

Dynamic Games with Almost Perfect Information*

Wei He[†] Yeneng Sun[‡]

This version: June 26, 2017

Abstract

This paper aims to solve two fundamental problems on finite or infinite horizon dynamic games with complete information. Under some mild conditions, we prove the existence of subgame-perfect equilibria and the upper hemicontinuity of equilibrium payoffs in general dynamic games with simultaneous moves (i.e., almost perfect information), which go beyond previous works in the sense that stagewise public randomization and the continuity requirement on the state variables are not needed. For alternating move (i.e., perfect-information) dynamic games with uncertainty, we show the existence of *pure-strategy* subgame-perfect equilibria as well as the upper hemicontinuity of equilibrium payoffs, extending the earlier results on perfect-information deterministic dynamic games. Two examples on dynamic stochastic oligopoly markets and intergenerational bequests are considered as illustrative applications.

*Earlier versions of this paper were presented in University of Iowa, November 2015; IMS Workshop on Stochastic Games, National University of Singapore, December 2015; Chinese University of Hong Kong, December 2015; Zhejiang University, December 2015; North American Summer Meeting of the Econometric Society, Philadelphia, June 2016; 5th World Congress of the Game Theory Society, Maastricht, July 2016; University of Minnesota, October 2016; Chinese University of Hong Kong-Shenzhen, January 2017; 2017 Economic Theory Workshop, National University of Singapore, February 2017; Hong Kong University of Science and Technology, March 2017; 8th Shanghai Microeconomics Workshop, June 2017. The paper was first circulated in March 2015 at <https://arxiv.org/abs/1503.08900>.

[†]Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. E-mail: he.wei2126@gmail.com.

[‡]Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570. Email: ynsun@nus.edu.sg.

Contents

1	Introduction	3
2	Model	6
3	Continuous dynamic games	8
3.1	Continuous dynamic games with almost perfect information	8
3.2	Continuous dynamic games with perfect information	9
3.3	An extension	10
4	Applications	11
4.1	A dynamic oligopoly model	11
4.2	Intergenerational bequests with stochastic production	13
5	Measurable dynamic games	14
5.1	Measurable dynamic games with almost perfect information	15
5.2	Measurable dynamic games with perfect information	17
6	Appendix A	17
6.1	Technical preparations	17
6.2	Proof of Theorem 1	28
6.2.1	Backward induction	28
6.2.2	Forward induction	29
6.2.3	Infinite horizon case	34
6.3	Proof of Theorem 2	40
6.4	Proof of Proposition 1	42
	References	43

1 Introduction

Dynamic games with complete information, where players observe the history and move simultaneously or alternately in finite or infinite horizon, arise naturally in many situations. As noted in Chapter 4 of [10] and [15], these games form a general class of regular dynamic games with many applications in economics, political science, and biology. The associated notion of subgame-perfect equilibrium is a fundamental game-theoretic concept. For alternating move games with finite actions, an early result on backward induction (subgame-perfect equilibrium) was presented in [35]. For simultaneous move games with finitely many actions and stages, the existence of subgame-perfect equilibria was shown in [31] while the infinite horizon but finite-action case was covered by [9].

Since the agents in many economic models need to make continuous choices, it is important to consider dynamic games with general action spaces. However, a simple example without any subgame-perfect equilibrium was presented in [15], where the game has two players in each of the two stages with only one player having a continuous choice set.¹ Thus, the existence of subgame-perfect equilibria under some suitable conditions remains an open problem even for two-stage dynamic games. The purpose of this paper is to prove the existence of subgame-perfect equilibria for general dynamic games with simultaneous or alternating moves in finite or infinite horizon under some suitable conditions. We shall adopt the general forms of intertemporal preferences, requiring neither stationarity nor additive separability.

For deterministic games with perfect information (i.e., the players move alternately), the existence of pure-strategy subgame-perfect equilibria was shown in [6, 7], [9], [13], [16], and [17] with the model parameters being continuous in actions, extending the early work of [35]. However, if the “deterministic” assumption is dropped by introducing a passive player - Nature, then a pure-strategy subgame-perfect equilibrium need not exist as shown by a four-stage game in [15]. In fact, the nonexistence of a mixed-strategy subgame-perfect equilibrium in a five-stage alternating move game was provided by [21].² Thus, it is still an open problem to show the existence of (pure or mixed strategy) subgame-perfect equilibria in (finite or infinite horizon) perfect-information dynamic games with uncertainty under some general conditions.

Continuous dynamic games with almost perfect information in the sense that

¹As noted in Section 2 of [15], such a dynamic game with continuous choices provides a minimal nontrivial counterexample; see also Exercise 13.4 in [10] for another counterexample of Harris.

²All those counterexamples show that various issues arise when one considers general dynamic games; see, for example, the discussion in [33]. In the setting with incomplete information, even the equilibrium notion needs to be carefully treated; see [26].

the players move simultaneously have been considered in [15]. In such games, there are a finite number of active players and a passive player (Nature), and all the relevant model parameters are assumed to be continuous in both action and state variables (i.e., Nature's moves). It was shown in [15] that subgame-perfect equilibria exist in those extended games obtained from the original games by introducing stagewise public randomization as a correlation device.³ As mentioned above, they also demonstrated the possible nonexistence of subgame-perfect equilibrium through a simple example.

The first aim of this paper is to resolve the above two open problems (in both finite and infinite horizon) for the class of continuous dynamic games. We assume the state transition in each period (except for those periods with one active player) to be an atomless probability measure for any given history.⁴ In Theorems 1 and 2 (and Proposition 1), we present the existence results for subgame-perfect equilibria, and also some regularity properties of the equilibrium payoff correspondences, including compactness and upper hemicontinuity.⁵ Note that our model allows the state history to fully influence all the model parameters, and hence covers the case with stagewise public randomization in the sense that the state transition has an additional atomless component that is independently and identically distributed across time, and does not enter the payoffs, state transitions and action correspondences. As a result, we obtain the existence result in [15] as a special case. In addition, we also provide a new existence result for continuous stochastic games in Proposition 2; see Remark 3 for discussions. As illustrative applications, we consider a dynamic oligopoly market in which firms face stochastic demand/cost, and an intergenerational bequest model with stochastic production.

For dynamic games with almost perfect information, our results allow the players to take mixed strategies. However, for the special class of continuous dynamic games with perfect information,⁶ we obtain the existence of *pure-strategy* subgame-perfect equilibria in Theorem 2. When Nature is present, there has been no general result on the existence of equilibria (even in mixed strategies) for continuous dynamic games with perfect information. Our Theorem 2 provides a new existence result in pure strategy, which extends the results of [6], [9], [13], [16], [17], and [35] to the case when Nature is present.

³See also [22] and [29].

⁴A probability measure on a separable metric space is atomless if every single point has measure zero.

⁵Such an upper hemicontinuity property in terms of correspondences of equilibrium payoffs, or outcomes, or correlated strategies has been the key for proving the relevant existence results in [6, 7], [13], [15], [16], [17], and [22].

⁶Dynamic games with perfect information also have wide applications. For example, see [2] and [28] for an intergenerational bequest game, and [11] and [27] for intrapersonal games in which consumers have changing preferences.

The condition of atomless transition is minimal in the particular sense that the existence results for continuous dynamic games may fail to hold if (1) the passive player, Nature, is not present in the model as shown in [15], or (2) with the presence of Nature, the state transition is not atomless as shown in [21].

The second aim of this paper is to consider an important extension in which the relevant model parameters are assumed to be continuous in actions, but measurable in states.⁷ In particular, we show the existence of a subgame-perfect equilibrium in a general dynamic game with almost perfect information under some suitable conditions on the state transitions. Theorems 3 and 4 below go beyond our results on continuous dynamic games by dropping the continuity requirement on the state variables.⁸ We work with the condition that the state transition in each period (except for those periods with one active player) has a component with a suitable density function with respect to some atomless reference measure.

In Appendix A, we provide a complete proof of Theorem 1, and point out those changes that are needed for proving Theorem 2 and Proposition 1. We follow the standard three-step procedure in obtaining subgame-perfect equilibria of dynamic games, namely, backward induction, forward induction, and approximation of infinite horizon by finite horizon. Because we drop the stagewise public randomization, new technical difficulties arise in the proofs. The main purpose of the step of backward induction is to show that if the payoff correspondence at a given stage satisfies certain regularity properties, then the equilibrium payoff correspondence at the previous stage is upper hemicontinuous. We notice that the condition of atomless transitions suffices for this purpose, and hence the exogenous stagewise public randomization is not needed for this step. For the step of forward induction, we need to obtain strategies that are jointly measurable in history. When there is a public randomization device, the joint measurability follows from the measurable version of Skorokhod’s representation theorem and implicit function theorem respectively as in [15] and [29]. Here we need to work with the deep “measurable” measurable choice theorem of [23].

In Appendix B, we prove Theorem 3 first, and then describe the needed changes for proving Theorem 4 and Proposition B.1. The proofs for the results in measurable dynamic games are much more difficult than those in the case of continuous dynamic games. In the step of backward induction, we obtain a new

⁷Since the agents need to make optimal choices, the continuity assumption in terms of actions is natural and widely adopted. However, the state variable is not a choice variable, and thus it is unnecessary to impose the state continuity requirement in a general model. Note that the state measurability assumption is the minimal regularity condition one would expect for the model parameters. We may also point out that the proof for the case with state continuity in Appendix A is much simpler than the proof for the general case in Appendix B.

⁸In Appendix B, we also present a new existence results on subgame-perfect equilibria for a general stochastic game.

existence result for discontinuous games with stochastic endogenous sharing rules, which extends the main result of [32] by allowing the payoff correspondence to be measurable (instead of upper hemicontinuous) in states. In order to extend the results to the infinite horizon setting, we need to handle various subtle measurability issues due to the lack of continuity in the state variables in the more general model, which is the most difficult part of the proof for Theorem 3.⁹

The rest of the paper is organized as follows. The model is presented in Section 2. The results for continuous dynamic games are given in Section 3. Two illustrative applications on dynamic oligopoly markets with stochastic demand/cost shock and intergenerational bequests with stochastic production are provided in Section 4. Section 5 extends continuous dynamic games to the setting in which the model parameters may only be measurable in the state variables. The proofs for the results of continuous dynamic games and measurable dynamic games are left in Appendices A and B, respectively.

2 Model

In this section, we shall present the model for an infinite-horizon dynamic game with almost perfect information.

The set of players is $I_0 = \{0, 1, \dots, n\}$, where the players in $I = \{1, \dots, n\}$ are active and player 0 is Nature. All the players move simultaneously. Time is discrete, and indexed by $t = 0, 1, 2, \dots$

The set of starting points is a closed set $H_0 = X_0 \times S_0$, where X_0 is a compact metric space and S_0 is a Polish space (that is, a complete separable metric space).¹⁰ At stage $t \geq 1$, player i 's action will be chosen from a subset of a Polish space X_{ti} for each $i \in I$, and $X_t = \prod_{i \in I} X_{ti}$. Nature's action is chosen from a Polish space S_t . Let $X^t = \prod_{0 \leq k \leq t} X_k$ and $S^t = \prod_{0 \leq k \leq t} S_k$. The Borel σ -algebras on X_t and S_t are denoted by $\mathcal{B}(X_t)$ and $\mathcal{B}(S_t)$, respectively. Given $t \geq 0$, a history up to the stage t is a vector

$$h_t = (x_0, s_0, x_1, s_1, \dots, x_t, s_t) \in X^t \times S^t.$$

The set of all such possible histories is denoted by H_t . For any $t \geq 0$, $H_t \subseteq X^t \times S^t$.

For any $t \geq 1$ and $i \in I$, let A_{ti} be a measurable, nonempty and compact valued

⁹Because our relevant model parameters are only measurable in the state variables, the usual method of approximating a limit continuous dynamic game by a sequence of finite games, as used in [7], [15] and [17], is not applicable in our setting.

¹⁰In each stage $t \geq 1$, there will be a set of action profiles X_t and a set of states S_t . Without loss of generality, we assume that the set of initial points is also a product space for notational consistency.

correspondence¹¹ from H_{t-1} to X_{ti} such that $A_{ti}(h_{t-1})$ is the set of available actions for player $i \in I$ given the history h_{t-1} . Let $A_t = \prod_{i \in I} A_{ti}$. Then $H_t = \text{Gr}(A^t) \times S_t$, where $\text{Gr}(A^t)$ is the graph of A^t .

For any $x = (x_0, x_1, \dots) \in X^\infty$, let $x^t = (x_0, \dots, x_t) \in X^t$ be the truncation of x up to the period t . Truncations for $s \in S^\infty$ can be defined similarly. Let H_∞ be 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 $t \geq 0$. Then H_∞ is the set of all possible histories in the game.¹² For any $t \geq 1$, Nature's action is given by a Borel measurable mapping f_{t0} from H_{t-1} to $\mathcal{M}(S_t)$,¹³ where $\mathcal{M}(S_t)$ is endowed with topology of weak convergence of measures on S_t .

For each $i \in I$, the payoff function u_i is a Borel measurable mapping from H_∞ to \mathbb{R}_{++} which is bounded by $\gamma > 0$.¹⁴

For player $i \in I$, a strategy f_i is a sequence $\{f_{ti}\}_{t \geq 1}$ such that f_{ti} is a Borel measurable mapping from H_{t-1} to $\mathcal{M}(X_{ti})$ with $f_{ti}(A_{ti}(h_{t-1})|h_{t-1}) = 1$ for all $h_{t-1} \in H_{t-1}$. A strategy profile $f = \{f_i\}_{i \in I}$ is a combination of strategies of all active players.

In any subgame, a strategy combination will generate a probability distribution over the set of possible histories. This probability distribution is called the path induced by the strategy combination in this subgame.

Definition 1. *Suppose that a strategy profile $f = \{f_i\}_{i \in I}$ and a history $h_t \in H_t$ are given for some $t \geq 0$. Let $\tau_t = \delta_{h_t}$, where δ_{h_t} is the probability measure concentrated at the one point h_t . If $\tau_{t'} \in \mathcal{M}(H_{t'})$ has already been defined for some $t' \geq t$, then let*

$$\tau_{t'+1} = \tau_{t'} \diamond (\otimes_{i \in I_0} f_{(t'+1)i}).^{15}$$

Finally, let $\tau \in \mathcal{M}(H_\infty)$ be the unique probability measure on H_∞ such that $\text{Marg}_{H_{t'}} \tau = \tau_{t'}$ for all $t' \geq t$. Then τ is called the path induced by f in the subgame h_t . For all $i \in I$, $\int_{H_\infty} u_i d\tau$ is the payoff of player i in this subgame.

¹¹Suppose that Y and Z are both Polish spaces, and Ψ is a correspondence from Y to Z . Hereafter, the measurability of Ψ , unless specifically indicated, is with respect to the Borel σ -algebra $\mathcal{B}(Y)$ on Y .

¹²A finite horizon dynamic game can be regarded as a special case of an infinite horizon dynamic game in the sense that the action correspondence A_{ti} is point-valued for each player $i \in I$ and $t \geq T$ for some stage $T \geq 1$; see, for example, [6] and [15].

¹³For a Polish space A , $\mathcal{M}(A)$ denotes the set of all Borel probability measures on A , and $\Delta(A)$ is the set of all finite Borel measures on A .

¹⁴Since u_i is bounded, we can assume that the value of the payoff function is strictly positive without loss of generality.

¹⁵Denote $\otimes_{i \in I_0} f_{(t'+1)i}$ as a transition probability from $H_{t'}$ to $\mathcal{M}(X_{t'+1})$. Note that the strategy profile is usually represented by a vector. For the notational simplicity later on, we assume that $\otimes_{i \in I_0} f_{(t'+1)i}(\cdot|h_{t'})$ represents the strategy profile in stage $t' + 1$ for a given history $h_{t'} \in H_{t'}$, where $\otimes_{i \in I_0} f_{(t'+1)i}(\cdot|h_{t'})$ is the product of the probability measures $f_{(t'+1)i}(\cdot|h_{t'})$, $i \in I_0$. If λ is a finite measure on X and ν is a transition probability from X to Y , then $\lambda \diamond \nu$ is a measure on $X \times Y$ such that $\lambda \diamond \nu(A \times B) = \int_A \nu(B|x) \lambda(dx)$ for any measurable subsets $A \subseteq X$ and $B \subseteq Y$.

