Supplement to

Markov distributional equilibrium dynamics in games with complementarities and no aggregate risk

Łukasz Balbus* Paweł Dziewulski† Kevin Reffett‡

Łukasz Woźny§

June 2020

Abstract

This supplement contains the proofs omitted from the main text of the paper as well as preliminaries on the law of large numbers and lattice theory.

A Preliminaries

In this section we introduce some mathematical notions in measure and lattice theory that are employed in our main analysis.

A.1 Fubini extensions and the law of large numbers

We begin by defining the notion of *super-atomless* probability space.[1] Let $(\Lambda, \mathcal{L}, \lambda)$ be a probability space. For any $E \in \mathcal{L}$ such that $\lambda(E) > 0$, let $\mathcal{L}^E := \{E \cap E' : E' \in \mathcal{L}\}$

* Faculty of Mathematics, Computer Sciences and Econometrics, University of Zielona Góra, Poland.

† Department of Economics, University of Sussex, UK.

‡ Department of Economics, Arizona State University, USA.

§ Department of Quantitative Economics, Warsaw School of Economics, Warsaw, Poland.

[1] The following definition is by Podczeck (2009, 2010), which we find to be the most convenient for our purposes. However, equivalent definitions are provided in Hoover and Keisler (1984), who call such spaces $\aleph_1$-*atomless*, and Keisler and Sun (2009), who dubbed such spaces *rich*.

and λ^E be the re-scaled measure from the restriction of λ to $\mathcal{L}^E$. Let $\mathcal{L}^E_\lambda$ be the set of equivalence classes of sets in $\mathcal{L}^E$ such that $\lambda^E(E_1 \triangle E_2) = 0$, for $E_1, E_2 \in \mathcal{L}^E$.[2] We endow the space with metric $d^E : \mathcal{L}^E_\lambda \times \mathcal{L}^E_\lambda \to \mathbb{R}$ given by $d^E(E_1, E_2) := \lambda^E(E_1 \triangle E_2)$.

Definition 1 (Super-atomless space)**.** A probability space $(\Lambda, \mathcal{L}, \lambda)$ is super-atomless if for any $E \in \mathcal{L}$ with $\lambda(E) > 0$, the space $(\mathcal{L}^E_\lambda, d^E)$ is non-separable.

Classical examples of super-atomless probability spaces include: $\{0, 1\}^I$ with its usual measure when I is an uncountable set; the product measure $[0, 1]^I$, where each factor is endowed with Lebesgue measure and I is uncountable;[3] subsets of these spaces with full outer measure when endowed with the subspace measure, or an atomless Loeb probability space. Furthermore, any atomless Borel probability measure on a Polish space can be extended to a super-atomless probability measure (see Podczeck, 2009).

Given a probability space $(\Lambda, \mathcal{L}, \lambda)$, a collection of random variables $(X_\alpha)_{\alpha\in\Lambda}$ is *essentially pairwise independent*, if for $(\lambda \otimes \lambda)$-almost every $(\alpha, \alpha') \in \Lambda \times \Lambda$, random variables X_α and $X_{\alpha'}$ are independent. For any set Ω and $E \subseteq (\Lambda \times \Omega)$, we denote its sections by $E_\alpha := \big\{\omega \in \Omega : (\alpha, \omega) \in E\big\}$ and $E_\omega := \big\{\alpha \in \Lambda : (\alpha, \omega) \in E\big\}$, for any $\alpha \in \Lambda$ and $\omega \in \Omega$. Similarly, for any function f defined over $\lambda \times \Omega$, let f_α and f_ω denote the section of f for a fixed α, ω, respectively. Consider the following definition.

Definition 2 (Fubini extension)**.** The probability space $(\Lambda \times \Omega, \mathcal{L} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ is a *Fubini extension* of the natural product of probability spaces $(\Lambda, \mathcal{L}, \lambda)$ and $(\Omega, \mathcal{F}, P)$ if:

(i) $\mathcal{L} \boxtimes \mathcal{F}$ includes all sets from $\mathcal{L} \otimes \mathcal{F}$;

(ii) for an arbitrary set $E \in \mathcal{L}\boxtimes\mathcal{F}$ and $(\lambda\otimes P)$-almost every $(\alpha, \omega) \in \Lambda\times\Omega$, the sections E_α and E_ω are $\mathcal{F}$- and $\mathcal{L}$-measurable, respectively, while

$$(\lambda \boxtimes P)(E) \;=\; \int_\Omega \lambda(E_\omega)P(d\omega) \;=\; \int_\Lambda P(E_\alpha)\lambda(d\alpha).$$

A Fubini extension is *rich*, if there is a $(\mathcal{L}\boxtimes\mathcal{F})$-measurable function $X : \Lambda\times\Omega \to \mathbb{R}$ such that the random variables $(X_\alpha)_{\alpha\in\Lambda}$ is essentially pairwise independent and the random variable X_α has the uniform distribution over $[0, 1]$, for λ-almost every $\alpha \in \Lambda$.

