SUPPLEMENT TO "UNCERTAIN REPEATED GAMES"

ILIA KRASIKOV
Department of Economics, Arizona State University

ROHIT LAMBA
Department of Economics, Cornell University

This Supplementary Appendix (SA) provides details and proofs omitted from Krasikov and Lamba (2026). Section 1 studies the statewise lower-bound condition in Theorem 2. Sections 2 and 3 supply, respectively, the formal results on symmetric games and the connection between XPE and dynamic variational preferences.

1. STATEWISE LOWER BOUNDS AND EX POST PAYOFFS

1.1. *A generalized SIR condition*

In this subsection, we introduce a generalization of ε-SIR used in the main text and discuss how Theorem 2 can be extended using this notion.

DEFINITION 1: Let $\underline{\boldsymbol{u}} = (\underline{u}(\theta))_{\theta\in\Theta} \in \mathbb{R}^{n\Theta}$. An *outcome* α is (ex post) $(\underline{\boldsymbol{u}},\varepsilon)$-*sequentially individually rational* ($(\underline{\boldsymbol{u}},\varepsilon)$-*SIR)* if for every environment $e \in \Theta^\infty$ and each on-path history $\widetilde{h}^t$ that is compatible with e,

$$\frac{1-\delta}{\delta} d_i(\alpha(\widetilde{h}^t)|\theta^t) \leq \mathbb{E}_t[U_i^\alpha(\widetilde{h}^{t+1}|e)] - \left[\varepsilon + (1-\delta)\sum_{s=t+1}^{\infty} \delta^{s-(t+1)} \underline{u}_i(\theta^s)\right] \ \forall i \in N. \quad (1)$$

The generalized notion termed $(\underline{\boldsymbol{u}},\varepsilon)$-SIR mirrors the logic of ε-SIR but allows for an additional set of tuning parameters, $\underline{\boldsymbol{u}}$, in the specification of punishment losses. Clearly, $(\underline{\boldsymbol{u}},\varepsilon)$-SIR coincides with ε-SIR whenever $\underline{u}(\theta) = 0$ for all stage games $\theta \in \Theta$. Examination of the proof of Theorem 2 reveals that it can be restated without the requirement of $\nu(\theta) = 0$ for all $\theta \in \Theta$ as follows:

Suppose that Int$(\mathscr{U}(\theta)) \cap \mathbb{R}^n_{++}$ is non-empty and let $\underline{\boldsymbol{u}} \in \mathbb{R}^{n\Theta}$ with $\underline{u}(\theta) \geq \nu(\theta)$ for all $\theta \in \Theta$. Then, for each sufficiently small $\varepsilon > 0$ there exists $\delta^\varepsilon < 1$ such that for every $\delta \geq \delta^\varepsilon$, an outcome α is justifiable if it is $(\underline{\boldsymbol{u}},\varepsilon)$-SIR.

According to this modification of Theorem 2, for each $\varepsilon > 0$ and $\underline{\boldsymbol{u}} \geq \boldsymbol{\nu}$, $(\underline{\boldsymbol{u}},\varepsilon)$-SIR outcomes are justifiable for large values of δ. However, in general, it may be impossible to push $\underline{\boldsymbol{u}}$ below $\boldsymbol{\nu}$ while retaining the conclusion of the claim that $(\underline{\boldsymbol{u}},\varepsilon)$-SIR outcomes are asymptotically justifiable.

The notable exception is the repeated games setting with a single stage game θ as established in the lemma that follows.

LEMMA 1: *Consider one stage game θ such that Int$(\mathscr{U}(\theta)) \cap \mathbb{R}^n_{++}$ is non-empty and let $\varepsilon > 0$. Then, there exist $\kappa^\varepsilon > 0$ and $\bar{\delta} < 1$ such that, for every $\delta \geq \bar{\delta}$, every ε-SIR outcome α is $(\nu(\theta), \frac{1}{2}\kappa^\varepsilon)$-SIR. As a result, the ε-SIR outcome α is justifiable.*

PROOF: To establish the result, we first show that $\nu_i^\varepsilon(\theta) := \inf_{w\in\mathscr{U}(\theta):w\geq\varepsilon} w_i > \nu_i(\theta)$ for all $i \in N$. By the way of contradiction, suppose that $\nu_i^\varepsilon(\theta)$ and $\nu_i(\theta)$ are equal for some player

i. Clearly, the infimum is attained by some $w^\varepsilon \in \mathscr{U}(\theta)$, which remains optimal even when ε is replaced by any $\varepsilon' < \varepsilon$, i.e., $w^\varepsilon = \nu_i^{\varepsilon'}(\theta)$. Since $\wedge \mathscr{U}(\theta) \leq 0$, there exists a point $w \in \mathscr{U}(\theta)$ such that $w_i \leq 0 < \nu_i(\theta)$. By convexity of $\mathscr{U}(\theta)$, the line segment connecting w and w^ε is contained in $\mathscr{U}(\theta)$, contradicting the optimality of w^ε for all $\varepsilon' < \varepsilon$ that is close enough to ε.