3 Continuous dynamic games

In this section, we shall show the existence of subgame-perfect equilibria for continuous dynamic games in which all the model parameters (the payoff functions, state transitions and action correspondences) are continuous in both action and state variables. In Subsections 3.1, we first consider dynamic games with almost perfect information, and show the existence of subgame-perfect equilibria. In Subsection 3.2, we consider dynamic games with perfect information in the sense that players move sequentially, and prove the existence of *pure-strategy* subgame-perfect equilibria. In Subsection 3.3, we extend the model so that the previous existence results for continuous dynamic games with perfect and almost perfect information are covered as special cases. As a byproduct, we provide a new existence result for stochastic games. The proofs of the results in this section are left in Appendix A.

3.1 Continuous dynamic games with almost perfect information

In this subsection, we study an infinite-horizon continuous dynamic game with almost perfect information. In particular, a dynamic game is said to be “continuous” if for each t and i ,

1. the action correspondence A_{ti} is continuous on H_{t-1} ;
2. the transition probability f_{t0} is a continuous mapping from H_{t-1} to $\mathcal{M}(S_t)$;
3. the payoff function u_i is continuous on H_∞ .

Below, we propose the condition of “atomless transitions” on the state space, which means that the state transition is an atomless probability measure in any stage.

Assumption 1 (Atomless Transitions). *For each $t \geq 1$, $f_{t0}(h_{t-1})$ is an atomless Borel probability measure for each $h_{t-1} \in H_{t-1}$.*

The notion of subgame-perfect equilibrium is given below. It requires each player’s strategy to be optimal in every subgame given the strategies of all other players.

Definition 2 (SPE). *A subgame-perfect equilibrium is a strategy profile f such that for all $i \in I$, $t \geq 0$, and all $h_t \in H_t$, player i cannot improve his payoff in the subgame h_t by a unilateral change in his strategy.*

Let $E_t(h_{t-1})$ be the set of subgame-perfect equilibrium payoffs in the subgame h_{t-1} . The following result shows that a subgame-perfect equilibrium exists,

and the correspondence E_t satisfies certain desirable compactness and upper hemicontinuity properties.

Theorem 1. *If a continuous dynamic game with almost perfect information has atomless transitions, then it possesses a subgame-perfect equilibrium. In addition, E_t is nonempty and compact valued, and upper hemicontinuous on H_{t-1} for any $t \geq 1$.*

Remark 1. *Theorem 1 goes beyond the main result of [15] for continuous dynamic games, where the existence of subgame-perfect equilibria was shown for those extended games obtained from the original games by introducing stagewise public randomization as a correlation device. Such a correlation device does not influence the payoffs, transitions or action correspondences. It is clear that the extended games with stagewise public randomization as in [15] automatically satisfy the condition of atomless transitions. The states in our model are completely endogenous in the sense that they affect all the model parameters such as payoffs, transitions, and action correspondences.*

3.2 Continuous dynamic games with perfect information

In this subsection, we consider another important class of continuous dynamic games, namely continuous dynamic games with perfect information (with or without Nature). In such games, players move sequentially. We show the existence of pure-strategy subgame perfect equilibria. In particular, the condition of atomless transitions is imposed only when Nature moves.

In a continuous dynamic game with perfect information, there is only one player moving in each stage. In stage t , if A_{ti} is not point valued for some player $i \in I$, then A_{tj} is point valued for any $j \in I$ as long as $j \neq i$, and $f_{t0}(h_{t-1}) \equiv \delta_{s_t}$ for some s_t . That is, only player i is strategic in stage t , while all the other players are inactive. If the state transition f_{t0} does not put probability 1 on some point, then A_{ti} must be point valued for any $i \in I$. That is, only Nature can move in stage t , and all the players $i \in I$ are inactive in this stage. A continuous dynamic game with perfect information is said to have atomless transitions if $f_{t0}(h_{t-1})$ is an atomless Borel probability measure when Nature moves in the stage t .

Theorem 2. *If a continuous dynamic game with perfect information has atomless transitions, then it possesses a pure-strategy subgame-perfect equilibrium. In addition, E_t is nonempty and compact valued, and upper hemicontinuous on H_{t-1} for any $t \geq 1$.*

Remark 2. *As shown in [6], [9], [13], [16], [17], and [35], pure-strategy subgame-perfect equilibria exist in deterministic (i.e., without Nature) continuous dynamic games with perfect information. Theorem 2 extends those existence results to the case with Nature. We may point out that the condition of atomless transitions in Theorem 2 is minimal. In particular, the counterexample provided by [21] is a five-stage continuous dynamic game with perfect information and Nature, but without any subgame-perfect equilibrium (even in mixed strategies). In that example, the state transition in the third period has an atom. Thus, our condition of atomless transitions is violated.*

3.3 An extension

In this subsection, we extend the model of continuous dynamic games as specified in the previous two subsections. The aim is to combine the models of dynamic games with perfect and almost perfect information, and cover an important class of dynamic games, namely stochastic games. We show the existence of a subgame-perfect equilibrium such that whenever there is only one active player at some stage, the player can play pure strategy as part of the equilibrium strategies. As a byproduct, we obtain a new existence result for continuous stochastic games.

For concreteness, we shall allow for the case in which (1) the state transition depends on the action profile in the current stage as well as on the previous history, and (2) the players may have perfect information in some stages. The first modification covers the model of stochastic games as a special case. The second change allows us to combine the models of dynamic games with perfect and almost perfect information.

1. For each $t \geq 1$, the choice of Nature depends not only on the history h_{t-1} , but also on the action profile x_t in this stage. For any $t \geq 1$, suppose that A_{t0} is a continuous, nonempty and closed valued correspondence from $\text{Gr}(A_t)$ to S_t . Then $H_t = \text{Gr}(A_{t0})$, and H_∞ 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 $t \geq 0$.
2. Nature's action is given by a continuous mapping f_{t0} from $\text{Gr}(A_t)$ to $\mathcal{M}(S_t)$ such that $f_{t0}(A_{t0}(h_{t-1}, x_t)|h_{t-1}, x_t) = 1$ for all $(h_{t-1}, x_t) \in \text{Gr}(A_t)$.
3. For each $t \geq 1$, let

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

where $|K|$ represents the number of points in the set K . 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. If $N_t = 0$, then Nature moves in this stage.

We will drop the condition of atomless transition in those periods with only one active player in I .

Assumption 2 (Atomless Transitions'). 1. For any $t \geq 1$ with $N_t = 1$, S_t is a singleton set $\{s_t\}$.
 2. For each $t \geq 1$ with $N_t = 0$, $f_{t0}(h_{t-1}, x_t)$ is an atomless Borel probability measure for each $h_{t-1} \in H_{t-1}$ and $x_t \in A_t(h_{t-1})$.

The result on the equilibrium existence is presented below.

Proposition 1. *If a continuous dynamic game (as described above) satisfies the condition of atomless transitions', 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_{tj} can be chosen to be deterministic. In addition, E_t is nonempty and compact valued, and upper hemicontinuous on H_{t-1} for any $t \geq 1$.*

Proposition 1 provides a new existence result for continuous stochastic games.

Proposition 2. *If a continuous stochastic game has atomless transitions, then it possesses a subgame-perfect equilibrium.*

Remark 3. *Consider a standard stochastic game with uncountable states as in [24], where the existence of a subgame-perfect equilibrium was shown by assuming the state transitions to be norm continuous with respect to the actions in the previous stage, but not necessarily atomless. We show that a subgame-perfect equilibrium exists in a continuous stochastic game by only requiring the state transitions to be weakly continuous.*

4 Applications

4.1 A dynamic oligopoly model

In this subsection, we study a dynamic oligopoly market in which firms face stochastic demand/cost. Such a model is a variant of the well-known dynamic oligopoly models as considered in [12] and [30], which examined the response of firms for demand fluctuations.

We consider a dynamic oligopoly market in which n firms produce l goods in an infinite-horizon setting, where all the firms could face the demand shock, and also some technology shock that influences their production costs. In period t ,

1. firm i 's output is $q_{ti} \in [0, \gamma]^l$, and $Q_t = \sum_{i=1}^n q_{ti}$ is the aggregate output;
2. the inverse demand function is denoted by $P(Q_t, s_t)$, where $s_t \in [0, 1]$ is the demand shock;
3. the cost of firm i is $c_i(q_{ti}, r_t)$, which is determined by its output q_{ti} and the technology shock $r_t \in [0, 1]$;
4. the discount factor of firm i is $\beta_i \in [0, 1)$.

We assume that the inverse demand function P and the cost function c_i are bounded and continuous, $1 \leq i \leq n$.

The timing of events is as follows.

1. At the beginning of period t , all firms learn the realization of $(s_t, r_t) \in [0, 1] \times [0, 1]$.
 - (a) The demand shock s_t is determined by

$$Prob(s_t = \psi_H(s_{t-1}, Q_{t-1}) | s_{t-1}, Q_{t-1}) = \alpha$$

and

$$Prob(s_t = \psi_L(s_{t-1}, Q_{t-1}) | s_{t-1}, Q_{t-1}) = 1 - \alpha,$$

where $0 \leq \alpha \leq 1$, ψ_H and ψ_L are two continuous functions with $0 \leq \psi_L \leq \psi_H \leq 1$, and H (resp. L) represents boom (resp. recession).

- (b) The technology shock r_t is determined by the cumulative distribution function $\kappa(\cdot | r_{t-1}, Q_{t-1})$. Suppose that $\kappa(\cdot | r_{t-1}, Q_{t-1})$ is a continuous function.
2. Firms then simultaneously choose the level of their output $q_t = (q_{t1}, \dots, q_{tn})$, where $q_{ti} \in A_{ti}(s_t, r_t) \subseteq \mathbb{R}^l$ for $i = 1, 2, \dots, n$. In particular, the correspondence A_{ti} is continuous, nonempty and compact valued, which gives the available actions of firm i .
3. The strategic choices of all the firms then become common knowledge and this one-period game is repeated.

In our model, the state in period t has two components s_t and r_t . The notation s_t denotes the demand shock that represents some exogenous shock to the market demand, and r_t is the technology shock that may affect firms' production efficiency. The market may be in two status, boom or recession, which is represented by H and L respectively. When the market is in boom (resp. recession), the demand shock is determined by the function ψ_H (resp. ψ_L). In addition, the probability for the market to be in boom is α .

1. In period t , given the shock (s_t, r_t) and the output $q_t = (q_{t1}, \dots, q_{tn})$, the payoff of firm i is

$$u_i(q_t, s_t, r_t) = [P(Q_t, s_t) - c_i(q_{ti}, r_t)] q_{ti}.$$

2. Given a sequence of outputs $\{q_t\}_{t \geq 1}$ and shocks $\{(s_t, r_t)\}_{t \geq 1}$, firm i receives the payoff

$$\sum_{t=1}^{\infty} \beta_i^{t-1} u_i(q_t, s_t, r_t).$$

Remark 4. *The dynamic oligopoly market model can be viewed as a stochastic game. Since the transition of the technology shock is determined by a continuous cumulative distribution function, the model has atomless transitions. However, as the transition of the demand shock concentrates on $\psi_H(s_{t-1}, Q_{t-1})$ and $\psi_L(s_{t-1}, Q_{t-1})$ which could be any points in the unit interval, the state transition may not be norm continuous. As a result, the previous equilibrium existence results (e.g., [24]) are not applicable. Since ψ_H , ψ_L and κ are all continuous functions, the state transition is weakly continuous.*

The following result is an immediate application of Proposition 1.

Proposition 3. *The dynamic oligopoly model has a subgame-perfect equilibrium.*

4.2 Intergenerational bequests with stochastic production

In this subsection, we consider a model of altruistic growth/intergenerational strategic bequests. This model was introduced in [28], and has been widely adopted to study economic growth and wealth distribution.

Consider the simplest case of a one-good economy with production, in which there is a sequence of generations. Each generation lives for one period, and cares about its own consumption as well as that of its immediate successor. Starting with an initial endowment e_t , a given generation t decides its consumption c_t . The rest, $e_t - c_t$, is the input into the production process. The output of the process is the next generation $t + 1$'s endowment e_{t+1} .

Let $U_t: \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ be the utility function of generation t , and $P: \mathbb{R}_+ \rightarrow \mathcal{M}(\mathbb{R}_+)$ be the probability distribution of the production output. Given the history $h_{t-1} = (e_1, c_1, \dots, e_{t-1}, c_{t-1})$ and endowment e_t , the expected payoff of the generation t consuming $c_t \in A(e_t) = [0, e_t]$, when the next generation $t + 1$ follows the pure

strategy (consumption policy) f_{t+1} , is given by

$$\int_{\mathbb{R}_+} U_t[c_t, f_{t+1}(h_t, e_{t+1})] dP(e_{t+1}|e_t - c_t),$$

where $h_t = (h_{t-1}, e_t, c_t)$.

We make the following assumptions.

Assumption 3. 1. For each t , U_t is a bounded continuous function.

2. The random production function $P(\cdot|e_t - c_t)$ takes atomless distributions as its values, and is continuous in $e_t - c_t$ with respect to the topology of weak convergence.

Remark 5. The model described above is an alternating move dynamic game with short-lived players. If the production function P is deterministic, then it can be modified as a dynamic game with perfect information and no Nature. In particular, given the history h_{t-1} , endowment e_t and consumption c_t , the endowment e_{t+1} is uniquely determined when P is deterministic, and hence the consumption policy f_{t+1} of the next generation $t + 1$ only depends on h_t . In this case, the payoff of generation t can be written as $U_t[c_t, f_{t+1}(h_t)]$. The equilibrium existence problem with such deterministic transitions has been studied by [4] and [20] in the stationary environment. In addition, the existence of pure-strategy subgame-perfect equilibria in such an intergenerational bequest game with perfect information and no Nature follows from the equilibrium existence results in [6], [13], [16] and [17]. The main departure of our model is to allow for the stochastic production.

The result below follows from Theorem 2.

Proposition 4. The intergenerational bequest model has a pure-strategy subgame-perfect equilibrium.

5 Measurable dynamic games

In this section, we consider the more general setting in which the model parameters are continuous in the action variables, but may only be measurable with respect to the state variables. The proofs of the results in this section are left in Appendix B.

In Subsection 5.1, we adopt the model specified in Subsection 3.1 and relax the continuity requirement in the state variable. To obtain the existence of subgame-perfect equilibria, we impose the condition ‘‘atomless reference measure (ARM)’’ on the state transitions. This condition means that in each stage, there is an atomless reference measure and the state transition is absolutely continuous with respect to this reference measure.

In Subsection 5.2, we consider dynamic games with perfect information whose parameters are continuous in actions and measurable in states. The ARM type condition is imposed only when Nature moves. We show the existence of *pure-strategy* subgame-perfect equilibria.

In Appendix B, we present a further extension by partially relaxing the ARM condition in two ways. First, we allow the possibility that there is only one active player (but no Nature) at some stages, where the ARM type condition is dropped. Second, we introduce an additional weakly continuous component on the state transitions at any other stages. In addition, we allow the state transition in each period to depend on the current actions as well as on the previous history. As the generalization of the model in Subsection 3.3, (1) we combine the models for measurable dynamic games with perfect and almost perfect information, (2) we show the existence of subgame perfect equilibria such that whenever there is only one active player at some stage, the player can play pure strategy as part of the equilibrium strategies, and (3) a new existence result is obtained for stochastic games.

5.1 Measurable dynamic games with almost perfect information

We will follow the setting and notations in Subsection 3.1 as closely as possible, and only describe the changes we need to make on the model.

1. For any $t \geq 1$ and $i \in I$, A_{ti} is sectionally continuous on X^{t-1} .¹⁶
2. For any $t \geq 1$, f_{t0} is sectionally continuous on X^{t-1} .
3. For each $i \in I$, the payoff function u_i is sectionally continuous on X^∞ .

For each $t \geq 0$, suppose that λ_t is a Borel probability measure on S_t and λ_t is atomless for $t \geq 1$. Let $\lambda^t = \otimes_{0 \leq k \leq t} \lambda_k$ for $t \geq 0$. We shall assume the following condition on the state transitions.

Assumption 4 (Atomless Reference Measure (ARM)). *A dynamic game is said to satisfy the “atomless reference measure (ARM)” condition if for each $t \geq 1$,*

1. *the probability $f_{t0}(\cdot|h_{t-1})$ is absolutely continuous with respect to λ_t on S_t with the Radon-Nikodym derivative $\varphi_{t0}(h_{t-1}, s_t)$ for all $h_{t-1} \in H_{t-1}$;*¹⁷

¹⁶Suppose that Y_1 , Y_2 and Y_3 are all Polish spaces, and $Z \subseteq Y_1 \times Y_2$. Denote $Z(y_1) = \{y_2 \in Y_2: (y_1, y_2) \in Z\}$ for any $y_1 \in Y_1$. A function (resp. correspondence) $f: Z \rightarrow Y_3$ is said to be sectionally continuous on Y_2 if $f(y_1, \cdot)$ is continuous on $Z(y_1)$ for all y_1 with $Z(y_1) \neq \emptyset$. Similarly, one can define the sectional upper hemicontinuity for a correspondence.