[2] We denote $E_1 \triangle E_2 := (E_1 \setminus E_2) \cup (E_2 \setminus E_1)$.

[3] Indeed, Maharam's theorem shows that the measure algebra of every super-atomless probability spaces must correspond to the countable convex combination of such spaces. See Maharam (1942).

Existence of a rich Fubini extension is proven in Proposition 5.6 of Sun (2006), for $\Lambda = [0,1]$. Moreover, $\mathcal{L}$ can not be a collection of Borel subsets of Λ (see Proposition 6.2 in Sun, 2006). In fact, Podczeck (2010) there exists a rich Fubini extension if and only if the space is super-atomless. Moreover, without loss, one may assume the random variables $(X_\alpha)_{\alpha\in\Lambda}$ to be independent, rather than pairwise-independent.

A *process* is a $(\mathcal{L} \boxtimes \mathcal{F})$-measurable function with values in a Polish space. For any process f and set $E \in \mathcal{L}$ such that $\lambda(E) > 0$, we denote the restriction of f to $E \times \Omega$ by f^E. Naturally, $\mathcal{L}^E \boxtimes \mathcal{F} := \big\{W \in \mathcal{L} \boxtimes \mathcal{F} : W \subseteq E \times \Omega\big\}$ and $(\lambda^E \boxtimes P)$ is a probability measure re-scaled from the restriction of $(\lambda \boxtimes P)$ to $(\mathcal{L}^E \boxtimes \mathcal{F})$. The following version of (exact) Law of Large Numbers is by Sun (2006).

Proposition 1 (Law of Large Numbers)**.** *Suppose that f is a process from a rich Fubini extension $(\Lambda \times \Omega, \mathcal{L} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ to some Polish space. Then, for all $E \in \mathcal{L}$ such that $\lambda(E) > 0$ and P-almost every $\omega \in \Omega$, we have $\lambda(f^E_\omega)^{-1} = (\lambda^E \boxtimes P)(f^E)^{-1}$.*[4]

A.2 Lattices, chains, and fixed points

A *partial order* $\geq_X$ over a set X is a reflexive, transitive, and antisymmetric binary relation. A partially ordered set, or a *poset*, is a pair $(X, \geq_X)$ consisting of a set X and a partial order $\geq_X$. Whenever it causes no confusion, we denote $(X, \geq_X)$ with X.

For any x, $x' \in X$, their *infimum* (the greatest lower bound) is denoted by $x \wedge x'$, and their *supremum* (the least upper bound) by $x \vee x'$. The poset X is a *lattice* if for any x, $x' \in X$ both $x \wedge x'$ and $x \vee x'$ belong to X. Set A is a *sublattice* of X, if $A \subseteq X$ and it is a lattice with the induced order, with $x \wedge x'$ and $x \vee x'$ defined with $\geq_X$.[5]

For any subset A of a poset X, we denote the *supremum* and *infimum* of A by $\bigvee A$ and $\bigwedge A$, respectively.[6] A lattice X is *complete*, if each both $\bigvee A$ and $\bigwedge A$ belong to X,

[4] Given the probability space $(\Lambda, \mathcal{L}, \lambda)$ and a measurable function $f : \Lambda \to Y$, we denote measure $\lambda f^{-1}(U) := \lambda\big(\{\alpha \in A : f(\alpha) \in U\}\big)$, for any measurable subset U of Y.

[5] A basic example of a lattice is the Euclidean space $\mathbb{R}^\ell$ endowed with the natural product order $\geq$, i.e., we have $x' \geq x$ if $x'_i \geq x_i$, for all $i = 1, \ldots, \ell$. In this case, we have $x \wedge x'$ and $x \vee x'$ are given by $(x \wedge x')_i = \min\{x_i, x'_i\}$ and $(x \vee x')_i = \max\{x_i, x'_i\}$, for all $i = 1, \ldots, \ell$.

[6] This is to say that, $\bigvee A$ is the least element of X such that $\bigvee A \geq a$, for all $a \in A$. Clearly, by definition, we have $x \vee x' = \bigvee\{x, x'\}$. We define $\bigwedge A$ analogously.

for any $A \subseteq X$. We define a *complete sublattice* analogously.

A function $f : X \to \mathbb{R}$ over a lattice X is *supermodular* in x if $f(x \wedge x') + f(x \vee x') \geq f(x)+f(x')$. If X and T are posets, then function $f : X \times T \to \mathbb{R}$ has *increasing differences* in (x, t) if, for any $x' \geq_X x$ and $t' \geq_T t$, we have $f(x', t') - f(x, t') \geq f(x', t) - f(x, t)$.

Finally, correspondence $\Gamma : X \times Y \to Z$, where X and Y are posets and Z is a lattice, satisfies *strict complementarities* if for any $x' \geq x$, $y' \geq y$, $z \in \Gamma(x, y')$, and $z' \in \Gamma(x', y)$, we have $z \wedge z' \in \Gamma(x, y)$ and $z \vee z' \in \Gamma(x', y')$.