By the previous argument, there exists $\kappa^\varepsilon > 0$ so that $w \geq \kappa^\varepsilon + \nu(\theta)$ whenever $w \in \mathscr{U}(\theta)$ and $w \geq \varepsilon$. Hence, since stage game payoffs are bounded, an ε-SIR outcome α is $(\nu(\theta), \frac{1}{2}\kappa^\varepsilon)$-SIR, provided that δ is large. By the extended version of Theorem 2 described in this subsection, there exists a threshold of δ so that every $(\nu(\theta), \frac{1}{2}\kappa^\varepsilon)$-SIR outcome is justifiable, which shows the claim. *Q.E.D.*

The first part of Lemma 1 should not be expected to extend to uncertain repeated games when there are multiple stage games because, in general, sequential individual rationality and feasibility imply a non-additive bound on ex-post payoffs. To see this, note that, for all XPE σ and for every environment e, we have $U^\sigma(e) \in \mathbb{R}^n_+$ due to Observation 1 in the main text. At the same time, $U^\sigma(e) \in \mathscr{U}(\mu(e))$, where $\mu(e) \in \Delta\Theta$ defined by $\mu(\theta|e) := (1-\delta)\sum_{t=0}^\infty \delta^t \mathbf{1}_{\{\theta^t = \theta\}}$ represents the discounted frequency of stages in e. Combining these two facts, we obtain

$$U^\sigma(e) \geq \nu(\mu(e)) = \bigwedge \Big(\mathscr{U}(\mu(e)) \cap \mathbb{R}^n_+ \Big).$$

The bound $(\nu(\mu(e)))_{e \in \Theta^\infty}$ is highly permissive, as it is defined separately for each environment e and relies on feasibility and individual rationality in the "average" sense, captured by frequencies $\mu(e)$. In general, this bound cannot be expressed in the "additive" form and can be strictly lower than $(1-\delta)\sum_{t=0}^\infty \delta^t \nu(\theta^t)$. The next subsection presents two examples: one in which the bound is additive and SIR remains sufficient, and one in which it is genuinely is nonlinear, creating an asymptotic gap between SIR and justifiability.

1.2. *Example with SIR asymptotically sufficient for justifiability*

Consider the two stage games in Table I. In both games, the pure minmax payoff is $(0,0)$, the feasible payoff set is full-dimensional and contains strictly positive interior payoffs, yet $\nu(L) = (1,0)$ and $\nu(R) = (2,0)$. Direct calculations yield that

$$\nu(\mu(e)) = (1 + \mu(R|e), 0),$$

and that it can be achieved by playing (A_1, A_2)—a static Nash equilibrium in both stages—irrespective of a past history.

	A_2	B_2	C_2
A_1	$1^\star, 0^\star$	$0^\star, -1$	$0, -2$
B_1	$1^\star, 0$	$0^\star, -1$	$3^\star, 1^\star$

Game L

	A_2	B_2	C_2
A_1	$2^\star, 0^\star$	$0^\star, -2$	$0, -3$
B_1	$2^\star, 0$	$0^\star, -2$	$4^\star, 1^\star$

Game R

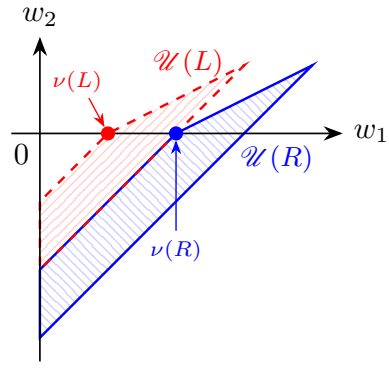


TABLE I

"$\star$" INDICATES THE STATIC BEST RESPONSE, $\mathscr{U}(L)$ IS THE RED DASHED POLYGON AND $\mathscr{U}(R)$ IS THE BLUE SOLID POLYGON, THE MARKED POINTS ARE $\nu(L) = (1, 0)$ AND $\nu(R) = (2, 0)$.

It is easy to see that 0-SIR is equivalent to justifiability irrespective of $\delta < 1$. Indeed, for any 0-SIR outcome α, complete it with the grim-trigger of (A_1, A_2). For player 2, grim reversion to (A_1, A_2) yields a punishment payoff of zero, so her no-deviation condition

$$\mathbb{E}_t[U_2^\alpha(\widetilde{h}^{t+1}|e)] \geq \frac{1-\delta}{\delta}\, d_2(\alpha(\widetilde{h}^t)|\theta^t)$$

is exactly her 0-SIR condition. As for player 1, since each stage game $\theta \in \Theta$ satisfies $d_1(a|\theta) \leq 2d_2(a|\theta)$ and $w_1 - \nu_1(\theta) \geq 2w_2$ for all $w \in \mathscr{U}(\theta)$, we obtain

$$\mathbb{E}_t[U_1^\alpha(\widetilde{h}^{t+1}|e)] - (1-\delta)\sum_{s=t+1}^{\infty} \delta^{s-(t+1)}\nu_1(\theta^s) \geq 2\,\mathbb{E}_t[U_2^\alpha(\widetilde{h}^{t+1}|e)]$$

$$\geq \frac{1-\delta}{\delta}\, 2d_2(\alpha(\widetilde{h}^t)|\theta^t) \geq \frac{1-\delta}{\delta}\, d_1(\alpha(\widetilde{h}^t)|\theta^t),$$

where the second inequality is player 2's 0-SIR condition.

1.3. *Example with an asymptotic gap between SIR and justifiability*

Consider the two stage games in Table II. In both games, the pure minmax payoff is $(0, 0)$, the feasible payoff set is full-dimensional and contains strictly positive interior payoffs, yet $\nu(L) = (1, 0)$ and $\nu(R) = (0, 1)$. Direct calculations reveal that

$$\nu(\mu) = \Big(\max\{0, 2\mu(L) - 1\}, \max\{0, 2\mu(R) - 1\}\Big),$$

which is strictly more permissive than $\mu(L)\nu(L) + \mu(R)\nu(R)$.

	x_2	y_2	z_2
x_1	$3^\star, 2^\star$	$0^\star, -1$	$2^\star, -1$
y_1	$1, 0^\star$	$-1, -2$	$2^\star, -1$

Game L

	x_2	y_2
x_1	$2^\star, 3^\star$	$0^\star, 1$
y_1	$-1, 0^\star$	$-2, -1$
z_1	$-1, 2^\star$	$-1, 2^\star$

Game R

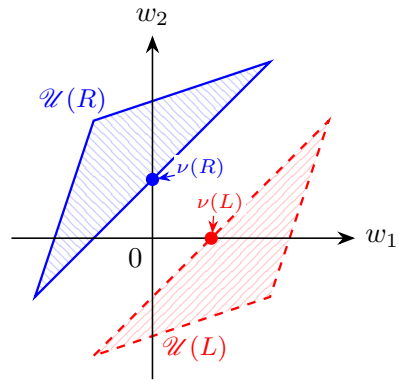


TABLE II

"$\star$" INDICATES THE STATIC BEST RESPONSE, $\mathscr{U}(L)$ IS THE RED DASHED POLYGON AND $\mathscr{U}(R)$ IS THE BLUE SOLID POLYGON, THE MARKED POINTS ARE $\nu(L) = (1, 0)$ AND $\nu(R) = (0, 1)$.