¹⁷It is common to have a reference measure when one considers a game with uncountable states. For example, if S_t is a subset of \mathbb{R}^l , then the Lebesgue measure is a natural reference measure.

2. the mapping φ_{t0} is Borel measurable and sectionally continuous on X^{t-1} , and integrably bounded in the sense that there is a λ_t -integrable function $\phi_t: S_t \rightarrow \mathbb{R}_+$ such that $\varphi_{t0}(h_{t-1}, s_t) \leq \phi_t(s_t)$ for any $h_{t-1} \in H_{t-1}$ and $s_t \in S_t$.

When one considers a dynamic game with infinite horizon, the following “continuity at infinity” condition is standard.¹⁸ In particular, all discounted repeated games or stochastic games satisfy this condition.

For any $T \geq 1$, let

$$w^T = \sup_{\substack{i \in I \\ (x,s) \in H_\infty \\ (\bar{x}, \bar{s}) \in H_\infty \\ x^{T-1} = \bar{x}^{T-1} \\ s^{T-1} = \bar{s}^{T-1}}} |u_i(x, s) - u_i(\bar{x}, \bar{s})|. \quad (1)$$

Assumption 5 (Continuity at Infinity). *A dynamic game is said to be “continuous at infinity” if $w^T \rightarrow 0$ as $T \rightarrow \infty$.*

We shall modify the notion of subgame-perfect equilibrium slightly. In particular, when the state space is uncountable and has a reference measure, it is natural to consider the optimality for almost all sub-histories in the probabilistic sense; see, for example, [1] and Footnote 4 therein.

Definition 3 (SPE). *A subgame-perfect equilibrium is a strategy profile f such that for all $i \in I$, $t \geq 0$, and λ^t -almost all $h_t \in H_t$,¹⁹ player i cannot improve his payoff in the subgame h_t by a unilateral change in his strategy.*

The following theorem shows the existence of a subgame-perfect equilibrium under the conditions of ARM and continuity at infinity. Recall that $E_t(h_{t-1})$ is the set of all subgame-perfect equilibrium payoffs in the subgame h_{t-1} . The theorem also shows the compactness and upper hemicontinuity properties of the correspondence E_t .

Theorem 3. *If a dynamic game with almost perfect information satisfies the ARM condition and is continuous at infinity, then it possesses a subgame-perfect equilibrium. In addition, E_t is nonempty and compact valued, and essentially sectionally upper hemicontinuous on X^{t-1} .²⁰*

¹⁸See, for example, [9].

¹⁹A property is said to hold for λ^t -almost all $h_t = (x^t, s^t) \in H_t$ if it is satisfied for λ^t -almost all $s^t \in S^t$ and all $x^t \in H_t(s^t)$.

²⁰Suppose that Y_1, Y_2 and Y_3 are all Polish spaces, and $Z \subseteq Y_1 \times Y_2$ and η is a Borel probability measure on Y_1 . Denote $Z(y_1) = \{y_2 \in Y_2: (y_1, y_2) \in Z\}$ for any $y_1 \in Y_1$. A function (resp. correspondence) $f: Z \rightarrow Y_3$ is said to be essentially sectionally continuous on Y_2 if $f(y_1, \cdot)$ is continuous on $Z(y_1)$ for η -almost all y_1 . Similarly, one can define the essential sectional upper hemicontinuity for a correspondence.

5.2 Measurable dynamic games with perfect information

In this subsection, we consider dynamic games with perfect information (with or without Nature). We will follow the setting and notations in Subsection 3.2, and make the same changes as those in Subsection 5.1. In particular, the continuity requirement in the state variables are dropped.

In dynamic games with perfect information where players move sequentially, we show the existence of pure-strategy subgame-perfect equilibria. The ARM condition is imposed when Nature moves, and is dropped in those periods with one active player from the set I .

Theorem 4. *If a dynamic game with perfect information satisfies the ARM condition and is continuous at infinity, then it possesses a pure-strategy subgame-perfect equilibrium. In addition, E_t is nonempty and compact valued, and essentially sectionally upper hemicontinuous on X^{t-1} for any $t \geq 1$.*

6 Appendix A

In Subsection 6.1, we present several lemmas as the mathematical preparations for proving Theorems 1, 2 and Proposition 1. Since correspondences will be used extensively in the proofs, we collect, for the convenience of the reader, several known results on various properties of correspondences.²¹ One can skip Subsection 6.1 first and go to the proofs in Subsections 6.2-6.4 directly, and refer to those technical lemmas in Subsection 6.1 whenever necessary.

The proof of Theorem 1 is provided in Subsection 6.2. In Subsections 6.3 and 6.4, we give the proofs of Theorem 2 and Proposition 1, respectively. We will only describe the necessary changes in comparison with the proofs presented in Subsections 6.2.1-6.2.3.

6.1 Technical preparations

Let (S, \mathcal{S}) be a measurable space and X a topological space with its Borel σ -algebra $\mathcal{B}(X)$. A correspondence Ψ from S to X is a function from S to the space of all subsets of X . The upper inverse Ψ^u of a subset $A \subseteq X$ is

$$\Psi^u(A) = \{s \in S : \Psi(s) \subseteq A\}.$$

²¹These technical lemmas are stated in a general form so that they can still be used in the proofs in Appendix B for case of measurable dynamic games.

The lower inverse Ψ^l of a subset $A \subseteq X$ is

$$\Psi^l(A) = \{s \in S : \Psi(s) \cap A \neq \emptyset\}.$$

The correspondence Ψ is

1. weakly measurable, if $\Psi^l(O) \in \mathcal{S}$ for each open subset $O \subseteq X$;
2. measurable, if $\Psi^l(K) \in \mathcal{S}$ for each closed subset $K \subseteq X$.

The graph of Ψ is denoted by $\text{Gr}(\Psi) = \{(s, x) \in S \times X : s \in S, x \in \Psi(s)\}$. The correspondence Ψ is said to have a measurable graph if $\text{Gr}(\Psi) \in \mathcal{S} \otimes \mathcal{B}(X)$.

If S is a topological space, then Ψ is

1. upper hemicontinuous, if $\Psi^u(O)$ is open for each open subset $O \subseteq X$;
2. lower hemicontinuous, if $\Psi^l(O)$ is open for each open subset $O \subseteq X$.

The following two lemmas present some basic measurability and continuity properties for correspondences.

Lemma 1. *Let (S, \mathcal{S}) be a measurable space, X a Polish space endowed with the Borel σ -algebra $\mathcal{B}(X)$, and \mathcal{K} the space of nonempty compact subsets of X endowed with its Hausdorff metric topology. Suppose that $\Psi: S \rightarrow X$ is a nonempty and closed valued correspondence.*

1. *If Ψ is weakly measurable, then it has a measurable graph.*
2. *If Ψ is compact valued, then the following statements are equivalent.*
 - (a) *The correspondence Ψ is weakly measurable.*
 - (b) *The correspondence Ψ is measurable.*
 - (c) *The function $f: S \rightarrow \mathcal{K}$, defined by $f(s) = \Psi(s)$, is Borel measurable.*
3. *Suppose that S is a topological space. If Ψ is compact valued, then the function $f: S \rightarrow \mathcal{K}$ defined by $f(s) = \Psi(s)$ is continuous if and only if the correspondence Ψ is continuous.*
4. *Suppose that $(S, \mathcal{S}, \lambda)$ is a complete probability space. Then Ψ is \mathcal{S} -measurable if and only if it has a measurable graph.*
5. *For a correspondence $\Psi: S \rightarrow X$ between two Polish spaces, the following statements are equivalent.*
 - (a) *The correspondence Ψ is upper hemicontinuous at a point $s \in S$ and $\Psi(s)$ is compact.*
 - (b) *If a sequence (s_n, x_n) in the graph of Ψ satisfies $s_n \rightarrow s$, then the sequence $\{x_n\}$ has a limit in $\Psi(s)$.*

6. For a correspondence $\Psi: S \rightarrow X$ between two Polish spaces, the following statements are equivalent.

- (a) The correspondence Ψ is lower hemicontinuous at a point $s \in S$.
- (b) If $s_n \rightarrow s$, then for each $x \in \Psi(s)$, there exist a subsequence $\{s_{n_k}\}$ of $\{s_n\}$ and elements $x_k \in \Psi(s_{n_k})$ for each k such that $x_k \rightarrow x$.

7. Given correspondences $F: X \rightarrow Y$ and $G: Y \rightarrow Z$, the composition F and G is defined by

$$G(F(x)) = \cup_{y \in F(x)} G(y).$$

The composition of upper hemicontinuous correspondences is upper hemicontinuous. The composition of lower hemicontinuous correspondences is lower hemicontinuous.

Proof. Properties (1), (2), (3), (5), (6) and (7) are Theorems 18.6, 18.10, 17.15, 17.20, 17.21 and 17.23 of [3], respectively. Property (4) follows from Proposition 4 in page 61 of [18]. \square

Lemma 2. 1. A correspondence Ψ from a measurable space (S, \mathcal{S}) into a topological space X is weakly measurable if and only if its closure correspondence $\overline{\Psi}$ is weakly measurable, where for each $s \in S$, $\overline{\Psi}(s) = \overline{\Psi(s)}$ and $\overline{\Psi(s)}$ is the closure of the set $\Psi(s)$ in X .

2. For a sequence $\{\Psi_m\}$ of correspondences from a measurable space (S, \mathcal{S}) into a Polish space, the union correspondence $\Psi(s) = \cup_{m \geq 1} \Psi_m(s)$ is weakly measurable if each Ψ_m is weakly measurable. If each Ψ_m is weakly measurable and compact valued, then the intersection correspondence $\Phi(s) = \cap_{m \geq 1} \Psi_m(s)$ is weakly measurable.

3. A weakly measurable, nonempty and closed valued correspondence from a measurable space into a Polish space admits a measurable selection.

4. A correspondence with closed graph between compact metric spaces is measurable.

5. A nonempty and compact valued correspondence Ψ from a measurable space (S, \mathcal{S}) into a Polish space is weakly measurable if and only if there exists a sequence $\{\psi_1, \psi_2, \dots\}$ of measurable selections of Ψ such that $\Psi(s) = \overline{\{\psi_1(s), \psi_2(s), \dots\}}$ for each $s \in S$.

6. The image of a compact set under a compact valued upper hemicontinuous correspondence is compact.²² If the domain is compact, then the graph of a compact valued upper hemicontinuous correspondence is compact.

²²Given a correspondence $F: X \rightarrow Y$ and a subset A of X , the image of A under F is defined to be the set $\cup_{x \in A} F(x)$.

7. *The intersection of a correspondence with closed graph and an upper hemicontinuous compact valued correspondence is upper hemicontinuous.*
8. *If the correspondence $\Psi: S \rightarrow \mathbb{R}^l$ is compact valued and upper hemicontinuous, then the convex hull of Ψ is also compact valued and upper hemicontinuous.*

Proof. Properties (1)-(7) are Lemmas 18.3 and 18.4, Theorems 18.13 and 18.20, Corollary 18.15, Lemma 17.8 and Theorem 17.25 in [3], respectively. Property 8 is Proposition 6 in page 26 of [18]. \square

Parts 1 and 2 of the following lemma are the standard Lusin's Theorem and Michael's continuous selection theorem, while the other two parts are about properties involving mixture of measurability and continuity of correspondences.

- Lemma 3.**
1. *Lusin's Theorem: Suppose that S is a Borel subset of a Polish space, λ is a Borel probability measure on S and \mathcal{S} is the completion of $\mathcal{B}(S)$ under λ . Let X be a Polish space. If f is an \mathcal{S} -measurable mapping from S to X , then for any $\epsilon > 0$, there exists a compact subset $S_1 \subseteq S$ with $\lambda(S \setminus S_1) < \epsilon$ such that the restriction of f to S_1 is continuous.*
 2. *Michael's continuous selection theorem: Let S be a metrizable space, and X a complete metrizable closed subset of some locally convex space. Suppose that $F: S \rightarrow X$ is a lower hemicontinuous, nonempty, convex and closed valued correspondence. Then there exists a continuous mapping $f: S \rightarrow X$ such that $f(s) \in F(s)$ for all $s \in S$.*
 3. *Let $(S, \mathcal{S}, \lambda)$ be a finite measure space, X a Polish space, and Y a locally convex linear topological space. Let $F: S \rightarrow X$ be a closed-valued correspondence such that $\text{Gr}(F) \in \mathcal{S} \otimes \mathcal{B}(X)$, and $f: \text{Gr}(F) \rightarrow Y$ a measurable function which is sectionally continuous on X . Then there exists a measurable function $f': S \times X \rightarrow Y$ such that (1) f' is sectionally continuous on X , (2) for λ -almost all $s \in S$, $f'(s, x) = f(s, x)$ for all $x \in F(s)$ and $f'(s, X) \subseteq \text{cof}(s, F(s))$.²³*
 4. *Let (S, \mathcal{S}) be a measurable space, and X and Y Polish spaces. Let $\Psi: S \times X \rightarrow \mathcal{M}(Y)$ be an $\mathcal{S} \otimes \mathcal{B}(X)$ -measurable, nonempty, convex and compact valued correspondence which is sectionally continuous on X , where the compactness and continuity are with respect to the weak* topology on $\mathcal{M}(Y)$. Then there exists an $\mathcal{S} \otimes \mathcal{B}(X)$ -measurable selection ψ of Ψ that is sectionally continuous on X .*

²³For any set A in a linear topological space, $\text{co}A$ denotes the convex hull of A .

Proof. Lusin's theorem is Theorem 7.1.13 in [5]. Michael's continuous selection theorem can be found in [25] and the last paragraph of page 228 of [5]. Property (3) is Theorem 2.7 in [8]. Property (4) follows from Theorem 1 and the main lemma of [19]. \square

The following lemma presents the convexity, compactness and continuity properties on integration of correspondences.

Lemma 4. *Let $(S, \mathcal{S}, \lambda)$ be an atomless probability space, X a Polish space, and F a correspondence from S to \mathbb{R}^l . Denote*

$$\int_S F(s)\lambda(ds) = \left\{ \int_S f(s)\lambda(ds) : f \text{ is an integrable selection of } F \text{ on } S \right\}.$$

1. *If F is measurable, nonempty and closed valued, and λ -integrably bounded by some integrable function $\psi: S \rightarrow \mathbb{R}_+$ in the sense that for λ -almost all $s \in S$, $\|y\| \leq \psi(s)$ for any $y \in F(s)$, then $\int_S F(s)\lambda(ds)$ is nonempty, convex and compact, and*

$$\int_S F(s)\lambda(ds) = \int_S \text{co}F(s)\lambda(ds).$$

2. *If G is a measurable, nonempty and closed valued correspondence from $S \times X \rightarrow \mathbb{R}^l$ such that (1) $G(s, \cdot)$ is upper (resp. lower) hemicontinuous on X for all $s \in S$, and (2) G is λ -integrably bounded by some integrable function $\psi: S \rightarrow \mathbb{R}_+$ in the sense that for λ -almost all $s \in S$, $\|y\| \leq \psi(s)$ for any $x \in X$ and $y \in G(s, x)$, then $\int_S G(s, x)\lambda(ds)$ is upper (resp. lower) hemicontinuous on X .*

Proof. See Theorems 2, 3 and 4, Propositions 7 and 8, and Problem 6 in Section D.II.4 of [18]. \square

The following result proves a measurable version of Lyapunov's theorem, which is taken from [23]. Let (S, \mathcal{S}) and (X, \mathcal{X}) be measurable spaces. A transition probability from S to X is a mapping f from S to the space $\mathcal{M}(X)$ of probability measures on (X, \mathcal{X}) such that $f(B|\cdot) : s \rightarrow f(B|s)$ is \mathcal{S} -measurable for each $B \in \mathcal{X}$.

Lemma 5. *Let $f(\cdot|s)$ be a transition probability from a measurable space (S, \mathcal{S}) to another measurable space (X, \mathcal{X}) (\mathcal{X} is separable).²⁴ Let Q be a measurable, nonempty and compact valued correspondence from $S \times X$ to \mathbb{R}^l , which is f -integrable in the sense that for any measurable selection q of Q , $q(\cdot, s)$ is $f(\cdot|s)$ -absolutely integrable for any $s \in S$. Let $\int Q df$ be the correspondence from S to*

²⁴A σ -algebra is said to be separable if it is generated by a countable collection of sets.

subsets of \mathbb{R}^l defined by

$$M(s) = \left(\int Q \, df \right) (s) = \left\{ \int_X q(s, x) f(dx|s) : q \text{ is a measurable selection of } Q \right\}.$$

Denote the graph of M by J . Let \mathcal{J} be the restriction of the product σ -algebra $\mathcal{S} \otimes \mathcal{B}(\mathbb{R}^l)$ to J .