B Auxiliary results

Lemma B.1. *Let $(\Xi, \geq)$ be a poset with its order topology, and $\{f_k\}$ be a sequence of increasing and monotone inf-preserving functions $f_k : \Xi \to \mathbb{R}$. Whenever $x_k \downarrow x$ in Ξ and $f_k \downarrow f$ (pointwise), then $f_k(x_k) \to f(x)$.*

Proof. Let $n \in \mathbb{N}$. Since $\{f_k\}$ is decreasing sequence of increasing functions and $x_k \downarrow x$, then $k \geq n$ implies $f(x) \leq f_k(x_k) \leq f_k(x_n)$. Thus, we have $f(x) \leq \liminf_{k\to\infty} f_k(x_k) \leq \limsup_{k\to\infty} f_k(x_k) \leq f(x_n)$. To finish the proof, let $n \to \infty$. $\square$

Lemma B.2. *Let $\{\nu_k\}$ be a sequence of probability measures on a Polish space S, and $\{h_k\}$ be a sequence of bounded, measurable functions $h_k : S \to \mathbb{R}$. If $\nu_k \downarrow \nu$ (stochastically and in weak topology) and $h_k \downarrow h$, then $\lim_{k\to\infty} \int h_k d\nu_k = \int h d\nu$.*

Proof. It is a consequence of Lemma B.1, where Ξ is a space of bounded, measurable, real valued functions on S, and $f_k(x) := \int_S x(s)\nu_k(ds)$, $x_k(s) = h_k(s)$. $\square$

Lemma B.3. *Let S_1, S_2 be topological spaces and $f : S_1 \times S_2 \mapsto \mathbb{R}$ be a continuous function. Let $\Gamma : S_1 \rightrightarrows S_2$ be a continuous, compact-valued correspondence and $\Gamma^*(x) := \arg\max_{y\in\Gamma(x)} f(x, y)$. If $x_k \to x$ in S_1, $y_k \to y$ in S_2, and $y_k \in \Gamma^*(x_k)$, then $y \in \Gamma^*(x)$.*

Proof. Let $y' \in \Gamma(x)$. By continuity of Γ, for any $k \in \mathbb{N}$, there is $y'_k \in \Gamma(x_k)$ such that $y'_k \to y'$. Since $y_k \in \Gamma^*(x_k)$, we have $f(x_k, y_k) \geq f(x_k, y'_k)$, for all $k \in \mathbb{N}$. By continuity of f, we have $f(x, y) \geq f(x, y')$. Since $y' \in \Gamma(x)$ is arbitrary, hence $y \in \Gamma^*(x)$. $\square$

C Omitted proofs

Proof of Proposition 4. This argument is analogous to Echenique (2005). Let $\bar{x}$ be the greatest element of X. Let $\mathscr{I}$ be a set of ordinal numbers with cardinality strictly greater than X. Define the following transfinite sequence with the initial element $x_0 = \bar{x}$ and $x_i = \bigwedge \{f(x_j) : j < i\}$, for $i \in \mathscr{I} \setminus \{0\}$. We claim that $\{x_i\}$ is a well-defined decreasing sequence. Clearly $x_1 = f(x_0) \leq x_0$. Suppose that $\{x_j\}_{j<i}$ is well-defined and decreasing for some i. Then $\{f(x_j)\}_{j<i}$ is a decreasing sequence, that has an infimum equal to x_i. Consequently x_j is well defined and decreasing on $[0, i]$. By transfinite induction, the transfinite sequence $\{x_i\}_{i \in \mathscr{I}}$ is well defined and decreasing. Since $\mathscr{I}$ has the cardinality strictly greater than X, there is no one-to-one mapping between $\mathscr{I}$ and X. Consequently, take the least element $\bar{i}$ in $\{i \in \mathscr{I} : x_i = x_{i+1}\}$. Then $x_{\bar{i}} = x_{\bar{i}+1} = f(x_{\bar{i}})$, and $e^* := x_{\bar{i}}$ is a fixed point of f. To show that $e^* = \bigvee \{x \in X : f(x) \geq x\}$, set $\mathcal{X} := \{x \in X : f(x) \geq x\}$. Obviously, we have $e^* \in \mathcal{X}$. For any other $y \in \mathcal{X}$, we have $y \leq x_0$. Suppose there is $i \in \mathscr{I}$ such that $y \leq x_j$, for any $j < i$. Since $y \in \mathcal{X}$, by transfinite induction, we have $y \leq f(y) \leq f(x_j)$. Thus, $y \leq \bigwedge \{f(x_j) : j \leq i\}$ and $y \leq x_i$, for any $i \in \mathscr{I}$, including $\bar{i}$. □