We show below that for all small enough $\varepsilon > 0$, there exists an outcome α^ε that is ε-SIR and yields

$$U^{\alpha^\varepsilon}(e) = \nu(\mu(e)) + (\varepsilon, \varepsilon) \tag{2}$$

provided that discounting is small. Yet, this outcome is not justifiable when δ is large. In fact, every justifiable outcome α must satisfy

$$U_i^\alpha(\theta^{0:k-1}, \theta, \theta, \ldots) \geq \delta^k \nu_i(\theta) \;\; \forall i \in N$$

for every $\theta^{0:k-1} \in \Theta^k$ and $\theta \in \Theta$. To see this, let player i best respond in each of the first k periods, collecting a nonnegative flow because the minmax payoff is zero. From period k onward, the continuation payoff of the completing XPE in the environment $(\theta, \theta, \ldots)$ is feasible and individually rational, hence at least $\nu_i(\theta)$. Now take $k = \lfloor \ln 2/(-\ln \delta) \rfloor$, so that $\delta^k \geq \frac{1}{2}$, which gives

$$U_1^\alpha(R^k, L^{k+1:\infty}) \geq \frac{\delta^k}{2} \geq \frac{1}{2}.$$

By contrast, since the discounted frequency of L in $(R^k, L^{k+1:\infty})$ equals δ^k,

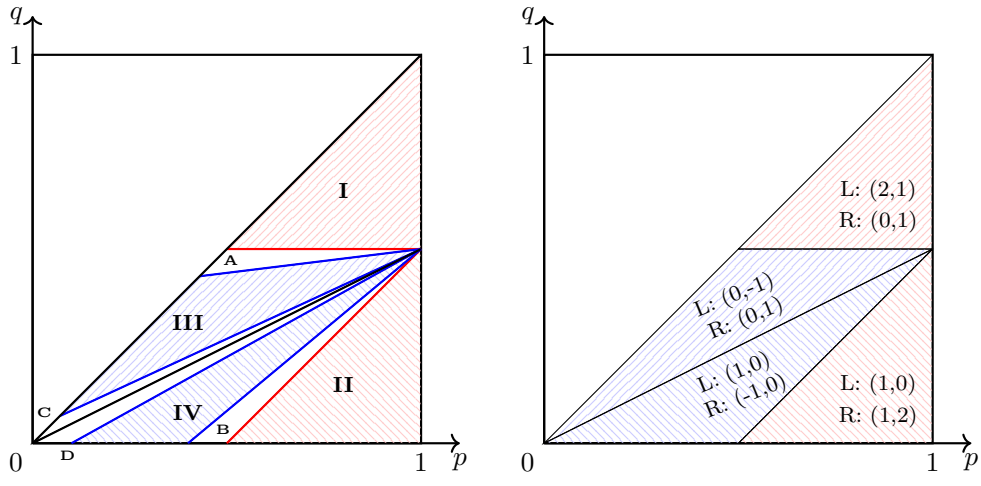
$$\nu_1(\mu(R^k, L^{k+1:\infty})) = \max\{0, 2\delta^k - 1\},$$

which converges to 0 as $\delta \to 1$.

Now, we construct an ε-SIR outcome that approximates $\{\nu(\mu(e))\}_{e \in \Theta^\infty}$.

Definition of a candidate outcome. The outcome α^ε depends on two state variables, $q \in [0, 1]$ and $p \in [0, 1]$, such that $q \leq p$. The former state variable q corresponds to the accumulated discounted probability of L along the given environment $e \in \Theta^\infty$, while p is the total accumulated discounted probability. Specifically, set $q^0 = p^0 = 0$ and define inductively

$$q^{t+1} = q^t + (1 - p^t)(1 - \delta)\iota^t, \; p^{t+1} = p^t + (1 - p^t)(1 - \delta),$$



(a) Partition for $\delta < 1$.

(b) Stage payoffs in the limit as $\delta \to 1$.

FIGURE 1.—Construction of the outcome α^ε.

where $\iota^t = 1$ if L occurs at time t and 0 otherwise, i.e., $\iota^t = \mathbf{1}_{\{\theta^t = L\}}$. As can be seen from these definitions, (q^t, p^t) are functions of the stage games realized before time t, that is, $\theta^0, ..., \theta^{t-1}$. We omit this dependence to simplify our notation.

Here is the description of α^ε. Under our proposed outcome, given the current stage game $\theta \in \Theta$ and the state $(q, p) \in [0, 1]^2$ with $q \leq p$, the players obtain a certain average payoff vector $x(q, p|\theta) + 2\varepsilon \in \mathscr{U}(\theta)$ when this period's sunspot is integrated out. To describe these payoffs as a function of θ and (q, p), we partition the state space $\{(q, p) \in [0, 1]^2 : q \leq p\}$ into 8 regions, termed I-IV and A-D.

The first set of regions (I-IV) covers all of the state space as $\delta \to 1$, and behavior in those regions is used to establish that the players can be pushed to their lower bound in Eq. (2) in the limit.