Then

1. M is a measurable, nonempty and compact valued correspondence;
2. there exists a measurable, \mathbb{R}^l -valued function g on $(X \times J, \mathcal{X} \otimes \mathcal{J})$ such that $g(x, e, s) \in Q(x, s)$ and $e = \int_X g(x, e, s) f(dx|s)$.

The following result is Lemma 6 in [29].

Lemma 6. *Suppose that H and X are compact metric spaces. Let $P: H \times X \rightarrow \mathbb{R}^n$ be a nonempty valued and upper hemicontinuous correspondence, and the mappings $f: H \rightarrow \mathcal{M}(X)$ and $\mu: H \rightarrow \Delta(X)$ be measurable. In addition, suppose that $\mu(\cdot|h) = p(h, \cdot) \circ f(\cdot|h)$ such that $p(h, \cdot)$ is a measurable selection of $P(h, \cdot)$.²⁵ Then there exists a jointly Borel measurable selection g of P such that $\mu(\cdot|h) = g(h, \cdot) \circ f(\cdot|h)$; that is, $g(h, x) = p(h, x)$ for $f(\cdot|h)$ -almost all x .*

Suppose that (S_1, \mathcal{S}_1) is a measurable space, S_2 is a Polish space endowed with the Borel σ -algebra, and $S = S_1 \times S_2$ which is endowed with the product σ -algebra \mathcal{S} . Let D be an \mathcal{S} -measurable subset of S such that $D(s_1)$ is compact for any $s_1 \in S_1$. The σ -algebra \mathcal{D} is the restriction of \mathcal{S} on D . Let X be a Polish space, and A a \mathcal{D} -measurable, nonempty and closed valued correspondence from D to X which is sectionally continuous on S_2 . The following lemma considers the property of upper hemicontinuity for the correspondence M as defined in Lemma 5.

Lemma 7. *Let $f(\cdot|s)$ be a transition probability from (D, \mathcal{D}) to $\mathcal{M}(X)$ such that $f(A(s)|s) = 1$ for any $s \in D$, which is sectionally continuous on S_2 . Let G be a bounded, measurable, nonempty, convex and compact valued correspondence from $Gr(A)$ to \mathbb{R}^l , which is sectionally upper hemicontinuous on $S_2 \times X$. Let $\int G \, df$ be the correspondence from D to subsets of \mathbb{R}^l defined by*

$$M(s) = \left(\int G \, df \right) (s) = \left\{ \int_X g(s, x) f(dx|s) : g \text{ is a measurable selection of } G \right\}.$$

Then M is \mathcal{S} -measurable, nonempty and compact valued, and sectionally upper hemicontinuous on S_2 .

²⁵The finite measure $\mu(\cdot|h) = p(h, \cdot) \circ f(\cdot|h)$ if $\mu(E|h) = \int_E p(h, x) f(dx|h)$ for any Borel set E .

Proof. Define a correspondence $\tilde{G}: S \times X \rightarrow \mathbb{R}^l$ as

$$\tilde{G} = \begin{cases} G(s, x), & \text{if } (s, x) \in \text{Gr}(A); \\ \{0\}, & \text{otherwise.} \end{cases}$$

Then $M(s) = \left(\int \tilde{G} \, d f \right) (s) = \left(\int G \, d f \right) (s)$. The measurability, nonemptiness and compactness follows from Lemma 5. Given $s_1 \in S_1$ such that (1) $D(s_1) \neq \emptyset$, (2) $f(s_1, \cdot)$ and $G(s_1, \cdot, \cdot)$ is upper hemicontinuous. The upper hemicontinuity of $M(s_1, \cdot)$ follows from Lemma 2 in [32] and Lemma 4 in [29]. \square

From now onwards, whenever we work with mappings taking values in the space $\mathcal{M}(X)$ of Borel probability measures on some separable metric space X , the relevant continuity or convergence is assumed to be in terms of the topology of weak convergence of measures on $\mathcal{M}(X)$ unless otherwise noted.

In the following lemma, we state some properties for transition correspondences.

Lemma 8. *Suppose that Y and Z are Polish spaces. Let G be a measurable, nonempty, convex and compact valued correspondence from Y to $\mathcal{M}(Z)$. Define a correspondence G' from $\mathcal{M}(Y)$ to $\mathcal{M}(Z)$ as*

$$G'(\nu) = \left\{ \int_Y g(y) \nu(dy) : g \text{ is a Borel measurable selection of } G \right\}.^{26}$$

1. *The correspondence G' is measurable, nonempty, convex and compact valued.*
2. *The correspondence G is upper hemicontinuous if and only if G' is upper hemicontinuous. In addition, if G is continuous, then G' is continuous.*

Proof. (1) is Lemma 19.29 of [3]. By Theorem 19.30 therein, G is upper hemicontinuous if and only if G' is upper hemicontinuous. We need to show that G' is lower hemicontinuous if G is lower hemicontinuous.

Let Z be endowed with a totally bounded metric, and $U(Z)$ the space of bounded, real-valued and uniformly continuous functions on Z endowed with the supremum norm, which is obviously separable. Pick a countable set $\{f_m\}_{m \geq 1} \subseteq U(Z)$ such that $\{f_m\}$ is dense in the unit ball of $U(Z)$. It follows from Theorem 15.2 of [3] that the weak* topology of $\mathcal{M}(Z)$ is metrizable by the metric d_z , where

$$d_z(\mu_1, \mu_2) = \sum_{m=1}^{\infty} \frac{1}{2^m} \left| \int_Z f_m(z) \mu_1(dz) - \int_Z f_m(z) \mu_2(dz) \right|$$

²⁶The integral $\int_Y g(y) \nu(dy)$ defines a Borel probability measure τ on Z such that for any Borel set C in Z , $\tau(C) = \int_Y g(y)(C) \nu(dy)$. The measure τ is also equal to the Gelfand integral of g with respect to the measure ν on Y , when $\mathcal{M}(Z)$ is viewed as a set in the dual space of the space of bounded continuous functions on Z ; see Definition 11.49 of [3].

for each pair of $\mu_1, \mu_2 \in \mathcal{M}(Z)$.

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 $\mu_0 \in G'(\nu_0)$. By the definition of G' , there exists a Borel measurable selection g of G such that $\mu_0 = \int_Y g(y)\nu_0(dy)$.

For each $k \geq 1$, by Lemma 3 (Lusin's theorem), there exists a compact subset $D_k \subseteq Y$ such that g is continuous on D_k and $\nu_0(Y \setminus D_k) < \frac{1}{3k}$. Define a correspondence $G_k: Y \rightarrow \mathcal{M}(X)$ as follows:

$$G_k(y) = \begin{cases} \{g(y)\}, & y \in D_k; \\ G(y), & y \in Y \setminus D_k. \end{cases}$$

Then G_k is nonempty, convex and compact valued, and lower hemicontinuous. By Theorem 3.22 in [3], Y is paracompact. Then by Lemma 3 (Michael's selection theorem), it has a continuous selection g_k .

For each k , since $\nu_j \rightarrow \nu_0$ and g_k is continuous, $\int_Y g_k(y)\nu_j(dy) \rightarrow \int_Y g_k(y)\nu_0(dy)$ in the sense that for any $m \geq 1$,

$$\int_Y \int_Z f_m(z)g_k(dz|y)\nu_j(dy) \rightarrow \int_Y \int_Z f_m(z)g_k(dz|y)\nu_0(dy).$$

Thus, there exists a point ν_{j_k} such that $\{j_k\}$ is an increasing sequence and

$$d_z \left(\int_Y g_k(y)\nu_{j_k}(dy), \int_Y g_k(y)\nu_0(dy) \right) < \frac{1}{3k}.$$

In addition, since g_k coincides with g on D_k and $\nu_0(Y \setminus D_k) < \frac{1}{3k}$,

$$d_z \left(\int_Y g_k(y)\nu_0(dy), \int_Y g(y)\nu_0(dy) \right) < \frac{2}{3k}.$$

Thus,

$$d_z \left(\int_Y g_k(y)\nu_{j_k}(dy), \int_Y g(y)\nu_0(dy) \right) < \frac{1}{k}.$$

Let $\mu_{j_k} = \int_Y g_k(y)\nu_{j_k}(dy)$ for each k . Then $\mu_{j_k} \in G'(\nu_{j_k})$ and $\mu_{j_k} \rightarrow \mu_0$ as $k \rightarrow \infty$. By Lemma 1, G' is lower hemicontinuous. \square

The next lemma presents some properties for the composition of two transition correspondences in terms of the product of transition probabilities.

Lemma 9. *Let X, Y and Z be Polish spaces, and G 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 $\mathcal{M}(Z)$. Define*

a correspondence Π from X to $\mathcal{M}(Y \times Z)$ as follows:

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

1. If F is sectionally continuous on Y , then Π is a measurable, nonempty and compact valued correspondence.
2. If there exists a function g from X to $\mathcal{M}(Y)$ such that $G(x) = \{g(x)\}$ for any $x \in X$, then Π is a measurable, nonempty and compact valued correspondence.
3. If both G and F are continuous correspondences, then Π is a nonempty and compact valued, and continuous correspondence.²⁷
4. If $G(x) \equiv \{\lambda\}$ for some fixed Borel probability measure $\lambda \in \mathcal{M}(Y)$ and F is sectionally continuous on X , then Π is a continuous, nonempty and compact valued correspondence.

Proof. (1) Define three correspondences $\tilde{F}: X \times Y \rightarrow \mathcal{M}(Y \times Z)$, $\hat{F}: \mathcal{M}(X \times Y) \rightarrow \mathcal{M}(Y \times Z)$ and $\check{F}: X \times \mathcal{M}(Y) \rightarrow \mathcal{M}(Y \times Z)$ as follows:

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

Since F is measurable, nonempty, convex and compact valued, \tilde{F} is measurable, nonempty, convex and compact valued. By Lemma 8, the correspondence \hat{F} is measurable, nonempty, convex and compact valued, and $\check{F}(x, \cdot)$ is continuous on $\mathcal{M}(Y)$ for any $x \in X$.

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 (5). For each $k \geq 1$, define a correspondence Π^k from X to $\mathcal{M}(Y \times Z)$ by letting $\Pi^k(x) = \check{F}(x, g_k(x)) = \hat{F}(x \otimes g_k(x))$. Then Π^k is measurable, nonempty, convex and compact valued.

Fix any $x \in X$. It is clear that $\Pi(x) = \check{F}(x, G(x))$ is a nonempty valued. Since $G(x)$ is compact, and $\check{F}(x, \cdot)$ is compact valued and continuous, $\Pi(x)$ is compact by Lemma 2. Thus, the closure $\overline{\bigcup_{k \geq 1} \Pi^k(x)}$ of $\bigcup_{k \geq 1} \Pi^k(x)$ is a subset of $\Pi(x)$.

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

²⁷In Lemma 29 of [15], they showed that Π is upper hemicontinuous if both G and F are upper hemicontinuous.

such that $g_{k_m}(x) \rightarrow \nu$. As $\check{F}(x, \cdot)$ is continuous, $\check{F}(x, g_{k_m}(x)) \rightarrow \check{F}(x, \nu)$. That is,

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

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

(2) As in (1), the correspondence \hat{F} is measurable, nonempty, convex and compact valued. If $G = \{g\}$ for some measurable function g , then $\Pi(x) = \hat{F}(\delta_x \otimes g(x))$, which is measurable, nonempty and compact valued.

(3) We continue to work with the two correspondences $\check{F}: X \times Y \rightarrow \mathcal{M}(Y \times Z)$ and $\hat{F}: \mathcal{M}(X \times Y) \rightarrow \mathcal{M}(Y \times Z)$ as in Part (1). By the condition on F , it is obvious that the correspondence \check{F} is continuous, nonempty, convex and compact valued. Lemma 8 implies the properties for the correspondence \hat{F} . Define a correspondence $\hat{G}: X \rightarrow \mathcal{M}(X \times Y)$ as $\hat{G}(x) = \delta_x \otimes G(x)$. 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, Π is compact valued by Lemma 2. As \hat{G} and \hat{F} are both continuous, Π is continuous by Lemma 1 (7).

(4) Let $Y' = Y$ and define a correspondence $\check{F}: X \times Y \rightarrow \mathcal{M}(Y' \times Z)$ as follows:

$$\check{F}(x, y) = \{\delta_y \otimes \mu : \mu \in F(x, y)\}.$$

Then \check{F} is also measurable, nonempty, convex and compact valued, and sectionally upper hemicontinuous on X .

Let d be a totally bounded metric on $Y' \times Z$, and $U(Y' \times Z)$ the space of bounded, real-valued and uniformly continuous functions on $Y' \times Z$ endowed with the supremum norm, which is obviously separable. It follows from Theorem 15.2 of [3] that the space of Borel probability measures on $Y' \times Z$ with the topology of weak convergence of measures can be viewed as a subspace of the dual space of $U(Y' \times Z)$ with the weak* topology. By Corollary 18.37 of [3], $\Pi(x) = \int_Y \check{F}(x, y) \lambda(dy)$ is nonempty, convex and compact for any $x \in X$.²⁸

Now we shall show the upper hemicontinuity. If $x_n \rightarrow x_0$ and $\mu_n \in \Pi(x_n)$, we need to prove that there exists some $\mu_0 \in \Pi(x_0)$ such that a subsequence of $\{\mu_n\}$ weakly converges to μ_0 . Suppose that for $n \geq 1$, f_n is a Borel measurable selection of $F(x_n, \cdot)$ such that $\mu_n = \lambda \diamond f_n$.

²⁸Note that the integral $\int_Y \check{F}(x, y) \lambda(dy)$ can be viewed as the Gelfand integral of the correspondence in the dual space of $U(Y' \times Z)$; see Definition 18.36 of [3].

Fix any $y \in Y$, let $J(y) = \overline{\text{co}}\{f_n(y) \otimes \delta_{x_n}\}_{n \geq 1}$, which is the closure of the convex hull of $\{f_n(y) \otimes \delta_{x_n}\}_{n \geq 1}$. It is obvious that $J(y)$ is nonempty and convex. It is also clear that $J(y)$ is the closure of the countable set

$$\left\{ \sum_{i=1}^n \alpha_i f_i(y) \otimes \delta_{x_i} : n \geq 1, \alpha_i \in \mathbb{Q}_+, i = 1, \dots, n, \sum_{i=1}^n \alpha_i = 1 \right\},$$

where \mathbb{Q}_+ is the set of non-negative rational numbers. Let $F'(x) = \{\mu \otimes \delta_x : \mu \in F(x, y)\}$ for any $x \in X$. Then, F' is continuous on X . Since $\{x_n : n \geq 0\}$ is a compact set, Lemma 2 (6) implies that $\bigcup_{n \geq 0} F'(x_n)$ is compact. Hence, $\{f_n(y) \otimes \delta_{x_n}\}_{n \geq 1}$ is relatively compact. By Theorem 5.22 of [3], $\{f_n(y) \otimes \delta_{x_n}\}_{n \geq 1}$ is tight. That is, for any positive real number ϵ , there is a compact set K_ϵ in $Z \times X$ such that for any $n \geq 1$, $f_n(y) \otimes \delta_{x_n}(K_\epsilon) > 1 - \epsilon$. Thus, $\sum_{i=1}^n \alpha_i f_i(y) \otimes \delta_{x_i}(K_\epsilon) > 1 - \epsilon$ for any $n \geq 1$, and for any $\alpha_i \in \mathbb{Q}_+$, $i = 1, \dots, n$ with $\sum_{i=1}^n \alpha_i = 1$. Hence, $J(y)$ is compact by Theorem 5.22 of [3] again.

For any $n \geq 1$, and for any $\alpha_i \in \mathbb{Q}_+$, $i = 1, \dots, n$ with $\sum_{i=1}^n \alpha_i = 1$, it is clear that $\sum_{i=1}^n \alpha_i f_i(y) \otimes \delta_{x_i}$ is measurable in $y \in Y$. Lemma 2 (5) implies that J is also a measurable correspondence from Y to $\mathcal{M}(Z \times X)$. By the argument in the second paragraph of the proof of Part (4), the set

$$\Lambda = \{\lambda \diamond \zeta : \zeta \text{ is a Borel measurable selection of } J\}$$

is compact.