Proof of Theorem 1. By Proposition 5.6 of Sun (2006) and Theorem 1 in Podczeck (2010) there is a probability space $(\Omega, \mathcal{F}, P)$ and a rich Fubini extension of a natural product space on $\Lambda \times \Omega$, denoted by $(\Lambda \times \Omega, \mathcal{L} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. Consequently, we can find a process $\eta : \Lambda \times \Omega \to [0, 1]$ such that the family $(\eta_\alpha)_{\alpha \in \Lambda}$ is essentially pairwise independent with the uniform distribution on $[0, 1]$. Define $(\eta_n)_{n \in \mathbb{N}}$ as a set of independent copies of η. Construct a sequence $(X_n)_{n=1}^{\infty}$ satisfying theses (i)–(iii). Let $(I, \mathcal{I}, \iota)$ be the standard interval $I = [0, 1]$, with Borel sets $\mathcal{I}$, and the Lebesgue measure ι. For any $\mu \in \mathcal{M}$, there is a $(\mathcal{I} \otimes \mathcal{T} \otimes \mathcal{A})$-measurable function $G^\mu : I \times T \times A \mapsto T$ such that

$$\iota\big(G^\mu_{(t,a)}\big)^{-1}(Z) \;=\; \iota\Big(\big\{l \in I : G^\mu(l, t, a) \in Z\big\}\Big) \;=\; q(Z|t, a, \mu),$$

for any $Z \in \mathcal{T}$.[7] For any initial distribution $\tau_1 \in \mathcal{M}_T$, there exists a T-valued $(I \otimes \mathcal{T})$-measurable function $\tilde{G}$ such that $\tau_0 = \iota \tilde{G}^{-1}$.[8] Put $X_1 := \tilde{G}(\eta_1)$. Having the initial random

[7] For example, see Lemma A5 in Sun (2006).

[8] Again, see Lemma A5 in Sun (2006).

variable X_1, define the following process $X_{n+1} = G^{\mu_n}(\eta_{n+1}, K_n)$, for $n > 1$, where $K_n := \big(X_n, \sigma(X_n, \tau_n)\big)$, $\tau_n := (\lambda \boxtimes P)X_n^{-1}$, and $\mu_n := (\lambda \boxtimes P)K_n^{-1}$. As usual, put $(K_n)_\alpha(\omega) := K_n(\alpha, \omega)$ for $(\alpha, \omega) \in \Lambda \times \Omega$. Let $\mathcal{S}_n$ by the sigma field generated by $\{\eta_k : k \leq n\}$. By definition of X_1 and X_{n+1}, we conclude that X_n is $\mathcal{S}_n$-measurable. Hence, $(X_n)_\alpha$ and $(\eta_{n+1})_\alpha$ are independent, for λ-almost every $\alpha \in \Lambda$. We show that (i)–(ii) are satisfied by induction on n. For $n = 1$, the claim holds by essential independence of η_1 and X_1. Moreover, by Proposition 1, for P-almost every $\omega \in \Omega$ the sampling distribution $\lambda(X_1)_\omega^{-1}$ of X_1, i.e., satisfies $\lambda(X_1)_\omega^{-1} = (\lambda \boxtimes P)X_1^{-1} = \tau$. Again by Proposition 1, for P-almost all $\omega \in \Omega$, we have $\lambda(K_1)_\omega^{-1} = (\lambda \boxtimes P)K_1^{-1} := \mu_1$. Hence, (ii) is satisfied for $n = 1$. Suppose that both (i) and (ii) hold, for some $n \geq 1$. Observe that $\big((\eta_{n+1})_\alpha, (X_n)_\alpha\big)_{\alpha \in \Lambda}$ is a family $(\lambda \otimes \lambda)$-almost everywhere pairwise conditionally independent random variables. This follows from induction hypothesis for $(X_n)_\alpha$, and the previous observation that random variables $(X_n)_\alpha$ and $(\eta_{n+1})_\alpha$ are independent λ-almost surely. Hence, by construction of X_{n+1}, the family $\big((X_{n+1})_\alpha\big)_{\alpha \in \Lambda}$ is $(\lambda \otimes \lambda)$-almost surely pairwise conditionally independent. Hence the property (i) is satisfied for $(n+1)$. By Proposition 1, we obtain (ii) for $(n+1)$. Thus, (i) and (ii) hold for all $n \geq 1$. To show (iii), let $(\mathcal{S}_n)_\alpha$ be the sigma field generated by $\{(\eta_k)_\alpha : k \leq n\}$ and similarly $(\Sigma_n)_\alpha$ by $\{(X_k)_\alpha : k \leq n\}$. By definition of X_n and $(\Sigma_n)_\alpha$ we conclude that $\sigma\big((X_n)_\alpha\big) \subseteq (\Sigma_n)_\alpha \subset (\mathcal{S}_n)_\alpha$. Let E be the standard expectation with respect to P. Hence the conditional distribution of $(X_{n+1})_\alpha$ with respect to $(\Sigma_n)_\alpha$ satisfies

$$
\begin{aligned}
P\big((X_{n+1})_\alpha \in Z \big| (\Sigma_n)_\alpha\big) \;&=\; E\Big[P\big((X_{n+1})_\alpha \in Z \big| (\mathcal{S}_n)_\alpha\big) | (\Sigma_n)_\alpha\Big] \\
&=\; E\Big[P\big(G^{\mu_n}((\eta_{n+1})_\alpha, (K_n)_\alpha) \in Z \big| (\mathcal{S}_n)_\alpha\big) | (\Sigma_n)_\alpha\Big] \\
&=\; E\Big[q\big(Z | (K_n)_\alpha, \mu_n\big) \big| (\Sigma_n)_\alpha\Big] \;=\; q\Big(Z \big| (X_n)_\alpha, \sigma^*\big((X_n)_\alpha, \tau_n\big), \mu_n\Big)
\end{aligned}
$$

for λ-almost all $\alpha \in \Lambda$ and all $Z \in \mathcal{T}$, where the last equality follows from independence of $(\eta_{n+1})_\alpha$ and $(X_n)_\alpha$. Hence, property (iii) is satisfied. $\square$

