Uncertain repeated games*

Ilia Krasikov Rohit Lamba

September 2022

Abstract

Multiple long run players play one amongst multiple possible stage games in each period. They observe and recall past play and are aware of the current stage game being played, but are maximally uncertain about the future evolution of stage games. This setup is termed an uncertain repeated game. The solution concept requires that a subgame perfect equilibrium be played no matter what sequence of stage games realize. The feasible set of payoffs is then so large and complex that it is not obvious how to frame standard results such as the folk theorem, and further how to construct credible rewards and punishments that work irrespective of the future evolution of games. The main goal of the paper is to build such a language and then to construct credible dynamic incentives that work generally for uncertain repeated games. The characterization of equilibrium outcomes is complete for large discount factors and strongly symmetric equilibria, and through an outer bound, shown to be tight for general discounting and asymmetric equilibria.

1 Introduction

The canonical repeated games model is a fulcrum of formal thinking on dynamic incentives. Its ubiquitous applicability in economics is apparent—industrial organization, macroeconomics, public finance, political economy and more. It has delivered the conceptual underpinnings of self-enforcing contracts, reputation building and understanding of institutions and cultural norms. A key, arguably restrictive, assumption in its formalization is that players repeatedly play the same game or that they have a common understanding of how and what games appear in the future. In this paper we attempt to expand the scope of this canon to repeated games with multiple long-run players, where multiple stage games are possible in each period, but the players are uncertain about which games will arrive in the future.

We consider the following model: A finite set of players interact repeatedly and play one amongst finitely many possible stage games in each period. At the start of every period they are informed of

---

*Krasikov: Higher School of Economics University Moscow, krasikov.is.main@gmail.com; Lamba: Pennsylvania State University, rlabma@psu.edu. We are indebted to Dilip Abreu, V. Bhaskar, Gabriel Carroll, Sylvain Chassang, Yuhta Ishii, Vijay Krishna, Elliot Lipnowski, Eilon Solan, Maxwell Stinchcombe, Bruno Strulovici, Takuo Sugaya, Caroline Thomas, Thomas Wiseman, Weijie Zhong, John Zhu, and seminar participants at University of Texas Austin, Stanford University, University of Kansas, Pennsylvania Economic Theory Conference and Stony Brook Conference on Game Theory for their comments and suggestions.

Electronic copy available at: https://ssrn.com/abstract=4222516
the game they are about to play; however they are maximally uncertain about the likelihood of how future stage games will be drawn. All past actions and realized games are observable. The players act optimally given the information they possess.\footnote{One substantial assumption is made, that the players' actions do no affect the evolution of stage games, and this is common knowledge. Extension to endogenous transitions is possible, but at a first pass, the attempt here is understand the complexity introduced by exogenous uncertainty.} The big picture question then is: What extent of conflict and cooperation is sustainable?

If the set of possible stage games is restricted to a singleton, the model collapses to the classical repeated games model (Friedman [1971] and Abreu [1988]). And, if we put a stochastic process on the evolution on the stage games, which is common knowledge amongst players who are Bayesian and time consistent in their beliefs, we are in the realm of the classical stochastic games setup (Shapley [1953] and Solan and Vieille [2015]). At the heart of repeated games (or dynamic and stochastic games more generally) is the observation that individuals can be made to take actions that may not be in their immediate best interest; this is done by promising a reward or punishment in the future, which is credible. So, conceptually the question we seek to ask is: What are such credible rewards and punishments when there may not be a common understanding of the future evolution of games?

There are at least two motivations in asking this question. First is to bring ideas of robustness of information from static games of incomplete information (Wilson [1987] and Bergemann and Morris [2005]) to the study of repeated (and stochastic) games. Relaxing the common knowledge of information on how stage games change over would produce predictions that are robust to the assumed stochastic process, and further on its common understanding amongst players or the outside analyst who is trying to rationalize behavior.

The second motivation is to dig deeper into the world of non-stationary dynamic games. Not much is known, in the form of general results, for arbitrary dynamic (and not necessarily repeated) games, even when these are deterministic. Considerable progress has been made by restoring stationary through Markov evolution of stage games, which is commonly understood by all players. To the best of our knowledge, the space outside the Bayesian realm is mostly unexplored.\footnote{In fact, even for the Bayesian realm, dynamic games can be hard to get a grip on, as Hörner, Sugaya, Takahashi, and Vieille [2011] point out in the opening lines of their paper: “Dynamic games are hard to solve.”}

A lot of focus in repeated games so far has been on testing the limits of cooperation when varying the monitoring structure—how much information on their opponent’s actions is accessible to the players (eg. Fudenberg and Yamamoto [2010] and Awaya and Krishna [2016]). We instead ask questions on the limits of cooperation and conflict while keeping monitoring in its simplest form but introducing uncertainty on the future state of play.\footnote{A literature starting at least with Rotemberg and Saloner [1986] provides theoretical underpinnings of price-wars during “boom” times. It is well documented that predictions of these models depend on the assumed information structure of how the underlying state evolves (see Kandori [1991]). A natural question then arises: to what extent can price-wars be sustained without making specific assumptions on a common understanding of how the state of the market evolves.}

In remarkable recent work, Carroll [2021] poses the precise question raised above, but for the case of one-long run player (in the sense of Fudenberg, Kreps, and Maskin [1990], wherein only one player
has a positive discount factor). This paper builds on Carroll’s analysis. While the basic motivations mentioned above are common to both papers, a fuller picture of the limits to cooperation and conflict introduced by uncertainty can only be rightfully addressed by studying repeated games with many long-run players. Most applications of the theory of repeated games, eg. study of collusion, price wars, risk sharing, sustainable policy plans, relational contracts and more, necessitate multiple long-run players. The multiple long-run player model poses non-trivial technical challenges, primary being that there is no one best or worst equilibrium payoff across all realizations of stage games, and hence the set of equilibrium payoffs has a complex geometry, which often renders it incalculable.

To describe our results, we first lay out the equilibrium concept: The strategies of players are defined in the usual way, as functions of past play to current choices. The analysis is restricted to pure strategies (mostly for simplicity, easily generalized to mixed), and players are allowed a public randomization device. Since we don’t impose beliefs on the evolution of information, a natural equilibrium concept is ex-post perfect equilibrium. It demands that for any realization of possible stage games, the strategy under consideration must be a subgame perfect equilibrium of that specific and standard dynamic game.

Ex-post equilibrium is a fairly demanding criterion, and so any cooperation (or conflict) attained under its guise is robust in its predictive power. It is robust to any mis-specifications that the players or the outside analyst may have about the information structure of the underlying stochastic game, and it naturally satisfies a non-regret condition for all players—no matter the future realization of stage games, no player will individually regret to not have deviated any any point.⁴

On to the results. Now, most of the classical results in repeated games, at least in the last three decades, work in the payoff space, because it allows the modeler to sidestep the curse of dimensionality in dealing with history dependent strategies. For example, the standard folk theorem goes something like this: the set of individually rational payoffs can be achieved in a subgame perfect equilibrium (under some conditions), as the players become arbitrarily patient.⁵ To visualize the result, the modeler would draw the set of equilibrium payoffs for a two-player game in a two dimensional space and show how the set expands in the discount factor, eventually assuming the set of all individually rational payoffs.

To state (or frame) results such for uncertain repeated games, we run into a fundamental difficulty. No such statement or visualization is possible for uncertain repeated games for even the set of feasible payoffs, is infinite dimensional and does not, generally, have an easily decipherable structure. So we go back to the classical approach of Abreu [1988] in describing equilibria through outcomes. An

⁴ As Carroll [2021] notes, ex-post perfect equilibrium is not necessarily steeped in a behavioral model of individual maximization. However, any plausible model of individual maximization that you could come up with, say Bayesian or ambiguity averse, dynamically consistent or inconsistent; its predictions will include the set of ex-post perfect equilibrium outcomes, but the converse may not be true. So this model of uncertain repeated games generates a robust ‘lower bound’ of equilibrium outcomes that can be rationalized by most standard dynamic strategic interaction frameworks. Importantly, this ‘lower bound’ is robust to the exact specification chosen by the modeler or the players.

⁵Individual rationality here means that the payoffs of each player are above their minmax, that is above the worst payoff the other players can keep this player at, given that she best responds to that worst strategy.
outcome is a sequence of actions taken by the players on-path as a function of past realized stage games. An outcome is justifiable if it can be selected on-path in an equilibrium.

A first lemma establishes a universal recursive condition for an outcome to be justifiable in terms of punishment outcomes that are initiated in case of a deviation, or equivalently the associated punishment (continuation) payoffs. For all current stage games, it states that the best possible deviation must be deterred by a robust notion of dynamic loss that can be executed irrespective of the realization of future stage games. This simple lemma is invoked repeatedly to characterize the set of equilibrium outcomes and hence payoffs.

An outcome is individually rational if the payoff associated with it is above the discounted sum of minmax payoffs at each node of the game tree. It is intuitive to see that individually rationality is a necessary condition for an outcome to be achieved in an equilibrium—punishments to deter deviations from an outcome that is not individually rational cannot be constructed, the lemma stated above will generate an empty set. The converse of this statement constitutes the folk theorem: under some conditions, all individually rational outcomes can be attained under some ex post perfect equilibrium above a large (but bounded) threshold discount factor.

The statement of the folk theorem here, when restricted even to the repetition of a single stage game, is stronger than the traditional folk theorem (eg. Fudenberg and Maskin [1986]). The standard result considers stationary outcomes, for instance cooperate in every period of a prisoner’s dilemma game. Payoffs corresponding to such stationary outcomes are sufficient to trace the entire equilibrium payoff set. So the traditional folk theorem doesn’t necessarily answer the question, can some arbitrary outcome (play cooperate for \( x \) periods than defect for \( y \) periods and so on) be an equilibrium for patient players?\(^6\)

Two types of conditions are required for the folk theorem. First, as is standard, there should be enough richness in the payoff structure (called interiority) to incentivize each player away from their static best responses. And, second, it must be possible to punish one player upon deviation, while providing some future incentives to the others to participate in this punishment, even if these are far in the future. Since players are maximally uncertain over future stages games, these conditions need to hold across environments. However, it must also be said that the conditions are almost negligible for two player games and when considering symmetric games; but they assume some bite for asymmetric games with three or more players.

To prove the folk theorem, or for that matter, to construct any equilibrium that goes beyond repetition of static Nash in each game, we use the aforementioned universal lemma. A fundamental difficulty here is how to provide incentives, viz. construct off-path punishments, that work for all realizations of future stage games. This idea is formalized in two steps: first, normalizing each stage game so that they are comparable in terms of payoffs, and then finding equilibria in the normalized game through a certain common gap recursion.

\(^6\)The standard folk theorems also typically state a limit discounting result, so individually rational payoffs can be attained in equilibrium as discount factor coverages to 1; whereas here, it is claimed that all individually rational outcomes are attained in some equilibria for large enough but bounded discount factor.
A gap is the difference between any two equilibrium payoffs offered to a player on and off path; it encodes all information required to provide dynamic incentives. Since the payoff space is infinite-dimensional, so is the ‘gap-space’. A common gap is a gap that remains constant for an individual player across environments. This brute forces a reduction in dimensionality. Analogous to Abreu, Pearce, and Stacchetti [1990], recursing over common gaps generates a largest fixed point, which characterizes the set of all equilibrium outcomes in which incentives are provided through the set of common gaps. Carroll [2021] introduces the technique to characterize equilibria for uncertain repeated games with one long-run player. For multiple long-run players, individual common gaps interact in a non-trivial way for what gap vectors can be jointly promised to multiple players are restricted by the structure of the available stage games.\footnote{In common gap recursion, the difference between the on-path and off-path continuation payoffs for each player in the normalized game is the same across future realizations of the stage game; though, the payoffs of other players to incentivize punishments to the candidate player can be a function the realized stage games. So, in general, this recursion will produce a larger set of equilibrium outcomes than if we imposed commonality directly on payoffs space across environments in the APS recursion.}

The normalization proves useful in adjusting the initial scale of payoffs, to make them comparable across stage games. The common gap recursion then does its thing on top of the normalization. So by varying the initial normalization, different sets of equilibrium common gaps can be produced, which gives flexibility in designing punishments. Several approaches to pick the initial normalization are discussed.

At a first pass, the reduction in dimensionality may seem extreme, but in a series of results, it is shown that dynamic incentives constructed through this method have little and often no loss in terms of achievability of equilibrium outcomes. For starters assume that the discount factor is large. Choosing the initial normalization to be the minmax payoff is each stage game, the approach is used to construct a payoff of ‘zero’ in equilibrium for patient players. Since this implies the minmax outcomes are attained in each stage game on-path in an equilibrium, it proves the folk theorem: Any individually rational outcome can be achieved on-path by using the minmax as the punishment outcome. So equilibrium incentives provided through the common gap recursion are sufficient to achieve all individual rational outcomes on-path. There may be other ways to achieve these outcomes, but the common gap approach is sufficient.

Second, for symmetric games, when studying strongly symmetric equilibria, preferences are perfectly aligned, since players play the same action on and off-path. This makes the space of continuation common gaps one-dimensional. As a consequence, there is a best and a worst (strongly symmetric) equilibrium payoff in future stage games. In fact, the best one is achieved through stationary outcomes, playing the best action profile corresponding to each stage game which sustains the largest continuation common gap. Analogous reasoning is used in Carroll [2021] to provide a complete characterization of equilibrium outcomes with one long-run player.

The reader may now (rightfully) ask, what about equilibrium outcomes for fixed (but small or intermediate) discounting in asymmetric games or simply asymmetric equilibria in otherwise sym-
metric games? The common gap recursion identifies a subset of equilibrium gaps and hence associated payoffs and outcomes (through the universal lemma) for any fixed discount factor. These sets of sustainable common gaps for uncertain repeated games with multiple long-run layers can be explicitly calculated using an iterative procedure. An algorithm to implement this procedure is provided, and the final result is visualized for two-players repeated Cournot and Bertrand models, where the state parameterizes different levels of market size and marginal cost for the firms.8

But, how ‘far’ are the subset of equilibria in covering the universe of all possibilities in uncertain repeated games with fixed discounting? In the final part of the paper, we construct an outer bound on the set of ex post perfect equilibrium payoffs. This is done by bounding from below the worst equilibrium payoff that can be delivered to the players in each environment. The argument follows a constructive approach using duality and support functions from convex analysis. The lower bound on the worst equilibrium payoffs then provides an outer bound for the set of equilibrium outcomes, because in general, more outcomes can be sustained with this stricter punishment than what is permissible on the equilibrium path. An algorithm is again presented that calculates the worst payoff for the outer bound. Theoretically this outer bound is tight for patient players and strongly symmetric equilibria.

The efficacy of providing incentives in this way is measured through the highest difference between the worst payoffs sustainable under the common gap recursion and the outer bound across all future realizations of stage games. The workings of the measure (and hence the outer bound) are then illustrated for the repeated Cournot and Bertrand model. For the Bertrand model, the measure is always zero, and hence characterization of equilibrium outcomes is exact for fixed \( \delta \). For the Cournot model, for low values of discounting, the bound is tight since static Nash is the only action profile that is sustainable both in the common gap recursion and the outer bound. As we increase the discount factor, some action profile may become enforceable in the outer bound but not in the common gap recursion, so the difference jumps but then drops back to zero again fairly quickly as discount factor increases a bit more for this action profile now becomes enforceable in the common gap recursion too. So, the measure is a line at or around zero with periodic jumps, which get progressively smaller. Eventually when the minmax is reached in both outer bound and common gap recursion, there are no more jumps as the folk theorem kicks in.