Since $\lambda \diamond (f_n \otimes \delta_{x_n}) \in \Lambda$ for each n , there exists some Borel measurable selection ζ of J such that a subsequence of $\{\lambda \diamond (f_n \otimes \delta_{x_n})\}$, say itself, weakly converges to $\lambda \diamond \zeta \in \Lambda$. Let $\zeta_X(y)$ be the marginal probability of $\zeta(y)$ on X for each y . Since x_n converges to x_0 , $\zeta_X(y) = \delta_{x_0}$ for λ -almost all $y \in Y$. As a result, there exists a Borel measurable function f_0 such that $\zeta = f_0 \otimes \delta_{x_0}$, where $f_0(y) \in \overline{\text{co}}\text{Ls}_n\{f_n(y)\}$ for λ -almost all $y \in Y$. Since F is convex and compact valued, and upper hemicontinuous on X , f_0 is a measurable selection of $F(x_0, \cdot)$. Let $\mu_0 = \lambda \diamond f_0$. Then μ_n weakly converges to μ_0 , which implies that Π is upper hemicontinuous.

Next we shall show the lower hemicontinuity of Π . Suppose that $x_n \rightarrow x_0$ and $\mu_0 \in \Pi(x_0)$. Then there exists a Borel measurable selection f_0 of $F(x_0, \cdot)$ such that $\mu_0 = \lambda \diamond f_0$. Since F is sectionally lower hemicontinuous on X and compact valued, for each $n \geq 1$, there exists a measurable selection f_n for $F(x_n, \cdot)$ such that $f_n(y)$ weakly converges to $f_0(y)$ for each $y \in Y$.²⁹ Denote $\mu_n = \lambda \diamond f_n$. For any bounded continuous function ψ on $Y \times Z$, $\int_Z \psi(y, z) f_n(dz|y)$ converges to $\int_Z \psi(y, z) f_0(dz|y)$

²⁹See Proposition 4.2 in [34]. Note that the atomless Loeb probability measurable space assumption is not needed for the result of lower hemicontinuity as in Theorem 10 therein.

for any $y \in Y$. Thus, we have

$$\begin{aligned} \int_{Y \times Z} \psi(y, z) \mu_n(\mathbf{d}(y, z)) &= \int_Y \int_Z \psi(y, z) f_n(\mathbf{d}z|y) \lambda(\mathbf{d}y) \\ &\rightarrow \int_Y \int_Z \psi(y, z) f_0(\mathbf{d}z|y) \lambda(\mathbf{d}y) \end{aligned}$$

by the Dominated Convergence Theorem. Therefore, Π is lower hemicontinuous. The proof is thus complete. \square

The following lemma is taken from [32] (see also Lemma 4 in [29]).

1. S and Y are Polish spaces, D is a closed subset of $S \times Y$, where $D(s)$ is compact for all $s \in S$;
2. $X = \prod_{1 \leq i \leq n} X_i$, where each X_i is a Polish space;
3. for each i , A_i is a nonempty and compact valued, continuous correspondence from D to X_i , $A = \prod_{1 \leq i \leq n} A_i$ and $E = \text{Gr}(A)$;
4. P is a bounded, nonempty, convex and compact valued, upper hemicontinuous correspondence from E to \mathbb{R}^n .

Lemma 10. *Consider the correspondence $\Phi: D \rightarrow \mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$ defined as follows: $(v, \alpha, \mu) \in \Phi(s, y)$ if*

1. $v = \int_X p(s, y, x) \alpha(\mathbf{d}x)$ such that $p(s, y, \cdot)$ is a Borel measurable selection of $P(s, y, \cdot)$;
2. $\alpha \in \otimes_{i \in I} \mathcal{M}(A_i(s, y))$ is a Nash equilibrium in the subgame (s, y) with payoff $p(s, y, \cdot)$ and action space $A_i(s, y)$ for each player i ;
3. $\mu = p(s, y, \cdot) \circ \alpha$.

Then Φ is nonempty and compact valued, and upper hemicontinuous on D .

6.2 Proof of Theorem 1

6.2.1 Backward induction

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

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

Since $f_{t0}(\cdot|h_{t-1})$ is atomless and Q_{t+1} is nonempty and compact valued, by Lemma 4,

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

where $\text{co}Q_{t+1}(h_{t-1}, x_t, s_t)$ is the convex hull of $Q_{t+1}(h_{t-1}, x_t, s_t)$. By Lemma 7, P_t is bounded, nonempty, convex and compact valued, and upper hemicontinuous.

By Lemma 10, there exists a bounded, measurable, nonempty and compact valued correspondence Φ_t from H_{t-1} to $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$ such that Φ_t is upper hemicontinuous, and for all $h_{t-1} \in H_{t-1}$, $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$ if

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

Denote the restriction of Φ_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 Lemma 10, $\Phi(Q_{t+1})$ is bounded, nonempty and compact valued, and upper hemicontinuous.

6.2.2 Forward induction

In this subsection, we prove the following proposition.

Proposition 5. *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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ such that for all $h_{t-1} \in H_{t-1}$,*

1. $f_t(h_{t-1}) \in \otimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$;
2. $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$;
3. $f_t(\cdot|h_{t-1})$ is a Nash equilibrium in the subgame h_{t-1} with action spaces $\{A_{ti}(h_{t-1})\}_{i \in I}$ and the 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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ and $\mu_t: H_{t-1} \rightarrow \Delta(X_t)$ such that (q_t, f_t, μ_t) is a selection of Φ_t . In step 2, we obtain a Borel measurable selection g_t of P_t such that for all $h_{t-1} \in H_{t-1}$,

1. $g_t(h_{t-1}) = \int_{A_t(h_{t-1})} g_t(h_{t-1}, x) f_t(dx|h_{t-1})$;

2. $f_t(h_{t-1})$ 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

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

For any $\{(h_{t-1}^k, v^k)\}_{1 \leq k \leq \infty} \subseteq \text{Gr}(\Phi_t(Q_{t+1}))$ such that (h_{t-1}^k, v^k) converges to $(h_{t-1}^\infty, v^\infty)$, pick (α^k, μ^k) such that $(v^k, \alpha^k, \mu^k) \in \Phi_t(h_{t-1}^k)$ for $1 \leq k < \infty$. Since Φ_t is upper hemicontinuous and compact valued, there exists a subsequence of (v^k, α^k, μ^k) , say itself, such that (v^k, α^k, μ^k) converges to some $(v^\infty, \alpha^\infty, \mu^\infty) \in \Phi_t(h_{t-1}^\infty)$ due to Lemma 1 (5). Thus, $(\alpha^\infty, \mu^\infty) \in \Psi_t(h_{t-1}^\infty, v^\infty)$, which implies that Ψ_t is also upper hemicontinuous and compact valued. By Lemma 2 (3), Ψ_t has a Borel measurable selection ψ_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 ϕ_t is a Borel measurable selection of Φ_t . Let f_t and μ_t be the second and third dimension of ϕ_t , respectively. By the construction of Φ_t , for all $h_{t-1} \in H_{t-1}$,

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

Step 2. Since P_t is upper hemicontinuous on $\{(h_{t-1}, x_t): h_{t-1} \in H_{t-1}, x_t \in A_t(h_{t-1})\}$, due to Lemma 6, there exists a Borel measurable mapping g such that (1) $g(h_{t-1}, x_t) \in P_t(h_{t-1}, x_t)$ for any $h_{t-1} \in H_{t-1}$ and $x_t \in A_t(h_{t-1})$, and (2) $g(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$ for $f_t(\cdot|h_{t-1})$ -almost all x_t .

In a subgame $h_{t-1} \in H_{t-1}$, let

$$B_{ti}(h_{t-1}) = \{y_i \in A_{ti}(h_{t-1})\}:$$

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

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 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 $h_{t-1} \in H_{t-1}$ 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 β_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 Λ'_i is a measurable and compact valued correspondence, and hence has a Borel measurable selection β'_i . Note that β'_i is a Borel measurable selection of P_t . Let

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

Note that

$$\begin{aligned} & \{(h_{t-1}, x_t) \in \text{Gr}(A_t) : h_{t-1} \in H_{t-1}, x_{ti} \in B_{ti}(h_{t-1}) \text{ and } x_{tj} \notin B_{tj}(h_{t-1}), \forall j \neq i; \} \\ &= \text{Gr}(A_t) \cap \cup_{i \in I} \left((\text{Gr}(B_{ti}) \times \prod_{j \neq i} X_{tj}) \setminus (\cup_{j \neq i} (\text{Gr}(B_{tj}) \times \prod_{k \neq j} X_{tk})) \right), \end{aligned}$$

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 $h_{t-1} \in H_{t-1}$ and $f_t(\cdot|h_{t-1})$ -almost all x_t .

Fix a subgame $h_{t-1} \in H_{t-1}$. We will verify 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

$$\begin{aligned} & \int_{A_{t(-i)}(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)}|h_{t-1}) \\ &= \int_{\prod_{j \neq i} B_{tj}^c(h_{t-1})} g_{ti}(h_{t-1}, \tilde{x}_{ti}, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)}|h_{t-1}) \end{aligned}$$

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

The first and the third equalities hold since $f_{tj}(B_{tj}(h_{t-1}) | h_{t-1}) = 0$ for each j , and hence $f_{t(-i)}(\prod_{j \neq i} B_{tj}^c(h_{t-1}) | h_{t-1}) = f_{t(-i)}(A_{t(-i)}(h_{t-1}) | h_{t-1})$. 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 $x_{t(-i)} \in \prod_{j \neq i} B_{tj}^c(h_{t-1})$. 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

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

The first and the third equalities hold since

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

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 $x_{t(-i)} \in \prod_{j \neq i} B_{tj}^c(h_{t-1})$. 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

1. $q(h_{t-1}, x_t, e, s_t) \in Q_{t+1}(h_{t-1}, x_t, s_t)$ for any $(h_{t-1}, x_t, e, s_t) \in \text{Gr}(P_t) \times S_t$;
2. $e = \int_{S_t} q(h_{t-1}, x_t, e, s_t) f_{t0}(ds_t|h_{t-1})$ 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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ such that for all $h_{t-1} \in H_{t-1}$, properties (1)-(3) 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, then run the forward induction from the initial period to the last period. We obtain the following corollary.

Corollary 1. *If a finite-horizon continuous dynamic game with almost perfect information satisfies the condition of atomless transitions, then it has a subgame-perfect equilibrium.*

6.2.3 Infinite horizon case

Pick a sequence $\xi = (\xi_1, \xi_2, \dots)$ such that (1) ξ_m is a transition probability from H_{m-1} to $\mathcal{M}(X_m)$ for any $m \geq 1$, and (2) $\xi_m(A_m(h_{m-1})|h_{m-1}) = 1$ for any $m \geq 1$ and $h_{m-1} \in H_{m-1}$. Denote the set of all such ξ as Υ .

Fix any $t \geq 1$, define a correspondence Δ_t^t as follows: in the subgame h_{t-1} ,

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

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

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

where $\mathcal{M}(A_{m_1})$ is regarded as the constant correspondence from H_{m_1-1} to the space of all Borel probability measures on X_{m_1} (by abusing the notations a little bit). For any $m_1 \geq t$, let $\varrho_{(h_{t-1}, \xi)}^{m_1} \in \Delta_t^{m_1}$ 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 $\Delta_t^{m_1}(h_{t-1})$ is the set of all such $\varrho_{(h_{t-1}, \xi)}^{m_1}$. Note that $\varrho_{(h_{t-1}, \xi)}^{m_1}$ can be regarded as a probability measure on $H_{m_1}(h_{t-1})$.

Similarly, let $\varrho_{(h_{t-1}, \xi)}$ be 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

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

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

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

Lemma 11. *1. For any $t \geq 1$, the correspondence $\Delta_t^{m_1}$ is nonempty and compact valued, and continuous for any $m_1 \geq t$.*

2. For any $t \geq 1$, the correspondence Δ_t is nonempty and compact valued, and continuous.

Proof. (1) Consider the case $m_1 = t \geq 1$, where

$$\Delta_t^{m_1}(h_{t-1}) = \mathcal{M}(A_t(h_{t-1})) \otimes f_{t0}(h_{t-1}).$$

Since A_t is nonempty and compact valued, and both A_t and f_{t0} are continuous,

$\Delta_t^{m_1}$ is nonempty and compact valued, and continuous.

Suppose that $\Delta_t^{m_2}$ is nonempty and compact valued, and continuous for some $m_2 \geq t \geq 1$. Note that

$$\begin{aligned} \Delta_t^{m_2+1}(h_{t-1}) &= \{g(h_{t-1}) \diamond (\xi_{m_2+1}(h_{t-1}, \cdot) \otimes f_{m_2+10}(h_{t-1}, \cdot))\}: \\ &g \text{ is a Borel measurable selection of } \Delta_t^{m_2}, \\ &\xi_{m_2+1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_2+1})\}. \end{aligned}$$

By Lemma 9 (3), $\Delta_t^{m_2+1}$ is nonempty and compact valued, and continuous.

(2) It is obvious that Δ_t is nonempty valued, we shall first prove that it is upper hemicontinuous and compact valued.

Given sequences $\{h_{t-1}^k\}$ and $\tau^k \subseteq \Delta_t(h_{t-1}^k)$, there exists a sequence of $\{\xi^k\}_{k \geq 1}$ such that $\xi^k = (\xi_1^k, \xi_2^k, \dots) \in \Upsilon$ and $\tau^k = \varrho_{(h_{t-1}^k, \xi^k)}$ for each $k \geq 1$. Suppose that $h_{t-1}^k \rightarrow h_{t-1}^\infty$. By (1), Δ_t^t is compact valued and upper hemicontinuous. Then there exists a measurable mapping g_t such that (1) $g^t = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, \xi_{t+1}^1, \dots) \in \Upsilon$, and (2) a subsequence of $\{\varrho_{(h_{t-1}^k, \xi^k)}^t\}$, say $\{\varrho_{(h_{t-1}^{k_{1l}}, \xi^{k_{1l}})}^t\}_{l \geq 1}$, weakly converges to $\varrho_{(h_{t-1}^\infty, g^t)}$. Note that $\{\xi_{t+1}^k\}$ is a Borel measurable selection of $\mathcal{M}(A_{t+1})$. By Lemma 9 (3), there is a Borel measurable selection g_{t+1} of $\mathcal{M}(A_{t+1})$ such that there is a subsequence of $\{\varrho_{(h_{t-1}^{k_{1l}}, \xi^{k_{1l}})}^{t+1}\}_{l \geq 1}$, say $\{\varrho_{(h_{t-1}^{k_{2l}}, \xi^{k_{2l}})}^{t+1}\}_{l \geq 1}$, weakly converges to $\varrho_{(h_{t-1}^\infty, g^{t+1})}$, where $g^{t+1} = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, g_{t+1}, \xi_{t+2}^1, \dots) \in \Upsilon$. Repeat this procedure, one can construct a Borel measurable mapping g such that $\varrho_{(h_{t-1}^{k_{11}}, \xi^{k_{11}})}, \varrho_{(h_{t-1}^{k_{22}}, \xi^{k_{22}})}, \varrho_{(h_{t-1}^{k_{33}}, \xi^{k_{33}})}, \dots$ weakly converges to $\varrho_{(h_{t-1}^\infty, g)}$. That is, $\varrho_{(h_{t-1}^\infty, g)}$ is a convergent point of $\{\varrho_{(h_{t-1}^k, \xi^k)}\}$. By Lemma 1 (5), Δ_t is compact valued and upper hemicontinuous.

Next, we consider the lower hemicontinuity of Δ_t . Suppose that $\tau^0 \in \Delta_t(h_{t-1}^0)$. Then there exists some $\xi \in \Upsilon$ 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 Δ_t^m is continuous, for each m , there exists some $\xi^m \in \Upsilon$ 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 $\tau^m = \varrho_{(h_{t-1}^{k_m}, \xi^m)}$. Then τ^m weakly converges to τ^0 , which implies that Δ_t is lower hemicontinuous. \square

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

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

The lemma below presents several desirable properties of Q_t^τ .

Lemma 12. *For any $t, \tau \geq 1$, Q_t^τ is bounded, measurable, nonempty and compact valued, and upper hemicontinuous.*

Proof. For $t > \tau$, Q_t^τ is bounded, measurable, nonempty and compact valued, and upper hemicontinuous because of the corresponding properties of u and Δ_t .

For $t \leq \tau$, we can start with $Q_{\tau+1}^\tau$. Repeating the backward induction in Subsection 6.2.1, Q_t^τ is bounded, measurable, nonempty and compact valued, and upper hemicontinuous. \square

Denote $Q_t^\infty = \bigcap_{\tau \geq 1} Q_t^\tau$. 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 all $h_{t-1} \in H_{t-1}$.³⁰

Lemma 13. 1. *The correspondence Q_t^∞ is bounded, measurable, nonempty and compact valued, and upper hemicontinuous.*

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

Proof. (1) It is obvious that Q_t^∞ is bounded. By the definition of Q_t^τ , for 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^τ is nonempty and compact valued, $Q_t^\infty = \bigcap_{\tau \geq 1} Q_t^\tau$ is nonempty and compact valued. By Lemma 2 (2), $\bigcap_{\tau \geq 1} Q_t^\tau$ is measurable, which implies that Q_t^∞ is measurable.

