma$  is  $\mathcal{S}$ -measurable. We can view  $\Gamma$  as a function from  $S$  into the space  $\mathcal{K}$  of nonempty compact subsets of  $Y \times X \times \mathbb{R}^n$ . By Lemma B.1,  $\mathcal{K}$  is a Polish space endowed with the Hausdorff metric topology. Then by Lemma 1(ii),  $\Gamma$  is an  $\mathcal{S}$ -measurable function from  $S$  to  $\mathcal{K}$ . One can also define a correspondence  $\tilde{A}_i$  from  $S$  to  $Y \times X$  as  $\tilde{A}_i(s) = \text{Gr}(A_i(s, \cdot))$ . It is easy to show that  $\tilde{A}_i$  can be viewed as an  $\mathcal{S}$ -measurable function from  $S$  to the space of nonempty compact subsets of  $Y \times X$ , which is endowed with the Hausdorff metric topology. By a similar argument,  $\tilde{D}$  can be viewed as an  $\mathcal{S}$ -measurable function from  $S$  to the space of nonempty compact subsets of  $Y$ .

By Lemma 3 (Lusin's theorem), there exists a compact subset  $S_1 \subseteq S$  such that  $\lambda(S \setminus S_1) < \frac{1}{2}$ ,  $\Gamma$ ,  $\tilde{D}$  and  $\{\tilde{A}_i\}_{1 \leq i \leq n}$  are continuous functions on  $S_1$ . By Lemma 1(iii),  $\Gamma$ ,  $\tilde{D}$ , and  $\tilde{A}_i$  are continuous correspondences on  $S_1$ . Let  $D_1 = \{(s, y) \in D : s \in S_1, y \in \tilde{D}(s)\}$ . Since  $S_1$  is compact and  $\tilde{D}$  is continuous,  $D_1$  is compact (see Lemma 2(vi)). Similarly,  $E_1 = E \cap (S_1 \times Y \times X)$  is also compact. Thus,  $P$  is an upper hemicontinuous correspondence on  $E_1$ . Define a correspondence  $\Phi_1$  from  $D_1$  to  $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$  as in Lemma 10; then it is nonempty and compact valued, and upper hemicontinuous on  $D_1$ .

Following the same procedure, for any  $m \geq 1$ , there exists a compact subset  $S_m \subseteq S$  such that (a)  $S_m \cap (\bigcup_{1 \leq k \leq m-1} S_k) = \emptyset$  and  $D_m = D \cap (S_m \times Y)$  is compact, (b)  $\lambda(S_m) > 0$  and  $\lambda(S \setminus (\bigcup_{1 \leq k \leq m} S_k)) < \frac{1}{2^m}$ , and (c) there is a nonempty and compact valued, upper hemicontinuous correspondence  $\Phi_m$  from  $D_m$  to  $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$  that satisfies

<sup>2</sup>Note that we require  $p(s, y, \cdot)$  to be measurable for each  $(s, y)$ , but  $p$  may not be jointly measurable.

<sup>3</sup>The finite measure  $\mu = p(s, y, \cdot) \circ \alpha$  if  $\mu(B) = \int_B p(s, y, x) \alpha(dx)$  for any Borel subset  $B \subseteq X$ .

conditions (i)–(iii) in Lemma 10. Thus, we have countably many disjoint sets  $\{S_m\}_{m \geq 1}$  such that (i)  $\lambda(\bigcup_{m \geq 1} S_m) = 1$ , and (ii)  $\Phi_m$  is nonempty and compact valued, and upper hemicontinuous on each  $D_m$ ,  $m \geq 1$ .

Since  $A_i$  is a  $\mathcal{B}(S) \otimes \mathcal{B}(Y)$ -measurable, nonempty and compact valued correspondence, it has a Borel measurable selection  $a_i$  by Lemma 2(iii). Fix a Borel measurable selection  $p$  of  $P$  (such a selection exists also due to Lemma 2(iii)). Define a mapping  $(v_0, \alpha_0, \mu_0)$  from  $D$  to  $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$  such that (a)  $\alpha_i(s, y) = \delta_{a_i(s, y)}$  and  $\alpha_0(s, y) = \bigotimes_{i \in I} \alpha_i(s, y)$ , (b)  $v_0(s, y) = p(s, y, a_1(s, y), \dots, a_n(s, y))$ , and (c)  $\mu_0(s, y) = p(s, y, \cdot) \circ \alpha_0$ . Let  $D_0 = D \setminus (\bigcup_{m \geq 1} D_m)$  and  $\Phi_0(s, y) = \{(v_0(s, y), \alpha_0(s, y), \mu_0(s, y))\}$  for  $(s, y) \in D_0$ . Then  $\Phi_0$  is  $\mathcal{B}(S) \otimes \mathcal{B}(Y)$ -measurable, nonempty and compact valued.

Let  $\Phi(s, y) = \Phi_m(s, y)$  if  $(s, y) \in D_m$  for some  $m \geq 0$ . Then  $\Phi(s, y)$  satisfies conditions (i)–(iii) if  $(s, y) \in D_m$  for  $m \geq 1$ . That is,  $\Phi$  is  $\mathcal{B}(D)$ -measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $Y$ , and satisfies conditions (i)–(iii) for  $\lambda$ -almost all  $s \in S$ .

Then consider  $\Phi|_{\mathbb{R}^n}$ , which is the restriction of  $\Phi$  on the first component  $\mathbb{R}^n$ . Let  $\Phi_m|_{\mathbb{R}^n}$  be the restriction of  $\Phi_m$  on the first component  $\mathbb{R}^n$  with the domain  $D_m$  for each  $m \geq 0$ . It is obvious that  $\Phi_0|_{\mathbb{R}^n}$  is measurable, nonempty and compact valued. For each  $m \geq 1$ ,  $D_m$  is compact, and  $\Phi_m$  is upper hemicontinuous and compact-valued. By Lemma 2(vi),  $\text{Gr}(\Phi_m)$  is compact. Thus,  $\text{Gr}(\Phi_m|_{\mathbb{R}^n})$  is also compact. By Lemma 2(iv),  $\Phi_m|_{\mathbb{R}^n}$  is measurable. In addition,  $\Phi_m|_{\mathbb{R}^n}$  is nonempty and compact valued, and upper hemicontinuous on  $D_m$ . Notice that  $\Phi|_{\mathbb{R}^n}(s, y) = \Phi_m|_{\mathbb{R}^n}(s, y)$  if  $(s, y) \in D_m$  for some  $m \geq 0$ . Thus,  $\Phi|_{\mathbb{R}^n}$  is measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $Y$ .

The proof is complete.  $\square$

#### B.4 Proof of Theorem 3

**B.4.1 Backward induction** For any  $t \geq 1$ , suppose that the correspondence  $Q_{t+1}$  from  $H_t$  to  $\mathbb{R}^n$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t$ . For any  $h_{t-1} \in H_{t-1}$  and  $x_t \in A_t(h_{t-1})$ , let

$$\begin{aligned} P_t(h_{t-1}, x_t) &= \int_{S_t} Q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t | h_{t-1}) \\ &= \int_{S_t} Q_{t+1}(h_{t-1}, x_t, s_t) \varphi_{t0}(h_{t-1}, s_t) \lambda_t(ds_t). \end{aligned}$$

It is obvious that the correspondence  $P_t$  is measurable and nonempty-valued. Since  $Q_{t+1}$  is bounded,  $P_t$  is bounded. For  $\lambda^t$ -almost all  $s^t \in S^t$ ,  $Q_{t+1}(\cdot, s^t)$  is bounded and upper hemicontinuous on  $H_t(s^t)$ , and  $\varphi_{t0}(s^t, \cdot)$  is continuous on  $\text{Gr}(A_0^t)(s^t)$ . As  $\varphi_{t0}$  is integrably bounded,  $P_t(s^{t-1}, \cdot)$  is also upper hemicontinuous on  $\text{Gr}(A^t)(s^{t-1})$  for  $\lambda^{t-1}$ -almost all  $s^{t-1} \in S^{t-1}$  (see Lemma 4); that is, the correspondence  $P_t$  is essentially sectionally upper hemicontinuous on  $X^t$ . Again by Lemma 4,  $P_t$  is convex and compact valued since  $\lambda_t$  is an atomless probability measure. That is,  $P_t: \text{Gr}(A^t) \rightarrow \mathbb{R}^n$  is a bounded, measurable, nonempty, convex and compact valued correspondence that is essentially sectionally upper hemicontinuous on  $X^t$ .

By [Proposition B.2](#), there exists a bounded, measurable, nonempty and compact valued correspondence  $\Phi_t$  from  $H_{t-1}$  to  $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$  such that  $\Phi_t$  is essentially sectionally upper hemicontinuous on  $X^{t-1}$ , and for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ ,  $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$  if the following conditions hold:

- (i) We have  $v = \int_{A_t(h_{t-1})} p_t(h_{t-1}, x) \alpha(dx)$  such that  $p_t(h_{t-1}, \cdot)$  is a Borel measurable selection of  $P_t(h_{t-1}, \cdot)$ .
- (ii) We have that  $\alpha \in \bigotimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$  is a Nash equilibrium in the subgame  $h_{t-1}$  with payoff  $p_t(h_{t-1}, \cdot)$  and action space  $\prod_{i \in I} A_{ti}(h_{t-1})$ .
- (iii) We have  $\mu = p_t(h_{t-1}, \cdot) \circ \alpha$ .

Denote the restriction of  $\Phi_t$  on the first component  $\mathbb{R}^n$  as  $\Phi(Q_{t+1})$ , which is a correspondence from  $H_{t-1}$  to  $\mathbb{R}^n$ . By [Proposition B.2](#),  $\Phi(Q_{t+1})$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^{t-1}$ .

**B.4.2 Forward induction** The following proposition presents the result on the step of forward induction.

**PROPOSITION B.3.** *For any  $t \geq 1$  and any Borel measurable selection  $q_t$  of  $\Phi(Q_{t+1})$ , there exists a Borel measurable selection  $q_{t+1}$  of  $Q_{t+1}$  and a Borel measurable mapping  $f_t: H_{t-1} \rightarrow \bigotimes_{i \in I} \mathcal{M}(X_{ti})$  such that for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ , the following statements hold:*

- (i) *We have  $f_t(h_{t-1}) \in \bigotimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$ .*
- (ii) *We have  $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$ .*
- (iii) *We have that  $f_t(\cdot|h_{t-1})$  is a Nash equilibrium in the subgame  $h_{t-1}$  with action spaces  $A_{ti}(h_{t-1})$ ,  $i \in I$ , and payoff functions*

$$\int_{S_t} q_{t+1}(h_{t-1}, \cdot, s_t) f_{t0}(ds_t|h_{t-1}).$$

**PROOF.** We divide the proof into three steps. In Step 1, we show that there exist Borel measurable mappings  $f_t: H_{t-1} \rightarrow \bigotimes_{i \in I} \mathcal{M}(X_{ti})$  and  $\mu_t: H_{t-1} \rightarrow \Delta(X_t)$  such that  $(q_t, f_t, \mu_t)$  is a selection of  $\Phi_t$ . In Step 2, we obtain a Borel measurable selection  $g_t$  of  $P_t$  such that for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ , the following statements hold:

- We have  $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} g_t(h_{t-1}, x) f_t(dx|h_{t-1})$ .
- We have that  $f_t(h_{t-1})$  is a Nash equilibrium in the subgame  $h_{t-1}$  with payoff  $g_t(h_{t-1}, \cdot)$  and action space  $A_t(h_{t-1})$ .

In Step 3, we show that there exists a Borel measurable selection  $q_{t+1}$  of  $Q_{t+1}$  such that for all  $h_{t-1} \in H_{t-1}$  and  $x_t \in A_t(h_{t-1})$ ,

$$g_t(h_{t-1}, x_t) = \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}).$$

Combining Steps 1–3, the proof is complete.