Now, there is an obvious normative appeal to providing dynamic incentives in a robust way through the common gap recursion. But, purely quantitatively speaking, why does the brute force method of reducing dimensionality perform so well in providing dynamic incentives in uncertain repeated games? To provide incentives across all realizations of future stage games, the construction of rewards and punishments cannot be too finely dependent on the exact realized sequence of stage games. The analysis here suggests that to take this logic to an extreme and mute this anyways limited flexibility proves fairly powerful, so the tractability of it trumps the minimal loss from its invocation.

Finally, for both the ‘lower bound’ of equilibrium outcomes constructed through the common

---

8The numerical implementation of the algorithm is based on variations of the APS recursion for standard repeated games (eg. Judd, Yeltekin, and Conklin [2003], Abreu and Sannikov [2014], and Abreu, Brooks, and Sannikov [2020])
gap recursion, and the outer bound described above, the explicit algorithms provided in the paper underscore the practical aspects of the construction. Since the set of equilibria in general can be complex and intractable, this machinery can give the applied economist an able tool to study dynamic strategic interactions (and incentive problems) with uncertainty.

This paper is a contribution to the theory of both dynamic decision problems and games, and robust approaches to games of incomplete information. While Carroll [2021] is the only paper we are aware of that rests within the rubric of uncertain repeated games, we also connect to other approaches in economics under the broad category of a robust approach to problems of dynamic incentives. For example, Fudenberg and Yamamoto [2010] studies repeated games (with a single stage game) where the monitoring structure is fixed but unknown and equilibrium notion is a version of ex-post perfect; it then establishes a folk theorem for this setup. Chassang [2013] studies dynamic contracts with limited liability constraints that are robust to underlying stochastic process that maps effort to output; it shows linear contracts do approximately well. Penta [2015] extends some of the ideas of robust mechanism design initiated by Bergemann and Morris [2005] from static to dynamic environments. de Oliveira and Lamba [2022] studies a model of dynamic choice where an analyst tries to rationalize a sequence of choices without having access to the dynamic information seen by the decision maker, and establishes a duality result to pin down these actions.

The main goal of the paper here is theoretical— set out to offer a general model of uncertain repeated games, provide a clear enough language to state standard results such as the folk theorem, and construct a reasonably exhaustive set of equilibrium outcomes through the common gap recursion. However, it must be stressed that a variety of applications seem imminent. For example, what kind of collusion (Green and Porter [1984]), price wars (Rotemberg and Saloner [1986]), sustainable policy plans (Chari and Kehoe [1990]), risk sharing (Kocherlakota [1996]) or relational contracts (Levin [2003]) are robustly sustainable, without making strong assumptions about how the future unfolds.

2 Setup

2.1 Primitives

Consider \( n \) players indexed by \( i \in \mathcal{N} := \{1, \ldots, n\} \) interacting in discrete time \( t = 1, \ldots \). The players discount the future with a common factor \( \delta \in [0,1) \).

Every period the players encounter one stage game chosen from a finite set \( \Theta \). A stage game \( \theta \in \Theta \) specifies a non-empty set of actions \( A_i(\theta) \), one for each player. An action profile \( a \in A(\theta) := A_1(\theta) \times \ldots \times A_n(\theta) \) generates a payoff of \( u_i(a|\theta) \in \mathbb{R} \) for player \( i \), written succinctly as the payoff vector \( u(a|\theta) \in \mathbb{R}^n \) for all players. For each \( i \in \mathcal{N} \) and \( a \in A(\theta) \), define the payoff from the best unilateral deviation as \( r_i(a|\theta) := \max_{\tilde{a}_i \in A_i(\theta)} u_i(\tilde{a}_i, a_{-i}|\theta) \), and let \( d_i(a|\theta) := r_i(a|\theta) - u_i(a|\theta) \) be the net gain from such deviation. Finally, for each \( i \in \mathcal{N} \), let \( r_i(\theta) := \min_{a \in A(\theta)} r_i(a|\theta) \) be player \( i \)'s (pure)
minmax payoff. We will often use the notation ∩ for the coordinate-wise infimum of a bounded Euclidean set, so we can write the minmax payoff vector in stage game $\theta$ by $r(\theta) := \cap_{a \in A(\theta)} r(a|\theta)$.

A sequence of realized or potentially realizable stage games, viz $e := \theta^{1:}\in \Theta^{\infty}$, is referred to as an environment. If $\Theta$ is a singleton, we are in a canonical repeated games setup, and we will often refer to the environment $e = (\theta, \theta, \ldots, \theta, \ldots)$ in such a situation as being stationary. To simplify notations, we will omit $\theta$ in $u, r, d$ and $r$ when $\Theta$ contains a single element.

Now, in each period, the players know the game about to be played and can recollect the history of all past stage games and actions. This fact is common knowledge. However, the players are maximally uncertainty about how future evolves — each player, in any period, finds every infinite sequence of future stage games possible, but there is no commonly held probability measure on how these stage games are drawn. Put differently, unlike the standard stochastic games setting, the players are unable or unwilling to attach any distribution over the evolution of future stage games.

2.2 Strategies, discounted payoffs and equilibrium

In this paper we only consider pure strategies but allow for a public correlation device, which is modelled through sunspots. Since all past actions are observable, this is a game of perfect monitoring.

**Definition 1.** A (pure) strategy $\sigma$; for player $i \in N$ prescribes for each date $t = 1, \ldots$, a pure action $\sigma_t^i(a^{1:t-1}, \omega^{1:t}|\theta^{1:t}) \in A_i(\theta_t)$ as a measurable function of

- a sequence of past actions $a^{1:t-1} := (a^1, \ldots, a^{t-1}) \in A(\theta^{1:t-1}) := A(\theta^1) \times \ldots \times A(\theta^{t-1})$,
- a sequence of sunspots $\omega^{1:t} := (\omega^1, \ldots, \omega^t) \in [0, 1]^t$,
- a sequence of stage games $\theta^{1:t} := (\theta^1, \ldots, \theta^t) \in \Theta^t$.

The sunspots are taken to be i.i.d. uniform random variables. Definition 1 writes down the usual notion of a strategy in stochastic games (Mertens, Sorin, and Zamir [2015]).

As is standard, for each node in the game tree, a strategy profile determines the stream of payoffs starting from that point onwards. In our setting, since the players are maximally uncertain about the stochastic process driving stage games, such streams of payoffs must be recorded for every possible infinite sequence of stage games. Given an environment $e = \theta^{1:}\in \Theta^{\infty}$, the discounted payoff under strategy profile $\sigma := (\sigma_1, \ldots, \sigma_N)$ is defined inductively as

$$
U^{t,\sigma}(a^{1:t-1}, \omega^{1:t}|e) = (1 - \delta)u(\sigma^t(a^{1:t-1}, \omega^{1:t}|\theta^{1:t})|\theta^t) + \\
+ \delta \mathbb{E}_t \left[ U^{t+1,\sigma}(a^{1:t-1}, \sigma^t(a^{1:t-1}, \omega^{1:t}|\theta^{1:t}), \omega^{1:t+1}|e) \right],
$$

9Recall that minmax payoff in games is the minimal payoff a player can guarantee herself when the other players are trying to minimize it.

10Analogously, $\cup$ will denote the coordinate-wise supremum of a bounded Euclidean set.

11The analysis of mixed strategies is almost identical when randomization chosen by players is observable.
where the expectation $\mathbb{E}_t [\cdot]$ is taken over the $(t + 1)$-th sunspot.

Denote the discounted payoff (at the outset) under $\sigma$ by $U^{\sigma}_t (e) := \mathbb{E} \left[ U^{\sigma}_t (\omega^1 | e) \right]$. Here $U^{\sigma}_t (e) \in \mathbb{R}^n$ is a finite dimensional vector, and $U^{\sigma} \in \mathbb{R}^{n \times \Theta^*}$, when viewed as a function of an environment $e \in \Theta^*$, is an infinite dimensional vector. In what follows we will often refer to elements of

$$F(\delta) := \left\{ w \in \mathbb{R}^{n \times \Theta^*} \mid \exists \sigma \text{ s.t. } w(e) = U^{\sigma}_t (e) \quad \forall e \in \Theta^* \right\}$$

as feasible discounted payoffs. In contrast to most of the literature, the set $F(\delta)$ is infinite-dimensional; moreover, it depends on $\delta$, because arbitrary non-stationary environments are allowed. We refer the reader to Example 1 for an illustration of this set.

For each fixed environment, the players face a dynamic game with perfect information. A standard equilibrium notion for such games is subgame perfect equilibrium. Formally, a strategy profile $\sigma$ is said to be a subgame perfect equilibrium (SPE) for environment $e \in \Theta^*$ if

$$U^{\sigma}_t (a^{1:t-1}, \omega^{1:t} | e) \geq U^{f, (\sigma_i, \sigma_{-i})}_t (a^{1:t-1}, \omega^{1:t} | e) \quad \forall \sigma_i, \forall i \in N.$$

In our setting of uncertain repeated games, there are multiple environments and the players consider each of them possible. A natural notion that respects subgame perfectness for each individual environment is ex-post perfect equilibrium.

**Definition 2.** A strategy profile $\sigma$ is said to be an ex-post perfect equilibrium (XPE) if it is a subgame perfect equilibrium (SPE) for every environment $e \in \Theta^*$.

Ex-post perfect equilibrium applies to situations in which the players agree that their actions cannot influence transitions between stage games but they are uncertain about the exact stochastic process that drives these transitions. In such contexts, XPE is a natural but demanding concept. It is based on the idea of no regret — for any possible realization of stage games, no player regrets having not deviated. It is also robust to heterogeneous beliefs and potential time inconsistencies in processing information by the players.

Our definition of XPE follows Carroll [2021], generalizing its functionality from uncertain repeated games with one long-run player to multiple long-run players. The concept is related to a similar equilibrium idea invoked in Fudenberg and Yamamoto [2010] that studies a different framework where the monitoring structure is uncertain. Spiritually, the concept is also related to the belief-free equilibrium which is used to gain some tractability and robustness to information in repeated games with private monitoring (Ely, Hörner, and Olszewski [2005]).

---

12 If only one stage game is available, then this set is independent of $\delta$, thus we simply write $F$. In fact, it is easy to see that $F = \text{co} \{ u(a) | a \in A \}$.

13 It is easy to see that when evaluating the incentive constraint for a particular strategy profile, since the one-shot deviation principle holds for SPE in any fixed environment, it also holds for XPE in the general uncertain repeated game.

14 As mentioned in the introduction (footnote 4), the idea of ex-post perfect equilibrium is not necessarily steeped in...
2.3 Describing XPE through discounted payoffs

If there is only one stage game, what we call the stationary environment, our model is equivalent to the standard repeated games setting in which monitoring is perfect. It is well-known, at least since Abreu, Pearce, and Stacchetti [1990], that any strategy profile in this case can be equivalently expressed as a certain automaton over (finite-dimensional) feasible discounted payoffs, i.e., elements of $F \subseteq \mathbb{R}^n$. Furthermore, the whole set of discounted XPE payoffs $E(\delta)$ can be identified as the largest fixed point of the so-called APS recursion defined over bounded subsets of $\mathbb{R}^n$.\footnote{The APS recursion has been successfully implemented by multiple numerical algorithms that approximate $E(\delta)$ for fixed $\delta < 1$; see, for example, Judd, Yeltekin, and Conklin [2003], Abreu and Sannikov [2014], and Abreu, Brooks, and Sannikov [2020]. Moreover, Fudenberg and Maskin [1986] and Fudenberg and Levine [1994] obtain an explicit characterization of the limit set, $\lim_{\delta \to 1} E(\delta)$, under a certain interiority assumption.}

The APS recursion has been successfully implemented by multiple numerical algorithms that approximate $E(\delta)$ for fixed $\delta < 1$; see, for example, Judd, Yeltekin, and Conklin [2003], Abreu and Sannikov [2014], and Abreu, Brooks, and Sannikov [2020]. Moreover, Fudenberg and Maskin [1986] and Fudenberg and Levine [1994] obtain an explicit characterization of the limit set, $\lim_{\delta \to 1} E(\delta)$, under a certain interiority assumption.

In the uncertain games setting with more than one stage game the set of feasible discounted payoffs $F(\delta)$ is necessarily infinite-dimensional, thus the set of discounted XPE payoffs, defined as

$$E(\delta) := \{ w \in \mathbb{R}^{n}\Theta^\omega \mid \exists \text{ XPE } \sigma \text{ s.t. } w(e) = U^\sigma(e) \ \forall e \in \Theta^\omega \},$$

is infinite-dimensional as well. Despite largeness of these objects, it is still true that any strategy profile can be written recursively as an automaton over elements of $F(\delta)$, whereas $E(\delta)$ admits the standard APS recursion. Formally, we have the following characterization, which is based on Abreu, Pearce, and Stacchetti [1990].

**Lemma 1.** $E(\delta)$ is the largest (in the sense of set-inclusion) subset of $F(\delta)$ that satisfies the following:

$$E(\delta) = \prod_{\theta \in \Theta} B_{\theta} E(\delta),$$

where for each $\theta \in \Theta$, the operator $B_{\theta}$ is defined over subsets of $\mathbb{R}^{n}\Theta^\omega$ as

$$B_{\theta} W := \operatorname{co} \left\{ w \in \mathbb{R}^{n}\Theta^\omega \mid \exists a \in A(\theta), \ \exists (v^i)_{i=0}^n \subset W \text{ s.t. } w = (1 - \delta) u(a|\theta) + \delta v^0, \right\},$$

where

$$\frac{1 - \delta}{\delta} d_i(a|\theta) \leq \inf_{e \in \Theta^\omega} \{ \psi_i^0(e) - \psi^i(e) \} \quad \forall i \in N.$$

As usual, the APS recursion (Eq. (2)), takes as an input some subset of discounted payoffs $W$, elements of which can be used as continuation values. It then returns the set of discounted payoffs attainable today at state $\theta$, that is $w$’s that respect the accounting identity with a certain on-path continuation $v^0 \in W$, and can be enforced by some punishment continuations $(v^i)_{i \in N} \subset W$, one for a (direct) model of individual maximization. However, any reasonable model you could come up with—Bayesian or ambiguity averse, dynamically consistent or inconsistent—will produce equilibria that encapsulate ex-post perfect equilibria. In that sense, this notion provides a robust ‘lower bound’ in predicting behavior in uncertain repeated games.

\footnote{See Mailath and Samuelson [2006] for a textbook introduction.}
each player.

In contrast to the case of one stage game, the APS recursion cannot be solved for explicitly in the simplest of examples, nor, more generally, numerically approximated for a fixed value of $\delta < 1$. There are at least two reasons: i) infinite-dimensionality and ii) lack of stationarity. First, the set $E(\delta) \subset \mathbb{R}^{n \cdot \Theta^t}$ has a complex geometry, because each period action profile cannot depend on a sequence of future stage games but needs to deter deviations irrespective of what stage games will arrive in the future, thus $\inf_{e \in \Theta^t}$ in Eq. (2). Second, even if the players knew ahead of time a specific environment that they are facing, i.e., some fixed $e \in \Theta^t$, the APS recursion is still not very useful. To the best of our knowledge, a computation of the set of discounted SPE payoffs for an arbitrary (non-stationary) sequence of stage games, is still an open problem.