Since Q_t^τ is upper hemicontinuous for any τ , by Lemma 2 (7), it has a closed graph for each τ , which implies that Q_t^∞ has a closed graph. Referring to Lemma 2 (7) again, Q_t^∞ is upper hemicontinuous.

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

Let $\{1, 2, \dots, \infty\}$ be a countable compact space endowed with the following metric: $d(k, m) = |\frac{1}{k} - \frac{1}{m}|$ for $1 \leq k, m \leq \infty$. 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 upper hemicontinuous on $H_t \times \{1, 2, \dots, \infty\}$. The step of backward induction in Subsection 6.2.1 shows that $\Phi(Q_{t+1})$ is measurable, nonempty and compact valued, and upper hemicontinuous on $H_t \times \{1, 2, \dots, \infty\}$. For $h_{t-1} \in H_{t-1}$ 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, $a \in \Phi(Q_{t+1}^\infty)(h_{t-1})$.

As a result, $Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$. \square

Though the definition of Q_t^τ involves correlated strategies for $\tau < t$, the following lemma shows that one can work with mixed strategies in terms of

³⁰The proofs of Lemmas 13 and 15 follow the standard argument with various modifications; see, for example, [14], [15] and [22].

equilibrium payoffs via the combination of backward and forward inductions in multiple steps.

Lemma 14. *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 any $h_{t-1} \in H_{t-1}$.*

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 5 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^∞ such that for all $h_{k-1} \in H_{k-1}$,

1. $f_k(h_{k-1})$ 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

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

- 2.

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

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

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

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

Fix a positive integer $M > k$. By Lemma 13, $c_k(h_{k-1}) \in Q_k^\infty(h_{k-1}) = \cap_{\tau \geq 1} Q_k^\tau(h_{k-1})$ for all $h_{k-1} \in H_{k-1}$. Since $Q_k^\tau = \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)$ for $k \leq \tau$, $c_k(h_{k-1}) \in \cap_{\tau \geq k} \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)(h_{k-1}) \subseteq \Phi^{M-k+1}(Q_{M+1}^M)(h_{k-1})$ for all $h_{k-1} \in H_{k-1}$. Thus, there exists a Borel measurable selection w of Q_{M+1}^M and some $\xi \in \Upsilon$ such that for all $h_{M-1} \in H_{M-1}$,

- i. $f_M(h_{M-1})$ 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.

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

iii. $w(h_M) = \int_{\prod_{m \geq M+1} (X_m \times S_m)} u(h_M, x, s) \varrho_{(h_M, \xi)}(\mathbf{d}(x, s))$.

Then for $h_{k-1} \in H_{k-1}$,

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

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

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

converges to

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

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

$$c_k(h_{k-1}) = \int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \varrho_{(h_{k-1}, f)}(\mathbf{d}(x, s)). \quad (2)$$

Step 2. Below, 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 follows: $\tilde{f}_i^k = (g_{1i}, \dots, g_{ki}, f_{(k+1)i}, f_{(k+2)i}, \dots)$. 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 $k \geq 1$. For any $h_k = (x^k, s^k)$, we have

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

The first and the last equalities follow from Equation (2) in the end of step 1. The second equality is due to (2) in step 1. The first inequality is based on (1) in step 1. The second inequality holds since by the choice of h_k and (1) in step 1, for $f_{(k+1)0}(h_k)$ -almost all $s_{k+1} \in S_{k+1}$ and all $x_{k+1} \in X_{k+1}$, we have

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

Repeating the above argument, one can show that

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

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

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

converges to

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

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

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

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

By Lemmas 10 and 13, the correspondence $\Phi(Q_{t+1}^\infty)$ is measurable, nonempty and compact valued. By Lemma 2 (3), it has a measurable selection. Then the equilibrium existence result in Theorem 1 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 the same as $Q_t^\infty(h_{t-1})$.

Lemma 15. For any $t \geq 1$, $E_t(h_{t-1}) = Q_t^\infty(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$.

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

By the construction of $\Phi(Q_{t+1}^\tau)$ in Subsection 6.2.1, for any c_t and $h_{t-1} = (x^{t-1}, s^{t-1}) \in H_{t-1}$, if

1. $c_t = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) \alpha(dx_t)$, 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 all $s_t \in S_t$ and $x_t \in A_t(h_{t-1})$;
2. $\alpha \in \otimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$ 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 $\prod_{i \in I} A_{ti}(h_{t-1})$,

then $c_t \in \Phi(Q_{t+1}^\tau)(h_{t-1})$.

Fix a subgame $h_{t-1} = (x^{t-1}, s^{t-1})$. Pick a point $c_t \in E_t(s^{t-1})$. 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 we have

1. $c_t = \int_{A_t(h_{t-1})} \int_{S_t} c_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$;
2. $f_t(\cdot|h_{t-1})$ 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 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 all $h_{t-1} \in H_{t-1}$.

(2) 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 (1), then we can show that $E_t(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$. Thus, $E_t(h_{t-1}) \subseteq Q_t^\infty(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$.

Suppose that c_t is a measurable selection from $\Phi(Q_{t+1}^\infty)$. Apply Proposition 5 recursively to obtain Borel measurable mappings $\{f_{ki}\}_{i \in I}$ for $k \geq t$. By Lemma 14, $c_t(h_{t-1})$ is a subgame-perfect equilibrium payoff vector for all $h_{t-1} \in H_{t-1}$. Consequently, $\Phi(Q_{t+1}^\infty)(h_{t-1}) \subseteq E_t(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$.

By Lemma 13, $E_t(h_{t-1}) = Q_t^\infty(h_{t-1}) = \Phi(Q_{t+1}^\infty)(h_{t-1})$ for all $h_{t-1} \in H_{t-1}$. \square

Therefore, we complete the proof of Theorem 1.

6.3 Proof of Theorem 2

Step 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 upper hemicontinuous on X^t .

If player j is the active player in stage t , then we assume that $S_t = \{s_t\}$. Thus, $P_t(h_{t-1}, x_t) = Q_{t+1}(h_{t-1}, x_t, s_t)$, which is nonempty and compact valued, and upper hemicontinuous. Note that P_t may not be convex valued. Then define the correspondence Φ_t from H_{t-1} to $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$ as $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$ if

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

If P_t is nonempty, convex and compact valued, and upper hemicontinuous, then we can use Lemma 10, the main result of [32], to prove the nonemptiness, compactness, and upper hemicontinuity of Φ_t . In [32], 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 paragraph, 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 $d_0 \in P(z_0)$.

Repeat the argument in the proof of the main theorem of [32], one can show that Φ_t is nonempty and compact valued, and upper hemicontinuous.

Next, we consider the case that Nature moves in stage t . That is, there is no active player in I moving in this stage and $A_t(h_{t-1})$ is a singleton set. Suppose that the correspondence Q_{t+1} from H_t to \mathbb{R}^n is bounded, measurable, nonempty and compact valued, and upper hemicontinuous. Let

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

where $A_t(h_{t-1}) = \{x_t\}$. Since $f_{t0}(\cdot | h_{t-1})$ is atomless, as in Subsection 6.2.1, P_t is nonempty, convex and compact valued, and upper hemicontinuous. The rest of the step remains the same as in Subsection 6.2.1.

In summary, Φ_t is nonempty and compact valued, and upper hemicontinuous.

Steps 2 and 3. Forward induction and the infinite horizon case.

These two steps are the same as that in Subsection 6.2, except the corresponding notations need to be changed to be consistent with the perfect information environment whenever necessary.

6.4 Proof of Proposition 1

The proof is essentially the combination of the proofs in Subsections 6.2 and 6.3. That is, when there is only one active player, we refer to the argument in Subsection 6.3. When there are more than one active players or Nature is the only player who moves, we modify the argument in Subsection 6.2.

Step 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 upper hemicontinuous on X^t .

If $N_t = 1$, then $S_t = \{s_t\}$. Thus, $P_t(h_{t-1}, x_t) = Q_{t+1}(h_{t-1}, x_t, s_t)$, which is nonempty and compact valued, and upper hemicontinuous. Then define the correspondence Φ_t from H_{t-1} to $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$ as $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$ if

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

As discussed in Subsection 6.3, Φ_t is nonempty and compact valued, and upper hemicontinuous.

When $N_t = 0$, for any $h_{t-1} \in H_{t-1}$ and $x_t \in A_t(h_{t-1})$,

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

Let $\text{co}Q_{t+1}(h_{t-1}, x_t, s_t)$ be the convex hull of $Q_{t+1}(h_{t-1}, x_t, s_t)$. Because Q_{t+1} is bounded, nonempty and compact valued, $\text{co}Q_{t+1}$ is bounded, nonempty, convex and compact valued. By Lemma 2 (8), $\text{co}Q_{t+1}$ is upper hemicontinuous.

Note that $f_{t0}(\cdot | h_{t-1}, x_t)$ is atomless and Q_{t+1} is nonempty and compact valued.

We have

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

By Lemma 7, P_t is bounded, nonempty, convex and compact valued, and upper hemicontinuous. Then by Lemma 10, one can conclude that Φ_t is bounded, nonempty and compact valued, and upper hemicontinuous.

Steps 2 and 3. Forward induction and the infinite horizon case.

These two steps are the same as that in Subsection 6.2. The only change is to modify the notations correspondingly.

References

- [1] D. Abreu, D. Pearce and E. Stacchetti, Toward A Theory of Discounted Repeated Games with Imperfect Monitoring, *Econometrica* **58** (1990), 1041–1063.
- [2] R. Amir, Strategic intergenerational bequests with stochastic convex production, *Economic Theory* **8** (1996), 367–376.
- [3] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin, 2006.
- [4] D. Bernheim and D. Ray, Economic Growth with Intergenerational Altruism, *Review of Economic Studies* **54** (1987), 227–242.
- [5] V. I. Bogachev, *Measure Theory, Volume II*, Springer-Verlag, Berlin Heidelberg, 2007.
- [6] T. Börgers, Perfect Equilibrium Histories of Finite and Infinite Horizon Games, *Journal of Economic Theory* **47** (1989), 218–227.
- [7] T. Börgers, Upper Hemicontinuity of The Correspondence of Subgame-Perfect Equilibrium Outcomes, *Journal of Mathematical Economics* **20** (1991), 89–106.
- [8] L. Brown and B. M. Schreiber, Approximation and Extension of Random Functions, *Monatshefte für Mathematik* **107** (1989), 111–123.
- [9] D. Fudenberg and D. Levine, Subgame-Perfect Equilibria of Finite and Infinite-Horizon Games, *Journal of Economic Theory* **31** (1983), 251–268.
- [10] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, 1991.
- [11] S. Goldman, Consistent Plans, *Review of Economic Studies* **47** (1980), 533–537.
- [12] E. J. Green and R. H. Porter, Noncooperative Collusion under Imperfect Price Information, *Econometrica* **52** (1984), 87–100.

- [13] C. Harris, Existence and Characterization of Perfect Equilibrium in Games of Perfect Information, *Econometrica* **53** (1985), 613–628.
- [14] C. Harris, The Existence of Subgame-Perfect Equilibrium with and without Markov Strategies: A Case for Extensive-Form Correlation, Discussion Paper No. 53, Nuffield College, Oxford, 1990.
- [15] C. Harris, P. J. Reny and A. Robson, The Existence of Subgame-Perfect Equilibrium in Continuous Games with Almost Perfect Information: A Case for Public Randomization, *Econometrica* **63** (1995), 507–544.
- [16] M. Hellwig and W. Leininger, On the Existence of Subgame-Perfect Equilibrium in Infinite-Action Games of Perfect Information, *Journal of Economic Theory* **43** (1987), 55–75.
- [17] M. Hellwig, W. Leininger, P. J. Reny and A. J. Robson, Subgame-Perfect Equilibrium in Continuous Games of Perfect Information: An Elementary Approach to Existence and Approximation by Discrete Games, *Journal of Economic Theory* **52** (1990), 406–422.
- [18] W. Hildenbrand, *Core and Equilibria of a Large Economy*, Princeton University Press, Princeton, NJ, 1974.
- [19] A. Kucia, Some Results on Carathéodory Selections and Extensions, *Journal of Mathematical Analysis and Applications* **223** (1998), 302–318.
- [20] W. Leininger, The Existence of Perfect Equilibria in A Model of Growth with Altruism between Generations, *Review of Economics Studies* **53** (1986), 349–367.
- [21] E. G. J. Luttmer and T. Mariotti, The Existence of Subgame-Perfect Equilibrium in Continuous Games with Almost Perfect Information: A Comment, *Econometrica* **71** (2003), 1909–1911.
- [22] T. Mariotti, Subgame-Perfect Equilibrium Outcomes in Continuous Games of Almost Perfect Information, *Journal of Mathematical Economics* **34** (2000), 99–128.
- [23] J.-F. Mertens, A Measurable ‘Measurable Choice’ Theorem, in *Proceedings of the NATO Advanced Study Institute on Stochastic Games and Applications*, ed. by A. Neyman and S. Sorin. NATO Science Series. Dordrecht: Kluwer Academic, 2003, 107–130.
- [24] J.-F. Mertens and T. Parthasarathy, Equilibria for Discounted Stochastic Games, in *Proceedings of the NATO Advanced Study Institute on Stochastic Games and Applications*, ed. by A. Neyman and S. Sorin. NATO Science Series. Dordrecht: Kluwer Academic, 2003, 131–172.
- [25] E. Michael, A Selection Theorem, *Proceedings of the American Mathematical Society* **17** (1966), 1404–1406.
- [26] R. B. Myerson and P. J. Reny, Sequential Equilibria of Multi-Stage Games with Infinite Sets of Types and Actions, working paper, University of Chicago, 2015.
- [27] B. Peleg and M. Yaari, On the Existence of A Consistent Course of Action

- When Tastes Are Changing, *Review of Economic Studies* **40** (1973), 391–401.
- [28] E. Phelps and R. Pollak, On Second Best National Savings and Game Equilibrium Growth, *Review of Economic Studies* **35** (1968), 185–199.
- [29] P. J. Reny and A. J. Robson, Existence of Subgame-Perfect Equilibrium with Public Randomization: A Short Proof, *Economics Bulletin* **24** (2002), 1–8.
- [30] J. J. Rotemberg and G. Saloner, A Supergame-Theoretic Model of Price Wars during Booms, *American Economic Review* **76** (1986), 390–407.
- [31] R. Selten, Spieltheoretische Behandlung Eines Oligopolmodells Mit Nachfragetraheit, *Z. Gesamte Staatswissenschaft* **12** (1965), 301–324.
- [32] L. K. Simon and W. R. Zame, Discontinuous Games and Endogenous Sharing Rules, *Econometrica* **58** (1990), 861–872.
- [33] M. B. Stinchcombe, Nash Equilibrium and Generalized Integration for Infinite Normal Form Games, *Games and Economic Behavior* **50** (2005), 332–365.
- [34] Y. N. Sun, Integration of Correspondences on Loeb Spaces, *Transactions of the American Mathematical Society* **349** (1997), 129–153.
- [35] E. Zermelo, Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels, *Proc. Fifth Congress Mathematicians*, Cambridge University Press, Cambridge, 1913, 501–504.

Dynamic Games with Almost Perfect Information: Appendix B

Wei He* Yeneng Sun†

This version: June 26, 2017

In Section B.1, we shall 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 Section B.2, we present Lemmas B.1-B.6 as the mathematical preparations for proving Theorem 3. We present in Section B.3 a new equilibrium existence result for discontinuous games with stochastic endogenous sharing rules. The proof of Theorem 3 is given in Section B.4. The proof of Proposition B.1 is provided in Section B.5, which covers Theorem 4 as a special case. One can skip Sections 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 will generalize the model of measurable dynamic games in three directions. The ARM condition is partially relaxed such that (1) perfect information may be allowed in some stages, (2) the state transitions could have a weakly continuous component in all other stages, and (3) 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

*Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. E-mail: he.wei2126@gmail.com.

†Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570. Email: ynsun@nus.edu.sg.

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.

1. The state space is a product space of two Polish spaces; that is, $S_t = \hat{S}_t \times \tilde{S}_t$ for each $t \geq 1$.
2. 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_∞ 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 $t \geq 0$.
3. 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 $\mathcal{M}(\tilde{S}_t)$.
4. For each $i \in I$, the payoff function u_i is a Borel measurable mapping from H_∞ to \mathbb{R}_{++} , which is sectionally continuous on $X^\infty \times \hat{S}^\infty$.