Proof of Lemma 1. Suppose that $v_n \in \mathcal{V}$, for all $n \in \mathbb{N}$, and $v_n \to v$. Furthermore, let (μ_k) and (Φ_k) be decreasing sequences in $\mathcal{M}$ and $\mathcal{D}$, respectively, such that $\mu_k \to \mu$ (weakly) and $\Phi_k \to \Phi$ (pointwise). Take any $t \in T$ and $\epsilon > 0$. There is $n_0 \in \mathbb{N}$ such that, for all $k \in \mathbb{N}$ and $n \geq n_0$, we have

$$\begin{aligned}|v(t,\mu_k,\Phi_k) - v(t,\mu,\Phi)| \;&\leq\; |v(t,\mu_k,\Phi_k) - v_n(t,\mu_k,\Phi_k)| + |v_n(t,\mu_k,\Phi_k) - v_n(t,\mu,\Phi)| \\ &+ |v_n(t,\mu,\Phi) - v(t,\mu,\Phi)| \;\leq\; \frac{2}{3}\epsilon + |v_n(t,\mu_k,\Phi_k) - v_n(t,\mu,\Phi)| \quad (1)\end{aligned}$$

Take any $n \in \mathbb{N}$ satisfying (1). Therefore, since $v_n \in \mathcal{V}$, for large enough k, we obtain $\big|v_n(t,\mu_k,\Phi_n) - v_n(t,\mu,\Phi)\big| \leq \epsilon/3$. Given (1), this implies $|v(t,\mu_k,\Phi_k) - v(t,\mu,\Phi)| < \epsilon$, for large k. Hence v is monotonically sup- and inf-preserving. Thus, $v \in \mathcal{V}$. $\square$

Continuation of the proof to Lemma 4. We prove (vi). Using Assumption 2, definition of $\mathcal{V}$, and Lemma 4, one can show that F is a Carathéodory function in (t,a), i.e., measurable in t and continuous in a. Hence, by Assumption 1 and Measurable Maximum Theorem (Theorem 18.19 in Aliprantis and Border, 2006) the correspondence $\Gamma(t,\mu;v,\Phi)$ is measurable in t, hence, weakly measurable.[9] For each $j = 1,2,\ldots,k$, the function $\pi_j(t) := \max_{a\in\Gamma(t,\mu;v,\Phi)} a_j$ is measurable (again, by Measurable Maximum Theorem). Thus, $t \to \overline{\gamma}(t,\mu,\Phi;v) = \big(\pi_1(t),\pi_2(t),\ldots,\pi_k(t)\big)$ is measurable. $\square$

Proof of Lemma 8. Suppose that $f : T\times A \mapsto \mathbb{R}$ belongs to the space of bounded and continuous function $C(T\times A)$. Clearly, we have $(1/N)f\big(\xi^N(\omega),\eta^N(\omega)\big) \to 0$, for all $\omega \in \Omega$. By the standard Kolmogorov Law of Large Numbers Theorem, we obtain

$$\lim_{N\to\infty}\frac{1}{N-1}\sum_{l\neq j} f\big(\tilde{T}_l,\sigma_n(\tilde{T}_l)\big) \;=\; \int_T f\big(t,\sigma_n(t)\big)\tau_n(dt) = \int_{T\times A} f(t,a)(\tau_n\star\sigma_n)(dt\times da),$$

$\mathbb{P}$-almost surely. Consequently, for $\mathbb{P}$-almost every $\omega \in \Omega$,

$$\lim_{N\to\infty}\int_{T\times A} f(t,a)\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big) \;=\; \int_{T\times A} f(t,a)(\tau_n\star\sigma_n)(dt\times da). \quad (2)$$

Let $\mathbf{F}$ be a countable, dense set in $C(T\times A)$. Let $\tilde{\Omega} \subseteq \Omega$ be such that any element of $\mathbf{F}$ obeys (2). Then, $\mathbb{P}(\tilde{\Omega}) = 1$. We claim that (2) holds for any $f \in C(T\times A)$ whenever $\omega \in \tilde{\Omega}$. Take any $\epsilon > 0$. Since $\mathbf{F}$ is dense in $C(T\times A)$, take $f_0 \in \mathbf{F}$ such that $\|f - f_0\|_\infty < \frac{\epsilon}{3}$. Then, $\int_{T\times A}|f(t,a) - f_0(t,a)|\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)(dt\times da) \leq \frac{\epsilon}{3}$ as well as

[9] See, e.g., Lemma 18.2 in Aliprantis and Border (2006).