As for the limit set of $E(\delta)$ as $\delta \nearrow 1$, the argument of Fudenberg and Maskin [1986] might fail to deliver a folk theorem type of result even when stage games are almost identical as happens in Example 4. The construction of Fudenberg and Levine [1994] depends critically on stationarity and finite-dimensionality, and we aren’t aware of how it can be applied to uncertain repeated games.

To sum up, the sets of XPE and associated discounted XPE payoffs are complex objects. Even for large values of $\delta$, to the best of our understanding, standard techniques will not deliver a folk theorem type of result. Moreover, due to infinite-dimensionality of the feasible set of payoffs, it isn’t immediately clear how such a result should be stated.

### 2.4 Outcome approach to XPE

In this paper we aim to provide a constructive characterization of a subset of XPE for arbitrary values of $\delta$, and a complete characterization of $E(\delta)$ (under certain conditions) for large values of $\delta$, which is reminiscent of the standard folk theorem. To this end, classical approach of representing strategy profiles through discounted payoffs is combined with an even classical-er idea of representing them through outcomes that record only on-path behavior.

**Definition 3.** An outcome $\alpha$ is a sequence of measurable functions $(\omega^1, \theta^1) \mapsto \alpha^t(\omega^1 | \theta^1) \in A(\theta^t)$ for $t = 1, \ldots.$

A specific class of outcomes that repeat an identical pattern will be particularly useful. We say an outcome $\alpha$ is stationary if there is a measurable function $(\omega, \theta) \mapsto \beta(\omega | \theta)$ such that for each $t = 1, \ldots$, $\alpha^t(\omega^1 | \theta^1) = \beta(\omega^t | \theta^t)$ for all $(\omega^1, \theta^1) \in [0, 1]^t \times \Theta^t$. So, a stationary outcome is time/history independent, it specifies the action to be taken as a function of the current stage game and the latest

---

16Unlike Fudenberg and Maskin [1986], that works with $T$-period punishment blocks, the proof strategy of Fudenberg and Levine [1994] is based on a geometric approach. In the context of repeated games, i.e., when $\Theta$ is a singleton, they show that every (smooth) convex subset $W$ in the interior of the set of feasible discounted payoffs larger than $r$ necessarily satisfies $W \subseteq \partial W$ for large values of $\delta$. In particular, each discounted payoff on the boundary of $W$ can be decomposed on the hyperplane tangent to $W$ at this point. This argument relies both on finite-dimensionality and stationary of discounted payoffs.
Outcomes are simply an alternative way to think about strategy profiles. A strategy profile induces an outcome that describes on-path behavior; as such, all information that is relevant for computing a discounted payoff under this strategy profile is contained in the associated outcome. For each environment \( e \in \Theta^\infty \), the discounted payoff under outcome \( \alpha \) is defined analogously to Eq. (1) as

\[
U_{t}^{1, \alpha}(\omega_{1:t}|e) = (1 - \delta)u(\alpha^t(\omega_{1:t}|\theta_{1:t})|\theta^t) + \delta \mathbb{E}_t \left[ U_{t+1}^{1, \alpha}(\omega_{1:t+1}|e) \right],
\]

where \( \mathbb{E}_t [\cdot] \) represents the expectation over the \((t+1)\)-th period sunspot. Set \( U_{\alpha}(e) \):=

Conversely, any strategy profile can equivalently be represented in the language of outcomes as follows:

- play according to an outcome \( \alpha \) as long as there hasn’t been any deviation in the past;
- if player \( i \) deviates at a history \((a_{1:t-1}, \omega_{1:t}, \theta_{1:t})\) to \( a_i \neq \alpha_i(\omega_{1:t}|\theta_{1:t})\), then switch to some outcome \( \tilde{\alpha}_i \), which may depend on both the history and the action taken by the deviator.

The idea of using outcomes in repeated games goes back to Abreu [1988]. Recently, Panov [2022] further developed this idea to study equilibria in continuous time to bypass the problem of defining strategies in continuous time. Carroll [2021] uses outcomes to describe XPE in uncertain repeated games with one long-run player. We push these ideas in the direction of a general description of equilibrium in terms of outcomes for uncertain repeated games.

Of course, not every outcome can arise on-path in some XPE.

**Definition 4.** An outcome \( \alpha \) is said to be justifiable if there exists an XPE \( \sigma \) such that \( \forall (\omega_{1:t}, \theta_{1:t}) \in [0,1]^t \times \Theta^t, \forall t = 1, \ldots, \)

\[
\alpha^t(\omega_{1:t}|\theta_{1:t}) = \sigma^t(\alpha^{t-1}(\omega_{1:t-1}|\theta_{1:t-1}), \omega_{1:t}|\theta_{1:t}).
\]

The advantage of describing XPE through justifiable outcomes is twofold. First, justifiable outcomes have a natural recursive structure, thanks to the one-shot deviation principle.

**Lemma 2.** An outcome \( \alpha \) is justifiable if and only if \( \forall (\omega_{1:t}, \theta_{1:t}) \in [0,1]^t \times \Theta^t, \forall t = 1, \ldots, \forall i \in N \) there exists a justifiable outcome \( \tilde{\alpha}_i \) such that

\[
(1 - \delta)d_i(\alpha^t(\omega_{1:t}|\theta_{1:t})|\theta^t) \leq \delta \inf_{e \in \Theta^t} \left\{ \mathbb{E}_t \left[ U_{t+1}^{i, \alpha}(\omega_{1:t+1}|\theta_{1:t+1}, e) \right] - U_{\tilde{\alpha}_i}(e) \right\}.
\]

Lemma 2 describes a recursive structure for the set of justifiable outcomes. An outcome \( \alpha \) is justifiable if and only if at every node in the game tree, each player \( i \)'s static gain from her best deviation can be overturned by a dynamic loss in an alternative punishment outcome \( \tilde{\alpha}_i \), that is also justifiable,
irrespective of future stage games. The result, thus, presents a universal condition that must be satisfied for any outcome to arise on-path in an equilibrium.

As can be seen from the condition in Eq. (4), the added complexity in uncertain repeated games with several stage games emanates from the necessity to provide incentives for all environments at the same time, hence \( \inf_{e \in \Theta^\omega} \). Such complexity is also absent in the standard stochastic games setting in which \( \inf_{e \in \Theta^\omega} \) is replaced by an expectation operator with a given measure.\(^{17}\)

Clearly, the set of XPE payoffs \( E(\delta) \) can be recovered from the set of justifiable outcomes through Eq. (3). As a result, Lemma 2 can be used to partially characterize the set \( E(\delta) \) if we can find some justifiable punishment outcomes \( (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n) \), one for each player. Alternatively, if we understand the discounted payoffs associated with the punishment outcomes, then too, substituting for \( U_e^{\tilde{\alpha}_i} \) in the right hand side of Eq. (4), we can test if the candidate outcome is justifiable, and hence, whether its associated payoff lies in \( E(\delta) \).

We end this section with some observations about justifiable outcomes that follow from Lemma 2. First, it is easy to see that if a static Nash equilibrium in pure strategies is available in each stage game, then an outcome in which a static Nash is played in every period is justifiable. To see it, specify the player \( i \)'s punishment outcome in Lemma 2 as follows. For each stage game \( \theta \), \( \tilde{\alpha}^i \) selects this player’s worst static Nash equilibrium in game \( \theta \) irrespective of a history. This sustains the original static Nash in every period as an XPE and hence makes the outcome corresponding to it justifiable.

The second observation is a necessary condition for justifiability. Recollect that \( r(\theta) = \bigwedge_{a \in A(\theta)} r(a|\theta) \) is the static minmax payoff vector. For each environment \( e \in \Theta^\omega \), define
\[
U(e) := (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r(\theta^t)
\]
(5) as the discounted minmax payoff.

**Definition 5.** An outcome \( \alpha \) is individually rational if \( \forall (\omega^{1:t}, \theta^{1:t}) \in [0, 1]^t \times \Theta^t, \forall t = 1, ..., \forall i \in N, \)
\[
(1 - \delta)d_i (\alpha^t(\omega^{1:t}|\theta^{1:t})|\theta^t) \leq \delta \inf_{e \in \Theta^\omega} \left\{ \mathbb{E}_e \left[ U_i^{t+1, \alpha} (\omega^{1:t+1}|\theta^{1:t}, e) \right] - U_i(e) \right\}.
\]
(6)

So, an outcome is individually rational if for each player, the discounted payoff under it at each date isn’t lower than what this player can obtain from the best one-shot deviation provided that the other players will minmax her in all future periods.

Individual rationality is necessary for an outcome to be justifiable. To see it, suppose that \( \alpha \) is justifiable. Then player \( i \)'s static payoff from the best deviation at time \( t \), given that no other player deviated in the past, is \( (1 - \delta)r(\alpha_i(\omega^{1:t}|\theta^{1:t})|\theta^t) \). Since this player can keep deviating ad infinitum,
she is guaranteed to receive at least her discounted minmax payoff from \( t + 1 \) onward, i.e., \( \delta \cdot (1 - \delta) \sum_{k=1}^{\infty} \delta^{k-1} r_i(\theta_i^k) \). Adding up these two terms and subtracting the player \( i \)'s payoff under \( \alpha \) we obtain Eq. (6).

**Lemma 3.** If an outcome \( \alpha \) is justifiable then it is individually rational.

The third observation is that the necessary condition in Lemma 3 combined with the first observation immediately gives a complete characterization of the set of justifiable outcomes for given \( \delta < 1 \) when for each stage game, every player’s minmax payoff is attained in some Nash equilibrium. We illustrate such situation in Example 3. Of course, often minmax payoffs cannot be attained at a Nash equilibrium, as, for instance, happens in Example 4.

### 3 Examples

Before digging into the main results, we provide five quick examples. These, we hope, communicate basic ideas of the setup and equilibrium notion, and provide a more easily digestible flavor of the results that will follow. The primary message of each example is stated at the beginning of its exposition.

**Example 1.** In this example we illustrate difficulties of computing the set of feasible discounted payoffs \( F(\delta) \), that is the range of \( \alpha \mapsto U^\alpha \). There is one player and two stage games, namely \( L \) and \( R \). Game \( L \) has two actions that yield \( \pm 1 \), whereas game \( R \) has only one action profile that gives 0.

\[
U^\alpha(e) = (1 - \delta)\mathbb{E}[ u(a^1(\omega^1|L)|L) ] ,
U^\alpha(e') = U^\alpha(e) + \delta(1 - \delta)\mathbb{E}[ u(a^2(\omega^{1:2}|L^{1:2})|L^{1:2}) ] ,
U^\alpha(e'') = U^\alpha(e') + \delta^2(1 - \delta)\mathbb{E}[ u(a^3(\omega^{1:3}|L^{1:3})|L^{1:3}) ] .
\]

Finding the set of feasible discounted payoffs is hard even for the simplest situation described here. Recollect \( F(\delta) \subset \mathbb{R}^{n \cdot \Theta^\omega} \), and note that the coordinates of this set are connected through an adaptedness condition—each period’s action cannot depend on future stage games. As a result, this
set isn’t a hypercube and it depends on $\delta$ in a non-trivial way. Figure ?? depicts the projection of $F(\delta)$ onto three different environments $e = (L^1, R^{2\infty})$, $e' = (L^{1:2}, R^{3\infty})$ and $e'' = (L^{1:3}, R^{4\infty})$ for $\delta = 1/2$. This projection can be found by varying $\alpha$ in the following relationship:  
\[ 18 \]

As we add more such environments, the set of payoffs keeps getting more complicated, both because of dimensionality, and because the adaptedness condition must be respected. And remember, in general, we would add more players, and so on.

**Example 2.** Is it always possible to sustain some cooperation in an uncertain repeated game, especially when it is possible to do so in a repetition of one of its individual stage games? The answer is quite obviously no and the objective of this example is to show this in a simple and stark way.

Game $L$ is a standard prisoner’s dilemma where the cooperative outcome can be sustained for $\delta \geq 1/2$. Game $R$ only has one action available for both players which yields them a payoff of 0 each.

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>3,3,</td>
<td>0.5*</td>
</tr>
<tr>
<td>$d$</td>
<td>5*, 0</td>
<td>1*, 1*</td>
</tr>
</tbody>
</table>

Table 1: Uncertain repeated game for Example 2, “\*” indicates the static best-response.

Suppose either game can be played in any period. It is pretty immediate that for any $\delta < 1$, the unique justifiable outcome here is to play the static Nash equilibrium in every period. Since game $R$ has only one action profile, any two outcomes give the same payoff in the stationary environment $R^{1\infty}$. In terms of Eq. (4), the static gain from a deviation is non-negative but the dynamic loss from being punished is zero for both players, because the infimum is achieved for the stationary environment $R^{1\infty}$. Thus, cooperation cannot be sustained for any value of $\delta$.

**Example 3.** Example 2 leaves us with a lingering thought: can any cooperation and/or conflict be sustained in XPE when multiple stage games are available to players? This example illustrates that it is indeed possible to provide dynamic incentives even under maximal uncertainty but the threshold value of $\delta$ required is typically larger than that for the repetition of individual stage games.

There are two players and two possible stage games, again $L$ and $R$. In each game player 1 has two available actions, $a$ and $b$, and player 2 has three actions denoted $x$, $y$ and $z$. The payoff matrices are as follows:

We now ask for what values of $\delta$, the stationary outcome $\alpha$, in which $(a, x)$ is played in every period, is justifiable. The reader can verify that in both games the players’ minmax payoffs, $r(L) = (2, 3)$ and $r(R) = (5/2, 2)$, are attained at a Nash equilibrium. Therefore, according to the discussion in the end of the previous section, justifiability of $\alpha$ is equivalent to individual rationality of this outcome.

---

18 Any outcome $\alpha$ must produce a (static) payoff of $\emptyset$ in stage game $R$, and a payoff in the set $[-1, 1]$ (through randomization between -1 and 1) in stage game $L$. 

Electronic copy available at: https://ssrn.com/abstract=4222516
Game $L$ | Game $R$
---|---
\( a \) | \( a \)
\( b \) | \( b \)

Table 2: Uncertain repeated game for Example (3), "⋆" indicates the static best-response.

Since $\alpha$ is stationary, the condition in Eq. (6) can be succinctly rewritten as

\[
(1 - \delta) \bigwedge_{\theta \in \Theta} d(a, x|\theta) \leq \delta \bigwedge_{\theta \in \Theta} (u(a, x|\theta) - z(\theta)).
\]

(7)

Here $\bigwedge$ calculates the worst possible environment to provide dynamic incentives, where the inf over all environments in Eq. (6) is replaced with the inf over current stages, because of stationarity of $\alpha$. The $\bigwedge$ on the left hand side identifies the on-path subgame with the tightest incentive constraint, i.e., the largest deviation payoffs across stage games.