As in Subsection 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 will drop the ARM condition in those periods with only one active player, and weaken the ARM condition in other periods.

Assumption B.1 (ARM'). 1. For any $t \geq 1$ with $N_t = 1$, S_t is a singleton set $\{s_t\}$ and $\lambda_t = \delta_{s_t}$.

2. 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 λ_t on \tilde{S}_t for all $(h_{t-1}, x_t, \hat{s}_t) \in Gr(\hat{A}_{t0})$, and $\varphi_{t0}(h_{t-1}, x_t, \hat{s}_t, \tilde{s}_t)$ is the corresponding density.¹

3. The mapping φ_{t0} is Borel measurable and sectionally continuous on $X^t \times \hat{S}^t$, and integrably bounded in the sense that there is a λ_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.

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_{tj} 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 [6], 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 \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 [1], \mathcal{K} is complete. In addition, Corollary 3.90 and Theorem 3.91 of [1] 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.

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

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 Π from X to Z as follows:*

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

If F is sectionally continuous on Y , then

1. 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 $\mathcal{M}(Y)$; and
2. Π is a measurable, nonempty and compact valued correspondence.
3. If F and G are both continuous, then Π is continuous.

Proof. (1) For any fixed $x \in X$, the upper hemicontinuity of $\tilde{F}(x, \cdot)$ follows from Lemma 7.

Next, we shall 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 follows:

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

Then F_k is nonempty, convex and compact valued, and lower hemicontinuous. By Theorem 3.22 in [1], 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.

(2) 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 (5). For each $k \geq 1$, define a correspondence Π^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 Π^k . By Lemma 5, Π^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, $\overline{\bigcup_{k \geq 1} \Pi^k(x)} \subseteq \Pi(x)$.

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 (1) and (2) imply that Π is measurable.

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

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

By (1), \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, Π is compact valued by Lemma 2. As \hat{G} and \hat{F} are both continuous, Π is continuous by Lemma 1 (7). \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 σ -algebras, and λ a Borel probability measure on S . Denote \mathcal{S} as the completion of the Borel σ -algebra $\mathcal{B}(S)$ of S under the probability measure λ . 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*

- (i) $\tilde{A}(s) = \text{Gr}(A(s, \cdot))$ 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 (1) $\lambda(\cup_{m \geq 1} S_m) = 1$, and (2) 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) $A(s, \cdot)$ 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 (4).

(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 $\lambda(S \setminus S_1) < \frac{1}{2}$, \tilde{D} and \tilde{A} are continuous functions on S_1 . By Lemma 1 (3), \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 (1) $S_m \cap (\cup_{1 \leq k \leq m-1} S_k) = \emptyset$ and $D_m = D \cap (S_m \times Y)$ is compact, (2) $\lambda(S_m) > 0$ and $\lambda(S \setminus (\cup_{1 \leq k \leq m} S_k)) < \frac{1}{2m}$, and (3) 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 λ a Borel probability measure on S . Suppose that $\{S_k\}_{k \geq 1}$ is a sequence of disjoint compact subsets of S such that $\lambda(\cup_{k \geq 1} S_k) = 1$. 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 τ_m weakly converges to τ_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 τ_m weakly converges to τ_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 τ_m weakly converges to τ_0 . \square

The following is a key lemma 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 Z is a compact metric space. Let λ be a Borel probability measure on S , and A a nonempty and compact valued correspondence from $Z \times S$ to X which 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, 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 τ on S is λ 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 Π 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 Π is nonempty and compact valued, and continuous.

Proof. Let \mathcal{S} be the completion of $\mathcal{B}(S)$ under the probability measure λ . 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' of 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 \mathcal{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 \mathcal{S}_1 -measurable function from S_1 to $C(K_1, \mathcal{K}_{\mathcal{M}(Y)})$ (see Theorem 4.55 in [1]). 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 $\text{Gr}(A) \cap (S_1 \times Z \times X)$, so is F . Let λ_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 ν 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, Π_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

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

Repeat this procedure, one can find a sequence of compact subsets $\{S_t\}$ such that (1) for any $t \geq 1$, $S_t \subseteq S'$, $S_t \cap (S_1 \cup \dots \cup S_{t-1}) = \emptyset$ and $\lambda(S_1 \cup \dots \cup S_t) \geq \frac{t}{t+1}$, (2) F is continuous on $\text{Gr}(A) \cap (S_t \times Z \times X)$, λ_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 (3) the correspondence

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

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 κ . 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 Π_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 $(\lambda_1 \diamond \mu^1) \diamond f^1(z_0, \cdot) \in \Pi_1(z_0)$. Repeat 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 Π is upper hemicontinuous.

Similarly, the compactness and lower hemicontinuity of Π follow from the compactness and lower hemicontinuity of Π_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 A a measurable, nonempty and compact valued correspondence from S to X . Suppose that λ is a Borel probability measure on S and $\{\nu_n\}_{1 \leq n \leq \infty}$ 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 τ_∞ on $S \times X$. Let $\{g_n\}_{1 \leq n \leq \infty}$ be a sequence of functions satisfying the following three properties.*

1. For each n between 1 and ∞ , $g_n: S \times X \rightarrow \mathbb{R}_+$ is measurable and sectionally continuous on X .

2. 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 $n \rightarrow \infty$.
3. The sequence $\{g_n\}_{1 \leq n \leq \infty}$ is integrably bounded in the sense that there exists a λ -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

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

Proof. By Theorem 2.1.3 in [2], for any integrably bounded function $g: S \times X \rightarrow \mathbb{R}_+$ which is sectionally continuous on X , we have

$$\int_{S \times X} g(s, x) \tau_n(\mathbf{d}(s, x)) \rightarrow \int_{S \times X} g(s, x) \tau_\infty(\mathbf{d}(s, x)). \quad (1)$$

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 (6), $G(s)$ is compact. By Lemma 1 (2), 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 ,

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

By Lemma 2 (6), 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 Equation (1), there exists a positive integer N_2 such that for any $n \geq N_2$,

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

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

$$\begin{aligned} & \left| \int_{S \times X} g_n(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(\mathbf{d}(s, x)) \right| \\ & \leq \left| \int_{S \times X} g_n(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) \right| \\ & + \left| \int_{S \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(\mathbf{d}(s, x)) \right| \\ & \leq \left| \int_{S_1 \times X} g_n(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S_1 \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) \right| \\ & + \left| \int_{(S \setminus S_1) \times X} g_n(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{(S \setminus S_1) \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) \right| \\ & + \left| \int_{S \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(\mathbf{d}(s, x)) \right| \\ & \leq \int_E |g_n(s, x) - g_\infty(s, x)| \tau_n(\mathbf{d}(s, x)) + 2 \cdot \int_{(S \setminus S_1) \times X} \psi(s) \tau_n(\mathbf{d}(s, x)) \\ & + \left| \int_{S \times X} g_\infty(s, x) \tau_n(\mathbf{d}(s, x)) - \int_{S \times X} g_\infty(s, x) \tau_\infty(\mathbf{d}(s, x)) \right| \\ & < \frac{\epsilon}{3} + 2 \cdot \frac{\epsilon}{6} + \frac{\epsilon}{3} \\ & = \epsilon. \end{aligned}$$

This completes the proof. \square

B.3 Discontinuous games with endogenous stochastic sharing rules

It was proved in [7] 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 α such that α is a Nash equilibrium when players take p as the payoff function (see Lemma 10).

In this section, we shall 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:

1. let S be a Borel subset of a Polish space, Y a Polish space, and λ a Borel probability measure on S ;
2. 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 $\lambda(\{s \in S: D(s) \neq \emptyset\}) > 0$;
3. $X = \prod_{1 \leq i \leq n} X_i$, where each X_i is a Polish space;
4. for each i , A_i is a measurable, nonempty and compact valued correspondence from D to X_i , which is sectionally continuous on Y ;
5. $A = \prod_{1 \leq i \leq n} A_i$ and $E = \text{Gr}(A)$;
6. P is a bounded, measurable, nonempty, convex and compact valued correspondence from E to \mathbb{R}^n which 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 shall prove the following proposition.

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

1. $v = \int_X p(s, y, x) \alpha(dx)$ such that $p(s, y, \cdot)$ is a Borel measurable selection of $P(s, y, \cdot)$;²
2. $\alpha \in \otimes_{i \in I} \mathcal{M}(A_i(s, y))$ is a Nash equilibrium in the subgame (s, y) with payoff $p(s, y, \cdot)$ and action space $A_i(s, y)$ for each player i ;
3. $\mu = p(s, y, \cdot) \circ \alpha$.³

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

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

In addition, denote the restriction of Φ 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 λ . 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 (4), \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 (6). Define a correspondence Γ 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, Γ is compact valued. By Lemma 1 (1), 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 (4), the correspondence Γ is \mathcal{S} -measurable. We can view Γ 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 (2), Γ 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 $\lambda(S \setminus S_1) < \frac{1}{2}$, Γ , \tilde{D} and $\{\tilde{A}_i\}_{1 \leq i \leq n}$ are continuous functions on S_1 . By Lemma 1 (3), Γ , \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 (6)). 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 Φ_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 (1) $S_m \cap (\cup_{1 \leq k \leq m-1} S_k) = \emptyset$ and $D_m = D \cap (S_m \times Y)$ is compact, (2) $\lambda(S_m) > 0$ and $\lambda(S \setminus (\cup_{1 \leq k \leq m} S_k)) < \frac{1}{2m}$, and (3) there is a nonempty and compact valued, upper

hemicontinuous correspondence Φ_m from D_m to $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$, which satisfies conditions (1)-(3) in Lemma 10. Thus, we have countably many disjoint sets $\{S_m\}_{m \geq 1}$ such that (1) $\lambda(\cup_{m \geq 1} S_m) = 1$, (2) Φ_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 (3). Fix a Borel measurable selection p of P (such a selection exists also due to Lemma 2 (3)). Define a mapping (v_0, α_0, μ_0) from D to $\mathbb{R}^n \times \mathcal{M}(X) \times \Delta(X)$ such that (1) $\alpha_i(s, y) = \delta_{a_i(s, y)}$ and $\alpha_0(s, y) = \otimes_{i \in I} \alpha_i(s, y)$; (2) $v_0(s, y) = p(s, y, a_1(s, y), \dots, a_n(s, y))$ and (3) $\mu_0(s, y) = p(s, y, \cdot) \circ \alpha_0$. Let $D_0 = D \setminus (\cup_{m \geq 1} D_m)$ and $\Phi_0(s, y) = \{(v_0(s, y), \alpha_0(s, y), \mu_0(s, y))\}$ for $(s, y) \in D_0$. Then, Φ_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 (1)-(3) if $(s, y) \in D_m$ for $m \geq 1$. That is, Φ is $\mathcal{B}(D)$ -measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on Y , and satisfies conditions (1)-(3) for λ -almost all $s \in S$.

Then consider $\Phi|_{\mathbb{R}^n}$, which is the restriction of Φ on the first component \mathbb{R}^n . Let $\Phi_m|_{\mathbb{R}^n}$ be the restriction of Φ_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 Φ_m is upper hemicontinuous and compact valued. By Lemma 2 (6), $\text{Gr}(\Phi_m)$ is compact. Thus, $\text{Gr}(\Phi_m|_{\mathbb{R}^n})$ is also compact. By Lemma 2 (4), $\Phi_m|_{\mathbb{R}^n}$ 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. □

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

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

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

It is obvious that the correspondence P_t is measurable and nonempty valued. Since Q_{t+1} is bounded, P_t is bounded. For λ^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 φ_{t0} is integrably bounded, $P_t(s^{t-1}, \cdot)$ is also upper hemicontinuous on $\text{Gr}(A^t)(s^{t-1})$ for λ^{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 λ_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 which is essentially sectionally upper hemicontinuous on X^t .

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

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

Denote the restriction of Φ_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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ such that for λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$,*

1. $f_t(h_{t-1}) \in \otimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$;

2. $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$;
3. $f_t(\cdot|h_{t-1})$ is a Nash equilibrium in the subgame h_{t-1} with action spaces $A_{ti}(h_{t-1})$, $i \in I$ and the 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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ and $\mu_t: H_{t-1} \rightarrow \Delta(X_t)$ such that (q_t, f_t, μ_t) is a selection of Φ_t . In step 2, we obtain a Borel measurable selection g_t of P_t such that for λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$,

1. $q_t(h_{t-1}) = \int_{A_t(h_{t-1})} g_t(h_{t-1}, x) f_t(dx|h_{t-1})$;
2. $f_t(h_{t-1})$ 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 the construction of Φ_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

1. $H_{t-1} = \cup_{m \geq 0} H_{t-1}^m$ and $\frac{\lambda^{t-1}(\cup_{m \geq 1} \text{proj}_{S^{t-1}}(H_{t-1}^m))}{\lambda^{t-1}(\text{proj}_{S^{t-1}}(H_{t-1}))} = 1$, 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} ;
2. for $m \geq 1$, H_{t-1}^m is compact, Φ_t is upper hemicontinuous on H_{t-1}^m , and P_t is upper hemicontinuous on

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

3. there exists a Borel measurable mapping (v_0, α_0, μ_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 Φ_t on H_{t-1}^m as Φ_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 $m \geq 1$, Ψ_t^m is measurable by Lemma 2 (4), and has a Borel measurable selection ψ_t^m due to Lemma 2 (3). 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 ψ_t is a Borel measurable selection of Ψ_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 ϕ_t is a Borel measurable selection of Φ_t . Denote $\tilde{H}_{t-1} = \cup_{m \geq 1} H_{t-1}^m$. By the construction of Φ_t , there exists Borel measurable mappings $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ and $\mu_t: H_{t-1} \rightarrow \Delta(X_t)$ such that for all $h_{t-1} \in \tilde{H}_{t-1}$,

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

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 (1) $g^m(h_{t-1}, x_t) \in P_t(h_{t-1}, x_t)$ for any $h_{t-1} \in H_{t-1}^m$ and $x_t \in A_t(h_{t-1})$, and (2) $g^m(h_{t-1}, x_t) = p_t(h_{t-1}, x_t)$ 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

$$B_{ti}(h_{t-1}) = \{y_i \in A_{ti}(h_{t-1}) :$$

$$\int_{A_t(-i)(h_{t-1})} g_i(h_{t-1}, y_i, x_{t(-i)}) f_{t(-i)}(dx_{t(-i)}|h_{t-1}) > \int_{A_t(h_{t-1})} p_{ti}(h_{t-1}, x_t) f_t(dx_t|h_{t-1}).$$

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 β_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 Λ'_i is a measurable and compact valued correspondence, and hence has a Borel measurable selection β'_i . Note that β'_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

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

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 will 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

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

The first and the third equalities hold since $f_{tj}(B_{tj}(h_{t-1}) | h_{t-1}) = 0$ for each j , and hence $f_{t(-i)}(\prod_{j \neq i} B_{tj}^c(h_{t-1}) | h_{t-1}) = f_{t(-i)}(A_{t(-i)}(h_{t-1}) | h_{t-1})$. 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 $x_{t(-i)} \in \prod_{j \neq i} B_{tj}^c(h_{t-1})$. 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

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

The first and the third equalities hold since

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

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 $x_{t(-i)} \in \prod_{j \neq i} B_{tj}^c(h_{t-1})$. 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

1. $q(h_{t-1}, x_t, e, s_t) \in Q_{t+1}(h_{t-1}, x_t, s_t)$ for any $(h_{t-1}, x_t, e, s_t) \in \text{Gr}(P_t) \times S_t$;
2. $e = \int_{S_t} q(h_{t-1}, x_t, e, s_t) f_{t0}(ds_t | h_{t-1})$ 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 $f_t: H_{t-1} \rightarrow \otimes_{i \in I} \mathcal{M}(X_{ti})$ such that for all $h_{t-1} \in \tilde{H}_{t-1}$, properties (1)-(3) 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, 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 (1) ξ_m is a transition probability from H_{m-1} to $\mathcal{M}(X_m)$ for any $m \geq 1$, and (2) $\xi_m(A_m(h_{m-1})|h_{m-1}) = 1$ for any $m \geq 1$ and $h_{m-1} \in H_{m-1}$. Denote the set of all such ξ as Υ .