$\int_{T\times A}|f(t,a)-f_0(t,a)|(\tau_n\star\sigma_n)(dt\times da)\le\frac{\epsilon}{3}$. This implies

$$\begin{aligned}\left|\int_{T\times A} f(t,a)\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)(dt\times da)-\int_{T\times A} f(t,a)(\tau_n\star\sigma_n)(dt\times da)\right|\\ \le \int_{T\times A}|f(t,a)-f_0(t,a)|\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)(dt\times da)\\ +\int_{T\times A}|f(t,a)-f_0(t,a)|(\tau_n\star\sigma_n)(dt\times da)\\ +\left|\int_{T\times A} f_0(t,a)\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)(dt\times da)-\int_{T\times A} f_0(t,a)(\tau_n\star\sigma_n)(dt\times da)\right|\le\\ \frac{2}{3}\epsilon+\left|\int_{T\times A} f_0(t,a)\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)(dt\times da)-\int_{T\times A} f_0(t,a)(\tau_n\star\sigma_n)(dt\times da)\right|.\end{aligned}\tag{3}$$

Since $\omega\in\tilde{\Omega}$, there exists an integer N_0 such that, for any $N>N_0$,

$$\left|\int_{T\times A} f_0(t,a)\hat{\mu}_n^N((\tilde{T}_{-j},\xi^N),\eta^N)(dt\times da)-\int_{T\times A} f_0(t,a)\tau_n\star\sigma_n(dt\times da)\right| < \frac{\epsilon}{3}.\tag{4}$$

Combining (3) and (4), for $N>N_0$, we have

$$\left|\int_{T\times A} f(t,a)\hat{\mu}_n^N((\tilde{T}_{-j},\xi^N),\eta^N)(dt\times da)-\int_{T\times A} f(t,a)\tau_n\star\sigma_n(dt\times da)\right| < \epsilon.\tag{5}$$

Since $\epsilon>0$, the (5) implies that (2) holds for f and $\omega\in\tilde{\Omega}$. Given that $f\in C(T\times A)$ is arbitrary and $\mathbb{P}(\tilde{\Omega})=1$, we have $\hat{\mu}_n^N\big((\tilde{T}_{-j},\xi^N),\eta^N\big)\to(\tau_n\star\sigma_n)$, almost surely. □

Recall that $\tilde{v}_1^N(t):=\sup_{\pi\in\Sigma}\mathcal{R}(\sigma^{-j},\pi)(t)$. Then, the Bellman equation for optimal value $\tilde{v}_n^N$, updated for any $n\in\mathbb{N}$, take the form of

$$\tilde{v}_n^N(t) = \max_{a\in\tilde{A}(t,\tau_n)}\left\{(1-\beta)r_n^N(t,a)+\beta\int_T \tilde{v}_{n+1}^N(t')q_n^N(ds'|t,a)\right\}.\tag{6}$$

Let $\mathbf{C}$ be the set of continuous real-valued functions on T, uniformly bounded by $\bar{r}$, which is a closed subset of a Banach space. The metric in product space $\mathcal{C}:=\mathbf{C}^\infty$ is embedded in the natural Banach space the following norm: For $v=(v_n)_{n\in\mathbb{N}}$, define

$$\|v\|^\zeta := \sum_{n=1}^{\infty}\frac{1}{\zeta^{n-1}}\sup_{t\in T}|v_n(t)|,$$

where $\zeta\in(0,1/\beta)$ is a fixed value. Clearly, $v^N\to v$ in $\|\cdot\|^\zeta$ if and only if $v_n^N\to v_n$, for any $n\in\mathbb{N}$. Let $v\in\mathcal{C}$, $t\in T$, and $B^N(v)(t):=\big(B_n^N(v)(t)\big)_{n\in\mathbb{N}}$ where

$$B_n^N(v)(t) := \max_{a\in\tilde{A}(t,\tau_n)}\left\{(1-\beta)r_n^N(t,a)+\beta\int_T v_{n+1}(t')q_n^N(dt'|t,a)\right\}.$$

Similarly, define $\mathcal{B}^N(v)(t) := \big(\mathcal{B}_n^N(v)(t)\big)_{n\in\mathbb{N}}$ where

$$\mathcal{B}_n^N(v)(t) \;:=\; (1-\beta)r_n^N\big(t,\sigma_n(t)\big) + \beta \int_T v_{n+1}(t')q_n^N\big(dt'|t,\sigma_n(t)\big).$$

For $v \in \mathcal{C}$, let $B^\infty(v)(t) := (B_n^\infty(v)(t))_{n\in\mathbb{N}}$ where

$$B_n^\infty(v) \;:=\; \max_{a\in\tilde{A}} \left\{(1-\beta)r_n(t,a) + \beta \int_T v_{n+1}(t')q_n(dt'|t,a)\right\},$$

where $r_n(t,a) := r(t,a,\tau_n \star \sigma_n)$ and $q_n(\cdot|t,a) := q(\cdot|t,a,\tau_n \star \sigma_n)$, for $(t,a) \in Gr\big(\tilde{A}(\cdot,\tau_n)\big)$. Similarly define $\mathcal{B}^\infty(v)(t) := \big(\mathcal{B}_n^\infty(v)(t)\big)_{n\in\mathbb{N}}$ where

$$\mathcal{B}_n^\infty(v)(t)' := \; (1-\beta)r_n\big(t,\sigma_n(t)\big) + \beta \int_T v_{n+1}(t')q_n\big(dt'|t,\sigma_n(t)\big).$$

Now we prove basic properties of B^N and B^∞.