Plugging the specific numbers into Eq. (7), we obtain a system of two inequalities, one for each player:

\[
\begin{align*}
(1 - \delta) \cdot \max\{3, 1\} &\leq \delta \cdot \min\{2, 3/2\}, \\
(1 - \delta) \cdot \max\{1, 2\} &\leq \delta \cdot \min\{1, 2\}.
\end{align*}
\]

\[
\iff \delta \geq \max\{3/5, 1/2\} = 3/5.
\]

To absorb the first inequality, note that the ‘on-path’ payoff for player 1 is 4, achieved at \((a, x)\). Her best deviation in game $L$ is to choose $b$ which gives her 7, a static gain of 3. And, the punishment in game $L$ would be to move to static Nash \((a, z)\) forever after, which leads to a loss of 2 in terms of average payoff. Analogously, the static gain in game $R$ is 1 and dynamic loss is $3/2$, scaled by $(1 - \delta)$ and $\delta$ respectively.

As can be seen from these calculations, player 1’s incentive to deviate is the highest whenever she faces game $L$ today and thinks that all future games will be $R$. The opposite is true for the second player. The precise threshold of $\delta$ for justifiability of \((a, x)\) in every period turns out to be $3/5$.

Can the players cooperate more often if they knew ahead what stage games would arrive in the future? One way to think about this question is to check if such cooperation is possible in SPE for some fixed environment. We say that an outcome is justifiable for environment $e \in \Theta^\infty$ if there exists an SPE for this environment satisfying the condition of Definition 4.

As an illustration, consider the stationary environment $L^\infty$ and evaluate Eq. (4) at it instead of taking the infimum. Following the same steps as above, we arrive at the following condition:

\[
\begin{align*}
(1 - \delta) \cdot 3 &\leq \delta \cdot 2, \\
(1 - \delta) \cdot 1 &\leq \delta \cdot 1.
\end{align*}
\]

\[
\iff \delta \geq \max\{1/3, 1/2\} = 1/3.
\]

So, the outcome in which \((a, x)\) is played in every period can be additionally sustained for $\delta \in$

Electronic copy available at: https://ssrn.com/abstract=4222516
\[ \frac{1}{3}, \frac{2}{3} \] in SPE for \( L^{1:\infty} \). Analogously, it can be shown that for the environment \( R^{1:\infty} \), the same outcome can be sustained for \( \delta \geq \frac{1}{3} \), so the interval \( \left[ \frac{1}{3}, \frac{2}{3} \right] \) gets added. In fact, a stronger claim here is true: for any fixed environment that cannot be expressed as either \( (\theta^{1:t-1}, L, R^{t+1:\infty}) \) or \( (\theta^{1:t-1}, R, L^{t+1:\infty}) \) for some \( t \geq 1 \), the outcome \( \alpha \) is justifiable for all \( \delta < \frac{2}{3} \) sufficiently close to \( \frac{2}{3} \). In other words, having more information about the environment will progressively reduce the threshold required to sustain the outcome \( \alpha \), and this threshold will always lie between \( \left[ \frac{1}{3}, \frac{2}{3} \right] \).

**Example 4.** We saw in Example 3 that a cooperative outcome, which is preferred by both players to the worst static Nash equilibria in every stage game, can be justified in an XPE when the players are sufficiently patient. Does that point towards a folk theorem type of result even when the minmax payoffs aren’t Nash? Perhaps. But what is its statement and how can it be proven?

The main objective of this example is to show when the minmax payoffs aren’t Nash, the standard \( T \)-period punishments blocks may fail to prove a folk theorem type of result for uncertain repeated games even when they work separately in the repetition of each individual stage game (in the sense of Fudenberg and Maskin [1986]).

There are two players and two stage games that differ only in the payoff from the mutual minmax action profile \((c, z)\), the minmax payoffs are identical and equal \((0, 0)\).

<table>
<thead>
<tr>
<th></th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>1, 1</td>
<td>-1, 3*</td>
<td>0*, 0</td>
</tr>
<tr>
<td>( b )</td>
<td>3*, -1</td>
<td>2*, 2*</td>
<td>-3, -2</td>
</tr>
<tr>
<td>( c )</td>
<td>0, 0*</td>
<td>-2, -3</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>

Table 3: Uncertain repeated game for Example (4), “*” indicates the static best-response.

The standard folk theorem argument with two players, which is due to Fudenberg and Maskin [1986], ensures that if only one stage (either \( L \) or \( R \)) is available, then the stationary outcome \( \alpha \) in which \((a, x)\) is played in every period is justifiable when the players are sufficiently patient. The argument is based on using the \( T \)-period punishment outcome \( \tilde{\alpha}^T \) for both players: the players play \((c, z)\) for \( T \) periods and then continue with playing \((a, x)\) for the rest of the game. The corresponding strategy profile \( \sigma^T \) in which the players start by playing \( \tilde{\alpha} \) and switch to \( \tilde{\alpha}^T \) upon any deviation is known to be a SPE for each individual stage game when both \( T \) and \( \delta \) are large enough.

We now show \( \sigma^T \) isn’t an XPE in the uncertain repeated game irrespective of \( T \) and \( \delta \). A necessary condition here is that player \( i \) has no incentive to deviate at the start of \( \tilde{\alpha}^T \) in which case Eq. (4) can be re-written as

\[
\bigvee_{\theta \in \Theta} d(c, z | \theta) \leq \delta T \bigwedge_{\theta \in \Theta} (u(a, x | \theta) - u(c, z | \theta)).
\]
Simplifying and plugging the specific numbers we obtain the following expression:

\[ \max\{1, 3\} \leq \delta T \cdot \min\{2, 4\}, \]

which is clearly violated for all \( \delta \) and \( T \).

To absorb this calculation, consider a standard repeated game in which only game \( L \) can be played each period. Then, either player wants to stick to the punishment outcome \( \tilde{a}^T \) if

\[ (-1) + \delta \cdot (-1) + \ldots + \delta^T \cdot (-1) + \delta^{T+1} \cdot 1 + \ldots \geq 0 + \delta \cdot (-1) + \delta^2 \cdot (-1) + \ldots + \delta^T \cdot (-1) + \delta^{T+1} \cdot 1 + \delta^{T+2} \cdot 1 + \ldots \]

which in terms of average payoffs can be re-written as

\[ (0 - (-1)) \leq \delta^T \cdot (1 - (-1)). \]

This is exactly Eq. (8) where \( \lor \) and \( \land \) drop out because only one stage game is possible. Clearly, the inequality is satisfied for large \( \delta \). Analogously, when only game \( R \) is possible, then we get

\[ (0 - (-3)) \leq \delta^T \cdot (1 - (-3)), \]

which holds for large \( \delta \).\(^{19}\)

For the uncertain repeated game, Eq. (8) takes the maximum of the left side of the inequalities for the standard repeated game with stage games \( L \) and \( R \), and the minimum of the right side of these inequalities. Hence, we get a violation.

Now, even though the standard argument breaks down, the reader might still expect that playing \((a, x)\) in every period can be justified for large values of \( \delta \). In the end, these two stage games are almost identical.

**Example 5.** We now revisit Example 4 and argue that the most effective punishments are sustainable in XPE when the players are sufficiently patient. Formally, we shall show that there is an XPE in which the players receive their discounted minmax payoffs \( U(e) \equiv 0 \) for \( \delta \) above a certain threshold. Then, the stationary outcome in which \((a, x)\) is played in every period can be justified for large values of \( \delta \).

The argument is based on the APS recursion described in Lemma 1. Consider \( W \subset F(\delta) \) defined by

\[ W = \left\{ w \in \mathbb{R}^{2^{4^\omega}} | \exists \gamma \in [0, 2] \text{ s.t. } w \equiv (\gamma, \gamma) \right\}. \]

Note that each element in \( W \) gives the same discounted payoff in all environments.\(^{20}\)

We now show that \( W \subset E(\delta) \) for \( \delta \geq \frac{3}{5} \). In particular, this set is self-generating in the language

\(^{19}\)Note that the instrument of large \( T \) is not used here, but is used to show that no player wants to deviate from \((a, x)\).

\(^{20}\)The fact that the player’ discounted payoffs in \( W \) coincide is merely for simplicity.
of Abreu, Pearce, and Stacchetti [1990], that is

$$ W \subseteq \prod_{\theta \in \Theta} B_\theta W, $$

where recollect that $B_\theta W$ is defined in Eq. (2). Indeed, for each $\theta \in \Theta$, unpacking the expression for $B_\theta W$, the reader can verify that

- the discounted payoff $w \equiv (2, 2)$ can be attained by playing $(b, \gamma)$ and using $v^0 \equiv (2, 2)$ and $(v^1, v^2) \equiv ((0, 0), (0, 0))$ as an on-path and off-path continuations, resp.;

- the discounted payoff $w \equiv (0, 0)$ can be attained by playing $(c, \gamma)$ and using $v^0 \equiv (\gamma, \gamma)$ with

  $$ \gamma = \frac{1-\delta}{\delta} \cdot 1_{\{\theta=L\}} + \frac{3(1-\delta)}{\delta} \cdot 1_{\{\theta=R\}}, $$

where $\theta$ is the current stage game, and $(v^1, v^2) \equiv ((0, 0), (0, 0))$ as an on-path and off-path continuations, resp.;

- the discounted payoff $w \equiv (\gamma, \gamma)$ with $\gamma \in (0, 2)$ can be attained by randomization between two extreme points corresponding to $\gamma = 0$ and $\gamma = 2$.

The key idea is that we are able to achieve fairly strong incentives even while killing the problem of dimensionality by asking for a discounted payoff set whose geometry is independent of the environment, thus $w(e) = (\gamma, \gamma)$ for all $e$. The bullet points above describe how the punishment payoffs are constructed to generate on-path payoffs in $[0, 2]$. The threshold on $\delta$ is generated simply from the fact that for the continuation payoff $\frac{3(1-\delta)}{\delta}$ to be self-generating, it must be less than 2, which gives $\delta \geq \frac{3}{5}$.

Once the minmax payoff is achieved in equilibrium, that is $U \in E(\delta)$, following Lemma 2 and the discussion after, we can completely characterize all justifiable outcome using this "worst" minmax punishment. Note that even though we constructed the discounted minmax payoff through symmetric strategies, once there is a way to achieve it, the characterization of justifiable outcomes through Lemma 2 and associated discounted XPE payoffs through Eq. (3) is complete, and not just limited to symmetric equilibria.

In Example 5, stage games are almost identical and there is a simple way to construct an XPE in which both players receive the same discounted payoffs irrespective of the environment. In general, stage games might differ substantially; as a result, recursing over common discounted payoffs may be not useful. However, the underlying principle of offering environment independent rewards and punishments has a more general appeal than may seem at a first thought. The main trick is to "normalize" the stage games appropriately, so that the idea developed in this example can be applied to general uncertain repeated games.

4 Folk theorem

In this section we state and prove a version of a folk theorem for XPE. Folk theorems for repeated (and stochastic games) are traditionally stated in terms of discounted payoffs. However, dealing with
discounted payoffs in the context of uncertain repeated games is rather difficult because of infinite-dimensionality and lack of stationarity. Hence, the question then arises: What should be the statement of a folk theorem kind of result in the first place?

4.1 The statement

As the reader might guess by now, we will state the folk theorem for uncertain repeated games in terms of outcomes. To motivate the statement that follows, recall that individual rationality is necessary for justifiability (Lemma 3). This assertion is the same in spirit to the fact that in the standard repeated games setting, any discounted SPE payoff is necessarily larger than a minmax payoff vector. Then, traditional folk theorems find conditions under which the (almost) converse statement is true for patient players — any discounted payoff that is above minmax payoffs can be attained in a SPE as \( \delta \to 1 \).

Building on these ideas, we show that under certain conditions individual rationality is also sufficient for justifiability when players are sufficiently patient. Our proof is constructive, and it is based on an existence of three objects: i) a positive vector of “gaps” \( (g^i)_{i \in N} \), one for each player and common across environments, ii) \( n \cdot |\Theta| \) punishment action profiles \((a^i(\theta))_{(i, \theta) \in N \times \Theta}\), and iii) \( n \cdot |\Theta| \) slack variables \((\xi^i(\theta))_{(i, \theta) \in N \times \Theta}\). The result is as follows.

**Theorem 1 (Folk theorem for uncertain repeated games).** Suppose that for each \( i \in N \), there exists a tuple \( (g^i, a^i, \xi^i) \in \mathbb{R}^n_+ \times \prod_{\theta \in \Theta} A(\theta) \times \mathbb{R}^n_+ \) such that for all \( \theta \in \Theta \), \( r_i(a^i(\theta)|\theta) - r_i(\theta) = \xi_i(\theta) = 0 \), and

(a) \( g^i \) belongs to the convex hull of \( \{u(a|\theta) - r(\theta)|a \in A(\theta)\} \);

(b) \( d(a^i(\theta)|\theta) + \xi^i(\theta) \) belongs to the conic hull of \( \{0, g^1, \ldots, g^n\} \);

(c) \( \lambda \cdot d(a^i(\theta)|\theta) + \xi^i(\theta) \geq \max_{(j, \theta) \in N \times \Theta} \lambda \cdot (r(a^j(\theta)|\theta) - r(\theta) + \xi^j(\theta)) \) for some \( \lambda \in \mathbb{R}^n_+ \) with \( \|\lambda\| = 1 \).

Then, there exists \( \bar{\delta} < 1 \) such that for every \( \delta > \bar{\delta} \), an outcome is justifiable iff it is individually rational.

Theorem 1 pins down precisely the set of justifiable outcomes when the players are sufficiently patient. In a nutshell, this theorem establishes that for large \( \delta \), for each player, there is an XPE in which this player is minmaxed irrespective of the environment, i.e., she obtains exactly her discounted minmax payoff \( U_i(e) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r_i(\theta^t) \). It can then be immediately concluded that all individually rational outcomes are justifiable.

In addition, it should be noted that Theorem 1 establishes a stronger claim than traditional folk theorems. To see this, we revisit the classical result for repeated games, i.e., when the same stage game is played in every period. The following result, due to Fudenberg and Maskin [1986], is originally stated in terms of feasible discounted payoffs and proven using stationary outcomes. We restate their result using our notation to stress the connection between stationary outcomes and discounted payoffs in the standard repeated games setting.

Electronic copy available at: https://ssrn.com/abstract=4222516
Theorem 2 (Folk theorem for repeated games (Fudenberg and Maskin [1986])). Suppose there is only one stage game and the interior of \( \text{co} \{ u(a) - r_i | a \in A \} \cap \mathbb{R}_n^+ \) is non-empty. For every feasible discounted payoff \( w \in F = \text{co} \{ u(a) | a \in A \} \) with \( w > r_i \), there exists a (stationary) outcome \( \alpha \) with \( w = U^\alpha \), irrespective of \( \delta \), and a threshold \( \delta > 1 \), such that for all \( \delta > \delta \), the outcome \( \alpha \) is justifiable.

As can be seen from Theorem 2, the traditional folk theorem for repeated games precisely characterizes the set of stationary outcomes as \( \delta \to 1 \). The equivalence is due to the fact that the set of feasible discounted payoffs \( F \) is independent of \( \delta \), and every discount payoff \( w \in F \) can be attained by the same stationary outcome for all values of \( \delta \).