Fix any $t \geq 1$, define correspondences Ξ_t^t and Δ_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 follows:

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

and

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

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 $\xi \in \Upsilon$, 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 $\varrho_{(h_{t-1}, \xi)}^t = \varphi_{t0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^t$.⁵

⁴For $m \geq t \geq 1$ 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 $\dot{\varphi}_{m0}(h_{t-1}, \cdot)$ 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 $\xi \in \Upsilon$, since $\rho_{(h_{t-1}, \xi)}^m$ concentrates on $H_m(h_{t-1})$, $\varphi_{m0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^m = \dot{\varphi}_{m0}(h_{t-1}, \cdot) \circ \rho_{(h_{t-1}, \xi)}^m$. For notational simplicity, we still use $\varphi_{m0}(h_{t-1}, \cdot)$, instead of $\dot{\varphi}_{m0}(h_{t-1}, \cdot)$, to denote the above extension. Similarly, we can work with a suitable extension of the payoff function u as needed.

⁵For a set A in a space X , $\mathbf{1}_A$ is the indicator function of A , which is one on A and zero on $X \setminus A$.

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

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

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 Δ_t are nonempty and compact valued, and sectionally continuous.

Lemma B.8. 1. 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 $m_1 \geq t$.

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

Proof. (1) 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} , Ξ_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

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

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.

1. Notice that $H_0(s_0) = X_0$ for any $s_0 \in S_0$, which is compact.
2. Suppose that $H_{m'}(s_0, s_1, \dots, s_{m'})$ is compact for some $0 \leq m' \leq t-1$ and any $(s_0, s_1, \dots, s_{m'}) \in S^{m'}$.
3. 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 (6), 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 $H_t(s_0, s_1, \dots, s_t)$ 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 τ on S_t is λ_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

1. 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;
2. 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;
3. for any $h_{t-1} \in H_{t-1}$ and $\tau \in \Xi_t^{m_1-1}(h_{t-1})$, the marginal of τ on $\prod_{t \leq m \leq m_1-1} S_m$ is $\otimes_{t \leq m \leq m_1-1} \lambda_m$ 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 follows:

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

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 $A_{m_1}(\cdot, s_0, s_1, \dots, s_{m_1-1})$ 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 (6), 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 τ on $\prod_{t \leq m \leq m_1} S_m$ is $\otimes_{t \leq m \leq m_1} \lambda_m$ 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 ξ^∞ 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 ξ^∞ 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 ψ 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.

1. For each k , χ_k is jointly measurable and sectionally continuous on $\prod_{t \leq m \leq m_1} X_m$.
2. 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 $k \rightarrow \infty$.
3. 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 χ' 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 $k \rightarrow \infty$,

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

Then by Lemma B.7,

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

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 $\varrho_{(h_{t-1}^\infty, \xi^\infty)}^{m_1}$.

The 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 $m_1 \geq t$.

(2) We show that Δ_t is nonempty and compact valued, and sectionally continuous on X^{t-1} .

It is obvious that Δ_t is nonempty valued, 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 $\xi^k = (\xi_1^k, \xi_2^k, \dots) \in \Upsilon$ and $\tau^k = \varrho_{(h_{t-1}, \xi^k)}$ for each k .

By (1), Ξ_t^t is compact. Then there exists a measurable mapping g_t such that (1) $g^t = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, \xi_{t+1}^1, \dots) \in \Upsilon$, and (2) a subsequence of $\{\rho_{(h_{t-1}, \xi^k)}^t\}$, say $\{\rho_{(h_{t-1}, \xi^{k_l})}^t\}_{l \geq 1}$, which 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 $\{\rho_{(h_{t-1}, \xi^{k_l})}^{t+1}\}_{l \geq 1}$, say $\{\rho_{(h_{t-1}, \xi^{k_{2l})}^{t+1})}\}_{l \geq 1}$, which weakly converges to $\rho_{(h_{t-1}, g^{t+1})}^{t+1}$, where $g^{t+1} = (\xi_1^1, \dots, \xi_{t-1}^1, g_t, g_{t+1}, \xi_{t+2}^1, \dots) \in \Upsilon$.

Repeat 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 Δ_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 Δ_t . Suppose that $\tau^0 \in \Delta_t(h_{t-1}^0)$. Then there exists some $\xi \in \Upsilon$ 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 Δ_t^m is continuous, for each m , there exists some $\xi^m \in \Upsilon$ such that $d(\varrho_{(h_{t-1}^{k_m}, \xi^{k_m})}^m, \tilde{\tau}^m) \leq \frac{1}{m}$ for k_m sufficiently large, where d is the Prokhorov metric. Let

$\tau^m = \varrho_{(h_{t-1}^{k_m}, \xi^m)}$. Then τ^m weakly converges to τ^0 , which implies that Δ_t is sectionally lower hemicontinuous. \square

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

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

The lemma below presents several properties of the correspondence Q_t^τ .

Lemma B.9. *For any $t, \tau \geq 1$, Q_t^τ 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 will show that Q_t^τ is bounded, nonempty and compact valued, and sectionally upper hemicontinuous on X^{t-1} .

The boundedness and nonemptiness of Q_t^τ are obvious. We shall prove that Q_t^τ 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

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

where $\xi^k = (\xi_1^k, \xi_2^k, \dots) \in \Upsilon$ for each k . As Δ_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, which weakly converges to $\varrho_{(h_{t-1}^0, \xi^0)}$ for $\xi^0 = (\xi_1^0, \xi_2^0, \dots) \in \Upsilon$.

We shall show that

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

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

$$\left| a^0 - \int_{\prod_{m \geq t}(X_m \times S_m)} u(h_{t-1}^0, x, s) \varrho_{(h_{t-1}^0, \xi^0)}(\mathrm{d}(x, s)) \right| < \delta. \quad (2)$$

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 ξ'_j of $\mathcal{M}(A_j)$ such that ξ'_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 which 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, μ is measurable and sectionally continuous on $X^{\tilde{M}}$. Let

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

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

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

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

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

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

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

Thus, we proved inequality (2), which implies that Q_t^τ is sectionally upper hemicontinuous on X^{t-1} for $t > \tau$.

Furthermore, to prove that Q_t^τ is compact valued, we only need to consider 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 will show that Q_t^τ is measurable.

Fix a sequence (ξ'_1, ξ'_2, \dots) , where ξ'_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

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

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 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 has been defined such that it is bounded, measurable, nonempty, convex and compact valued, and sectionally continuous on X^j . Let

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

Let $\check{S}_j = S_j$.⁶ Since

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

we have

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

Let

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

⁶We will 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.

w_M^j is a Borel measurable selection of W_M^j .

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} .

It is easy to see that

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

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 $W^{t-1} = \overline{\cup_{M \geq t} W_M^{t-1}}$. That is, W^{t-1} is the closure of $\cup_{M \geq t} W_M^{t-1}$, 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^τ is compact valued. Second, fix h_{t-1} and $q \in Q_t^\tau(h_{t-1})$. Then there exists a mapping $\xi \in \Upsilon$ such that

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

For $M \geq t$, let

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

and

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

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^τ is measurable for $t > \tau$.

Step 3. For $t \leq \tau$, we can start with $Q_{\tau+1}^\tau$. Repeating the backward induction in Subsection B.4.1, we have that Q_t^τ 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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$.⁷

Lemma B.10. 1. *The correspondence Q_t^∞ is bounded, measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on X^{t-1} .*

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

Proof. (1) It is obvious that Q_t^∞ is bounded. By the definition of Q_t^τ , for λ^{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^τ is nonempty and compact valued, $Q_t^\infty = \bigcap_{\tau \geq 1} Q_t^\tau$ is nonempty and compact valued for λ^{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 (2), $\bigcap_{\tau \geq 1} Q_t^\tau$ is measurable, which implies that Q_t^∞ 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 τ . By Lemma 2 (7), $Q_t^\tau(\cdot, s^{t-1})$ has a closed graph for each τ , which implies that $Q_t^\infty(\cdot, s^{t-1})$ has a closed graph. Referring to Lemma 2 (7) again, $Q_t^\infty(\cdot, s^{t-1})$ is upper hemicontinuous on $H_{t-1}(s^{t-1})$. Since Q_t^τ is essentially upper hemicontinuous on X^{t-1} for each τ , Q_t^∞ is essentially upper upper hemicontinuous on X^{t-1} .

(2) For any $\tau \geq 1$ and λ^{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 following 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 Subsection B.4.1 shows that $\Phi(Q_{t+1})$ is measurable, nonempty and compact valued, and essentially sectionally upper hemicontinuous on $X^t \times \{1, 2, \dots, \infty\}$.

⁷The proofs for Lemmas B.10 and B.12 follow the standard ideas with various modifications; see, for example, [3], [4] and [5].

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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$. \square

Though the definition of Q_t^τ 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 λ^{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^∞ such that for λ_{k-1} -almost all $h_{k-1} \in H_{k-1}$,

1. $f_k(h_{k-1})$ 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

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

- 2.

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

We need to show that $c_1(h_0)$ is a subgame-perfect equilibrium payoff vector for λ_0 -almost all $h_0 \in H_0$.

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

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

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, $c_k(h_{k-1}) \in Q_k^\infty(h_{k-1}) = \cap_{\tau \geq 1} Q_k^\tau(h_{k-1})$

for λ_{k-1} -almost all $h_{k-1} \in H_{k-1}$. Since $Q_k^\tau = \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)$ for $k \leq \tau$, $c_k(h_{k-1}) \in \cap_{\tau \geq k} \Phi^{\tau-k+1}(Q_{\tau+1}^\tau)(h_{k-1}) \subseteq \Phi^{M-k+1}(Q_{M+1}^M)(h_{k-1})$ for λ_{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 $\xi \in \Upsilon$ such that for λ_{M-1} -almost all $h_{M-1} \in H_{M-1}$,

- i. $f_M(h_{M-1})$ 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.

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

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

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

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

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

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

converges to

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

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

$$c_k(h_{k-1}) = \int_{\prod_{m \geq k} (X_m \times S_m)} u(h_{k-1}, x, s) \varrho_{(h_{k-1}, f)}(d(x, s)). \quad (3)$$

Step 2. Below, 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 follows: $\tilde{f}_i^k = (g_{1i}, \dots, g_{ki}, f_{(k+1)i}, f_{(k+2)i}, \dots)$. 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 (1) and (2) of step 1 and Equation (3) 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 $\otimes_{k+1 \leq j \leq \tilde{M}} \lambda_j(D^{\tilde{M}}(s^k)) = 1$ 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 $\hat{D}^k = (\cap_{\tilde{M} > k} E_k^{\tilde{M}}) \cap D^k$. 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

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

The first and the last equalities follow from Equation (3) in the end of step 1. The second equality is due to (2) in step 1. The first inequality is based on (1) in step 1. The second inequality holds by the following arguments:

- i. by the choice of h_k and (1) in step 1, for λ_{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

$$\begin{aligned} & \int_{A_{k+2}(h_k, x_{k+1}, s_{k+1})} \int_{S_{k+2}} c_{(k+3)i}(h_k, x_{k+1}, s_{k+1}, x_{k+2}, s_{k+2}) \\ & f_{(k+2)0}(\mathbf{d}s_{k+2}|h_k, x_{k+1}, s_{k+1}) f_{(k+2)(-i)} \otimes f_{(k+2)i}(\mathbf{d}x_{k+2}|h_k, x_{k+1}, s_{k+1}) \end{aligned}$$

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

- ii. since $f_{(k+1)0}$ is absolutely continuous with respect to λ_{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

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

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

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

converges to

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

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

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

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 (3), 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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$.*

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

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

1. $c_t = \int_{A_t(h_{t-1})} \int_{S_t} q_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) \alpha(dx_t)$, 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 λ_t -almost all $s_t \in S_t$ and $x_t \in A_t(h_{t-1})$;
2. $\alpha \in \otimes_{i \in I} \mathcal{M}(A_{ti}(h_{t-1}))$ 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 $\prod_{i \in I} A_{ti}(h_{t-1})$,

then $c_t \in \Phi(Q_{t+1}^\tau)(h_{t-1})$.

Fix a subgame $h_{t-1} = (x^{t-1}, \acute{s}^{t-1})$ such that $\acute{s}^{t-1} \in \acute{S}^{t-1}$. Pick a point $c_t \in E_t(\acute{s}^{t-1})$. 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 we have

1. $c_t = \int_{A_t(h_{t-1})} \int_{S_t} c_{t+1}(h_{t-1}, x_t, s_t) f_{t0}(ds_t|h_{t-1}) f_t(dx_t|h_{t-1})$;
2. $f_t(\cdot|h_{t-1})$ 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 λ_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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$.

(2) 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 (1), then we can show that $E_t(h_{t-1}) \subseteq Q_t^\tau(h_{t-1})$ for λ^{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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$.

(3) 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 λ^{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 λ^{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 λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$. \square

B.5 Proof of Proposition B.1

We will highlight the needed changes in comparison with the proofs presented in Subsections B.4.1-B.4.3.

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

If $N_t = 1$, then $S_t = \{s'_t\}$. Thus, $P_t(h_{t-1}, x_t) = Q_{t+1}(h_{t-1}, x_t, s'_t)$, 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 follows: $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$ if

1. $v = p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{tj}^*)$ such that $p_t(h_{t-1}, \cdot)$ is a measurable selection of $P_t(h_{t-1}, \cdot)$;⁸
2. $x_{tj}^* \in A_{tj}(h_{t-1})$ is a maximization point of player j given the payoff function $p_{tj}(h_{t-1}, A_{t(-j)}(h_{t-1}), \cdot)$ and the action space $A_{tj}(h_{t-1})$, $\alpha_i = \delta_{A_{ti}(h_{t-1})}$ for $i \neq j$ and $\alpha_j = \delta_{x_{tj}^*}$;
3. $\mu = \delta_{p_t(h_{t-1}, A_{t(-j)}(h_{t-1}), x_{tj}^*)}$.

This is a single agent problem. We need to show that Φ_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 [7], to prove the nonemptiness, compactness, and upper hemicontinuity of Φ_t . In [7], 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 $d_0 \in P(z_0)$.

Repeat the argument in the proof of the main theorem of [7], one can show that Φ_t is nonempty and compact valued, and upper hemicontinuous.

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

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 Sections B.3: there exists a bounded, measurable, nonempty and compact valued correspondence Φ_t from H_{t-1} to $\mathbb{R}^n \times \mathcal{M}(X_t) \times \Delta(X_t)$ such that Φ_t is essentially sectionally upper hemicontinuous on $X^{t-1} \times \hat{S}^{t-1}$, and for λ^{t-1} -almost all $h_{t-1} \in H_{t-1}$, $(v, \alpha, \mu) \in \Phi_t(h_{t-1})$ if

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

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 Subsection 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 Subsection B.4.1.

2. Forward induction: unchanged.

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 Ξ_t^t as follows: in the subgame h_{t-1} ,

$$\Xi_t^t(h_{t-1}) = (\mathcal{M}(A_t(h_{t-1})) \diamond \hat{f}_{t0}(h_{t-1}, \cdot)) \otimes \lambda_t.$$

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 follows:

$$\begin{aligned} \Xi_t^{m_1}(h_{t-1}) = & \{g(h_{t-1}) \diamond ((\xi_{m_1}(h_{t-1}, \cdot) \diamond \hat{f}_{m_10}(h_{t-1}, \cdot)) \otimes \lambda_{m_1}) : \\ & g \text{ is a Borel measurable selection of } \Xi_t^{m_1-1}, \\ & \xi_{m_1} \text{ is a Borel measurable selection of } \mathcal{M}(A_{m_1}) \}. \end{aligned}$$

Then the result in Subsection B.4.3 is true with the above $\Xi_t^{m_1}$.

Consequently, a subgame-perfect equilibrium exists.

References

- [1] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin, 2006.
- [2] C. Castaing, P. R. De Fitte and M. Valadier, *Young Measures on Topological Spaces: with Applications in Control Theory and Probability Theory*, Vol. 571, Springer Science & Business Media, 2004.
- [3] C. Harris, The Existence of Subgame-Perfect Equilibrium with and without Markov Strategies: A Case for Extensive-Form Correlation, Discussion Paper No. 53, Nuffield College, Oxford, 1990.
- [4] C. Harris, P. J. Reny and A. Robson, The Existence of Subgame-Perfect Equilibrium in Continuous Games with Almost Perfect Information: A Case for Public Randomization, *Econometrica* **63** (1995), 507–544.
- [5] T. Mariotti, Subgame-Perfect Equilibrium Outcomes in Continuous Games of Almost Perfect Information, *Journal of Mathematical Economics* **34** (2000), 99–128.
- [6] J.-F. Mertens and T. Parthasarathy, Equilibria for Discounted Stochastic Games, in *Proceedings of the NATO Advanced Study Institute on Stochastic Games and Applications*, ed. by A. Neyman and S. Sorin. NATO Science Series. Dordrecht: Kluwer Academic, 2003, 131–172.
- [7] L. K. Simon and W. R. Zame, Discontinuous Games and Endogenous Sharing Rules, *Econometrica* **58** (1990), 861–872.