Step 1. Let  $\Psi_t: \text{Gr}(\Phi_t(Q_{t+1})) \rightarrow \mathcal{M}(X_t) \times \Delta(X_t)$  be

$$\Psi_t(h_{t-1}, v) = \{(\alpha, \mu): (v, \alpha, \mu) \in \Phi_t(h_{t-1})\}.$$

Recall that the construction of  $\Phi_t$  and the proof of [Proposition B.2](#),  $H_{t-1}$  can be divided into countably many Borel subsets  $\{H_{t-1}^m\}_{m \geq 0}$  such that the following statements hold:

- (i) We have  $H_{t-1} = \bigcup_{m \geq 0} H_{t-1}^m$  and  $\frac{\lambda^{t-1}(\bigcup_{m \geq 1} \text{proj}_{S^{t-1}}(H_{t-1}^m))}{\lambda^{t-1}(\text{proj}_{S^{t-1}}(H_{t-1}))} = 1$ , where  $\text{proj}_{S^{t-1}}(H_{t-1}^m)$  and  $\text{proj}_{S^{t-1}}(H_{t-1})$  are projections of  $H_{t-1}^m$  and  $H_{t-1}$  on  $S^{t-1}$ .
- (ii) For  $m \geq 1$ ,  $H_{t-1}^m$  is compact,  $\Phi_t$  is upper hemicontinuous on  $H_{t-1}^m$ , and  $P_t$  is upper hemicontinuous on

$$\{(h_{t-1}, x_t): h_{t-1} \in H_{t-1}^m, x_t \in A_t(h_{t-1})\}.$$

- (iii) There exists a Borel measurable mapping  $(v_0, \alpha_0, \mu_0)$  from  $H_{t-1}^0$  to  $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$  such that  $\Phi_t(h_{t-1}) \equiv \{(v_0(h_{t-1}), \alpha_0(h_{t-1}), \mu_0(h_{t-1}))\}$  for any  $h_{t-1} \in H_{t-1}^0$ .

Denote the restriction of  $\Phi_t$  on  $H_{t-1}^m$  as  $\Phi_t^m$ . For  $m \geq 1$ ,  $\text{Gr}(\Phi_t^m)$  is compact, and hence the correspondence  $\Psi_t^m(h_{t-1}, v) = \{(\alpha, \mu): (v, \alpha, \mu) \in \Phi_t^m(h_{t-1})\}$  has a compact graph. For  $m \geq 1$ ,  $\Psi_t^m$  is measurable by Lemma 2(iv), and has a Borel measurable selection  $\psi_t^m$  due to Lemma 2(iii). Define  $\psi_t^0(h_{t-1}, v_0(h_{t-1})) = (\alpha_0(h_{t-1}), \mu_0(h_{t-1}))$  for  $h_{t-1} \in H_{t-1}^0$ . For  $(h_{t-1}, v) \in \text{Gr}(\Phi(Q_{t+1}))$ , let  $\psi_t(h_{t-1}, v) = \psi_t^m(h_{t-1}, v)$  if  $h_{t-1} \in H_{t-1}^m$ . Then  $\psi_t$  is a Borel measurable selection of  $\Psi_t$ .

Given a Borel measurable selection  $q_t$  of  $\Phi(Q_{t+1})$ , let

$$\phi_t(h_{t-1}) = (q_t(h_{t-1}), \psi_t(h_{t-1}, q_t(h_{t-1}))).$$

Then  $\phi_t$  is a Borel measurable selection of  $\Phi_t$ . Denote  $\tilde{H}_{t-1} = \bigcup_{m \geq 1} H_{t-1}^m$ . By the construction of  $\Phi_t$ , there exist Borel measurable mappings  $f_t: H_{t-1} \rightarrow \bigotimes_{i \in I} \mathcal{M}(X_{ti})$  and  $\mu_t: H_{t-1} \rightarrow \Delta(X_t)$  such that for all  $h_{t-1} \in \tilde{H}_{t-1}$ , the following relationships hold:

- (a) We have  $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} p_t(h_{t-1}, x) f_t(dx|h_{t-1})$  such that  $p_t(h_{t-1}, \cdot)$  is a Borel measurable selection of  $P_t(h_{t-1}, \cdot)$ .
- (b) We have that  $f_t(h_{t-1}) \in \bigotimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$  is a Nash equilibrium in the subgame  $h_{t-1}$  with payoff  $p_t(h_{t-1}, \cdot)$  and action space  $\prod_{i \in I} A_{ti}(h_{t-1})$ .
- (c) We have  $\mu_t(\cdot|h_{t-1}) = p_t(h_{t-1}, \cdot) \circ f_t(\cdot|h_{t-1})$ .

Step 2. Since  $P_t$  is upper hemicontinuous on  $\{(h_{t-1}, x_t): h_{t-1} \in H_{t-1}^m, x_t \in A_t(h_{t-1})\}$ , due to Lemma 6, there exists a Borel measurable mapping  $g^m$  such that (i)  $g^m(h_{t-1}, x_t) \in P_t(h_{t-1}, x_t)$  for any  $h_{t-1} \in H_{t-1}^m$  and  $x_t \in A_t(h_{t-1})$ , and (ii)  $g^m(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$  for  $f_t(\cdot|h_{t-1})$ -almost all  $x_t$ . Fix an arbitrary Borel measurable selection  $g'$  of  $P_t$ . Define a Borel measurable mapping from  $\text{Gr}(A_t)$  to  $\mathbb{R}^n$  as

$$g(h_{t-1}, x_t) = \begin{cases} g^m(h_{t-1}, x_t) & \text{if } h_{t-1} \in H_{t-1}^m \text{ for } m \geq 1, \\ g'(h_{t-1}, x_t) & \text{otherwise.} \end{cases}$$

Then  $g$  is a Borel measurable selection of  $P_t$ .

In a subgame  $h_{t-1} \in \tilde{H}_{t-1}$ , let

$$B_{ti}(h_{t-1}) = \left\{ y_i \in A_{ti}(h_{t-1}) : \int_{A_{t(-i)}(h_{t-1})} g_i(h_{t-1}, y_i, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \right. \\ \left. > \int_{A_t(h_{t-1})} p_{ti}(h_{t-1}, x_t) f_t(dx_t | h_{t-1}) \right\}.$$

Since  $g(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$  for  $f_t(\cdot | h_{t-1})$ -almost all  $x_t$ ,

$$\int_{A_t(h_{t-1})} g(h_{t-1}, x_t) f_t(dx_t | h_{t-1}) = \int_{A_t(h_{t-1})} p_t(h_{t-1}, x_t) f_t(dx_t | h_{t-1}).$$

Thus,  $B_{ti}$  is a measurable correspondence from  $\tilde{H}_{t-1}$  to  $A_{ti}(h_{t-1})$ . Let  $B_{ti}^c(h_{t-1}) = A_{ti}(h_{t-1}) \setminus B_{ti}(h_{t-1})$  for each  $h_{t-1} \in H_{t-1}$ . Then  $B_{ti}^c$  is a measurable and closed-valued correspondence, which has a Borel measurable graph by Lemma 1. As a result,  $B_{ti}$  also has a Borel measurable graph. As  $f_t(h_{t-1})$  is a Nash equilibrium in the subgame  $h_{t-1} \in \tilde{H}_{t-1}$  with payoff  $p_t(h_{t-1}, \cdot)$ ,  $f_{ti}(B_{ti}(h_{t-1}) | h_{t-1}) = 0$ .

Denote  $\beta_i(h_{t-1}, x_t) = \min P_{ti}(h_{t-1}, x_t)$ , where  $P_{ti}(h_{t-1}, x_t)$  is the projection of  $P_t(h_{t-1}, x_t)$  on the  $i$ th dimension. Then the correspondence  $P_{ti}$  is measurable and compact-valued, and  $\beta_i$  is Borel measurable. Let  $\Lambda_i(h_{t-1}, x_t) = \{\beta_i(h_{t-1}, x_t)\} \times [0, \gamma]^{n-1}$ , where  $\gamma > 0$  is the upper bound of  $P_t$ . Denote  $\Lambda'_i(h_{t-1}, x_t) = \Lambda_i(h_{t-1}, x_t) \cap P_t(h_{t-1}, x_t)$ . Then  $\Lambda'_i$  is a measurable and compact-valued correspondence, and hence it has a Borel measurable selection  $\beta'_i$ . Note that  $\beta'_i$  is a Borel measurable selection of  $P_t$ . Let

$$g_t(h_{t-1}, x_t) = \begin{cases} \beta'_i(h_{t-1}, x_t) & \text{if } h_{t-1} \in \tilde{H}_{t-1}, x_{ti} \in B_{ti}(h_{t-1}) \text{ and } x_{tj} \notin B_{tj}(h_{t-1}), \forall j \neq i, \\ g(h_{t-1}, x_t) & \text{otherwise.} \end{cases}$$

Notice that

$$\{(h_{t-1}, x_t) \in \text{Gr}(A_t) : h_{t-1} \in \tilde{H}_{t-1}, x_{ti} \in B_{ti}(h_{t-1}) \text{ and } x_{tj} \notin B_{tj}(h_{t-1}), \forall j \neq i; \} \\ = \text{Gr}(A_t) \cap \bigcup_{i \in I} \left( \left( \text{Gr}(B_{ti}) \times \prod_{j \neq i} X_{tj} \right) \setminus \left( \bigcup_{j \neq i} \left( \text{Gr}(B_{tj}) \times \prod_{k \neq j} X_{tk} \right) \right) \right),$$

which is a Borel set. As a result,  $g_t$  is a Borel measurable selection of  $P_t$ . Moreover,  $g_t(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$  for all  $h_{t-1} \in \tilde{H}_{t-1}$  and  $f_t(\cdot | h_{t-1})$ -almost all  $x_t$ .

Fix a subgame  $h_{t-1} \in \tilde{H}_{t-1}$ . We show that  $f_t(\cdot | h_{t-1})$  is a Nash equilibrium given the payoff  $g_t(h_{t-1}, \cdot)$  in the subgame  $h_{t-1}$ . Suppose that player  $i$  deviates to some action  $\tilde{x}_{ti}$ .

If  $\tilde{x}_{ti} \in B_{ti}(h_{t-1})$ , then player  $i$ 's expected payoff is

$$\int_{A_{t(-i)}(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1})$$

$$\begin{aligned}
&= \int_{\prod_{j \neq i} B_{ij}^c(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&= \int_{\prod_{j \neq i} B_{ij}^c(h_{t-1})} \beta_i(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&\leq \int_{\prod_{j \neq i} B_{ij}^c(h_{t-1})} p_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&= \int_{A_{t(-i)}(h_{t-1})} p_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&\leq \int_{A_t(h_{t-1})} p_{ti}(h_{t-1}, x_t) f_t(dx_t | h_{t-1}) \\
&= \int_{A_t(h_{t-1})} g_{ti}(h_{t-1}, x_t) f_t(dx_t | h_{t-1}).
\end{aligned}$$

The first and the third equalities hold since  $f_{ij}(B_{ij}(h_{t-1}) | h_{t-1}) = 0$  for each  $j$ , and hence,  $f_{t(-i)}(\prod_{j \neq i} B_{ij}^c(h_{t-1}) | h_{t-1}) = f_{t(-i)}(A_{t(-i)}(h_{t-1}) | h_{t-1})$ . The second equality and the first inequality are due to the fact that  $g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) = \beta_i(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) = \min P_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) \leq p_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)})$  for  $x_{t(-i)} \in \prod_{j \neq i} B_{ij}^c(h_{t-1})$ . The second inequality holds since  $f_t(\cdot | h_{t-1})$  is a Nash equilibrium given the payoff  $p_t(h_{t-1}, \cdot)$  in the subgame  $h_{t-1}$ . The fourth equality follows from the fact that  $g_t(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$  for  $f_t(\cdot | h_{t-1})$ -almost all  $x_t$ .

If  $\tilde{x}_{ti} \notin B_{ti}(h_{t-1})$ , then player  $i$ 's expected payoff is

$$\begin{aligned}
&\int_{A_{t(-i)}(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&= \int_{\prod_{j \neq i} B_{ij}^c(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&= \int_{\prod_{j \neq i} B_{ij}^c(h_{t-1})} g_i(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&= \int_{A_{t(-i)}(h_{t-1})} g_i(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)} | h_{t-1}) \\
&\leq \int_{A_t(h_{t-1})} p_{ti}(h_{t-1}, x_t) f_t(dx_t | h_{t-1}) \\
&= \int_{A_t(h_{t-1})} g_{ti}(h_{t-1}, x_t) f_t(dx_t | h_{t-1}).
\end{aligned}$$

The first and the third equalities hold since

$$f_{t(-i)}\left(\prod_{j \neq i} B_{ij}^c(h_{t-1}) | h_{t-1}\right) = f_{t(-i)}(A_{t(-i)}(h_{t-1}) | h_{t-1}).$$

The second equality is due to the fact that  $g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) = g_i(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)})$  for  $x_{t(-i)} \in \prod_{j \neq i} B_{ij}^c(h_{t-1})$ . The first inequality follows from the definition of  $B_{ti}$ , and the fourth equality holds since  $g_t(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$  for  $f_t(\cdot|h_{t-1})$ -almost all  $x_t$ .