- In Region I, $\frac{1}{2} \leq q \leq p$: the players obtain $x(q, p|L) = (2, 1)$ and $x(q, p|R) = (0, 1)$.
- In Region II, $0 \leq q \leq p - \frac{1}{2}$: the players obtain $x(q, p|L) = (1, 0)$ and $x(q, p|R) = (1, 2)$.
- Region III corresponds to $\frac{p}{2} + \frac{(1-\delta)(1-p)}{2} \leq q \leq \frac{1}{2} - (1 - \delta)(1 - p)$ that simplifies to $\frac{p}{2} \leq q \leq \frac{1}{2}$ for $\delta \to 1$. The players obtain $x(q, p|L) = (0, -1)$ and $x(q, p|R) = (0, 1)$.
- Region IV corresponds to $p - \frac{1}{2} + (1 - \delta)(1 - p) \leq q \leq \frac{p}{2} - \frac{(1-\delta)(1-p)}{2}$ that simplifies to $p - \frac{1}{2} \leq q \leq \frac{p}{2}$ for $\delta \to 1$. The players obtain $x(q, p|L) = (1, 0)$ and $x(q, p|R) = (-1, 0)$.

The left panel of Figure 1 illustrates these regions for $\delta < 1$, while the right panel shows them in the limit as $\delta \to 1$.

We use the second set of regions (A-D) to show that the players can be pushed to their lower bound in Eq. (2) not only in the limit as $\delta \to 1$ but for fixed values of δ. For fixed $\delta < 1$, the state process (q^t, p^t) moves in increments depending on realized stage games and potentially crossing over region boundaries in a zig-zag manner. The sole role of Regions A-D, which are defined to fill a bit of space between I-IV, is to ensure that such environments in which (q^t, p^t) crosses boundaries of various regions multiple times do not lead to violations of Eq. (2).

- In Region A, $\frac{1}{2} - (1 - \delta)(1 - p) < q < \frac{1}{2}$: the players obtain $x(q, p|L) = r_A(0, -1) + (1 - r_A)(2, 1)$ and $x(q, p|R) = (0, 1)$, where $r_A = \frac{\frac{1}{2} - q}{(1-\delta)(1-p)} \in [0, 1]$.

- In Region B, $p-\frac{1}{2}<q<p-\frac{1}{2}+(1-\delta)(1-p)$: the players obtain $x(q,p|L)=(1,0)$ and $x(q,p|R)=r_B(-1,0)+(1-r_B)(1,2)$, where $r_B=\frac{q-(p-\frac{1}{2})}{(1-\delta)(1-p)}\in[0,1]$.
- In Region C, $\frac{p}{2}<q<\frac{p}{2}+\frac{(1-\delta)(1-p)}{2}$: the players obtain $x(q,p|L)=(0,-1)$ and $x(q,p|R)=r_C(0,1)+(1-r_C)(-1,0)$, where $r_C=\frac{q-\frac{p}{2}}{(1-\delta)(1-p)/2}\in[0,1]$.
- In Region D, $\frac{p}{2}-\frac{(1-\delta)(1-p)}{2}<q<\frac{p}{2}$: the players obtain $x(q,p|L)=r_D(1,0)+(1-r_D)(0,-1)$ and $x(q,p|R)=(-1,0)$, where $r_D=\frac{\frac{p}{2}-q}{(1-\delta)(1-p)/2}\in[0,1]$.

To recap, on the equilibrium path, $\iota^t=\mathbf{1}_{\{\theta^t=L\}}$, and $(q^t,p^t)_{t=0}^{\infty}$ evolves as described above. As a function of the history summarized by the current state (q^t,p^t) and this period's game ι^t, the players obtain the following stage payoff

$$\iota^t x(q^t,p^t|L)+(1-\iota^t)x(q^t,p^t|R)+2\varepsilon$$

after integrating out this period's sunspot. The definition of x's does not depend on a specific value of ε, and we can always find some correlated action profiles to implement these x's as long as ε is small enough, i.e., $x(q,p|\theta)+2\varepsilon\in\mathscr{U}(\theta)$ for all $\theta\in\Theta$ and for each state combination $(q,p)\in[0,1]^2$ with $q\leq p$. However, x's are functions of δ, and we write $x(q,p|\theta;\delta)$ when we need to stress this dependence.

Attaining the bound in Eq. (2). We now show that our candidate outcome α^{ε} can approximate $\{\nu(\mu(e))\}_{e\in\Theta^{\infty}}$ for all environments at the same time.

To begin, let $(\iota^t)_{t=0}^{\infty}$ be the sequence of 1's and 0's that describes the environment, where 1 means that $\theta^t=L$ and 0 stands for $\theta^t=R$, and set $\tau^t:=t(-\ln\delta)$ for each $t=0,1,....$ Consider the following continuous-time dynamics parameterized by $\tau\geq 0$. First, set $\chi(\tau)=\iota^t$ whenever $\tau\in[\tau^t,\tau^{t+1})$, and

$$Q'(\tau)=(1-P(\tau))\chi(\tau),\ \ P'(\tau)=1-P(\tau),$$
$$V'(\tau)=(1-P(\tau))[\chi(\tau)x(Q(\tau),P(\tau)|L;1)+(1-\chi(\tau))x(Q(\tau),P(\tau)|R;1)+2\varepsilon],$$

with $Q(0)=P(0)=V(0)=0$. The first two functions $(Q,P):\mathbb{R}_+\to[0,1]^2$ can be thought of as a continuous time approximation of the on-path state dynamics in the limit as $\delta\to 1$, while the function $V:\mathbb{R}_+\to\mathbb{R}^2$ corresponds to the accumulated payoffs by time τ. The discrete analogue of the function V for some fixed $\delta<1$ is given by the following sequence:

$$v^t=(1-\delta)\sum_{s=0}^{t-1}\delta^s[\iota^s x(q^s,p^s|L;\delta)+(1-\iota^s)x(q^s,p^s|R;\delta)+2\varepsilon].$$