Lemma C.1. *Let σ be a Borel measurable function. Then,*

(i) mappings $B^N, \mathcal{B}_n^N$, B^∞, and $\mathcal{B}_n^\infty$ map $\mathcal{C}$ into itself;

(ii) B^N, $\mathcal{B}_n^N$, B^∞, and $\mathcal{B}_n^N$ are $\beta\zeta$-contraction mappings on $\mathcal{C}$;

(iii) if $v^N \to v$ in $\mathcal{C}$, then $B^N(v^N) \to B^\infty(v)$ and $\mathcal{B}^N(v^N) \to \mathcal{B}^\infty(v)$ in $\mathcal{C}$;

(iv) we have $\|\tilde{v}^N - \tilde{v}^\infty\|_\infty \to 0$, where $\tilde{v}^N$, $\tilde{v}^\infty$ in $\mathbf{C}$ is a fixed point of B^N, B^∞;

(v) we have $\|\check{v}^N - \check{v}^\infty\|_\infty \to 0$, where $\check{v}^N$, $\check{v}^\infty$ in $\mathbf{C}$ is a fixed point of $\mathcal{B}^N$, $\mathcal{B}^\infty$.

Proof. In order to prove (i), take any $v \in \mathcal{C}$. Given Assumptions 6, for any n and N, the following functions $\Pi_n^N(t,a,v) = (1-\beta)r_n^N(t,a) + \beta\int_T v_{n+1}(t')q_n^N(dt'|t,a)$ and $\Pi_n^\infty(t,a,v) = (1-\beta)r_n(t,a) + \beta\int_T v_{n+1}(t')q_n(dt'|t,a)$, are both continuous in (t,a). Since $B_n^N(v)(t) = \max_{a\in\tilde{A}(t,\tau_n)} \Pi_n^N(t,a,v)$ and $B_n^\infty(v)(t) = \max_{a\in\tilde{A}(t,\tau_n)} \Pi_n^\infty(t,a,v)$, statement (i) follows immediately from Berge Maximum Theorem. We show (ii). It is routine to verify $\|B_n^N(v) - B_n^N(w)\|_\infty \le \beta||v_{n+1} - w_{n+1}||_\infty$, for $v,w \in \mathcal{C}$. By dividing both sides by ζ^{n-1} and summing over n, we obtain

$$\|B_n^N(v) - B_n^N(w)\|^\zeta \;=\; \sum_{n=1}^\infty \frac{\|B_n^N(v) - B_n^N(w)\|_\infty}{\zeta^{n-1}} \;\le\; \beta\zeta\sum_{n=1}^\infty \|v_n - w_n\|_\infty \;=\; \beta\zeta\|v_n - w_n\|_\infty.$$

An analogous argument can be applied to prove the property for B^∞. In order to show (iii), suppose that $v^N \to v$ in $(\mathcal{C}, ||\cdot||_\infty)$ and $(t^N,a^N) \to (t,a)$, for $(t^N,a^N) \in \tilde{A}(t^N,\tau_n)$.

We claim that $\Pi_n^N(t^N, a^N, v^N) \to \Pi_n^\infty(t, a, v)$. By Lemma 8and Assumption 6 we have that $r_n^N(t^N, a^N) \to r_n(t, a)$ and $q_n^N(\cdot|t^N, a^N) \to q_n(\cdot|t, a)$. This proves the claim. Furthermore, by (i), there is t^N such that

$$\sup_{t\in T} |B_n^N(v^N)(t) - B_n^\infty(v)(t)| \;=\; \|B_n^N(v^N)(t^N) - B_n^\infty(v)(t^N)\|.$$