In contrast, Theorem 1 establishes that under conditions (a)-(c) all (potentially non-stationary) individually rational outcome are justifiable for large values of \( \delta \). It is stronger than the traditional folk theorem even for the standard repeated games setting for it pins down when non-stationary outcomes too are justifiable for \( \delta \) above a certain threshold.

In addition, this theorem also applies to dynamic games in which the players know exactly the environment they are facing, i.e., the environment is fixed and known, but it is not stationary. As such, this can constitute any arbitrary (finite action) dynamic game with perfect information, and to be best of our knowledge, a 'folk theorem' for such games hasn’t been formalized.

Now, if several stage games are available, then the analogue of Theorem 2 might not be possible, because the equivalence of the set of feasible discounted payoffs and the set of stationary outcomes breaks down (see Example 1). First, stationary outcomes are now restrictive and cannot trace the whole set \( F(\delta) \); and, in general, \( \lim_{\delta \to 1} E(\delta) \neq \lim_{\delta \to 1} \{ U^\alpha | \alpha \text{ is justifiable and stationary} \} \).

Second, even for a stationary outcome \( \alpha \), the associated discounted payoff \( U^\alpha(e) \) is a non-trivial function of \( \delta \) when the underlying environment \( e \) is non-stationary, and \( \lim_{\delta \to 1} U^\alpha(e) \) depends on frequencies of stage games in the tail of this environment \( e \). On the other hand, Theorem 1 still applies in that it gives the sets of justifiable outcomes and associated discounted XPE payoffs for large values of \( \delta \). In this sense, Theorem 1 provides a conventional reading of folk theorem type of results that "everything goes for \( \delta \) close enough to 1".

Since the conclusion of Theorem 1 is stronger than the traditional result, it is natural to expect that a more demanding set of conditions is needed. The three technical conditions invoked in the theorem are exposited in Figure 2.

The first condition demands that each vector \( g^i \gg 0 \) can be expressed as a convex combination of the stage game payoffs normalized by the minmax payoff vectors. As shown in Example 2, some condition like this is necessary to provide incentives in uncertain repeated games. It is our version of the interiority assumption, which demands that some incentives can be provided across stage games.

The second condition guarantees that each vector \( \frac{1-\delta}{\delta} \{ d(a^i(\theta)) + \xi^i(\theta) \} \) belongs to a convex hull of \( 0 \) and \( g^i \)'s for all sufficiently large \( \delta \). We view condition (b) as rather mild when \( g^i \)'s can be chosen sufficiently close to the axes, that is the conic hull of \( 0 \) and non-negative points in \( \cap_{\theta \in \Theta} \text{co} \{ u(a) - r_i \} \).
\( r(\theta) \) is all of \( \mathbb{R}^n_+ \). Even if it is not the case, the slack variables can be used to relax the constraint as shown in Figure 2b.

The third condition ensures that each vector \( d(a^i(\theta)|\theta) + \xi^i(\theta) \) lies above the hyperplane through the stage game deviation payoffs from the minmax action profiles in the normalized uncertain repeated game. Condition (c) isn’t restrictive at all when there are two players.\(^{22}\) For uncertain repeated games with more than two players, this condition can be seen as a weaker version of the requirement that in each stage game there is a way to punish a deviator without hurting the other players too much uniformly across stage games.

Finally, it can be noted that (b) and (c) put no restrictions if every stage game is symmetric and the interiority condition holds. Symmetry and condition (a) ensure that there exists some \( \gamma > 0 \) such that \( (\gamma, \ldots, \gamma) \) belongs to the convex hull of \( \{u(a|\theta) - r(\theta) | a \in A(\theta)\} \) for all \( \theta \in \Theta \). So, both conditions (b) and (c) are satisfied when each \( a_i^i(\theta) \) is the joint minmax in game \( \theta \) and \( \xi^i \equiv 0 \), because \( r(a_i^i(\theta)|\theta) - r(\theta) = 0 \) and \( d(a_i^i(\theta)|\theta) \) lies on the diagonal.

It is, however, important to note that it is not true that such symmetric gaps \( (\gamma, \ldots, \gamma) \) are necessarily attained by symmetric action profiles. Asymmetric action profiles are needed when every symmetric action profile are worse than a minmax payoff vector. For example, suppose there is one stage game and two players:

<table>
<thead>
<tr>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>1*, 2*</td>
</tr>
<tr>
<td>2*, 1*</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

So even for symmetric games, the folk theorem cannot, in general, be proven using only strongly symmetric equilibria.

To formally prove the folk theorem for uncertain repeated games, we will now develop a recursive technique to characterize certain subsets of discounted XPE payoffs in a robust (and tractable) way. Then, we will be able to complete the proof of Theorem 1 and explain the role of the assumptions in more details.

\(^{22}\)For each player \( i \), let the action profile \( a^i(\theta) \) to be the joint minmax in game \( \theta \) and set \( \xi^i \equiv 0 \). The reader can verify that condition (c) is then always satisfied.
4.2 Recursive technique

Here we characterize certain subsets of the set of discounted XPE payoffs in which it is possible to recurse over environment-independent gaps.

Formally, a gap \( g \in \mathbb{R}^n \cdot \Theta^\infty \) in a set \( W \subseteq \mathbb{R}^n \cdot \Theta^\infty \) is a vector that can be expressed as

\[
g_i = \nu_i^0 - \nu_i^1 \quad \forall i \in N
\]

for some tuple \((\nu^i)_{i=0}^n \subset W\). This essentially captures the right hand side of the incentive constraint in the APS recursion, Eq. (2) and definition of justifiability, Eq. (4). So, gaps encode all the relevant information about dynamic losses from deviations. And, this gap is said to be environment-independent (or common) if \( g(e) \) is the same for all environments \( e \in \Theta^\infty \). Common gaps assert that such losses are constant across different environments, thus \( \inf_{e \in \Theta^\infty} \) in Eq. (2) and Eq. (4) are made redundant.

We now develop a technique to recurse over common gaps in a tractable way. Since stage games might differ substantially, it is useful to first pre-normalize them by some fixed payoffs \( p \in \prod_{\theta \in \Theta} \text{co} \{ u(a|\theta) | a \in A(\theta) \} \). In what follows we consider the normalized uncertain repeated game in which utilities are given by \( u(\cdot|\theta) - p(\theta) \). And, to keep track of the normalization, we write \( \mathcal{B}^p_\theta \) for the APS operator and \( E^p(\delta) \) for the set of discounted XPE payoffs that correspond to the chosen value of \( p \).

We seek to identify self-generating sets \( W \subseteq E^p(\delta) \) that can be written as

\[
W = \left\{ w \in \mathbb{R}^n \cdot \Theta^\infty \mid \exists \gamma \in \bigcap_{\theta \in \Theta} \Gamma(\theta) \text{ s.t. } w(\theta, \cdot) \equiv \gamma(\theta) \quad \forall \theta \in \Theta \right\}, \tag{9}
\]

where \((\Gamma(\theta))_{\theta \in \Theta}\) are some compact subsets of \( \mathbb{R}^n \). So every discounted payoff in \( W \) may depend only on this period’s game but not on the sequence of stage games in the future.

The functional form shown in Eq. (9) is tractable in that \( W \) has a manageable boundary that can be found explicitly as opposed to the whole set of infinite-dimensional discounted XPE payoffs, which can potentially have a significantly complex geometry. Then, the elements of \( W \) can be used in Lemma 2 as punishments. As we will show later, such punishments can deliver the folk theorem for uncertain repeated games as well as provide an exact characterization of strongly symmetric XPE.

We will make use of the following result in the proof of the folk theorem.

**Lemma 4.** Let \( p \in \prod_{\theta \in \Theta} \text{co} \{ u(a|\theta) | a \in A(\theta) \} \) be given. Then, \( W \) defined in Eq. (9) is a self-generating subset of \( E^p(\delta) \) if \( \bigcap_{\theta \in \Theta} \Gamma(\theta) \neq \emptyset \) and

\[
\Gamma(\theta) \subseteq T^p_\theta \left( \bigcap_{\theta \in \Theta} \Gamma(\theta) - \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(\theta) \right) + \delta \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(\theta) \quad \forall \theta \in \Theta, \tag{10}
\]

For now \( p \) is taken to be exogenous, ways to determine this normalization are discussed in Section 5.2.

23\(^\text{12}\) \( E^p(\delta) \) is nothing but a translation of the set of discounted XPE payoffs in the original uncertain repeated game by \( (1-\delta) \sum_{t=1}^\infty \delta^{t-1} p(\theta^t) \).

24
where for each \( \theta \in \Theta \), the operator \( T_\theta \) is defined over bounded subsets of \( \mathbb{R}^n \) as

\[
T_\theta G := \text{co} \left\{ \gamma \in \mathbb{R}^n | \exists (a, g) \in A(\theta) \times G \text{ s.t. } \gamma = (1 - \delta)(u(a|\theta) - p(\theta)) + \delta g, \quad \frac{1-\delta}{\delta} d(a|\theta) \leq g \right\}. \quad (11)
\]

To absorb the notation here, note that \( \hat{\theta} \in \Theta \) first registers the pointwise infimum of a compact Euclidean set \( \Gamma(\hat{\theta}) \subset \mathbb{R}^n \), and then takes the coordinatewise supremum over these for the set of all stage games in \( \Theta \). So, \( \bigvee_{\theta \in \Theta} \Gamma(\hat{\theta}) \) is the vector of lowest payoffs in \( (\Gamma(\theta))_{\theta \in \Theta} \), one for each player, that can be attained irrespective of a stage game. And, \( \bigwedge_{\theta \in \Theta} \Gamma(\tilde{\theta}) \) is simply the set of common vectors in this family of sets. Putting these together, the reader should think of \( \bigwedge_{\theta \in \Theta} \Gamma(\tilde{\theta}) - \bigvee_{\theta \in \Theta} \Gamma(\hat{\theta}) \) as a set of common gaps for the family \( (\Gamma(\theta))_{\theta \in \Theta} \).

With this in mind, Lemma 4 provides a sufficient condition for a set \( W \) of the form described in Eq. (9) to be self-generating with respect to the APS recursion in the normalized uncertain repeated game, i.e.,

\[
W \subseteq \prod_{\theta \in \Theta} B^\theta W.
\]

We show in the appendix that every element of \( W \) can be obtained by recursing over environment-independent gaps in the following fashion:

- for each player \( i \), her punishment continuation \( v^i \in W \) delivers this player the same discounted payoff irrespective of the future realization of stage games; in fact, \( v^i_i \equiv \bigvee_{\theta \in \Theta} \bigwedge_i \Gamma(\theta) \) is the smallest number that can be guaranteed to be available across stage games;

- a constant gap \( g \in \bigwedge_{\theta \in \Theta} \Gamma(\tilde{\theta}) - \bigvee_{\theta \in \Theta} \Gamma(\hat{\theta}) \) over the punishments continuations \( (v^i_i)_{i=1}^n \) is offered on-path.

The right-hand side of Eq. (10) precisely pins down the set of discounted payoffs that can be obtained in this way. The first term computes the set of discounted payoffs that can be attained with common gaps in the normalized game \( \theta \) using the (finite-dimensional) gap recursion corresponding to this game, given by Eq. (11). The operator \( T_\theta \) resembles the standard APS operator, that is used in the repeated games setting, in which the players are punished by zero. The second term is an adjustment that adds the punishment payoffs, which were used to compute common gaps, to map the gaps back into payoffs.

### 4.3 Proof of the folk theorem

We now have all ingredients to prove the folk theorem for uncertain repeated games. Necessity of individual rationality for outcomes to be justifiable is already discussed in the text (Lemma 3), we therefore focus here on sufficiency.

---

24The other players’ discounted payoffs might depend on the future, i.e., \( v^j_i \) for \( j \neq i \) need not be constant across environments.

---

Electronic copy available at: https://ssrn.com/abstract=4222516
Proof of Theorem 1. Without further ado, normalize each stage game by its minmax payoffs, i.e., \( p(\theta) := \gamma(\theta) \) for all \( \theta \in \Theta \). Fix the three objects: i) common gaps \(( g^i)_{i \in N} \), ii) punishment action profiles \(( a'((\theta))_{(i,\theta) \in N \times \Theta} \) and iii) slack variables \(( \varepsilon((\theta))_{(i,\theta) \in N \times \Theta} \) that satisfy conditions (a), (b) and (c). Define \( \Gamma(\theta) \) as follows:

\[
\Gamma(\theta) := \{ g^1, \ldots, g^n, (1-\delta)(g(a^1(\theta)|\theta)-\gamma(\theta)+\varepsilon((\theta))), \ldots, (1-\delta)(g(a^n(\theta)|\theta)-\gamma(\theta)+\varepsilon((\theta))) \}.
\]

We claim that for sufficiently large values of \( \delta \), the conditions of Lemma 4 are met; thus,

\[
\mathcal{W} := \left\{ w \in \mathbb{R}^n : \exists \gamma \in \prod_{\theta \in \Theta} \Gamma(\theta) \text{ s.t. } w(e) = \gamma(\theta_1) + U_e(\theta) \forall e \in \Theta^{\infty} \right\}
\]

is a subset of \( \mathcal{E}(\delta) \). Note that for all \( \theta \in \Theta \) we have \( \bigwedge \Gamma(\theta) = 0 \) because \( \varepsilon((\theta)) = 0 \) for every \( i \in N \). As a result, Eq. (12) implies that for every player \( i \), there is a justifiable outcome in which this player receives her discounted minmax payoff \( U_i \). Then the folk theorem immediately follows because if the worst individually rational payoff is achieved by a justifiable outcome, then this can be used as punishment to achieve all other individually rational outcomes.

Set \( G := \bigcap_{\theta \in \Theta} \Gamma(\theta) \). By construction, the set \( G \) is non-empty because it contains at least the convex hull of \( \{ g^1, \ldots, g^n \} \), thanks to condition (a). Furthermore, conditions (b) and (c) jointly ensure that for all sufficiently large values of \( \delta \), each vector \( \frac{1-\delta}{\delta}(d(a^i(\theta)|\theta)+\varepsilon^i((\theta))) \) is in \( G \), see the figure below.

![Figure 3: Illustrating the construction of G.](https://ssrn.com/abstract=4222516)

In view of Lemma 4, since \( \Gamma(\theta) \) is convex, it is sufficient to establish that every extreme point of \( \Gamma(\theta) \) is in \( T^g_G(G) \). This is nothing but the inclusion in Eq. (10) combined with the fact that \( \bigvee_{\theta \in \Theta} \bigwedge \Gamma(\theta) = 0 \). To this end, consider the following construction.

- Condition (a) ensures that there is a way to generate the gap \( g^i \) in each normalized stage game \( \theta \) by randomizing according to some probabilities \( \mu(\theta) \in \Delta(\theta) \). Since all coordinates of \( g^i \) are positive, \( \frac{1-\delta}{\delta} \sum_{a \in A(\theta)} d(a|\theta) \leq g \) for all sufficiently large \( \delta \). It follows that the gap \( g^i \) belongs...
to $T^\delta_\theta(G)$ when $\delta$ is large enough.