Understanding $(v^t)_{t=0}^{\infty}$ is crucial for checking Eq. (2) and sequential individual rationality as the discounted (ex post) on-path payoff in period t can be expressed

$$\frac{v^{\infty}-v^t}{1-p^t}=(1-\delta)\sum_{s=t}^{\infty}\delta^{s-t}[\iota^s x(q^s,p^s|L;\delta)+(1-\iota^s)x(q^s,p^s|R;\delta)+2\varepsilon].$$

However, due to discreteness of this process, it is hard to study its behavior across various environments and also as $\delta\to 1$ directly. We use the continuous time dynamics to overcome this challenge. As we explain below, (Q,P,V) continuously interpolates the sequence $(q^t,p^t,v^t)_{t=0}^{\infty}$ to all of $\mathbb{R}_+$ in the sense that the values of (Q,P,V) on the grid of points $(\tau^t)_{t=0}^{\infty}$

coincide with $(q^t, p^t, v^t)_{t=0}^{\infty}$. Note that all dependence of (Q, P, V) on δ is built into the definition of the piecewise constant function χ, and so, it is much more convenient for exploring the limit of $\delta \to 1$ as opposed to its discrete counterpart.

Interpolation of state variables. By construction, we have $P(\tau) = 1 - e^{-\tau}$, thus

$$P(\tau^t) = 1 - e^{-\tau^t} = 1 - \delta^t = p^t \; \forall t.$$

As a result,

$$Q(\tau^{t+1}) - Q(\tau^t) = \int_{\tau^t}^{\tau^{t+1}} e^{-s} ds \iota^t = (p^{t+1} - p^t)\iota^t = q^{t+1} - q^t \; \forall t,$$

which implies $Q(\tau^t) = q^t$ for all t. Having established that (Q, P) continuously interpolate the sequence of state variables $(q^t, p^t)_{t=0}^{\infty}$, we now show that the same is true for V and $(v^t)_{t=0}^{\infty}$, respectively. To establish this assertion, we go over multiple cases one by one.

Interpolation of discounted payoffs in Regions I and II. Suppose (q^t, p^t) is in Region I. This region is absorbing, i.e., (q^{t+1}, p^{t+1}) is also in Region I irrespective of ι^t. Clearly,

$$x(Q(\tau), P(\tau)|\theta; 1) = x(q^t, p^t|\theta; 1) \; \forall \tau \in [\tau^t, \tau^{t+1}),$$

which gives

$$\begin{aligned} V_1(\tau^{t+1}) - V_1(\tau^t) &= \int_{\tau^t}^{\tau^{t+1}} e^{-s}[\iota^t x_1(Q(s), P(s)|L; 1) + (1 - \iota^t) x_1(Q(s), P(s)|R; 1) + 2\varepsilon] ds \\ &= \int_{\tau^t}^{\tau^{t+1}} e^{-s} ds [\iota^t x_1(q^t, p^t|L; \delta) + (1 - \iota^t) x_1(q^t, p^t|R; \delta) + 2\varepsilon] \\ &= (p^{t+1} - p^t)[\iota^t x_1(q^t, p^t|L; \delta) + (1 - \iota^t) x_1(q^t, p^t|R; \delta) + 2\varepsilon] \\ &= v_1^{t+1} - v_1^t. \end{aligned}$$

Exactly the same argument applies to Region II and player 2 due to symmetry.

Interpolation of discounted payoffs in Regions III and IV. Suppose (q^t, p^t) is in Region III. Even though it is not absorbing, the outcome is constructed in such a way that (q^{t+1}, p^{t+1}) is not in either of I and IV. Remark that

$$x(Q(\tau), P(\tau)|\theta; 1) = x(q^t, p^t|\theta; 1) \; \forall \tau \in [\tau^t, \tau^{t+1}),$$

which shows that our earlier argument goes as is. Of course, exactly the same reasoning applies to Region IV.

Interpolation of discounted payoffs in Regions A-D. Suppose (q^t, p^t) is in Region A. If this period's stage game is R, then both the continuous time dynamics and the discrete time dynamics assign the same stage game payoff of $(2\varepsilon, 1 + 2\varepsilon)$, i.e.,

$$x(Q(\tau), P(\tau)|R; 1) = x(q^t, p^t|R; 1) \; \forall \tau \in [\tau^t, \tau^{t+1}),$$

and we may use our earlier argument. If this period's stage game is L, then, by construction, (q^{t+1}, p^{t+1}) is in Region I. In this case,

$$V_1(\tau^{t+1}) - V_1(\tau^t) = 2\varepsilon \int_{\tau^t}^{\tau^{t+1}} e^{-s} \mathbf{1}_{\{Q(s) < \frac{1}{2}\}} ds + (2 + 2\varepsilon) \int_{\tau^t}^{\tau^{t+1}} e^{-s} \mathbf{1}_{\{Q(s) \geq \frac{1}{2}\}} ds$$

$$
\begin{aligned}
&= 2\varepsilon \int_{\tau^t}^{\tau^{t+1}} \mathbf{1}_{\{Q(s)<\frac{1}{2}\}} dQ(s) + (2+2\varepsilon) \int_{\tau^t}^{\tau^{t+1}} \mathbf{1}_{\{Q(s)\geq\frac{1}{2}\}} dQ(s) \\
&= 2\varepsilon \left(\frac{1}{2} - q^t \right) + (2+2\varepsilon) \left(q^{t+1} - \frac{1}{2} \right) \\
&= (p^{t+1} - p^t)[2(1-r_A) + 2\varepsilon] \\
&= v_1^{t+1} - v_1^t,
\end{aligned}
$$

where we used the definition of r_A and the fact that $q^{t+1} - q^t = p^{t+1} - p^t$ whenever $\iota^t = 1$. The argument for the remaining cases and player 2 is identical, thus omitted. Taking all pieces together, we conclude that V continuously interpolates $(v^t)_{t=0}^{\infty}$. One important implication is that the (ex post) payoff at the outset v^∞ equals the value of $V(\infty)$.