Thus, player  $i$  cannot improve his payoff in the subgame  $h_t$  by a unilateral change in his strategy for any  $i \in I$ , which implies that  $f_t(\cdot|h_{t-1})$  is a Nash equilibrium given the payoff  $g_t(h_{t-1}, \cdot)$  in the subgame  $h_{t-1}$ .

Step 3. For any  $(h_{t-1}, x_t) \in \text{Gr}(A_t)$ ,

$$P_t(h_{t-1}, x_t) = \int_{S_t} Q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}).$$

By Lemma 5, there exists a Borel measurable mapping  $q$  from  $\text{Gr}(P_t) \times S_t$  to  $\mathbb{R}^n$  such that the following relationships hold:

- (i) We have  $q(h_{t-1}, x_t, e, s_t) \in Q_{t+1}(h_{t-1}, x_t, s_t)$  for any  $(h_{t-1}, x_t, e, s_t) \in \text{Gr}(P_t) \times S_t$ .
- (ii) We have  $e = \int_{S_t} q(h_{t-1}, x_t, e, s_t) f_{t0}(ds_t|h_{t-1})$  for any  $(h_{t-1}, x_t, e) \in \text{Gr}(P_t)$ , where  $(h_{t-1}, x_t) \in \text{Gr}(A_t)$ .

Let

$$q_{t+1}(h_{t-1}, x_t, s_t) = q(h_{t-1}, x_t, g_t(h_{t-1}, x_t), s_t)$$

for any  $(h_{t-1}, x_t, s_t) \in H_t$ . Then  $q_{t+1}$  is a Borel measurable selection of  $Q_{t+1}$ .

For  $(h_{t-1}, x_t) \in \text{Gr}(A_t)$ ,

$$\begin{aligned} g_t(h_{t-1}, x_t) &= \int_{S_t} q(h_{t-1}, x_t, g_t(h_{t-1}, x_t), s_t) f_{t0}(ds_t|h_{t-1}) \\ &= \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}). \end{aligned}$$

Therefore, we have a Borel measurable selection  $q_{t+1}$  of  $Q_{t+1}$  and a Borel-measurable mapping  $f_t: H_{t-1} \rightarrow \bigotimes_{i \in I} \mathcal{M}(X_{ti})$  such that for all  $h_{t-1} \in \tilde{H}_{t-1}$ , properties (i)–(iii) are satisfied. The proof is complete.  $\square$

If a dynamic game has only  $T$  stages for some positive integer  $T \geq 1$ , then let  $Q_{T+1}(h_T) = \{u(h_T)\}$  for any  $h_T \in H_T$  and  $Q_t = \Phi(Q_{t+1})$  for  $1 \leq t \leq T-1$ . We can start with the backward induction from the last period and stop at the initial period, and then run the forward induction from the initial period to the last period. Thus, the following result is immediate.

**PROPOSITION B.4.** *Any finite-horizon dynamic game with the ARM condition has a subgame-perfect equilibrium.*

**B.4.3 Infinite-horizon case** Pick a sequence  $\xi = (\xi_1, \xi_2, \dots)$  such that (a)  $\xi_m$  is a transition probability from  $H_{m-1}$  to  $\mathcal{M}(X_m)$  for any  $m \geq 1$ , and (b)  $\xi_m(A_m(h_{m-1})|h_{m-1}) = 1$  for any  $m \geq 1$  and  $h_{m-1} \in H_{m-1}$ . Denote the set of all such  $\xi$  as  $Y$ .

Fix any  $t \geq 1$ , and define correspondences  $\Xi_t^t$  and  $\Delta_t^t$  as follows: in the subgame  $h_{t-1}$ ,

$$\Xi_t^t(h_{t-1}) = \mathcal{M}(A_t(h_{t-1})) \otimes \lambda_t$$

and

$$\Delta_t^t(h_{t-1}) = \mathcal{M}(A_t(h_{t-1})) \otimes f_{t0}(h_{t-1}).$$

For any  $m_1 > t$ , suppose that the correspondences  $\Xi_t^{m_1-1}$  and  $\Delta_t^{m_1-1}$  have been defined. Then we can define correspondences  $\Xi_t^{m_1}: H_{t-1} \rightarrow \mathcal{M}(\prod_{t \leq m \leq m_1} (X_m \times S_m))$  and  $\Delta_t^{m_1}: H_{t-1} \rightarrow \mathcal{M}(\prod_{t \leq m \leq m_1} (X_m \times S_m))$  as

$$\begin{aligned} \Xi_t^{m_1}(h_{t-1}) &= \{g(h_{t-1}) \diamond (\xi_{m_1}(h_{t-1}, \cdot) \otimes \lambda_{m_1}) : \\ &\quad g \text{ is a Borel measurable selection of } \Xi_t^{m_1-1}, \\ &\quad \xi_{m_1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_1})\} \end{aligned}$$

and

$$\begin{aligned} \Delta_t^{m_1}(h_{t-1}) &= \{g(h_{t-1}) \diamond (\xi_{m_1}(h_{t-1}, \cdot) \otimes f_{m_1 0}(h_{t-1}, \cdot)) : \\ &\quad g \text{ is a Borel measurable selection of } \Delta_t^{m_1-1}, \\ &\quad \xi_{m_1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_1})\}, \end{aligned}$$

where  $\mathcal{M}(A_{m_1})$  is regarded as a correspondence from  $H_{m_1-1}$  to the space of Borel probability measures on  $X_{m_1}$ . For any  $m_1 \geq t$ , let  $\rho_{(h_{t-1}, \xi)}^{m_1} \in \Xi_t^{m_1}$  be the probability measure on  $\prod_{t \leq m \leq m_1} (X_m \times S_m)$  induced by  $\{\lambda_m\}_{t \leq m \leq m_1}$  and  $\{\xi_m\}_{t \leq m \leq m_1}$ , and  $\varrho_{(h_{t-1}, \xi)}^{m_1} \in \Delta_t^{m_1}$  be the probability measure on  $\prod_{t \leq m \leq m_1} (X_m \times S_m)$  induced by  $\{f_{m0}\}_{t \leq m \leq m_1}$  and  $\{\xi_m\}_{t \leq m \leq m_1}$ . Then  $\Xi_t^{m_1}(h_{t-1})$  is the set of all such  $\rho_{(h_{t-1}, \xi)}^{m_1}$ , while  $\Delta_t^{m_1}(h_{t-1})$  is the set of all such  $\varrho_{(h_{t-1}, \xi)}^{m_1}$ . Note that  $\varrho_{(h_{t-1}, \xi)}^{m_1} \in \Delta_t^{m_1}(h_{t-1})$  if and only if  $\rho_{(h_{t-1}, \xi)}^{m_1} \in \Xi_t^{m_1}(h_{t-1})$ . Both  $\varrho_{(h_{t-1}, \xi)}^{m_1}$  and  $\rho_{(h_{t-1}, \xi)}^{m_1}$  can be regarded as probability measures on  $H_{m_1}(h_{t-1})$ .

Similarly, let  $\rho_{(h_{t-1}, \xi)}$  be the probability measure on  $\prod_{m \geq t} (X_m \times S_m)$  induced by  $\{\lambda_m\}_{m \geq t}$  and  $\{\xi_m\}_{m \geq t}$ , and  $\varrho_{(h_{t-1}, \xi)}$  the probability measure on  $\prod_{m \geq t} (X_m \times S_m)$  induced by  $\{f_{m0}\}_{m \geq t}$  and  $\{\xi_m\}_{m \geq t}$ . Denote the correspondence

$$\Xi_t: H_{t-1} \rightarrow \mathcal{M}\left(\prod_{m \geq t} (X_m \times S_m)\right)$$

as the set of all such  $\rho_{(h_{t-1}, \xi)}$  and

$$\Delta_t: H_{t-1} \rightarrow \mathcal{M}\left(\prod_{m \geq t} (X_m \times S_m)\right)$$

as the set of all such  $\varrho_{(h_{t-1}, \xi)}$ .

The following lemma demonstrates the relationship between  $\varrho_{(h_{t-1}, \xi)}^{m_1}$  and  $\rho_{(h_{t-1}, \xi)}^{m_1}$ .

LEMMA B.7. *For any  $m_1 \geq t$  and  $h_{t-1} \in H_{t-1}$ ,*

$$\varrho_{(h_{t-1}, \xi)}^{m_1} = \left( \prod_{t \leq m \leq m_1} \varphi_{m0}(h_{t-1}, \cdot) \right) \circ \rho_{(h_{t-1}, \xi)}^{m_1}.^4$$

PROOF. Fix  $\xi \in Y$ , and Borel subsets  $C_m \subseteq X_m$  and  $D_m \subseteq S_m$  for  $m \geq t$ . First, we have

$$\begin{aligned} & \varrho_{(h_{t-1}, \xi)}^t(C_t \times D_t) \\ &= \xi_t(C_t | h_{t-1}) \cdot f_{t0}(D_t | h_{t-1}) \\ &= \int_{X_t \times S_t} \mathbf{1}_{C_t \times D_t}(x_t, s_t) \varphi_{t0}(h_{t-1}, s_t) (\xi_t(h_{t-1}) \otimes \lambda_t)(d(x_t, s_t)), \end{aligned}$$

which implies that  $\varrho_{(h_{t-1}, \xi)}^t = \varphi_{t0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^t$ .<sup>5</sup>

Suppose that  $\varrho_{(h_{t-1}, \xi)}^{m_2} = (\prod_{t \leq m \leq m_2} \varphi_{m0}(h_{t-1}, \cdot)) \circ \rho_{(h_{t-1}, \xi)}^{m_2}$  for some  $m_2 \geq t$ . Then

$$\begin{aligned} & \varrho_{(h_{t-1}, \xi)}^{m_2+1} \left( \prod_{t \leq m \leq m_2+1} (C_m \times D_m) \right) \\ &= \varrho_{(h_{t-1}, \xi)}^{m_2} \diamond (\xi_{m_2+1}(h_{t-1}, \cdot) \otimes f_{(m_2+1)0}(h_{t-1}, \cdot)) \left( \prod_{t \leq m \leq m_2+1} (C_m \times D_m) \right) \\ &= \int_{\prod_{t \leq m \leq m_2} (X_m \times S_m)} \int_{X_{m_2+1} \times S_{m_2+1}} \mathbf{1}_{\prod_{t \leq m \leq m_2+1} (C_m \times D_m)}(x_t, \dots, x_{m_2+1}, s_t, \dots, s_{m_2+1}) \cdot \\ & \quad \xi_{m_2+1} \otimes f_{(m_2+1)0}(d(x_{m_2+1}, s_{m_2+1}) | h_{t-1}, x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2}) \\ & \quad \varrho_{(h_{t-1}, \xi)}^{m_2}(d(x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2}) | h_{t-1}) \\ &= \int_{\prod_{t \leq m \leq m_2} (X_m \times S_m)} \int_{S_{m_2+1}} \int_{X_{m_2+1}} \mathbf{1}_{\prod_{t \leq m \leq m_2+1} (C_m \times D_m)}(x_t, \dots, x_{m_2+1}, s_t, \dots, s_{m_2+1}) \cdot \\ & \quad \times \varphi_{(m_2+1)0}(h_{t-1}, x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2+1}) \\ & \quad \xi_{m_2+1}(dx_{m_2+1} | h_{t-1}, x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2}) \\ & \quad \lambda_{(m_2+1)0}(ds_{m_2+1}) \prod_{t \leq m \leq m_2} \varphi_{m0}(h_{t-1}, x_t, \dots, x_{m-1}, s_t, \dots, s_m) \\ & \quad \rho_{(h_{t-1}, \xi)}^{m_2}(d(x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2}) | h_{t-1}) \end{aligned}$$

<sup>4</sup>For  $m \geq t \geq 1$  and  $h_{t-1} \in H_{t-1}$ , the function  $\varphi_{m0}(h_{t-1}, \cdot)$  is defined on  $H_{m-1}(h_{t-1}) \times S_m$ , which is measurable and sectionally continuous on  $\prod_{t \leq k \leq m-1} X_k$ . By Lemma 3,  $\varphi_{m0}(h_{t-1}, \cdot)$  can be extended to be a measurable function  $\hat{\varphi}_{m0}(h_{t-1}, \cdot)$  on the product space  $\left( \prod_{t \leq k \leq m-1} X_k \right) \times \left( \prod_{t \leq k \leq m} S_k \right)$ , which is also sectionally continuous on  $\prod_{t \leq k \leq m-1} X_k$ . Given any  $\xi \in Y$ , since  $\rho_{(h_{t-1}, \xi)}^m$  concentrates on  $H_m(h_{t-1})$ ,  $\varphi_{m0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^m = \hat{\varphi}_{m0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^m$ . For notational simplicity, we still use  $\varphi_{m0}(h_{t-1}, \cdot)$ , instead of  $\hat{\varphi}_{m0}(h_{t-1}, \cdot)$ , to denote the above extension. Similarly, we can work with a suitable extension of the payoff function  $u$  as needed.