- The gap $(1-\delta)(r(a'(\theta)|\theta) - r(\theta) + \xi^i(\theta))$ can be attained by playing $a'(\theta)$ with certainty and using $\frac{1-\delta}{\delta}(d_i(a'(\theta)|\theta) + \xi^i(\theta))$ as a continuation gap, i.e.,

$$(1-\delta)(r(a'(\theta)|\theta) - r(\theta) + \xi^i(\theta)) = (1-\delta)(u(a'(\theta)|\theta) - r(\theta)) + \delta \cdot \frac{1-\delta}{\delta}(d_i(a'(\theta)|\theta) + \xi^i(\theta)).$$

It is easy to see that, since $\xi^i(\theta) \geq 0$, no player has any incentive to deviate. As argued above, conditions (b) and (c) jointly ensure that this continuation is in $G$ when the players are patient enough.

To sum up, the construction here establishes that there is $\bar{\delta}$ such that for any $\delta > \bar{\delta}$, the set $W$ defined in Eq. (12) is a subset of $E(\delta)$. Thus, for such large values of $\delta$, each player can be minmaxed irrespective of the environment. □

5 Equilibria for fixed $\delta$

Many applications of the theory of standard repeated games seek, in addition to multiple long-run players, equilibria for a fixed $\delta$; see for example the study collusion in Green and Porter [1984]), price wars in Rotemberg and Saloner [1986], sustainable policy plans in Chari and Kehoe [1990], risk sharing in Kocherlakota [1996], and relational contracts in Levin [2003]. In potentially extending these and other applications to uncertain environments of the form studied here, it is natural to ask how the theory developed thus far can be used to study equilibria for fixed (and not just large) $\delta$.

In the previous section, we saw that recursing over common gaps in suitably normalized uncertain repeated games can sustain the discounted minmax payoffs in XPE when $\delta$ is large. This recursive technique of constructing punishments is useful (and quite powerful) even for intermediate values of $\delta$. First, it provides an exact characterization of strongly symmetric XPE, the fact that will be established later in Section 6.3. Second, for general asymmetric XPE, it gives a partial, yet reasonably tight, characterization of justifiable outcomes through Lemma 2, where the meaning of words “partial” and “tight” will be formalized.

5.1 The common gap recursion

We now present a systematic approach to compute punishments using the common gap recursion, inspired by Lemma 4. Recall that a normalizing vector $p \in \prod_{\theta \in \Theta} co\{u(a|\theta)|a \in A(\theta)\}$ is fixed at the outset. The following result establishes that for each such $p$, the family of sets defined by Eq. (9) that satisfy the condition of Lemma 4 admit the largest element, and moreover, it provides an iterative procedure to find this largest element.

Proposition 1. Let $p \in \prod_{\theta \in \Theta} co\{u(a|\theta)|a \in A(\theta)\}$ be given. Set $G^0 := \bigcap_{\theta \in \Theta} co\{u(a|\theta) - p(\theta)|a \in A(\theta)\}$ and $G^{n+1}$ be the family of closed subsets of $A(\theta)$ that satisfy the conditions of Lemma 4. Then, $G := \bigcap_{\theta \in \Theta} co\{u(a|\theta) - p(\theta)|a \in A(\theta)\}$ is the largest element of $G^0$.
\( A(\theta) - \bigvee_{\theta \in \Theta} \wedge_{a \in A(\theta)} (u(a|\theta) - p(\theta)) \) and define inductively \((G^k)^\infty_{k=1}\) as follows:

\[
G^{k+1} := \begin{cases} 
\cap_{\theta \in \Theta} T_{\theta}^p G^k - \bigvee_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G^k & \text{if } \cap_{\theta \in \Theta} T_{\theta}^p G^k \neq \emptyset, \\
\emptyset & \text{if } \cap_{\theta \in \Theta} T_{\theta}^p G^k = \emptyset.
\end{cases}
\] (13)

Then, the sequence \((G^k)\) converges (in the Hausdorff distance) to some compact (potentially empty) set \(G(p)\). Furthermore:

- If \(G(p)\) is non-empty, then \((\Gamma(\theta|p))_{\theta \in \Theta}\), defined as \(\Gamma(\theta|p) \equiv T_{\theta}^p G(p) + \frac{\delta}{1-\delta} \cap_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G(p)\), is the largest (in the sense of set-inclusion) collection of sets satisfying Eq. (10) in Lemma 4;

- If \(G(p)\) is empty, then there is no collection of sets satisfying Eq. (10) in Lemma 4.

To reiterate, Proposition 1 delivers a necessary and sufficient condition for a collection of sets \((\Gamma(\theta|p))_{\theta \in \Theta}\), satisfying the conditions of Lemma 4 to exist, that is non-emptiness of the limit set \(G(p)\). The proposition also asserts the existence of the largest collection of such sets, and describes an iterative procedure to obtain it.

The iterative approach recurses over common gaps, as described in Eq. (13). The recursion admits a set of common gaps \(G\) that can be used as continuation gaps. It then produces as an output, the set of common gaps that are attainable today in the normalized uncertain repeated game, that is

\[
\cap_{\theta \in \Theta} T_{\theta}^p G - \bigvee_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G.
\]

The first term corresponds to the set of common discounted payoffs that can be attained with common gaps, and the second term records the lowest common discounted payoffs, one for each player. The difference between them is precisely the set of common gaps over the lowest common discounted payoffs, that closes the recursive loop.

The iterations in Eq. (13) start with the “largest” possible set of common gaps \(G^0\), and then inductively construct “smaller” and “smaller” sets of common gaps. The sets \((G^k)\) aren’t necessarily ordered by inclusion because at each inductive step the common gap recursion also recomputes the "origin", i.e., \(\cap_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G^k\), translating common gaps in the south-west direction. Nevertheless, these sets are guaranteed to converge to a certain limit \(G(p)\), which might be empty.

If the limit set \(G(p)\) is non-empty, we obtain the largest collection of sets \((\Gamma(\theta|p))_{\theta \in \Theta}\) satisfying the condition of Lemma 4. Specifically, for each stage game \(\theta \in \Theta\), the set \(\Gamma(\theta|p)\) is simply the set of discounted payoffs that can be attained today with common gaps in \(G(p)\) plus the lowest common discounted payoffs \(\frac{\delta}{1-\delta} \cap_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G(p)\) from the next period onward, thus pre-multiplied by \(\delta\).

---

26 We show in the proof of this proposition that \(\cap_{\theta \in \Theta} \cap_{\theta \in \Theta} T_{\theta}^p G^k\) is non-decreasing.

27 Convergence here is meant in the sense of Hausdorff distance \(d_H\), with convention \(d_H(\emptyset, \emptyset) = 0\).
5.2 Constructing punishments and choosing normalization

Having described the general common gap recursion that works for all \( \delta \), we now operationalize it to find a subset of justifiable outcomes. They key to that end is to identify the worst element in the set of discounted XPE payoffs generated by the common gap recursion, which when substituted in Eq. (4), characterize justifiable outcomes.

Proposition 1 delivers a simple approach to construct a subset of XPE. In particular, it gives the sets \( \{\Gamma(\theta|p)\}_{\theta \in \Theta} \) that satisfy the conditions of Lemma 4, thus

\[
W(p) := \left\{ w \in \mathbb{R}^{n \times \Theta} \mid \exists y \in \prod_{\theta \in \Theta} \Gamma(\theta|p) \text{ s.t. } w(e) = \gamma(\theta_1) + (1 - \delta) \sum_{i=1}^{\infty} \delta^{i-1} p(\theta^i) \ \forall e \in \Theta^n \right\} \tag{14}
\]

is a subset of \( E(\delta) \). By varying \( p \) and using the worst payoff in \( W(p) \) for each player as punishment, we can identify the (sub)set of justifiable outcomes.

In fact, as a quick sanity check note that if there is only one stage game only, and we are back to the standard repeated games setting, then the set \( W(p) \) is independent of \( p \) and it coincides with \( E(\delta) \)\(^{28}\) The sets \( \{\Gamma(\theta|p)\}_{\theta \in \Theta} \) computed in Proposition 1, as a function of \( p \), will adjust so that when calculating \( w(e) = \gamma(\theta_1) + (1 - \delta) \sum_{i=1}^{\infty} \delta^{i-1} p(\theta^i) \) in the definition of \( W(p) \) above, the value turns out to be independent of \( p \).

More generally, we obtain a partial characterization of the set of justifiable outcomes. Recollect that \( \wedge \) records the pointwise infimum of a bounded Euclidean set, and \( \wedge_i \) returns a scalar, the projection of this vector onto the \( i \)-th coordinate of the set.

**Corollary 1.** An outcome \( \alpha \) is justifiable if there exists \( (p^i)_{i=1}^n \subset \prod A(\theta) \) such that \( \forall (\omega^{1:t}, \theta^{1:t}) \in [0, 1]^t \times \Theta^t, \ \forall t = 1, ..., t \in N, \)

\[
(1 - \delta) d_i(\omega^{1:t} | \theta^{1:t}) \leq \inf_{\tilde{\alpha} \in \Theta^n} \left\{ \mathbb{E}_t \left[ U^{t+1}_{i+1}(\omega^{1:t+1} | \theta^{1:t}, \tilde{\alpha}) \right] - \right. \\
- \left. \sum_{i} \Gamma(\tilde{\alpha}^i | p^i) - (1 - \delta) \sum_{k=1}^{\infty} \delta^{k-1} p^k_i(\tilde{\alpha}^k) \right\}. \tag{15}
\]

The corollary provides a sufficient condition for justifiability, punishing each player \( i \) by her worst discounted payoffs in \( W(p^i) \) for some fixed normalizing vector of payoffs \( (p^i)_{i=0}^n \). As we saw in Theorem 1, such a characterization is guaranteed to be tight (under the three conditions of the theorem) when the players are patient enough and each normalizing vector equals the minmax payoff, i.e., \( p^i(\theta) \equiv r(\theta) \). For intermediate values of \( \delta \), it might not be possible to attain the discounted minmax payoffs in an XPE in general or more specifically with common gaps. So, a normalization distinct from the minmax payoff vector and different for different players might be useful.

\(^{28}\)Mathematically, the common gap recursion described in Proposition 1 is isomorphic to the standard APS recursion defined in Lemma 1.
Corollary 1 opens a question on how to select the normalizing payoffs \((p^i)_{i=0}^n\). We leave this to the reader to decide and only point to some choices.

To begin, let \(p^i\) select the player \(i\)'s worst static (pure) Nash equilibrium in each stage, assuming that such equilibria exist, of course. It is easy to see that the collection \((\Gamma(\theta))_{\theta \in \Theta}\) with \(\Gamma(\theta) = \{0\}\) satisfies the condition of Lemma 4 for this fixed normalization \(p^i\). It follows then from Proposition 1 that the set \(G(p^i)\) is non-empty, thus the associated largest collection \((\Gamma(\theta|p))_{\theta \in \Theta}\) identified in this proposition is non-empty as well. We conclude that the player \(i\)'s worst discounted payoff in \(W(p^i)\) is (weakly) smaller than her discounted payoff from playing her worst static (pure) Nash equilibrium in each stage. So, we have punishments that are at least as good as Nash reversion, maybe even better. For completeness, we record this observation in the corollary that follows.

**Corollary 2.** Suppose every stage game admits a pure strategy Nash equilibrium. For every player, there exists a normalization \(p \in \prod_{\theta \in \Theta} \text{co}(u(a|\theta)|a \in A(\theta))\) such that this player’s worst discounted XPE payoff in \(W(p)\) as defined in Eq. (14) is (weakly) smaller than her discounted payoff under any outcome that plays some static Nash equilibrium in every period.

Corollary 2 shows that recursing over common gaps in normalized uncertain repeated games delivers punishments that are at least as tight as Nash reversion. In general, there is no “tightest punishment” for player \(i\), i.e., the normalization \(p\) that minimizes

\[
\bigwedge_i \Gamma(\theta^i|p) + (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} p_i(\theta^t)
\]

simultaneously for all environments. In specific applications however the collection of sets \((\Gamma(\theta|p))_{\theta \in \Theta}\) can be approximated for each fixed \(p \in \prod \text{co}(u(a|\theta)|a \in A(\theta))\) on some finite grid using the iterative procedure of Proposition 1 and variations of numerical algorithms in Judd et al. [2003], Abreu and Sannikov [2014] and Abreu et al. [2020]. It is then possible to select those \(p\) that deliver lower values of \((\bigwedge_i \Gamma(\theta|p))_{\theta \in \Theta}\) and \((p_i(\theta))_{\theta \in \Theta}\) in a Pareto sense. Alternatively, \(p\) can be selected by minimizing some (universal or global) measure of tightness of punishments with common gaps in the uncertain repeated game normalized by \(p\). We propose one such measure in Section 6.2.

If one has to take a call on \(p\) without any reference to numerics, then a possible choice is to determine the normalization endogenously, by setting \(p = \overline{p}\), where \(\bigwedge_i \Gamma(\theta|\overline{p}) = 0\). In other words, the normalization is chosen in a way that the lowest discounted payoffs satisfy

\[
\bigwedge W(\overline{p}) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \overline{p}(\theta^t).
\]

That is, the uncertain repeated game is normalized by exactly the lowest payoff that can be attained with common gaps over these lowest payoffs.

To be clear, this endogenous normalization isn’t guaranteed to be the “best”. Yet, as we will see later, it provides an exact description of the set of justifiable outcomes for strongly symmetric XPE for every \(\delta < 1\). We shall establish this fact by deriving a bound on how much each player can be
credibly punished in any XPE. More generally, the endogenous normalization performs well with respect to the aforementioned bound for intermediate values of \( \delta \) even when asymmetric equilibria are included.

5.3 Two examples

We now illustrate the workings of the common gap recursion using two classical models, namely Cournot and (differentiated) Bertrand. In both models there are two symmetric players and two stage games, \( \theta = L, R \). The stage game payoffs are parametrized by a demand parameter \( D \) and marginal cost \( C \).

In what follows, we find the endogenous normalization \( p \) by performing iterations described in Proposition 1 and updating \( p \) at each step so that \( \bigwedge \mathcal{T}^p_{\theta} G^k = \bigwedge \mathcal{T}^0_{\theta} G^k - (1 - \delta)p \equiv 0 \). The iterations are performed in Julia using exact rational arithmetic. Specifically, we represent convex sets through their extreme points and iterate over such extreme points using Polyhedra, LazySets packages. In the end we round all results to \( 1/100 \).

**Example 6** (Cournot). Player \( i \)’s payoff is given by

\[
 u_i = \left( \max \left\{ 0, D - a_i - a_{-i} \right\} - C \right)a_i.
\]

The two games are parametrized as follows: \((D(L), C(L)) = (\nicefrac{9}{2}, 2)\) and \((D(R), C(R)) = (3, 1)\). Each player \( i \)’s set of quantities, \( A_i(\theta) \), is the grid of 11 evenly spaced points from 0 to \( D(\theta) \).

Figure 4: Set of common gaps in the Cournot model: dashed — only \( L \) available; dotted, light gray — only \( R \) available and dark gray — both stage games available.