Solving for discounted payoffs using continuous time dynamics. Fix $\tau \leq \infty$. We now solve for the value of $V(\tau)$ in a closed form. For simplicity, we focus only on player 1's payoff, since player 2's payoff can be found symmetrically. In continuous time, the payoff function $x(Q, P|\theta; 1)$ is the same in the three Regions III, A, C, thus we shall refer to them simply as Region III. Similarly, this function is identical for Regions IV, B, D, and we shall refer to them as Region IV. We distinguish below between two cases that, as we explain, cover all possibilities.

Case 1. Suppose we continue in Region IV for some time, and then possibly transit to Region II. Let $0 \leq \tau^* \leq \tau$ be the first time at which we leave this region, where $\tau^* = \tau$ means that we stay there until the end and $\tau^* = 0$ means that we leave immediately. Since, if we enter Region II, then we shall stay there ad infinitum,

$$
\begin{aligned}
V_1(\tau) &= 2\varepsilon(1 - e^{-\tau}) + \int_0^{\tau^*} e^{-s}[2\chi(s) - 1]ds + \int_{\tau^*}^{\tau} e^{-s} ds \\
&= 2\varepsilon(1 - e^{-\tau}) + \int_0^{\tau^*} [2dQ(s) - dP(s)] + \int_{\tau^*}^{\tau} dP(s) \\
&= 2\varepsilon(1 - e^{-\tau}) + 2Q(\tau^*) - P(\tau^*) + P(\tau) - P(\tau^*).
\end{aligned}
$$

Case 2. Suppose we continue in Region IV for some time (potentially 0), and then possibly transit to Region III. Let τ^* be the first time at which we leave this region, where $\tau^* = \tau$ means that we stay there until the end and $\tau^* = 0$ means that we leave immediately. If we come back to Region IV at some later time τ^{**}, then, since player 1's incremental payoff equals zero on $[\tau^*, \tau^{**}]$, the system will be at exactly the same position as if we never left Region IV at time τ^* and instead moved along the diagonal till τ^{**}. So, it is without loss of generality to assume that we never come back to Region IV. Let $\tau^{**} \leq \tau$ be the first time at which we enter Region I, where $\tau^{**} = \tau$ means that we stay in Region III till the end. Then,

$$
\begin{aligned}
V_1(\tau) &= 2\varepsilon(1 - e^{-\tau}) + \int_0^{\tau^*} e^{-s}[2\chi(s) - 1]ds + \int_{\tau^{**}}^{\tau} e^{-s} 2\chi(s) ds \\
&= 2\varepsilon(1 - e^{-\tau}) + \int_0^{\tau^*} [2dQ(s) - dP(s)] + \int_{\tau^{**}}^{\tau} 2dQ(s) \\
&= 2\varepsilon(1 - e^{-\tau}) + 2Q(\tau^*) - P(\tau^*) + 2Q(\tau) - 2Q(\tau^{**}).
\end{aligned}
$$

Using the boundary conditions between the regions, we obtain that irrespective of the function χ, i.e., for any environment, we have

$$v_1^t = V_1(\tau^t) = 2\varepsilon(1-e^{-\tau^t}) + \begin{cases} 2q^t - 1 & (q^t,p^t) \text{ in Region I,} \\ p^t - 1 & (q^t,p^t) \text{ in Region II,} \\ 0 & (q^t,p^t) \text{ in Region III,} \\ 2q^t - p^t & (q^t,p^t) \text{ in Region IV.} \end{cases}$$

This gives the following closed-form expression for the on-path discounted payoffs:

$$\frac{v_1^\infty - v_1^t}{1-p^t} = 2\varepsilon + \begin{cases} \frac{2(q^\infty - q^t)}{1-p^t} & (q^t,p^t) \text{ in Region I,} \\ 1 & (q^t,p^t) \text{ in Region II,} \\ \frac{(2q^\infty - 1)^+}{1-p^t} & (q^t,p^t) \text{ in Region III,} \\ \frac{(2q^\infty - 1)^+ + (p^t - 2q^t)}{1-p^t} & (q^t,p^t) \text{ in Region IV,} \end{cases}$$

which shows that

$$\frac{v_1^\infty - v_1^t}{1-p^t} \geq 2\varepsilon \; \forall t.$$

Since $Q(\infty) = q^\infty$, which is the total discounted probability of game L, this result shows that

$$U_1^{\alpha^\varepsilon}(e) = \nu_1(\mu(e)) + 2\varepsilon \;\; \forall e \in \Theta^\infty.$$

The argument for the other player is identical.

Verification of ε-SIR. By the previous argument, at every on-path history $\widetilde{h}^t$ and in every compatible environment e, the continuation payoff satisfies

$$\mathbb{E}_t[U_i^{\alpha^\varepsilon}(\widetilde{h}^{t+1}|e)] \geq 2\varepsilon \;\; \forall i \in N.$$

Since the stage payoffs are bounded, for δ close enough to 1, we have

$$(1-\delta)d_i(\alpha^\varepsilon(\widetilde{h}^t)|\theta^t) \leq \delta\varepsilon \leq \delta\left(\mathbb{E}_t[U_i^{\alpha^\varepsilon}(\widetilde{h}^{t+1}|e)] - \varepsilon\right),$$

so α^ε is ε-SIR.

2. SYMMETRIC GAMES

2.1. *Feasible and IR payoffs in symmetric games*

LEMMA 2: *Consider a symmetric uncertain repeated game with $n \geq 2$ such that*

$$\textit{Int}(\mathscr{U}(\theta)) \cap \mathbb{R}^n_{++} \neq \emptyset \qquad \textit{for all } \theta \in \Theta.$$

Then, $\nu(\theta) = 0$ for all $\theta \in \Theta$.