Without loss of generality suppose that $t^N \to t$. Combining the definition of r_n and q_n, Lemma 8, and the above claim, it follows that the right hand-side above tends to 0. Hence, $\|B^N(v^N) - B^\infty(v^\infty)\|^\zeta \to 0$. Finally, to prove (iv), observe that

$$\begin{aligned}
\|\tilde{v}^N - \tilde{v}^\infty\|^\kappa \;&=\; \|B^N(\tilde{v}^N) - B^\infty(\tilde{v}^\infty)\|^\zeta \\
&\leq\; \|B^N(\tilde{v}^N) - B^N(\tilde{v}^\infty)\|^\zeta + \|B^N(\tilde{v}^\infty) - B^\infty(\tilde{v}^\infty)\|^\zeta \\
&\leq\; \beta\zeta||\tilde{v}^N - \tilde{v}^\infty||^\zeta + ||B^N(\tilde{v}^\infty) - B^\infty(\tilde{v}^\infty)||^\zeta,
\end{aligned}$$

where the last inequality is by (ii). Thus, $\|\tilde{v}^N - \tilde{v}^\infty\|^\kappa \leq \|B^N(\tilde{v}^\infty) - B^\infty(\tilde{v}^\infty)\|^\zeta/(1-\beta\zeta)$. To finish the proof, we only take $N \to \infty$, since by (iii) the right hand-side above tends to 0. The proof of (v) is analogous to (iv). $\square$

Lemma C.2. *Consider MDP, where $(\tau_n)_{n\in\mathbb{N}}$ and $(\sigma_n)_{n\in\mathbb{N}}$ are implied by sequences of distribution on types-policies for some MSDE (μ^*, Φ^*). Then, the sequences of value functions $\bar{v}$ for (μ^*, Φ^*) is a common fixed point of B^∞ and $\mathcal{B}^\infty$. As a result, $\bar{v} = \tilde{v}^\infty = \check{v}^\infty$.*

Proof. By Lemma C.1, it follows that B^∞ and $\mathcal{B}^\infty$ are both contractions on $\mathcal{C}$. Hence, we only need to show $\bar{v}$ is the fixed point of B^∞ and $\mathcal{B}^\infty$. By definition of $\bar{v}$, v^*, μ_n, and τ_n, for any $t \in T$, we have $\bar{v}_n(t) = v^*(t, \tau_n, \Phi^*)$ and

$$\begin{aligned}
\bar{v}_n(t) \;&=\; \max_{a\in\tilde{A}(t,\tau_n)} \left\{(1-\beta)r(t, a, \mu_n) + \beta\int_T v^*(t', \mu_{n+1}, \Phi^*)q(dt'|t, a, \mu_n)\right\} \\
&=\; \max_{a\in\tilde{A}(t,\tau_n)} \left\{(1-\beta)r(t, a, \mu_n) + \beta\int_T \bar{v}_{n+1}(t')q(dt'|t, a, \mu_n)\right\} \\
&=\; \max_{a\in\tilde{A}(t,\tau_n)} \left\{(1-\beta)r(t, a, \tau_n \star \sigma_n) + \beta\int_T \bar{v}_{n+1}(t')q(dt'|t, a, \tau_n \star \sigma_n)\right\} = B_n^\infty(\bar{v}_{n+1})(t).
\end{aligned}$$

Hence $\bar{v} = B^\infty(\bar{v})$ and by uniqueness of the fixed point of B^∞, $\bar{v} = \tilde{v}^\infty$. By the same argument we obtain $\bar{v} = \mathcal{B}^\infty(\bar{v})$, and $\bar{v} = \check{v}$. $\square$

Proof of Theorem 4. Let $\epsilon > 0$ and $\sigma = (\sigma_n)_{n\in\mathbb{N}}$ be a sequential policy function associated with (μ^*, Φ^*). If player j unilaterally deviates from σ using π then, for any $t \in T$, we have $\mathcal{R}^N((\sigma)^{-j}, \pi)(t) - \check{v}_1^N(t) \leq \tilde{v}_1^N(t_1^j) - \check{v}_1^N(t_1^j) \leq \|\tilde{v}_1^N - \check{v}_1^N\|_\infty$. By Lemma C.1, $\tilde{v}_1^N \to v_1^\infty$ and $\check{v}_1^N \to \check{v}_1^\infty$. Since the policy is $\sigma = \sigma^*$ and the initial state is $\tau_1 = \tau^*$, then $\check{v}_1^\infty = v^\infty$, by Lemma C.2. Thus, for large enough N, $\|\tilde{v}_1^N - \check{v}_1^N\|_\infty < \epsilon$. $\square$

References

Aliprantis, C. D. and K. C. Border (2006): *Infinite dimentional analysis. A hitchhiker's guide*, Springer Verlag: Heilbelberg.

Echenique, F. (2005): "A short and constructive proof of Tarski's fixed-point theorem," *International Journal of Game Theory*, 33, 215Ű218.

Hoover, D. and J. Keisler (1984): "Adapted probability distributions," *Transactions of the American Mathematical Society*, 286, 159–201.

Keisler, H. J. and Y. Sun (2009): "Why saturated probability spaces are necessary," *Advances in Mathematics*, 221, 1584 – 1607.

Maharam, D. (1942): "On homogeneous measure algebra," *Proceedings of the N.A.S.*, 28, 108–111.

Podczeck, K. (2009): "On purification of measure valued maps," *Economic Theory*, 38, 399–418.

——— (2010): "On existence of rich Fubini extensions," *Economic Theory*, 45, 1–22.

Sun, Y. (2006): "The exact law of large numbers via Fubini extension and characterization of insurable risks," *Journal of Economic Theory*, 126, 31–69.