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p(L) ), ( L ) only</td>
<td>0.63</td>
<td>0.32</td>
<td>0.11</td>
<td>0.11</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>( p(R) ), ( R ) only</td>
<td>0.48</td>
<td>0.16</td>
<td>0.06</td>
<td>0.06</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>( p(L) ), both ( L ) and ( R )</td>
<td>0.63</td>
<td>0.34</td>
<td>0.32</td>
<td>0.11</td>
<td>0.11</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>( p(R) ), both ( L ) and ( R )</td>
<td>0.48</td>
<td>0.17</td>
<td>0.06</td>
<td>0.06</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 4: Endogenous normalization associated to set of of common gaps in the Cournot model.
Figure 4 plots the set of gaps for the repeated games with stages games L and R and the set of common gaps for the uncertain repeated game in which both games are possible. As we increase the value of δ, each individual set expands, gets closer to the axes, and the dark region of the common gaps converges to the intersection of gaps of the two individual stage games. So more can be sustained as we increase δ, and common gaps alone can provide powerful incentives for large δ.

Indeed, Table 4 makes this point formal, documenting the lowest (average) payoff that can be generated in an XPE with common gaps and an endogenous normalization. The first two rows look at the individual repeated games, and the third and fourth rows look at the realization of the two stage games separately for the uncertain repeated game in which both games are possible.

Note that 0.63 and 0.48 are (approximately) the Nash equilibrium payoffs in games L and R resp., and the (static) minmax payoff in both games is 0. So when δ is small, the worst (and indeed the best) equilibrium payoff is static Nash. At the other extreme when δ is reasonably high, the minmax is attained, which gives us the folk theorem. The action in the middle is the gradual decrease in the worst attainable payoff.

The value in first row is lower than third always and that in the second two is lower than the fourth. In fact around δ = 0.5, the minmax has almost been attained in the two individual repeated games but not in the uncertain repeated game for as we saw in Example 3, providing incentives is harder in the uncertain repeated game.

Example 7 (Bertrand). Player i’s payoff is given by

\[ u_i = \max \left\{ 0, D - a_i + \frac{d}{2} \right\} (a_i - C). \]

As before, the two games are parametrized by \((D(L), C(L)) = (\frac{3}{2}, 2)\) and \((D(R), C(R)) = (3, 1)\). For each player i, set of price choices, \(A_i(\theta)\), is the grid of 11 evenly spaced points from 0 to \(D(\theta)\).

![Figure 5: Set of common gaps in the Bertrand model: dashed — only L available; dotted, light gray — only R available and dark gray — both stage games available.](https://ssrn.com/abstract=4222516)
from the origin in the set of common gaps from small values of $\delta$. So if we restricted the analysis to strongly symmetric equilibria, then much more can be done in Bertrand than in Cournot.

The reading of Table 5 is also analogous to Table 4: 5.07, 2.81 are (approximately) the static Nash payoffs and 1.55, 0.99 are (approximately) the static minmax payoffs in games $L$ and $R$ resp. Here again when $\delta$ is small, the worst (and indeed the best) equilibrium payoff is static Nash. As $\delta$ increases, the players can sustain more efficient punishments in an XPE, and for large $\delta$, the discounted minmax payoffs are attained. The threshold value of $\delta$ required for the “hardest” punishments to be the minmax payoffs is larger in the uncertain repeated game than in each individual stage games.

6 Tightness of common gap XPE

In the previous section, a subset of XPE was constructed which uses common gaps to provide incentives and met out punishments. A natural question to then ask is this: How much are we missing by focusing on this subset of equilibria?

We already know from Theorem 1 that these class of incentives are sufficient to sustain all equilibria (under some conditions) for large enough discount factors. We also noted in Corollary 2 that Nash reversion is covered by the common gap recursion. Intuitively, for a small enough $\delta$, (static) Nash equilibria are all that can be sustained in standard repeated games, and hence also in the associated uncertain repeated game. So for small $\delta$ too, there is no loss from using these smaller class of incentives.

In this section we will provide a tight bound for intermediate values of $\delta$. We devise a lower bound on the worst XPE payoff across environments, thereby delivering an outer bound on the (infinite-dimensional) discounted XPE payoffs. Then, we formulate a measure of tightness to quantify the distances between the common gap equilibrium set and the outer bound. Finally, we show that the bound is exact for strongly symmetric equilibria.

6.1 An outer bound

One possible bound on discounted XPE payoffs, since any justifiable outcome is necessarily individually rational, is the discounted minmax payoffs, that is for every justifiable $\alpha$,

$$U^{\alpha}(e) \geq \underline{U}(e) := (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \underline{r}(\theta^t) \quad \forall e \in \Theta^\infty.$$  

(17)

Table 5: Endogenous normalization associated to the set of of common gaps in the Bertrand model.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0(L)$, $L$ only</td>
<td>5.07</td>
<td>5.07</td>
<td>4.61</td>
<td>4.15</td>
<td>3.69</td>
<td>2.88</td>
<td>1.81</td>
<td>1.55</td>
<td>1.55</td>
</tr>
<tr>
<td>$p_0(R)$, $R$ only</td>
<td>2.81</td>
<td>2.31</td>
<td>2.31</td>
<td>2.10</td>
<td>1.68</td>
<td>1.32</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>$\bar{p}_0(L)$, both $L$ and $R$</td>
<td>5.07</td>
<td>5.07</td>
<td>4.61</td>
<td>4.61</td>
<td>4.15</td>
<td>3.24</td>
<td>2.88</td>
<td>2.16</td>
<td>1.55</td>
</tr>
<tr>
<td>$\bar{p}_0(R)$, both $L$ and $R$</td>
<td>2.81</td>
<td>2.55</td>
<td>2.31</td>
<td>2.10</td>
<td>1.68</td>
<td>1.32</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Electronic copy available at: https://ssrn.com/abstract=4222516
As we argued in Theorem 1, the discounted minmax payoffs give the exact bound for large values of $\delta$ (if conditions (a)-(c) hold); however, this bound mostly may not be exact, or even tight enough, for intermediate values of $\delta$. We therefore aim to obtain a tighter bound $\bar{U} \geq U$ such that Eq. (17) still holds with $U$ replaced by $\bar{U}$. As we will see, for intermediate values of $\delta$, the tighter bound $\bar{U}$ isn’t the same as the discounted minmax payoffs, and it is always exact for strongly symmetric XPE.

To begin note that $U$ is too permissive precisely because it looks for minmax payoffs across all action profiles, even those which cannot ever be played on path. This observation suggests to strengthen the condition in Eq. (17) by focusing on action profiles, deviations from which can be deterred by some gaps in $E(\delta)$.

Formally, define the set of enforceable action profiles $S(\theta)$ in game $\theta \in \Theta$ as

\[
S(\theta) := \left\{ a \in A(\theta) | \exists (\psi^i)_{i=0}^n \subset E(\delta) \text{ s.t. } \frac{1-\delta}{\delta} d(a|\theta) \leq \inf_{e \in \Theta^{\omega}} \{ \psi^{i}(e) - \psi_i^{i}(e) \} \right\}
\]  

(18)

Then, the following result is an immediate consequence of Lemma 2.

Lemma 5. For every justifiable outcome $\alpha$, each action profile played in period $t$ in a game $\theta^t$ is an element of $S(\theta^t)$. As a result, the discounted payoff under $\alpha$ satisfies

\[
U^\alpha(e) \geq (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r_S(\theta^t) \quad \forall e \in \Theta^{\omega},
\]

where $r_S(\theta^t) := \min_{a \in S(\theta^t)} r(a|\theta) \in \text{co} \{ u(a|\theta) | a \in A(\theta) \}$ for all $\theta^t \in \Theta$.

Lemma 5 is the first step towards our construction of the tighter lower bound $\bar{U}$. It is useful because for intermediate values of $\delta$, the sets $(S(\theta))_{\theta \in \Theta}$ could exclude some action profiles; thus, the resulting “quasi-minmax” payoffs $r_S$ could be different than the original minmax payoffs.

However, Lemma 5 cannot be applied as is, because the sets of enforceable action profiles described in Eq. (18) depend the set of discounted XPE payoffs that is endogenous. To overcome this difficulty we will identify another collection of sets $(S(\theta))_{\theta \in \Theta}$ that is known to contain $(S(\theta))_{\theta \in \Theta}$, but, still potentially exclude some action profiles (from the set of all action profiles). Thus, the quasi-minmax can now be written as: for each $\theta \in \Theta$,

\[
\bar{r}(\theta) := \min_{a \in S(\theta)} r(a|\theta) \geq r_S(\theta) = \min_{a \in S(\theta)} r(a|\theta).
\]  

(19)

Then, by Lemma 5, for every justifiable $\alpha$,

\[
U^\alpha(e) \geq \bar{U}(e) := (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \bar{r}(\theta^t) \quad \forall e \in \Theta^{\omega},
\]

(20)

which is a tighter bound than $U$. 

33
To formally identify the collection of sets \((\bar{S}(\theta))_{\theta \in \Theta}\) mentioned above, we need one more auxiliary object, that is a function \(\pi : \mathbb{R}_+^n \rightarrow \mathbb{R}\) defined by

$$\pi(\lambda) := \inf_{e \in \Theta} \sup_{v \in \mathcal{E}(\delta)} \lambda \cdot \left( \psi(e) - (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r_{\lambda}(\theta^t) \right),$$  \hspace{1cm} (21)

The function is well-defined when \(\mathcal{E}(\delta)\) is non-empty, which is assumed throughout this section.

To motivate this object, note that in the uncertain repeated game normalized by \(r\), every discounted XPE payoff is necessarily non-negative. It follows that the sets of enforceable action profiles can be bounded from above as follows. Start by rewriting Eq. (18) in terms of payoffs normalized by \(v\), as a function of \(\pi\) outlines an iterative procedure to find a function using the auxiliary function \(Q\) where the operator \(\mathcal{Q}\) inductively as follows:

$$Q(\pi) = \{ a \in A(\theta) \mid \exists (v^i)_{i=1}^n \subseteq \mathcal{E}(\delta) \text{ s.t. } \frac{1 - \delta}{\delta} d(a|\theta) \leq \inf_{e \in \Theta} \{ \psi(e) - v^i(e) \} \} \subseteq \{ a \in A(\theta) \mid \exists v \in \mathcal{E}(\delta) \text{ s.t. } \frac{1 - \delta}{\delta} d(a|\theta) \leq \inf_{e \in \Theta} \psi(e) \} \subseteq \{ a \in A(\theta) \mid \frac{1 - \delta}{\delta} \cdot d(a|\theta) \leq \sup_{v \in \mathcal{E}(\delta)} \inf_{e \in \Theta} \lambda \cdot \psi(e) \ \forall \lambda \in \mathbb{R}_+^n \text{ s.t. } \|\lambda\| = 1 \} \subseteq \{ a \in A(\theta) \mid \frac{1 - \delta}{\delta} \cdot d(a|\theta) \leq \pi(\lambda) \ \forall \lambda \in \mathbb{R}_+^n \text{ s.t. } \|\lambda\| = 1 \} =: \bar{S}(\pi).$$  \hspace{1cm} (22)

The second line in Eq. (22) is an upper bound because it assumes that punishments \((v^i)_{i=1}^n\) to be \(v^i \equiv 0\) even though such points may not be attainable in an XPE and hence not available in \(\mathcal{E}(\delta)\). The third line identifies the largest score in a (non-negative) direction \(\lambda\) that is attainable in an XPE in the normalized uncertain repeated game irrespective of the future stage games. Finally, the last line interchanges the order of infimum and supremum thereby relaxing the inequality, and substituting \(\inf_{e \in \Theta} \sup_{v \in \mathcal{E}(\delta)} \lambda \cdot \psi(e) \equiv \pi(\lambda)\). In what follows, it will be convenient to treat the set of actions satisfying the inequality in the last line of Eq. (22) as a function of \(\pi\), which we denote by \(S(\cdot|\theta)\).

As can be seen from Eq. (22), the sets of enforceable action profiles can be bounded from above using the auxiliary function \(\pi\). Proposition 2 that follows makes this idea formal. Specifically, it outlines an iterative procedure to find a function \(\bar{\pi} : \mathbb{R}_+^n \rightarrow \mathbb{R}\) that majorizes \(\pi\) defined in Eq. (21). Then, it uses this function and Eq. (22) to construct the collection of sets \((\bar{S}(\theta))_{\theta \in \Theta}\), replacing \(\pi\) by \(\bar{\pi}\) in the third line. Finally, it sets \(\bar{U}\) using Eq. (19) and (20).

**Proposition 2.** For each \(\lambda \in \mathbb{R}_+^n\), set \(\pi^0(\lambda) := \min_{\theta \in \Theta} \max_{a \in A(\theta)} \lambda \cdot (u(a|\theta) - r(\theta))\) and define \(\pi^{k+1} = \mathcal{Q}(\pi^k)\) inductively as follows:

$$\pi^{k+1} := \mathcal{Q}(\pi^k),$$

where the operator \(\mathcal{Q}\) maps the set of linearly homogeneous, bounded functions on \(\mathbb{R}_+^n\) to itself and is defined\(^{29}\)

\(^{29}\)In fact, this function is linearly homogeneous, thus only its values on the boundary of the unit ball, i.e., \(\|\lambda\| = 1\), are relevant. In what follows we identify each linearly homogeneous function \(q : \mathbb{R}_+^n \rightarrow \mathbb{R}\) with its values on the unit ball, and say that such \(q\) is bounded if \(\sup_{\lambda \in \mathbb{R}_+^n, \|\lambda\| = 1} |q(\lambda)| < \infty\).
by

\[ \mathcal{D}(q)(\lambda) := \min_{\theta \in \Theta} \inf_{\mu \in \mathbb{R}^n_+} \max_{a \in S(q|\theta)} (1 - \delta) \left( \lambda \cdot \left( u(a|\theta) - \sum_{\tilde{a} \in S(q|\theta)} r(\tilde{a}|\theta) \right) - \mu \cdot d(a|\theta) \right) + \delta q(\lambda + \mu). \]  

(23)

If the set of XPE is non-empty, then the sequence \((\pi^k)\) monotonically converges (pointwise) to some linearly homogeneous function \(\overline{\pi}\). Furthermore:

- For each stage game \(\theta \in \Theta\),
  \[ S(\theta) \subseteq \overline{S}(\theta) := S(\overline{\pi}|\theta); \]

- \(\overline{U}(e)\) defined using \((\overline{S}(\theta))_{\theta \in \Theta}\) by Eq. (19) and (20) is a lower bound on the set of discounted XPE payoffs.

Proposition 2 describe an explicit procedure to obtain the tighter bound \(\overline{U}\). In the proof we consider projections of \(E(\delta)\) onto individual environments and use a duality argument to show that \(\pi\) satisfies

\[ \pi \leq \mathcal{D}(\pi). \]

Moreover, we establish that the operator \(\mathcal{D}\) is monotone and \(\mathcal{D}(\pi^0) \leq \pi^0\). As a result, successive applications of \(\mathcal{D}\) yield a monotone sequence of functions that converges to the largest fixed point \(\overline{\pi}\) of \(\mathcal{D}\) amongst the set of functions majorized by \(\pi^0\). It is easy to see that \(\pi\) defined in Eq. (21) is necessarily majorized by \(\pi^0\), that is, \(\pi\) is not larger than \(\overline{\pi}\). The construction of the collection of sets \((\overline{S}(\theta))_{\theta \in \Theta}\) and the bound \(\overline{U}(e)\) then follows.