PROOF: Fix a stage game θ. By the pure-minmax normalization, there exists $y \in \mathscr{U}(\theta)$ with $y_1 \leq 0$. The interiority hypothesis provides $z \in \text{Int}(\mathscr{U}(\theta)) \cap \mathbb{R}^n_{++}$. By convexity, the segment joining y and z contains a payoff $w \in \mathscr{U}(\theta)$ with $w_1 = 0$. The result is immediate if w is

nonnegative. Suppose instead that it is not. Let $\beta \in \Delta A(\theta)$ give w, i.e., $w = u(\beta|\theta)$. Denote the set of permutations of players by $\mathcal{I}$; then, by symmetry,

$$\sum_{a\in A}\beta(a)\frac{\sum_{\iota\in\mathcal{I}:\iota(1)=1}u_i(a_{\iota(i)},a_{\iota(-i)}|\theta)}{\sum_{\iota\in\mathcal{I}:\iota(1)=1}} = \begin{cases}0 & i=1,\\ \frac{1}{n-1}\sum_{j=2}^{n}w_j & i\neq 1.\end{cases} \tag{3}$$

It follows from Eq. (3) that $(0,x,\ldots,x)$ is an element of $\mathscr{U}(\theta)$ with $x := \frac{1}{n-1}\sum_{j=2}^{n}w_j$. Note that $x<0$ due to our assumption that no nonnegative point in $\mathscr{U}(\theta)\cap\{w\in\mathbb{R}^n|w_1=0\}$ exists.

The interiority hypothesis implies existence of a point $\widetilde{w}\in\mathscr{U}(\theta)\cap\mathbb{R}^n_{++}$. Repeating the same construction as in Eq. (3) but now averaging over all permutations, we obtain that $(\widetilde{x},\widetilde{x},\ldots,\widetilde{x})$ with $\widetilde{x} := \frac{1}{n}\sum_{i=1}^{n}\widetilde{w}_i$ is an element of $\mathscr{U}(\theta)$. By construction, $\widetilde{x}>0$, which implies that 0 is in $\mathscr{U}(\theta)$ as a convex combination of $(\widetilde{x},\widetilde{x},\ldots,\widetilde{x})$, $(0,x,\ldots,x)$, $(x,0,\ldots,x)$, ..., $(x,x,\ldots,0)$. Since $0\in\mathscr{U}(\theta)$ and pure minmax payoffs are normalized to zero, $\nu(\theta)=0$. *Q.E.D.*

2.2. *Equivalence of XPE with one long-run player and SSXPE*

Consider an uncertain repeated game in which player 1 is long-lived and discounts the future with $\delta<1$, while the other players are short-lived. Let σ be an XPE. Clearly, the short-lived players must be playing a static best response at each history h^t, meaning that $\sigma(h^t)\in A^*(\theta^t)$, which is defined as

$$A^*(\theta) := \{a\in A(\theta)|d_i(a|\theta)=0\ \forall i\neq 1\}\ \ \forall\theta\in\Theta.$$

Now, define a symmetric uncertain repeated game $\left\langle\Theta,(\widetilde{A}(\theta),\widetilde{u}(\cdot|\theta))_{\theta\in\Theta},\delta\right\rangle$ with two long-lived players as follows:

$$\widetilde{A}_i(\theta) := A^*(\theta),\ \ u_i(\widetilde{a}|\theta) := \widetilde{u}_1(\widetilde{a}_i|\theta)\mathbf{1}_{\{\widetilde{a}_1=\widetilde{a}_2\}} + r_1(\widetilde{a}_{-i}|\theta)\mathbf{1}_{\{\widetilde{a}_1\neq\widetilde{a}_2\}}\ \ \forall i\in\{1,2\}.$$

Since the long-lived player cannot profitably deviate in the original game, and the two players in the new game are effectively replicas of her, (σ,σ) is a SSXPE of this symmetric uncertain repeated game.

Conversely, let σ be a SSXPE of an uncertain repeated game in which every player is long-lived. Define an uncertain repeated game $\left\langle\Theta,(\widetilde{A}(\theta),\widetilde{u}(\cdot|\theta))_{\theta\in\Theta},\delta\right\rangle$ with long-lived player 1 and short-lived player 2 as follows: $\widetilde{A}(\theta) := A(\theta)$, and

$$\widetilde{u}_1(a|\theta) := u_1(a|\theta),\ \ \widetilde{u}_2(a|\theta) := -\mathbf{1}_{\{a_1\neq a_2\}}.$$

In this formulation, it is always in player 2's interest to match player 1's action. Since σ is an SSXPE of the original game, player 1 cannot profitably deviate provided that player 2 matches her action. Consequently, (σ_1,σ_2) is an XPE of the new game.

2.3. *SSXPE can be restrictive*

EXAMPLE 1: Consider a repeated game with three players, each of whom has two actions. As usual, player 1 chooses a row, player 2 selects a column, and player 3 chooses a table.

The reader can verify that the interiority assumption holds, meaning that the set $\mathscr{U}\cap\mathbb{R}^n_+$ forms the convex hull of $\{w|\exists i\in N \text{ s.t. } w_i\in\{\frac{1}{2},2\},w_j=0\ \forall j\neq i\}$. At either identical-action profile, every player receives $\frac{1}{2}$, whereas the smallest best-deviation payoff across the two profiles is 2. Thus,

$$\overline{m}(\infty)=\tfrac{1}{2}<2=\underline{m}(\infty),$$

a_3'	a_2'	a_2''
a_1'	$\frac{1}{2},\frac{1}{2},\frac{1}{2}$	$0^\star,2^\star,0^\star$
a_1''	$2^\star,0^\star,0^\star$	$-1,-1,\frac{5}{2}^\star$

a_3''	a_2'	a_2''
a_1'	$0^\star,0^\star,2^\star$	$\frac{5}{2}^\star,-1,-1$
a_1''	$-1,\frac{5}{2}^\star,-1$	$\frac{1}{2},\frac{1}{2},\frac{1}{2}$

TABLE III

REPEATED GAME FOR EXAMPLE 1. A STAR INDICATES A STATIC BEST RESPONSE.

and no SSXPE exists, regardless of the discount factor $\delta < 1$. However, the best feasible symmetric payoff vector, $(\frac{2}{3},\frac{2}{3},\frac{2}{3})$, can be sustained in a subgame-perfect equilibrium by randomizing in a history-independent manner across the three static Nash equilibria with equal probability. This is the largest symmetric feasible payoff because the sum of payoffs at every pure action profile is at most 2. Additionally, each player's minmax payoff can also be achieved in subgame-perfect equilibrium by playing the corresponding static Nash equilibrium.