<sup>5</sup>For a set  $A$  in a space  $X$ ,  $\mathbf{1}_A$  is the indicator function of  $A$ , which is 1 on  $A$  and is 0 on  $X \setminus A$ .

$$\begin{aligned}
&= \int_{\prod_{t \leq m \leq m_2+1} (X_m \times S_m)} \mathbf{1}_{\prod_{t \leq m \leq m_2+1} (C_m \times D_m)}(x_t, \dots, x_{m_2+1}, s_t, \dots, s_{m_2+1}) \cdot \\
&\quad \times \prod_{t \leq m \leq m_2+1} \varphi_{m0}(h_{t-1}, x_t, \dots, x_{m-1}, s_t, \dots, s_m) \\
&\quad \rho_{(h_{t-1}, \xi)}^{m_2+1}(d(x_t, \dots, x_{m_2}, s_t, \dots, s_{m_2}) | h_{t-1}),
\end{aligned}$$

which implies that

$$\varrho_{(h_{t-1}, \xi)}^{m_2+1} = \left( \prod_{t \leq m \leq m_2+1} \varphi_{m0}(h_{t-1}, \cdot) \right) \circ \rho_{(h_{t-1}, \xi)}^{m_2+1}.$$

The proof is thus complete.  $\square$

The next lemma shows that the correspondences  $\Delta_t^{m_1}$  and  $\Delta_t$  are nonempty and compact valued, and sectionally continuous.

LEMMA B.8. (i) For any  $t \geq 1$ , the correspondence  $\Delta_t^{m_1}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$  for any  $m_1 \geq t$ .

(ii) For any  $t \geq 1$ , the correspondence  $\Delta_t$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ .

PROOF. (i) We first show that the correspondence  $\Xi_t^{m_1}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$  for any  $m_1 \geq t$ .

Consider the case  $m_1 = t \geq 1$ , where

$$\Xi_t^t(h_{t-1}) = \mathcal{M}(A_t(h_{t-1})) \otimes \lambda_t.$$

Since  $A_{ti}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ ,  $\Xi_t^t$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ .

Now suppose that  $\Xi_t^{m_2}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$  for some  $m_2 \geq t \geq 1$ . Notice that

$$\begin{aligned}
\Xi_t^{m_2+1}(h_{t-1}) &= \{g(h_{t-1}) \diamond (\xi_{m_2+1}(h_{t-1}, \cdot) \otimes \lambda_{(m_2+1)}) : \\
&\quad g \text{ is a Borel measurable selection of } \Xi_t^{m_2}, \\
&\quad \xi_{m_2+1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_2+1})\}.
\end{aligned}$$

First, we claim that  $H_t(s_0, s_1, \dots, s_t)$  is compact for any  $(s_0, s_1, \dots, s_t) \in S^t$ . We prove this claim by induction.

- (a) Notice that  $H_0(s_0) = X_0$  for any  $s_0 \in S_0$ , which is compact.
- (b) Suppose that  $H_{m'}(s_0, s_1, \dots, s_{m'})$  is compact for some  $0 \leq m' \leq t-1$  and any  $(s_0, s_1, \dots, s_{m'}) \in S^{m'}$ .

- (c) Since  $A_{m'+1}(\cdot, s_0, s_1, \dots, s_{m'})$  is continuous and compact-valued, it has a compact graph by Lemma 2(vi), which is  $H_{m'+1}(s_0, s_1, \dots, s_{m'+1})$  for any  $(s_0, s_1, \dots, s_{m'+1}) \in S^{m'+1}$ .

Thus, we prove the claim.

Define a correspondence  $A_t^t$  from  $H_{t-1} \times S_t$  to  $X_t$  as  $A_t^t(h_{t-1}, s_t) = A_t(h_{t-1})$ . Then  $A_t^t$  is nonempty and compact valued, sectionally continuous on  $X^{t-1}$ , and has a  $\mathcal{B}(X^t \times S^t)$ -measurable graph. Since the graph of  $A_t^t(\cdot, s_0, s_1, \dots, s_t)$  is  $H_t(s_0, s_1, \dots, s_t)$  and since  $H_t(s_0, s_1, \dots, s_t)$  is compact,  $A_t^t(\cdot, s_0, s_1, \dots, s_t)$  has a compact graph. For any  $h_{t-1} \in H_{t-1}$  and  $\tau \in \Xi_t^t(h_{t-1})$ , the marginal of  $\tau$  on  $S_t$  is  $\lambda_t$  and  $\tau(\text{Gr}(A_t^t(h_{t-1}, \cdot))) = 1$ .

For any  $m_1 > t$ , suppose that the correspondence

$$A_t^{m_1-1}: H_{t-1} \times \prod_{t \leq m \leq m_1-1} S_m \rightarrow \prod_{t \leq m \leq m_1-1} X_m$$

has been defined such that the following statements hold:

- It is nonempty and compact valued, sectionally upper hemicontinuous on  $X^{t-1}$ , and has a  $\mathcal{B}(X^{m_1-1} \times S^{m_1-1})$ -measurable graph.
- For any  $(s_0, s_1, \dots, s_{m_1-1})$ ,  $A_t^{m_1-1}(\cdot, s_0, s_1, \dots, s_{m_1-1})$  has a compact graph.
- For any  $h_{t-1} \in H_{t-1}$  and  $\tau \in \Xi_t^{m_1-1}(h_{t-1})$ , the marginal of  $\tau$  on  $\prod_{t \leq m \leq m_1-1} S_m$  is  $\bigotimes_{t \leq m \leq m_1-1} \lambda_m$  and  $\tau(\text{Gr}(A_t^{m_1-1}(h_{t-1}, \cdot))) = 1$ .

We define a correspondence  $A_t^{m_1}: H_{t-1} \times \prod_{t \leq m \leq m_1} S_m \rightarrow \prod_{t \leq m \leq m_1} X_m$  as

$$\begin{aligned} A_t^{m_1}(h_{t-1}, s_t, \dots, s_{m_1}) &= \{(x_t, \dots, x_{m_1}) : \\ &\quad x_{m_1} \in A_{m_1}(h_{t-1}, x_t, \dots, x_{m_1-1}, s_t, \dots, s_{m_1-1}), \\ &\quad (x_t, \dots, x_{m_1-1}) \in A_t^{m_1-1}(h_{t-1}, s_t, \dots, s_{m_1-1})\}. \end{aligned}$$

It is obvious that  $A_t^{m_1}$  is nonempty-valued. For any  $(s_0, s_1, \dots, s_{m_1})$ , since  $A_t^{m_1-1}(\cdot, s_0, s_1, \dots, s_{m_1-1})$  has a compact graph, and since  $A_{m_1}(\cdot, s_0, s_1, \dots, s_{m_1-1})$  is continuous and compact-valued,  $A_t^{m_1}(\cdot, s_0, s_1, \dots, s_{m_1})$  has a compact graph by Lemma 2(vi), which implies that  $A_t^{m_1}$  is compact-valued and sectionally upper hemicontinuous on  $X^{t-1}$ . In addition,  $\text{Gr}(A_t^{m_1}) = \text{Gr}(A_{m_1}) \times S_{m_1}$ , which is  $\mathcal{B}(X^{m_1} \times S^{m_1})$ -measurable. For any  $h_{t-1} \in H_{t-1}$  and  $\tau \in \Xi_t^{m_1}(h_{t-1})$ , it is obvious that the marginal of  $\tau$  on  $\prod_{t \leq m \leq m_1} S_m$  is  $\bigotimes_{t \leq m \leq m_1} \lambda_m$  and  $\tau(\text{Gr}(A_t^{m_1}(h_{t-1}, \cdot))) = 1$ .

By Lemma B.5,  $\Xi_t^{m_2+1}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ .

Now we show that the correspondence  $\Delta_t^{m_1}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$  for any  $m_1 \geq t$ .

Given  $s^{t-1}$  and a sequence  $\{x_0^k, x_1^k, \dots, x_{t-1}^k\} \in H_{t-1}(s^{t-1})$  for  $1 \leq k \leq \infty$ . Let  $h_{t-1}^k = (s^{t-1}, (x_0^k, x_1^k, \dots, x_{t-1}^k))$ . It is obvious that  $\Delta_t^{m_1}$  is nonempty-valued. We first show that  $\Delta_t^{m_1}$  is sectionally upper hemicontinuous on  $X^{t-1}$ . Suppose that  $\varrho_{(h_{t-1}^k, \xi^k)}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^k)$

for  $1 \leq k < \infty$  and  $(x_0^k, x_1^k, \dots, x_{t-1}^k) \rightarrow (x_0^\infty, x_1^\infty, \dots, x_{t-1}^\infty)$ . We need to show that there exists some  $\xi^\infty$  such that a subsequence of  $\varrho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$  and  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^\infty)$ .

Since  $\Xi_t^{m_1}$  is sectionally upper hemicontinuous on  $X^{t-1}$ , there exists some  $\xi^\infty$  such that a subsequence of  $\rho_{(h_{t-1}^k, \xi^k)}^{m_1}$ , say itself, weakly converges to  $\rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$  and  $\rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Xi_t^{m_1}(h_{t-1}^\infty)$ . Then  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^\infty)$ .

For any bounded continuous function  $\psi$  on  $\prod_{t \leq m \leq m_1} (X_m \times S_m)$ , let

$$\begin{aligned} & \chi_k(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \\ &= \psi(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \cdot \prod_{t \leq m \leq m_1} \varphi_{m0}(h_{t-1}^k, x_t, \dots, x_{m-1}, s_t, \dots, s_m). \end{aligned}$$

Then  $\{\chi_k\}$  is a sequence of functions satisfying the following three properties.

- (i) For each  $k$ ,  $\chi_k$  is jointly measurable and sectionally continuous on  $\prod_{t \leq m \leq m_1} X_m$ .
- (ii) For any  $(s_t, \dots, s_{m_1})$  and any sequence  $(x_t^k, \dots, x_{m_1}^k) \rightarrow (x_t^\infty, \dots, x_{m_1}^\infty)$  in  $\prod_{t \leq m \leq m_1} X_m$ ,  $\chi_k(x_t^k, \dots, x_{m_1}^k, s_t, \dots, s_{m_1}) \rightarrow \chi_\infty(x_t^\infty, \dots, x_{m_1}^\infty, s_t, \dots, s_{m_1})$  as  $k \rightarrow \infty$ .
- (iii) The sequence  $\{\chi_k\}_{1 \leq k \leq \infty}$  is integrably bounded in the sense that there exists a function  $\chi': \prod_{t \leq m \leq m_1} S_m \rightarrow \mathbb{R}_+$  such that  $\chi'$  is  $\otimes_{t \leq m \leq m_1} \lambda_m$ -integrable, and for any  $k$  and  $(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1})$ ,  $\chi_k(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \leq \chi'(s_t, \dots, s_{m_1})$ .

By [Lemma B.6](#), as  $k \rightarrow \infty$ ,

$$\begin{aligned} & \int_{\prod_{t \leq m \leq m_1} (X_m \times S_m)} \chi_k(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \rho_{(h_{t-1}^k, \xi^k)}^{m_1}(d(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1})) \\ & \rightarrow \int_{\prod_{t \leq m \leq m_1} (X_m \times S_m)} \chi_\infty(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}(d(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1})). \end{aligned}$$

Then by [Lemma B.7](#),

$$\begin{aligned} & \int_{\prod_{t \leq m \leq m_1} (X_m \times S_m)} \psi(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \varrho_{(h_{t-1}^k, \xi^k)}^{m_1}(d(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1})) \\ & \rightarrow \int_{\prod_{t \leq m \leq m_1} (X_m \times S_m)} \psi(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1}) \varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}(d(x_t, \dots, x_{m_1}, s_t, \dots, s_{m_1})), \end{aligned}$$

which implies that  $\varrho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ . Therefore,  $\Delta_t^{m_1}$  is sectionally upper hemicontinuous on  $X^{t-1}$ . If one chooses  $h_{t-1}^1 = h_{t-1}^2 = \dots = h_{t-1}^\infty$ , then we indeed show that  $\Delta_t^{m_1}$  is compact-valued.

In the argument above, we indeed proved that if  $\rho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ , then  $\varrho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ .

All that is left is to show that  $\Delta_t^{m_1}$  is sectionally lower hemicontinuous on  $X^{t-1}$ . Suppose that  $(x_0^k, x_1^k, \dots, x_{t-1}^k) \rightarrow (x_0^\infty, x_1^\infty, \dots, x_{t-1}^\infty)$  and  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^\infty)$ , we need to show that there exists a subsequence  $\{(x_0^{k_m}, x_1^{k_m}, \dots, x_{t-1}^{k_m})\}$  of  $\{(x_0^k, x_1^k, \dots, x_{t-1}^k)\}$  and  $\varrho_{(h_{t-1}^{k_m}, \xi^{k_m})}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^{k_m})$  for each  $k_m$  such that  $\varrho_{(h_{t-1}^{k_m}, \xi^{k_m})}^{m_1}$  weakly converges to  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ .

Since  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Delta_t^{m_1}(h_{t-1}^\infty)$ , we have  $\rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1} \in \Xi_t^{m_1}(h_{t-1}^\infty)$ . Because  $\Xi_t^{m_1}$  is sectionally lower hemicontinuous on  $X^{t-1}$ , there exists a subsequence of  $\{(x_0^k, x_1^k, \dots, x_{t-1}^k)\}$ , say itself, and  $\rho_{(h_{t-1}^k, \xi^k)}^{m_1} \in \Xi_t^{m_1}(h_{t-1}^k)$  for each  $k$  such that  $\rho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\rho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ . As a result,  $\varrho_{(h_{t-1}^k, \xi^k)}^{m_1}$  weakly converges to  $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$ , which implies that  $\Delta_t^{m_1}$  is sectionally lower hemicontinuous on  $X^{t-1}$ .

Therefore,  $\Delta_t^{m_1}$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$  for any  $m_1 \geq t$ .

(ii) We show that  $\Delta_t$  is nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ .

It is obvious that  $\Delta_t$  is nonempty-valued, so we first prove that it is compact-valued.

Given  $h_{t-1}$  and a sequence  $\{\tau^k\} \subseteq \Delta_t(h_{t-1})$ , there exists a sequence of  $\{\xi^k\}_{k \geq 1}$  such that  $\xi^k = (\xi_1^k, \xi_2^k, \dots) \in Y$  and  $\tau^k = \varrho_{(h_{t-1}, \xi^k)}$  for each  $k$ .

By (i),  $\Xi_t^1$  is compact. Then there exists a measurable mapping  $g_t$  such that (a)  $g^t = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, \xi_{t+1}^1, \dots) \in Y$  and (b) a subsequence of  $\{\rho_{(h_{t-1}, \xi^k)}^t\}$ , say  $\{\rho_{(h_{t-1}, \xi^{k_{1l}})}^t\}_{l \geq 1}$ , that weakly converges to  $\rho_{(h_{t-1}, g^t)}^t$ . Note that  $\{\xi_{t+1}^k\}$  is a Borel measurable selection of  $\mathcal{M}(A_{t+1})$ . By Lemma B.5, there is a Borel measurable selection  $g_{t+1}$  of  $\mathcal{M}(A_{t+1})$  such that there is a subsequence of  $\{\rho_{(h_{t-1}, \xi^{k_{1l}})}^{t+1}\}_{l \geq 1}$ , say  $\{\rho_{(h_{t-1}, \xi^{k_{2l}})}^{t+1}\}_{l \geq 1}$ , that weakly converges to  $\rho_{(h_{t-1}, g^{t+1})}^{t+1}$ , where  $g^{t+1} = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, g_{t+1}, \xi_{t+2}^1, \dots) \in Y$ .

By repeating this procedure, one can construct a Borel measurable mapping  $g$  such that  $\rho_{(h_{t-1}, \xi^{k_{11}})}, \rho_{(h_{t-1}, \xi^{k_{22}})}, \rho_{(h_{t-1}, \xi^{k_{33}})}, \dots$  weakly converges to  $\rho_{(h_{t-1}, g)}$ . That is,  $\rho_{(h_{t-1}, g)}$  is a convergent point of  $\{\rho_{(h_{t-1}, \xi^k)}\}$ , which implies that  $\varrho_{(h_{t-1}, g)}$  is a convergent point of  $\{\varrho_{(h_{t-1}, \xi^k)}\}$ .

The sectional upper hemicontinuity of  $\Delta_t$  follows a similar argument as above. In particular, given  $s^{t-1}$  and a sequence  $\{x_0^k, x_1^k, \dots, x_{t-1}^k\} \subseteq H_{t-1}(s^{t-1})$  for  $k \geq 0$ . Let  $h_{t-1}^k = (s^{t-1}, (x_0^k, x_1^k, \dots, x_{t-1}^k))$ . Suppose that  $(x_0^k, x_1^k, \dots, x_{t-1}^k) \rightarrow (x_0^0, x_1^0, \dots, x_{t-1}^0)$ . If  $\{\tau^k\} \subseteq \Delta_t(h_{t-1}^k)$  for  $k \geq 1$  and  $\tau^k \rightarrow \tau^0$ , then one can show that  $\tau^0 \in \Delta_t(h_{t-1}^0)$  by repeating a similar argument as in the proof above.

Finally, we consider the sectional lower hemicontinuity of  $\Delta_t$ . Suppose that  $\tau^0 \in \Delta_t(h_{t-1}^0)$ . Then there exists some  $\xi \in Y$  such that  $\tau^0 = \varrho_{(h_{t-1}^0, \xi)}$ . Denote  $\tilde{\tau}^m = \varrho_{(h_{t-1}^0, \xi)}^m \in \Delta_t^m(h_{t-1}^0)$  for  $m \geq t$ . As  $\Delta_t^m$  is sectionally continuous, for each  $m$ , there exists some  $\xi^m \in Y$  such that  $d(\varrho_{(h_{t-1}^m, \xi^m)}^m, \tilde{\tau}^m) \leq \frac{1}{m}$  for  $k_m$  sufficiently large, where  $d$  is the Prokhorov metric. Let  $\tau^m = \varrho_{(h_{t-1}^m, \xi^m)}$ . Then  $\tau^m$  weakly converges to  $\tau^0$ , which implies that  $\Delta_t$  is sectionally lower hemicontinuous.  $\square$

Define a correspondence  $Q_t^\tau: H_{t-1} \rightarrow \mathbb{R}_{++}^n$  as

$$Q_t^\tau(h_{t-1}) = \begin{cases} \left\{ \int_{\prod_{m \geq t}(X_m \times S_m)} u(h_{t-1}, x, s) \varrho_{(h_{t-1}, \xi)}(d(x, s)) : \varrho_{(h_{t-1}, \xi)} \in \Delta_t(h_{t-1}) \right\}, & t > \tau, \\ \Phi(Q_{t+1}^\tau)(h_{t-1}), & t \leq \tau. \end{cases}$$

The lemma below presents several properties of the correspondence  $Q_t^\tau$ .

**LEMMA B.9.** *For any  $t, \tau \geq 1$ ,  $Q_t^\tau$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^{t-1}$ .*

**PROOF.** We prove the lemma in three steps.

Step 1. Fix  $t > \tau$ . We show that  $Q_t^\tau$  is bounded, nonempty and compact valued, and sectionally upper hemicontinuous on  $X^{t-1}$ .

The boundedness and nonemptiness of  $Q_t^\tau$  are obvious. We prove that  $Q_t^\tau$  is sectionally upper hemicontinuous on  $X^{t-1}$ . Given  $s^{t-1}$  and a sequence  $\{x_0^k, x_1^k, \dots, x_{t-1}^k\} \subseteq H_{t-1}(s^{t-1})$  for  $k \geq 0$ , let  $h_{t-1}^k = (s^{t-1}, (x_0^k, x_1^k, \dots, x_{t-1}^k))$ . Suppose that  $a^k \in Q_t^\tau(h_{t-1}^k)$  for  $k \geq 1$ ,  $(x_0^k, x_1^k, \dots, x_{t-1}^k) \rightarrow (x_0^0, x_1^0, \dots, x_{t-1}^0)$ , and  $a^k \rightarrow a^0$ , we need to show that  $a^0 \in Q_t^\tau(h_{t-1}^0)$ .

By the definition, there exists a sequence  $\{\xi^k\}_{k \geq 1}$  such that

$$a^k = \int_{\prod_{m \geq t}(X_m \times S_m)} u(h_{t-1}^k, x, s) \varrho_{(h_{t-1}^k, \xi^k)}(d(x, s)),$$

where  $\xi^k = (\xi_1^k, \xi_2^k, \dots) \in Y$  for each  $k$ . As  $\Delta_t$  is compact-valued and sectionally continuous on  $X^{t-1}$ , there exist some  $\varrho_{(h_{t-1}^0, \xi^0)} \in \Delta_t(h_{t-1}^0)$  and a subsequence of  $\varrho_{(h_{t-1}^k, \xi^k)}$ , say itself, that weakly converges to  $\varrho_{(h_{t-1}^0, \xi^0)}$  for  $\xi^0 = (\xi_1^0, \xi_2^0, \dots) \in Y$ .

We show that

$$a^0 = \int_{\prod_{m \geq t}(X_m \times S_m)} u(h_{t-1}^0, x, s) \varrho_{(h_{t-1}^0, \xi^0)}(d(x, s)).$$

Toward this aim, we need to show only that for any  $\delta > 0$ ,

$$\left| a^0 - \int_{\prod_{m \geq t}(X_m \times S_m)} u(h_{t-1}^0, x, s) \varrho_{(h_{t-1}^0, \xi^0)}(d(x, s)) \right| < \delta. \quad (\text{B2})$$

Since the game is continuous at infinity, there exists a positive integer  $\tilde{M} \geq t$  such that  $w^m < \frac{1}{5}\delta$  for any  $m > \tilde{M}$ .

For each  $j > \tilde{M}$ , by Lemma 3, there exists a measurable selection  $\xi_j'$  of  $\mathcal{M}(A_j)$  such that  $\xi_j'$  is sectionally continuous on  $X^{j-1}$ . Let  $\mu: H_{\tilde{M}} \rightarrow \prod_{m > \tilde{M}}(X_m \times S_m)$  be the transition probability that is induced by  $(\xi_{\tilde{M}+1}', \xi_{\tilde{M}+2}', \dots)$  and  $\{f_{(\tilde{M}+1)0}, f_{(\tilde{M}+2)0}, \dots\}$ . By Lemma 9,  $\mu$  is measurable and sectionally continuous on  $X^{\tilde{M}}$ . Let

$$V_{\tilde{M}}(h_{t-1}, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})$$

$$\begin{aligned}
&= \int_{\prod_{m \geq \tilde{M}} (X_m \times S_m)} u(h_{t-1}, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}, x, s) \\
&\quad d\mu(x, s | h_{t-1}, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}).
\end{aligned}$$

Then  $V_{\tilde{M}}$  is bounded and measurable. In addition,  $V_{\tilde{M}}$  is sectionally continuous on  $X^{\tilde{M}}$  by [Lemma B.6](#).

For any  $k \geq 0$ , we have

$$\begin{aligned}
&\left| \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}^k, x, s) \varrho_{(h_{t-1}^k, \xi^k)}(d(x, s)) \right. \\
&\quad \left. - \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^k, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \right. \\
&\quad \left. \varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \right| \\
&\leq w^{\tilde{M}+1} \\
&< \frac{1}{5} \delta.
\end{aligned}$$

Since  $\varrho_{(h_{t-1}^k, \xi^k)}$  weakly converges to  $\varrho_{(h_{t-1}^0, \xi^0)}$  and  $\varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}$  is the marginal of  $\varrho_{(h_{t-1}^k, \xi^k)}$  on  $\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)$  for any  $k \geq 0$ , the sequence  $\varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}$  also weakly converges to  $\varrho_{(h_{t-1}^0, \xi^0)}^{\tilde{M}}$ . By [Lemma B.6](#), we have

$$\begin{aligned}
&\left| \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^k, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \right. \\
&\quad \left. - \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^0, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \right. \\
&\quad \left. \varrho_{(h_{t-1}^0, \xi^0)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \right| \\
&< \frac{1}{5} \delta
\end{aligned}$$

for  $k \geq K_1$ , where  $K_1$  is a sufficiently large positive integer. In addition, there exists a positive integer  $K_2$  such that  $|a^k - a^0| < \frac{1}{5} \delta$  for  $k \geq K_2$ .