3. SUBGAME PERFECTION UNDER DYNAMIC VARIATIONAL PREFERENCES

In this section, we elaborate on the connection between the criterion of ex post perfectness and subgame perfection when players have dynamic variational preferences as in Maccheroni, Marinacci, and Rustichini (2006a,b). A similar point in the context of outcome justifiability is made in Carroll (2026).

According to this model, player i's conditional preferences over strategy profiles after observing $\theta^{0:t}$ are described by the so-called ambiguity index $\pi \in \Delta\Theta^\infty \mapsto C_i(\pi|\theta^{0:t}) \in \mathbb{R}_+ \cup \{\infty\}$. The ambiguity index represents player i's cost associated with selecting a probability distribution over future stage games. It is required to satisfy three properties: it must be closed, lower-semicontinuous, and grounded, meaning that it is finite for at least some probability distributions.[1] Given this framework, player i evaluates strategy profiles based on the following criterion:

$$\inf_{\pi\in\Delta\Theta^\infty}(\mathbb{E}^\pi[U_i^\sigma(h^t|\theta^{0:t},e)] + C_i(\pi|\theta^{0:t})), \tag{4}$$

where the expectation is taken with respect to $e \sim \pi$.

Now, let σ be ex post perfect. We claim that player i plays a best response at each history, regardless of her dynamic ambiguity index. To see this, assume for contradiction that there exist an ambiguity index C_i, a history h^t, and an alternative strategy $\widetilde{\sigma}_i$ such that deviating from σ_i is profitable:

$$\inf_{\pi\in\Delta\Theta^\infty}\Big(\mathbb{E}^\pi[U_i^\sigma(h^t|\theta^{0:t},e)] + C_i(\pi|\theta^{0:t})\Big) < \inf_{\pi\in\Delta\Theta^\infty}\Big(\mathbb{E}^\pi\big[U_i^{(\widetilde{\sigma}_i,\sigma_{-i})}(h^t|\theta^{0:t},e)\big] + C_i(\pi|\theta^{0:t})\Big).$$

This implies that there exists some probability distribution $\widetilde{\pi}$ so that

$$\begin{aligned}\mathbb{E}^{\widetilde{\pi}}[U_i^\sigma(h^t|\theta^{0:t},e)] + C_i(\widetilde{\pi}|\theta^{0:t}) &< \inf_{\pi\in\Delta\Theta^\infty}\Big(\mathbb{E}^\pi[U_i^{(\widetilde{\sigma}_i,\sigma_{-i})}(h^t|\theta^{0:t},e)] + C_i(\pi|\theta^{0:t})\Big)\\ &\le \mathbb{E}^{\widetilde{\pi}}[U_i^{(\widetilde{\sigma}_i,\sigma_{-i})}(h^t|\theta^{0:t},e)] + C_i(\widetilde{\pi}|\theta^{0:t}),\end{aligned}$$

which contradicts ex post perfectness of σ.

[1]Here, Θ^∞ is endowed with the natural algebra generated by sets of the form $\{\theta^0\}\times\ldots\times\{\theta^t\}\times\Theta^\infty$, and $\Delta\Theta^\infty$ stands for the space of finitely additive probability measures equipped with the weak-* topology as a subset of the topological dual of the space of simple functions endowed with the sup norm.

Conversely, we claim that σ is ex post perfect if each player plays a best response at every history for all possible dynamic variational preferences they might have. To establish this, assume for contradiction that there exists a player i who has a profitable deviation in some environment e at some history h^t. Define player i's ambiguity index C_i as follows. For each $s = 0, 1, \ldots$, set $C_i(\pi|\widetilde{\theta}^{0:s})$ to zero when $\widetilde{\theta}^{0:s} = \theta^{0:s}$ and π assigns probability one to every event containing the continuation environment $\theta^{s+1:\infty}$; otherwise, $C_i(\pi|\widetilde{\theta}^{0:s})$ is set to B, where B is large. Since players' stage-game payoffs are bounded, at h^t, for any player i's deviation $\widetilde{\sigma}_i$, the infimum in Eq. (4) is attained by a distribution π that assigns probability one to $\theta^{t+1:\infty}$. Consequently, this infimum simplifies to $U_i^{(\widetilde{\sigma}_i, \sigma_{-i})}(h^t|e)$. As a result, since σ_i is not an ex post best response at h^t, player i can profitably deviate at this history.

REFERENCES

CARROLL, GABRIEL (2026): "Dynamic Incentives in Incompletely Specified Environments," *Econometrica*, 94 (2), 375–406. [11]

KRASIKOV, ILIA AND LAMBA, ROHIT (2026): "Uncertain Repeated Games," forthcoming, *Econometrica*. [1]

MACCHERONI, FABIO, MASSIMO MARINACCI AND ALDO RUSTICHINI (2006a): "Dynamic variational preferences," *Journal of Economic Theory*, 128 (1), 4–44. [11]

MACCHERONI, FABIO, MASSIMO MARINACCI AND ALDO RUSTICHINI (2006b): "Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica*, 74 (6), 1447–1498. [11]