Fix  $k > \max\{K_1, K_2\}$ . Combining the inequalities above, we have

$$\begin{aligned}
&\left| \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}^0, x, s) \varrho_{(h_{t-1}^0, \xi^0)}(d(x, s)) - a^0 \right| \\
&\leq \left| \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}^0, x, s) \varrho_{(h_{t-1}^0, \xi^0)}(d(x, s)) \right|
\end{aligned}$$

$$\begin{aligned}
& - \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^0, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \\
& \varrho_{(h_{t-1}^0, \xi^0)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \Big| \\
& + \left| \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^0, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \right. \\
& \varrho_{(h_{t-1}^0, \xi^0)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \\
& - \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^k, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \\
& \varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \Big| \\
& + \left| \int_{\prod_{t \leq m \leq \tilde{M}} (X_m \times S_m)} V_{\tilde{M}}(h_{t-1}^k, x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}}) \right. \\
& \varrho_{(h_{t-1}^k, \xi^k)}^{\tilde{M}}(d(x_t, \dots, x_{\tilde{M}}, s_t, \dots, s_{\tilde{M}})) \\
& - \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}^k, x, s) \varrho_{(h_{t-1}^k, \xi^k)}(d(x, s)) \Big| \\
& + \left| \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}^k, x, s) \varrho_{(h_{t-1}^k, \xi^k)}(d(x, s)) - a^0 \right| \\
& < \delta.
\end{aligned}$$

Thus, we proved inequality (B2), which implies that  $Q_t^\tau$  is sectionally upper hemicontinuous on  $X^{t-1}$  for  $t > \tau$ .

Furthermore, to prove that  $Q_t^\tau$  is compact-valued, we need to consider only the case that  $\{x_0^k, x_1^k, \dots, x_{t-1}^k\} = \{x_0^0, x_1^0, \dots, x_{t-1}^0\}$  for any  $k \geq 0$  and repeat the above proof.

Step 2. Fix  $t > \tau$ . We show that  $Q_t^\tau$  is measurable.

Fix a sequence  $(\xi'_1, \xi'_2, \dots)$ , where  $\xi'_j$  is a selection of  $\mathcal{M}(A_j)$  measurable in  $s^{j-1}$  and continuous in  $x^{j-1}$  for each  $j$ . For any  $M \geq t$ , let

$$\begin{aligned}
& W_M^M(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M) \\
& = \left\{ \int_{\prod_{m > M} (X_m \times S_m)} u(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M, x, s) \varrho_{(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M, \xi')} (d(x, s)) \right\}.
\end{aligned}$$

By Lemma 9,  $\varrho_{(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M, \xi')}$  is measurable from  $H_M$  to  $\mathcal{M}(\prod_{m > M} (X_m \times S_m))$  and is sectionally continuous on  $X^M$ . Thus,  $W_M^M$  is bounded, measurable, nonempty, convex and compact valued. By Lemma B.6,  $W_M^M$  is sectionally continuous on  $X^M$ .

Suppose that for some  $t \leq j \leq M$ ,  $W_M^j$  is defined such that it is bounded, measurable, nonempty, convex and compact valued, and sectionally continuous on  $X^j$ . Let

$$W_M^{j-1}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1})$$

$$\begin{aligned}
&= \left\{ \int_{X_j \times S_j} w_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j) \varrho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j(d(x_j, s_j)) : \right. \\
&\quad \varrho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j \in \Delta_j^j(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}), \\
&\quad \left. w_M^j \text{ is a Borel measurable selection of } W_M^j \right\}.
\end{aligned}$$

Let  $\check{S}_j = S_j$ .<sup>6</sup> Since

$$\begin{aligned}
&\int_{X_j \times S_j} W_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j) \varrho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j(d(x_j, s_j)) \\
&= \int_{S_j} \int_{X_j \times \check{S}_j} W_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j) \rho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j(d(x_j, \check{s}_j)) \\
&\quad \times \varphi_{j0}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_j) \lambda_j(ds_j),
\end{aligned}$$

we have

$$\begin{aligned}
&W_M^{j-1}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}) \\
&= \left\{ \int_{S_j} \int_{X_j \times \check{S}_j} w_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j) \rho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j(d(x_j, \check{s}_j)) \right. \\
&\quad \times \varphi_{j0}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_j) \lambda_j(ds_j) : \\
&\quad \rho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j \in \Xi_j^j(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}), \\
&\quad \left. w_M^j \text{ is a Borel measurable selection of } W_M^j \right\}.
\end{aligned}$$

Let

$$\begin{aligned}
&\check{W}_M^j(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_j) \\
&= \left\{ \int_{X_j \times \check{S}_j} w_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j) \cdot \rho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j(d(x_j, \check{s}_j)) : \right. \\
&\quad \rho_{(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}, \xi)}^j \in \Xi_j^j(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}), \\
&\quad \left. w_M^j \text{ is a Borel measurable selection of } W_M^j \right\}.
\end{aligned}$$

Since  $W_M^j(h_{t-1}, x_t, \dots, x_j, s_t, \dots, s_j)$  is continuous in  $x_j$  and does not depend on  $\check{s}_j$ , it is continuous in  $(x_j, \check{s}_j)$ . In addition,  $W_M^j$  is bounded, measurable, nonempty, convex and compact valued. By [Lemma B.2](#),  $\check{W}_M^j$  is bounded, measurable, nonempty and compact valued, and sectionally continuous on  $X^{j-1}$ .

---

<sup>6</sup>We need to use [Lemma B.2](#) below, which requires the continuity of the correspondences in terms of the integrated variables. Since  $W_M^j$  is only measurable, but not continuous, in  $s_j$ , we add a dummy variable  $\check{s}_j$  so that  $W_M^j$  is trivially continuous in such a variable.

It is easy to see that

$$\begin{aligned} W_M^{j-1}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_{j-1}) \\ = \int_{S_j} \check{W}_M^j(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_j) \varphi_{j0}(h_{t-1}, x_t, \dots, x_{j-1}, s_t, \dots, s_j) \lambda_j(ds_j). \end{aligned}$$

By Lemma 4, it is bounded, measurable, nonempty and compact valued, and sectionally continuous on  $X^{j-1}$ . By induction, one can show that  $W_M^{t-1}$  is bounded, measurable, nonempty and compact valued, and sectionally continuous on  $X^{t-1}$ .

Let  $W^{t-1} = \overline{\bigcup_{M \geq t} W_M^{t-1}}$ . That is,  $W^{t-1}$  is the closure of  $\bigcup_{M \geq t} W_M^{t-1}$ , which is measurable due to Lemma 2.

First,  $W^{t-1} \subseteq Q_t^\tau$  because  $W_M^{t-1} \subseteq Q_t^\tau$  for each  $M \geq t$  and  $Q_t^\tau$  is compact-valued. Second, fix  $h_{t-1}$  and  $q \in Q_t^\tau(h_{t-1})$ . Then there exists a mapping  $\xi \in Y$  such that

$$q = \int_{\prod_{m \geq t} (X_m \times S_m)} u(h_{t-1}, x, s) Q_{(h_{t-1}, \xi)}(d(x, s)).$$

For  $M \geq t$ , let

$$\begin{aligned} V_M(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M) \\ = \int_{\prod_{m \geq M} (X_m \times S_m)} u(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M, x, s) Q_{(h_{t-1}, x_t, \dots, x_M, s_t, \dots, s_M, \xi)}(x, s) \end{aligned}$$

and

$$q_M = \int_{\prod_{t \leq m \leq M} (X_m \times S_m)} V_M(h_{t-1}, x, s) Q_{(h_{t-1}, \xi)}^M(d(x, s)).$$

Hence,  $q_M \in W_M^{t-1}$ . Because the dynamic game is continuous at infinity,  $q_M \rightarrow q$ , which implies that  $q \in W^{t-1}(h_{t-1})$  and  $Q_t^\tau \subseteq W^{t-1}$ .

Therefore,  $W^{t-1} = Q_t^\tau$ , and hence  $Q_t^\tau$  is measurable for  $t > \tau$ .

Step 3. For  $t \leq \tau$ , we can start with  $Q_{\tau+1}^\tau$ . Repeating the backward induction in [Appendix B.4.1](#), we have that  $Q_t^\tau$  is also bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^{t-1}$ .  $\square$

Denote

$$Q_t^\infty = \begin{cases} Q_t^{t-1}, & \text{if } \bigcap_{\tau \geq 1} Q_t^\tau = \emptyset, \\ \bigcap_{\tau \geq 1} Q_t^\tau, & \text{otherwise.} \end{cases}$$

The following three lemmas show that  $Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1}) = E_t(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .<sup>7</sup>

<sup>7</sup>The proofs for [Lemmas B.10](#) and [B.12](#) follow the standard ideas with various modifications; see, for example, [Harris \(1990\)](#), [Harris et al. \(1995\)](#), and [Mariotti \(2000\)](#).

LEMMA B.10. (i) *The correspondence  $Q_t^\infty$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^{t-1}$ .*

(ii) *For any  $t \geq 1$ ,  $Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .*

PROOF. (i) It is obvious that  $Q_t^\infty$  is bounded. By the definition of  $Q_t^\tau$ , for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ ,  $Q_t^{\tau_1}(h_{t-1}) \subseteq Q_t^{\tau_2}(h_{t-1})$  for  $\tau_1 \geq \tau_2$ . Since  $Q_t^\tau$  is nonempty and compact valued,  $Q_t^\infty = \bigcap_{\tau \geq 1} Q_t^\tau$  is nonempty and compact valued for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ . If  $\bigcap_{\tau \geq 1} Q_t^\tau = \emptyset$ , then  $Q_t^\infty = Q_t^{t-1}$ . Thus,  $Q_t^\infty(h_{t-1})$  is nonempty and compact valued for all  $h_{t-1} \in H_{t-1}$ . By Lemma 2(ii),  $\bigcap_{\tau \geq 1} Q_t^\tau$  is measurable, which implies that  $Q_t^\infty$  is measurable.

Fix any  $s^{t-1} \in S^{t-1}$  such that  $Q_t^\tau(\cdot, s^{t-1})$  is upper hemicontinuous on  $H_{t-1}(s^{t-1})$  for any  $\tau$ . By Lemma 2(vii),  $Q_t^\tau(\cdot, s^{t-1})$  has a closed graph for each  $\tau$ , which implies that  $Q_t^\infty(\cdot, s^{t-1})$  has a closed graph. Referring to Lemma 2(vii) again,  $Q_t^\infty(\cdot, s^{t-1})$  is upper hemicontinuous on  $H_{t-1}(s^{t-1})$ . Since  $Q_t^\tau$  is essentially upper hemicontinuous on  $X^{t-1}$  for each  $\tau$ ,  $Q_t^\infty$  is essentially upper upper hemicontinuous on  $X^{t-1}$ .

(ii) For any  $\tau \geq 1$  and  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ ,  $\Phi(Q_{t+1}^\infty)(h_{t-1}) \subseteq \Phi(Q_{t+1}^\tau)(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$ , and hence,  $\Phi(Q_{t+1}^\infty)(h_{t-1}) \subseteq Q_t^\infty(h_{t-1})$ .

The space  $\{1, 2, \dots, \infty\}$  is a countable compact set endowed with the metric  $d(k, m) = |\frac{1}{k} - \frac{1}{m}|$  for any  $1 \leq k, m \leq \infty$ . The sequence  $\{Q_{t+1}^\tau\}_{1 \leq \tau \leq \infty}$  can be regarded as a correspondence  $Q_{t+1}$  from  $H_t \times \{1, 2, \dots, \infty\}$  to  $\mathbb{R}^n$ , which is measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \{1, 2, \dots, \infty\}$ . The backward induction in [Appendix B.4.1](#) shows that  $\Phi(Q_{t+1})$  is measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \{1, 2, \dots, \infty\}$ .

Since  $\Phi(Q_{t+1})$  is essentially sectionally upper hemicontinuous on  $X^t \times \{1, 2, \dots, \infty\}$ , there exists a measurable subset  $\check{S}^{t-1} \subseteq S^{t-1}$  such that  $\lambda^{t-1}(\check{S}^{t-1}) = 1$ , and  $\Phi(Q_{t+1})(\cdot, \cdot, \check{s}^{t-1})$  is upper hemicontinuous for any  $\check{s}^{t-1} \in \check{S}^{t-1}$ . Fix  $\check{s}^{t-1} \in \check{S}^{t-1}$ . For  $h_{t-1} = (x^{t-1}, \check{s}^{t-1}) \in H_{t-1}$  and  $a \in Q_t^\infty(h_{t-1})$ , by its definition,  $a \in Q_t^\tau(h_{t-1}) = \Phi(Q_{t+1}^\tau)(h_{t-1})$  for  $\tau \geq t$ . Thus,  $a \in \Phi(Q_{t+1}^\infty)(h_{t-1})$ .

In summary,  $Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .  $\square$

Though the definition of  $Q_t^\tau$  involves correlated strategies for  $\tau < t$ , the following lemma shows that one can work with mixed strategies in terms of equilibrium payoffs via the combination of backward and forward inductions in multiple steps.

LEMMA B.11. *If  $c_t$  is a measurable selection of  $\Phi(Q_{t+1}^\infty)$ , then  $c_t(h_{t-1})$  is a subgame-perfect equilibrium payoff vector for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .*

PROOF. Without loss of generality, we only prove the case  $t = 1$ .

Suppose that  $c_1$  is a measurable selection of  $\Phi(Q_2^\infty)$ . Apply [Proposition B.3](#) recursively to obtain Borel measurable mappings  $\{f_{ki}\}_{i \in I}$  for  $k \geq 1$ . That is, for any  $k \geq 1$ , there exists a Borel measurable selection  $c_k$  of  $Q_k^\infty$  such that for  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ , the following statements hold:

- (i) We have that  $f_k(h_{k-1})$  is a Nash equilibrium in the subgame  $h_{k-1}$ , where the action space is  $A_{ki}(h_{k-1})$  for player  $i \in I$  and the payoff function is given by

$$\int_{S_k} c_{k+1}(h_{k-1}, \cdot, s_k) f_{k0}(ds_k | h_{k-1}).$$

- (ii) We have

$$c_k(h_{k-1}) = \int_{A_k(h_{k-1})} \int_{S_k} c_{k+1}(h_{k-1}, x_k, s_k) f_{k0}(ds_k | h_{k-1}) f_k(dx_k | h_{k-1}).$$

We need to show that  $c_1(h_0)$  is a subgame-perfect equilibrium payoff vector for  $\lambda_0$ -almost all  $h_0 \in H_0$ .

Step 1. We show that for any  $k \geq 1$  and  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ ,

$$c_k(h_{k-1}) = \int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \mathcal{Q}_{(h_{k-1}, f)}(d(x, s)).$$

Since the game is continuous at infinity, there exists some positive integer  $M > k$  such that  $w^M$  is sufficiently small. By Lemma B.10,  $c_k(h_{k-1}) \in Q_k^\infty(h_{k-1}) = \bigcap_{\tau \geq 1} Q_k^\tau(h_{k-1})$  for  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ . Since  $Q_k^\tau = \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)$  for  $k \leq \tau$ ,  $c_k(h_{k-1}) \in \bigcap_{\tau \geq k} \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)(h_{k-1}) \subseteq \Phi^{M-k+1}(Q_{M+1}^M)(h_{k-1})$  for  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ . Thus, there exists a Borel measurable selection  $w$  of  $Q_{M+1}^M$  and some  $\xi \in Y$  such that for  $\lambda^{M-1}$ -almost all  $h_{M-1} \in H_{M-1}$ , the following statements hold:

- (i) We have that  $f_M(h_{M-1})$  is a Nash equilibrium in the subgame  $h_{M-1}$ , where the action space is  $A_{Mi}(h_{M-1})$  for player  $i \in I$  and the payoff function is given by

$$\int_{S_M} w(h_{M-1}, \cdot, s_M) f_{M0}(ds_M | h_{M-1}).$$

- (ii) We have

$$c_M(h_{M-1}) = \int_{A_M(h_{M-1})} \int_{S_M} w(h_{M-1}, x_M, s_M) f_{M0}(ds_M | h_{M-1}) f_M(dx_M | h_{M-1}).$$

- (iii) We have  $w(h_M) = \int_{\prod_{m \geq M+1} (X_m \times S_m)} u(h_M, x, s) \mathcal{Q}_{(h_M, \xi)}(d(x, s))$ .

Then for  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ ,

$$c_k(h_{k-1}) = \int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \mathcal{Q}_{(h_{k-1}, f^M)}(d(x, s)),$$

where  $f_k^M$  is  $f_k$  if  $k \leq M$  and is  $\xi_k$  if  $k \geq M+1$ . Since the game is continuous at infinity,

$$\int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \mathcal{Q}_{(h_{k-1}, f^M)}(d(x, s))$$

converges to

$$\int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \mathcal{Q}_{(h_{k-1}, f)}(d(x, s))$$

when  $M$  goes to infinity. Thus, for  $\lambda^{k-1}$ -almost all  $h_{k-1} \in H_{k-1}$ ,

$$c_k(h_{k-1}) = \int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \mathcal{Q}_{(h_{k-1}, f)}(d(x, s)). \quad (\text{B3})$$

Step 2. We show that  $\{f_{ki}\}_{i \in I}$  is a subgame-perfect equilibrium.

Fix a player  $i$  and a strategy  $g_i = \{g_{ki}\}_{k \geq 1}$ . For each  $k \geq 1$ , define a new strategy  $\tilde{f}_i^k$  as  $\tilde{f}_i^k = (g_{1i}, \dots, g_{ki}, f_{(k+1)i}, f_{(k+2)i}, \dots)$ . That is, we simply replace the initial  $k$  stages of  $f_i$  by  $g_i$ . Denote  $\tilde{f}^k = (\tilde{f}_i^k, f_{-i})$ .

Fix  $k \geq 1$  and a measurable subset  $D^k \subseteq S^k$  such that (i) and (ii) of Step 1 and (B3) hold for all  $s_k \in D^k$  and  $x^k \in H_k(s^k)$ , and  $\lambda^k(D^k) = 1$ . For each  $\tilde{M} > k$ , by the Fubini property, there exists a measurable subset  $E_k^{\tilde{M}} \subseteq S^k$  such that  $\lambda^k(E_k^{\tilde{M}}) = 1$  and  $\bigotimes_{k+1 \leq j \leq \tilde{M}} \lambda_j(D^{\tilde{M}}(s^k)) = 1$  for all  $s^k \in E_k^{\tilde{M}}$ , where

$$D^{\tilde{M}}(s^k) = \{(s_{k+1}, \dots, s_{\tilde{M}}) : (s^k, s_{k+1}, \dots, s_{\tilde{M}}) \in D^{\tilde{M}}\}.$$

Let  $\hat{D}^k = (\bigcap_{\tilde{M} > k} E_k^{\tilde{M}}) \cap D^k$ . Then  $\lambda^k(\hat{D}^k) = 1$ .

For any  $h_k = (x^k, s^k)$  such that  $s^k \in \hat{D}^k$  and  $x^k \in H_k(s^k)$ , we have

$$\begin{aligned} & \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, f)}(d(x, s)) \\ &= \int_{A_{k+1}(h_k)} \int_{S_{k+1}} c_{(k+2)i}(h_k, x_{k+1}, s_{k+1}) f_{(k+1)0}(ds_{k+1}|h_k) f_{k+1}(dx_{k+1}|h_k) \\ &\geq \int_{A_{k+1}(h_k)} \int_{S_{k+1}} c_{(k+2)i}(h_k, x_{k+1}, s_{k+1}) f_{(k+1)0}(ds_{k+1}|h_k) \\ &\quad (f_{(k+1)(-i)} \otimes g_{(k+1)i})(dx_{k+1}|h_k) \\ &= \int_{A_{k+1}(h_k)} \int_{S_{k+1}} \int_{A_{k+2}(h_k, x_{k+1}, s_{k+1})} \int_{S_{k+2}} c_{(k+3)i}(h_k, x_{k+1}, s_{k+1}, x_{k+2}, s_{k+2}) \\ &\quad f_{(k+2)0}(ds_{k+2}|h_k, x_{k+1}, s_{k+1}) f_{(k+2)(-i)} \otimes f_{(k+2)i}(dx_{k+2}|h_k, x_{k+1}, s_{k+1}) \\ &\quad f_{(k+1)0}(ds_{k+1}|h_k) f_{(k+1)(-i)} \otimes g_{(k+1)i}(dx_{k+1}|h_k) \\ &\geq \int_{A_{k+1}(h_k)} \int_{S_{k+1}} \int_{A_{k+2}(h_k, x_{k+1}, s_{k+1})} \int_{S_{k+2}} c_{(k+3)i}(h_k, x_{k+1}, s_{k+1}, x_{k+2}, s_{k+2}) \\ &\quad f_{(k+2)0}(ds_{k+2}|h_k, x_{k+1}, s_{k+1}) f_{(k+2)(-i)} \otimes g_{(k+2)i}(dx_{k+2}|h_k, x_{k+1}, s_{k+1}) \\ &\quad f_{(k+1)0}(ds_{k+1}|h_k) f_{(k+1)(-i)} \otimes g_{(k+1)i}(dx_{k+1}|h_k) \\ &= \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, \tilde{f}^{k+2})}(d(x, s)). \end{aligned}$$

The first and the last equalities follow from (B3) in the end of Step 1. The second equality is due to (ii) in Step 1. The first inequality is based on (i) in Step 1. The second inequality holds by the following arguments:

- (i) By the choice of  $h_k$  and (i) in Step 1, for  $\lambda_{k+1}$ -almost all  $s_{k+1} \in S_{k+1}$  and all  $x_{k+1} \in X_{k+1}$  such that  $(h_k, x_{k+1}, s_{k+1}) \in H_{k+1}$ , we have

$$\begin{aligned}
& \int_{A_{k+2}(h_k, x_{k+1}, s_{k+1})} \int_{S_{k+2}} c_{(k+3)i}(h_k, x_{k+1}, s_{k+1}, x_{k+2}, s_{k+2}) \\
& \quad f_{(k+2)0}(ds_{k+2}|h_k, x_{k+1}, s_{k+1}) f_{(k+2)(-i)} \otimes f_{(k+2)i}(dx_{k+2}|h_k, x_{k+1}, s_{k+1}) \\
& \geq \int_{A_{k+2}(h_k, x_{k+1}, s_{k+1})} \int_{S_{k+2}} c_{(k+3)i}(h_k, x_{k+1}, s_{k+1}, x_{k+2}, s_{k+2}) \\
& \quad f_{(k+2)0}(ds_{k+2}|h_k, x_{k+1}, s_{k+1}) f_{(k+2)(-i)} \otimes g_{(k+2)i}(dx_{k+2}|h_k, x_{k+1}, s_{k+1}).
\end{aligned}$$

- (ii) Since  $f_{(k+1)0}$  is absolutely continuous with respect to  $\lambda_{k+1}$ , the above inequality also holds for  $f_{(k+1)0}(h_k)$ -almost all  $s_{k+1} \in S_{k+1}$  and all  $x_{k+1} \in X_{k+1}$  such that  $(h_k, x_{k+1}, s_{k+1}) \in H_{k+1}$ .

Repeating the above argument, one can show that

$$\begin{aligned}
& \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, f)}(d(x, s)) \\
& \geq \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, \tilde{f}^{\tilde{M}+1})}(d(x, s))
\end{aligned}$$

for any  $\tilde{M} > k$ . Since

$$\int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, \tilde{f}^{\tilde{M}+1})}(d(x, s))$$

converges to

$$\int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, (g_i, f_{-i}))}(d(x, s))$$

as  $\tilde{M}$  goes to infinity, we have

$$\begin{aligned}
& \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, f)}(d(x, s)) \\
& \geq \int_{\prod_{m \geq k+1} (X_m \times S_m)} u(h_k, x, s) \mathcal{Q}_{(h_k, (g_i, f_{-i}))}(d(x, s)).
\end{aligned}$$

Therefore,  $\{f_{ki}\}_{i \in I}$  is a subgame-perfect equilibrium.  $\square$

By Lemma B.10 and Proposition B.2, the correspondence  $\Phi(Q_{t+1}^\infty)$  is measurable, nonempty and compact valued; by Lemma 2(iii), it has a measurable selection. Then Theorem 3 follows from the above lemma.

For  $t \geq 1$  and  $h_{t-1} \in H_{t-1}$ , recall that  $E_t(h_{t-1})$  is the set of payoff vectors of subgame-perfect equilibria in the subgame  $h_{t-1}$ . The following lemma shows that  $E_t(h_{t-1})$  is essentially the same as  $Q_t^\infty(h_{t-1})$ .

LEMMA B.12. For any  $t \geq 1$ ,  $E_t(h_{t-1}) = Q_t^\infty(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .

PROOF. (a) We first prove the following claim: for any  $t$  and  $\tau$ , if  $E_{t+1}(h_t) \subseteq Q_{t+1}^\tau(h_t)$  for  $\lambda^t$ -almost all  $h_t \in H_t$ , then  $E_t(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ . We need to consider only the case that  $t \leq \tau$ .

By the construction of  $\Phi(Q_{t+1}^\tau)$  in [Appendix B.4.1](#), there exists a measurable subset  $\hat{S}^{t-1} \subseteq S^{t-1}$  with  $\lambda^{t-1}(\hat{S}^{t-1}) = 1$  such that for any  $c_t$  and  $h_{t-1} = (x^{t-1}, \hat{s}^{t-1}) \in H_{t-1}$  with  $\hat{s}^{t-1} \in \hat{S}^{t-1}$ , if the following relationships hold, then  $c_t \in \Phi(Q_{t+1}^\tau)(h_{t-1})$ .

- We have  $c_t = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) \alpha(dx_t)$ , where  $q_{t+1}(h_{t-1}, \cdot)$  is measurable and  $q_{t+1}(h_{t-1}, x_t, s_t) \in Q_{t+1}^\tau(h_{t-1}, x_t, s_t)$  for  $\lambda_t$ -almost all  $s_t \in S_t$  and  $x_t \in A_t(h_{t-1})$ .
- We have that  $\alpha \in \bigotimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$  is a Nash equilibrium in the subgame  $h_{t-1}$  with payoff  $\int_{S_t} q_{t+1}(h_{t-1}, \cdot, s_t) f_{t0}(ds_t|h_{t-1})$  and action space  $\prod_{i \in I} A_{ti}(h_{t-1})$ .

Fix a subgame  $h_{t-1} = (x^{t-1}, \hat{s}^{t-1})$  such that  $\hat{s}^{t-1} \in \hat{S}^{t-1}$ . Pick a point  $c_t \in E_t(h_{t-1})$ . There exists a strategy profile  $f$  such that  $f$  is a subgame-perfect equilibrium in the subgame  $h_{t-1}$  and the payoff is  $c_t$ . Let  $c_{t+1}(h_{t-1}, x_t, s_t)$  be the payoff vector induced by  $\{f_{ti}\}_{i \in I}$  in the subgame  $(h_t, x_t, s_t) \in \text{Gr}(A_t) \times S_t$ . Then the following relationships hold:

- We have  $c_t = \int_{A_t(h_{t-1})} \int_{S_t} c_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$ .
- We have that  $f_t(\cdot|h_{t-1})$  is a Nash equilibrium in the subgame  $h_{t-1}$  with action space  $A_t(h_{t-1})$  and payoff  $\int_{S_t} c_{t+1}(h_{t-1}, \cdot, s_t) f_{t0}(ds_t|h_{t-1})$ .

Since  $f$  is a subgame-perfect equilibrium in the subgame  $h_{t-1}$ ,  $c_{t+1}(h_{t-1}, x_t, s_t) \in E_{t+1}(h_{t-1}, x_t, s_t) \subseteq Q_{t+1}^\tau(h_{t-1}, x_t, s_t)$  for  $\lambda_t$ -almost all  $s_t \in S_t$  and  $x_t \in A_t(h_{t-1})$ , which implies that  $c_t \in \Phi(Q_{t+1}^\tau)(h_{t-1}) = Q_t^\tau(h_{t-1})$ .

Therefore,  $E_t(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .

(b) For any  $t > \tau$ ,  $E_t \subseteq Q_t^\tau$ . If  $t \leq \tau$ , we can start with  $E_{\tau+1} \subseteq Q_{\tau+1}^\tau$  and repeat the argument in (a). Then we can show that  $E_t(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ . Thus,  $E_t(h_{t-1}) \subseteq Q_t^\infty(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .

(c) Suppose that  $c_t$  is a measurable selection from  $\Phi(Q_{t+1}^\infty)$ . Apply [Proposition B.3](#) recursively to obtain Borel measurable mappings  $\{f_{ki}\}_{i \in I}$  for  $k \geq t$ . By [Lemma B.11](#),  $c_t(h_{t-1})$  is a subgame-perfect equilibrium payoff vector for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ . Consequently,  $\Phi(Q_{t+1}^\infty)(h_{t-1}) \subseteq E_t(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .

By [Lemma B.10](#),  $E_t(h_{t-1}) = Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1})$  for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ .  $\square$

### B.5 Proof of [Proposition B.1](#)

We highlight the needed changes in comparison with the proofs presented in [Appendices B.4.1–B.4.3](#).

Step 1: Backward induction. We first consider stage  $t$  with  $N_t = 1$ .

If  $N_t = 1$ , then  $S_t = \{\hat{s}_t\}$ . Thus,  $P_t(h_{t-1}, x_t) = Q_{t+1}(h_{t-1}, x_t, \hat{s}_t)$ , which is nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \hat{S}^{t-1}$ . Notice that  $P_t$  may not be convex-valued.

We first assume that  $P_t$  is upper hemicontinuous. Suppose that  $j$  is the player who is active in this period. Consider the correspondence  $\Phi_t: H_{t-1} \rightarrow \mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$  defined as  $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$  if the following relationships hold:

- (i) We have  $v = p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{ij}^*)$  such that  $p_t(h_{t-1}, \cdot)$  is a measurable selection of  $P_t(h_{t-1}, \cdot)$ .<sup>8</sup>
- (ii) We have that  $x_{ij}^* \in A_{ij}(h_{t-1})$  is a maximization point of player  $j$  given the payoff function  $p_{ij}(h_{t-1}, A_{t(-j)}(h_{t-1}), \cdot)$  and the action space  $A_{ij}(h_{t-1})$ ,  $\alpha_i = \delta_{A_{ii}(h_{t-1})}$  for  $i \neq j$  and  $\alpha_j = \delta_{x_{ij}^*}$ .
- (iii) We have  $\mu = p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{ij}^*) \circ \alpha$ .

This is a single agent problem. We need to show that  $\Phi_t$  is nonempty and compact valued, and upper hemicontinuous.

If  $P_t$  is nonempty, convex and compact valued, and upper hemicontinuous, then we can use Lemma 10, the main result of [Simon and Zame \(1990\)](#), to prove the nonemptiness, compactness, and upper hemicontinuity of  $\Phi_t$ . In [Simon and Zame \(1990\)](#), the only step they need the convexity of  $P_t$  for the proof of their main theorem is Lemma 2 therein. However, the one-player pure-strategy version of their Lemma 2, stated in the following, directly follows from the upper hemicontinuity of  $P_t$  without requiring the convexity.

Let  $Z$  be a compact metric space, and  $\{z_n\}_{n \geq 0} \subseteq Z$ . Let  $P: Z \rightarrow \mathbb{R}_+$  be a bounded, upper hemicontinuous correspondence with nonempty and compact values. For each  $n \geq 1$ , let  $q_n$  be a Borel measurable selection of  $P$  such that  $q_n(z_n) = d_n$ . If  $z_n$  converges to  $z_0$  and  $d_n$  converges to some  $d_0$ , then  $d_0 \in P(z_0)$ .

Repeating the argument in the proof of the main theorem of [Simon and Zame \(1990\)](#), one can show that  $\Phi_t$  is nonempty and compact valued, and upper hemicontinuous.

Then we go back to the case that  $P_t$  is nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \hat{S}^{t-1}$ . Recall that we proved [Proposition B.2](#) based on Lemma 10. If  $P_t$  is essentially sectionally upper hemicontinuous on  $X^t \times \hat{S}^{t-1}$ , we can show the following result based on a similar argument as in [Appendix B.3](#): there exists a bounded, measurable, nonempty and compact valued correspondence  $\Phi_t$  from  $H_{t-1}$  to  $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$  such that  $\Phi_t$  is essentially sectionally upper hemicontinuous on  $X^{t-1} \times \hat{S}^{t-1}$  and for  $\lambda^{t-1}$ -almost all  $h_{t-1} \in H_{t-1}$ ,  $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$  if the following relationships hold:

- (i) We have  $v = p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{ij}^*)$  such that  $p_t(h_{t-1}, \cdot)$  is a measurable selection of  $P_t(h_{t-1}, \cdot)$ .
- (ii) We that  $x_{ij}^* \in A_{ij}(h_{t-1})$  is a maximization point of player  $j$  given the payoff function  $p_{ij}(h_{t-1}, A_{t(-j)}(h_{t-1}), \cdot)$  and the action space  $A_{ij}(h_{t-1})$ ,  $\alpha_i = \delta_{A_{ii}(h_{t-1})}$  for  $i \neq j$  and  $\alpha_j = \delta_{x_{ij}^*}$ .
- (iii) We have  $\mu = p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{ij}^*) \circ \alpha$ .

<sup>8</sup>Note that  $A_{t(-j)}$  is point-valued since all players other than  $j$  are inactive.

Next we consider the case that  $N_t = 0$ . Suppose that the correspondence  $Q_{t+1}$  from  $H_t$  to  $\mathbb{R}^n$  is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \hat{S}^t$ . For any  $(h_{t-1}, x_t, \hat{s}_t) \in \text{Gr}(\hat{A}_t)$ , let

$$\begin{aligned} R_t(h_{t-1}, x_t, \hat{s}_t) &= \int_{\tilde{S}_t} Q_{t+1}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t) \tilde{f}_{t0}(d\tilde{s}_t | h_{t-1}, x_t, \hat{s}_t) \\ &= \int_{\tilde{S}_t} Q_{t+1}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t) \varphi_{t0}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t) \lambda_t(d\tilde{s}_t). \end{aligned}$$

Then following the same argument as in [Appendix B.4.1](#), one can show that  $R_t$  is a nonempty, convex and compact valued, and essentially sectionally upper hemicontinuous correspondence on  $X^t \times \hat{S}^t$ .

For any  $h_{t-1} \in H_{t-1}$  and  $x_t \in A_t(h_{t-1})$ , let

$$P_t(h_{t-1}, x_t) = \int_{\hat{A}_{t0}(h_{t-1}, x_t)} R_t(h_{t-1}, x_t, \hat{s}_t) \hat{f}_{t0}(d\hat{s}_t | h_{t-1}, x_t).$$

By Lemma 7,  $P_t$  is nonempty, convex and compact valued, and essentially sectionally upper hemicontinuous on  $X^t \times \hat{S}^{t-1}$ . The rest of the step remains the same as in [Appendix B.4.1](#).

Step 2: Forward induction. This step is unchanged.

Step 3: Infinite horizon. We need to slightly modify the definition of  $\Xi_t^{m_1}$  for any  $m_1 \geq t \geq 1$ . Fix any  $t \geq 1$ . Define a correspondence  $\Xi_t'$  as follows: in the subgame  $h_{t-1}$ ,

$$\Xi_t'(h_{t-1}) = (\mathcal{M}(A_t(h_{t-1})) \diamond \hat{f}_{t0}(h_{t-1}, \cdot)) \otimes \lambda_t.$$

For any  $m_1 > t$ , suppose that the correspondence  $\Xi_t^{m_1-1}$  has been defined. Then we can define a correspondence  $\Xi_t^{m_1} : H_{t-1} \rightarrow \mathcal{M}(\prod_{t \leq m \leq m_1} (X_m \times S_m))$  as

$$\begin{aligned} \Xi_t^{m_1}(h_{t-1}) &= \{g(h_{t-1}) \diamond ((\xi_{m_1}(h_{t-1}, \cdot) \diamond \hat{f}_{m_1 0}(h_{t-1}, \cdot)) \otimes \lambda_{m_1}) : \\ &\quad g \text{ is a Borel measurable selection of } \Xi_t^{m_1-1}, \\ &\quad \xi_{m_1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_1})\}. \end{aligned}$$

Then the result in [Appendix B.4.3](#) is true with the above  $\Xi_t^{m_1}$ .

Consequently, a subgame-perfect equilibrium exists.

## REFERENCES

- Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third edition. Springer, Berlin. [3, 4, 7]
- Castaing, Charles, Paul Raynaud De Fitte, and Michel Valadier (2004), *Young Measures on Topological Spaces: With Applications in Control Theory and Probability Theory*, volume 571 of *Mathematics and Its Applications*. Springer Science & Business Media, Netherlands. [8]

Harris, Christopher (1990), *The Existence of Subgame-Perfect Equilibrium With and Without Markov Strategies: A Case for Extensive-Form Correlation*. Discussion Paper, Nuffield College, Oxford, Discussion Paper No. 53. [28]

Harris, Christopher, Philip Reny, and Arthur Robson (1995), “The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization.” *Econometrica*, 63, 507–544. [28]

Mariotti, Thomas (2000), “Subgame-perfect equilibrium outcomes in continuous games of almost perfect information.” *Journal of Mathematical Economics*, 34, 99–128. [28]

Mertens, Jean-François and T. Parthasarathy (2003), “Equilibria for discounted stochastic games.” In *Stochastic Games and Applications*. NATO Science Series (Abraham Neyman and Sylvain Sorin, eds.), 131–172, Springer Science+Business Media, New York. [3]

Simon, Leo K. and William R. Zame (1990), “Discontinuous games and endogenous sharing rules.” *Econometrica*, 58, 861–872. [10, 34]

---

Co-editor Thomas Mariotti handled this manuscript.

Manuscript received 26 June, 2017; final version accepted 16 September, 2019; available online 14 October, 2019.