Proposition 2 assumes that the set of XPE is non-empty, for which a sufficient condition is simply that each stage game admits a pure strategy Nash equilibrium. In that case the iterations in Proposition 2 are particularly easy to implement numerically. We show in the appendix that every element of \((\overline{\pi}^k)_{k=0}^{\infty}\) is necessarily continuous when static Nash equilibria exist. Since each function in the sequence \((\overline{\pi}^k)_{k=0}^{\infty}\) is also linearly homogeneous, the whole sequence can be approximated on a grid over the unit ball in \(\mathbb{R}^n_+\). We present the operator \(\mathcal{D}\) in a form suitable for numerical computations. The applicability of this proposition is illustrated in the numerical example that follows.

6.2 Measure of tightness

We now have all the ingredients to assess tightness of punishments with common gaps developed in the previous section. Recollect that for each fixed normalizing vector \(p\), if the limit set of common gaps \(G(p)\), which is identified in Proposition 1, is non-empty, then for every player \(i\) there is an XPE in which this player receives the discounted payoff of

\[ \int \Gamma(\theta^t|p) + (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} p_i(\theta^t). \]  

(24)
In fact, this is the “harshest” punishment for player $i$ in the set $\mathcal{W}(p)$. By Propositions 1 and 2, the discounted payoff in Eq. (24) is not lower than one from our bound $\bar{U}_i$. This suggests to use the difference between the discounted payoff in Eq. (24) and $\bar{U}_i$ to measure tightness of our construction. Formally, let $\eta(p)$ be the maximal such difference across environments:

$$\eta(p) := (1 - \delta) \cdot \min_{\theta \in \Theta} \left( p(\theta) + \frac{\Lambda(\theta|p)}{1 - \delta} - \bar{\tau}(\theta) \right) + \delta \cdot \max_{\theta \in \Theta} \left( p(\theta) - \bar{\tau}(\theta) \right).$$

(25)

Extend $\eta$ to all possible $p \in \prod_{\theta \in \Theta} \{u(a|\theta)\mid a \in A(\theta)\}$ by setting $\eta(p) = (\infty, \ldots, \infty)$ whenever $G(p) = \emptyset$.

By construction, $\eta(p)$ is non-negative for all $p$. If $\eta_i(p) = 0$ for some normalization $p$, then there is some XPE, which can be sustained by recursing over common gaps, that attains the lower bound $\bar{U}_i$ irrespective of the environment. In other words, we have successfully identified the “harshest” punishment for player $i$ amongst all possible XPE. In the next section we will show this is always the case for strongly symmetric XPE, that is there is a certain $p$ such that the measure $\eta(p)$ defined in Eq. (25) is identically zero. If we allow for asymmetric XPE, then $\eta(p)$ isn’t necessarily zero; however, it turns out to be the case in the Bertrand model introduced in Section 5.3.

As for the Cournot model, the measure of tightness evaluated at the endogenous normalization is zero for small and large values of $\delta$, see Figure 6. The former confirms that only static Nash action profiles can be played in any XPE when the players are impatient. The latter illustrates the workings of the folk theorem result, i.e., the discounted minmax payoffs are attainable in an XPE for large $\delta$. For most intermediate values of $\delta$, the value of $\eta_i(p)$ is zero (or almost zero), which means that the outer bound $\bar{U}_i$ is tight. Occasionally, the measure jumps in value: As we increase the discount factor, some action profile may become enforceable in the outer bound but not in the common gap recursion, so the difference temporarily increases but drops back to zero again fairly quickly as discount factor increases a bit more for this action profile now becomes enforceable in the common gap recursion too. Thus, overall, we end up having a reasonably complete characterization of justifiable outcomes even when the discounted minmax payoffs cannot be attained in an XPE.

![Figure 6: Measure of tightness $\eta_i(p)$ in the Cournot model](https://ssrn.com/abstract=4222516)
6.3 Strongly symmetric XPE

A frequently studied special case of repeated interactions is when a stage game is symmetric, i.e., the players have the same action sets \( A_1 = \ldots = A_n = B \), and there exists a function \( \phi : B \times B^n \to \mathbb{R} \) such that \( u_i(a) = \phi(a, a) \) for every \( a \in A \) and for all \( i = 1, \ldots, n \).

In symmetric games it is natural, at a first pass, to focus on situations in which the players play an identical action (see, for example, Kandori [1991], Athey, Bagwell, and Sanchirico [2004], and Sannikov and Skrzypacz [2007]). For every common action \( b \in B \), define the representative player’s payoff by \( \hat{u}(b) := \phi(b, (b, \ldots, b)) \), the best deviation payoff by \( \hat{v}(b) := \max_{b' \in B} \phi(b', (b', b, \ldots, b)) \) and her net gains by \( \hat{d}(b) := \hat{v}(b) - \hat{u}(b) \). Then, we say that a profile of strategies \( \sigma \) is strongly symmetric XPE (SSXPE) if \( \sigma \) is XPE, and the players play an identical action after every history.

Clearly, characterizing SSXPE is equivalent to XPE in a single-player game; thus, the methods and results we developed above can be directly applied. In what follows we use the same notations as above treating the uncertain repeated game as a strategic situation with one representative player. The main insight of this section is tightness of our characterization of justifiability in the sense that the lower bound \( \overline{U} \) is exact, formally stated below.

**Corollary 3.** Suppose the set of SSXPE is non-empty. Then, there exists a number \( p \in \prod_{\theta \in \Theta} \min_{b \in B(\theta)} u(b|\theta) \),

\[
\max_{b \in B(\theta)} u(b|\theta) \left| \right. \theta(p) = 0. \text{ In other words, the worst discounted payoff in } W(p), \text{ which is defined in Proposition 1, coincides with the lower bound constructed in Proposition 2.}
\]

Corollary 3 asserts that the “harshest” punishments in SSXPE can be constructed by recurring over common gaps; moreover, it gives the exact characterization of such punishments through \( \overline{U} \). The proof of this corollary is rather simple yet instructive, we therefore provide it here.

**Proof of Corollary 3.** We first take a closer look at the function \( \pi \) identified in Proposition 2 and then use to construct the normalization mentioned in the corollary.

Note that any linearly homogeneous function \( q : \mathbb{R}_+ \to \mathbb{R} \) in 1D is necessarily of the form \( q(\lambda) = \lambda \cdot q(1) \). It follows that each set \( S(q|\theta) \) is completely determined by \( q(1) \):

\[
S(q|\theta) = \left\{ b \in B(\theta) \mid \frac{1-\delta}{\delta} \hat{d}(b|\theta) \leq q(1) \right\}.
\]

The same is true for the operator \( \mathcal{D}(q) \), and it can be simplified as

\[
\mathcal{D}(q)(1) = \min_{\theta \in \Theta} \inf_{\mu \in \mathbb{R}_+} \max_{b \in S(q|\theta)} (1-\delta) \left( \hat{u}(b|\theta) - \min_{b' \in S(q|\theta)} \hat{v}(b|\theta) - \mu \cdot \hat{d}(b|\theta) \right) + \delta q(1 + \mu) = (26)
\]

\[
= \min_{\theta \in \Theta} (1-\delta) \left( \max_{b \in S(q|\theta)} \hat{u}(b|\theta) - \min_{b \in S(q|\theta)} \hat{v}(b|\theta) \right) + \delta q(1), \quad (27)
\]

where we unpacked \( q(1 + \mu) = (1 + \mu) \cdot q(1) \) and optimally set \( \mu = 0 \). This value of \( \mu \) is optimal because for each \( b \in S(q|\theta) \), we have \( \frac{1-\delta}{\delta} \mu \cdot \hat{d}(b|\theta) \leq \mu \cdot q(1) \).
By definition, \( \pi(1) \) is the largest fixed point of Eq. (26), and \( \bar{\tau}(\theta) = \min_{b \in S(\pi|\theta)} \bar{r}(b|\theta) \) for every \( \theta \in \Theta \). We now show that the set \([0, \pi(1)]\) is the set of common gaps for the normalization \( \bar{\tau} \) as defined in Proposition 1. Indeed, for each stage game, the smallest element of \( T_{\bar{\tau}}[0, \pi(1)] \) is

\[
\min_{(b, g) \in B(\theta) \times [0, \pi(1)]} (1 - \delta)(\bar{u}(b|\theta) - \bar{\tau}(\theta)) + \delta g = \min_{b \in S(\bar{\tau}|\theta)} (1 - \delta)(\bar{r}(b|\theta) - \bar{\tau}(\theta)) = 0. \tag{28}
\]

where we used the fact that \( b_d(\cdot|\theta) \geq 0 \) and \( 0 \in G \). Similarly, the largest element is

\[
\max_{(b, g) \in B(\theta) \times [0, \pi(1)]} (1 - \delta)(\bar{u}(b|\theta) - \bar{\tau}(\theta)) + \delta g = \max_{b \in S(\bar{\tau}|\theta)} (1 - \delta)(\bar{u}(b|\theta) - \bar{\tau}(\theta)) + \delta \pi(1), \tag{29}
\]

where we set \( g = \pi(1) \) because it relaxes the constraint and makes the objective larger at the same time.

Combine Eq. (28) and (29) to conclude that \( G(\bar{r}) = [0, \pi(1)] \) for the normalization defined above. Moreover, by Proposition 1, the worst discounted payoff that is attainable in this case is exactly the same as \( U(e) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \pi(\theta^t) \), which concludes the proof. \( \square \)

The characterization of the worst SSXPE through common gap is analogous to main results in Carroll [2021], which looks at XPE in uncertain repeated games with one long run player—the two problems are mathematically isomorphic. The approach of Carroll [2021] relies on the fact that for each fixed environment, discounted payoffs are one-dimensional, which simplifies the problem.

First ask: What is the largest dynamic loss that can be attained in a SSXPE irrespective of the environment, i.e., the term on the right hand side in Eq. (4)? Given such gap \( g \), compute the “best” SSXPE by playing the same symmetric action profile in every period, see Eq. (29). There is also the “worst” SSXPE in which randomization is used to ensure the players obtain their quasi-minmax payoffs in every period, see Eq. (28). Finally, close the loop and find the largest common gap \( g \) that can be obtained as a difference between the “best” SSXPE and the “worst” SSXPE.

This exact construction isn’t possible with several long run players when general asymmetric XPE are considered. The “best”/“worst” XPE, i.e., ones that are always preferred or least preferred by all players irrespective of \( e \), because different players have different preferences over action profiles. Yet, as shown in this paper, strong incentives can be provided by recursing over common gaps even when there are several long-run players and the “best”/“worst” XPE aren’t available.

7 Final remarks

This paper studies a model of repeated games with multiple long-run players where one of multiple possible stage games can realize in any period, and the players (or analyst) are maximally uncertain about how future stage games are drawn. The framework is termed an uncertain repeated game, and
it provides a non-Bayesian approach to thinking about dynamic incentives in strategic frameworks. A complete characterization of equilibrium outcomes is presented for patient players and strongly symmetric equilibria, and through an outer bound, the characterization is shown to be tight for asymmetric equilibria and fixed discounting.

Expanding the scope of the analysis to different monitoring structures, incorporating endogenous transitions between stage games, and putting more structure on information and behavior of players, all seem like promising next steps. Understanding the implications of uncertainty in other applied settings such as relational contracting or doing inference in models of dynamic incentives without making strong assumptions on the information structure also seem like fruitful areas for future work. There is much more here to build and explore.

8 Appendix

8.1 Appendix A

Proof of Lemma 4. Take \( p \in \prod_{\theta \in \Theta} \text{co} \{ u(a|\theta) | a \in A(\theta) \} \) and the family \( (\Gamma(\theta))_{\theta \in \Theta} \) as described in the lemma. We shall show that the condition in Eq. (10) is sufficient for the set \( W \) defined in Eq. (9) to be self-generating with respect to the APS recursion, that is

\[
W \subseteq \prod_{\theta \in \Theta} B_{\theta} W.
\]

The result then will follow from Lemma 1.

By definition, for each \( \theta \in \Theta \),

\[
B_{\theta} W = \text{co} \left\{ w \in \mathbb{R}^{n-\Theta} | \exists a \in A(\theta), \exists (\gamma^i)_{i=0}^n \subseteq \prod_{\tilde{\theta} \in \Theta} \Gamma(\tilde{\theta}) \text{ s.t.} \right. \\
\left. (PK) \ w(\tilde{\theta}, \cdot) \equiv (1 - \delta) (u(a|\theta) - p(\theta)) + \delta \gamma^0(\tilde{\theta}) \ \forall \tilde{\theta} \in \Theta, \right. \\
\left. (IC) \ \frac{1 - \delta}{\delta} d_i(a|\theta) \leq \min_{\tilde{\theta} \in \Theta} \{ \gamma^0_i(\tilde{\theta}) - \gamma^i(\tilde{\theta}) \} \ \forall i \in N \right\} \supseteq \\
\supseteq \text{co} \left\{ w \in \mathbb{R}^{n-\Theta} | \exists a \in A(\theta), \exists (\gamma^i)_{i=0}^n \subseteq \prod_{\tilde{\theta} \in \Theta} \exists (g, \gamma) \in \mathbb{R}^n \times \mathbb{R}^n \text{ s.t.} (PK), (IC), \\
(WP) \ \gamma^i(\tilde{\theta}) = \gamma^i, \ \forall \tilde{\theta} \in \Theta, \ \forall i \in N \\
(CG) \ \gamma^0_i(\tilde{\theta}) - \gamma^i(\tilde{\theta}) = g_i, \ \forall \tilde{\theta} \in \Theta, \ \forall i \in N \right\},
\]

where the inclusion is due to the fact that the second operator imposes two additional restrictions. First, each player \( i \) must receive the same discounted payoff \( \gamma_i \) irrespective of \( \tilde{\theta} \) upon any deviation \( WP \). Second, the players are promised some common gap \( g \) over their punishment discounted payoffs \( CG \).

Since the on-path and off-path continuations must belong to the collection \( (\Gamma(\theta))_{\theta \in \Theta} \), the pair
\((\vartheta, \gamma)\) cannot be chosen freely. It is easy to see that each \(\gamma\) is an element of \([\bigvee_{\vartheta \in \Theta} \Gamma(\vartheta), \bigwedge_{\vartheta \in \Theta} \Gamma(\vartheta)]\), which is a non-empty interval, because of \(\bigcap_{\vartheta \in \Theta} \Gamma(\vartheta) \neq \emptyset\). By \((CG)\), for some fixed \(\gamma\), the gap \(\vartheta\) must belong to \(\bigcap_{\vartheta \in \Theta} \Gamma(\vartheta) - \gamma\), and it yields

\[
w = \left(1 - \delta\right)\left(u(a|\theta) - p(\theta)\right) + \delta \gamma.
\]

Clearly, choosing a smaller value of \(\gamma\) relaxes \((IC)\) but doesn’t affect the space of attainable \(w\)’s, setting \(\gamma = \bigvee_{\vartheta \in \Theta} \Gamma(\vartheta)\) is optimal. Thus, for each \(\theta \in \Theta\),

\[
\mathcal{B}_p^\theta W \supseteq \left\{ w \in \mathbb{R}^n: \exists y \in T^\theta_p \left(\bigvee_{\vartheta \in \Theta} \Gamma(\vartheta) - \bigwedge_{\vartheta \in \Theta} \Gamma(\vartheta)\right) + \delta \bigvee_{\vartheta \in \Theta} \Gamma(\vartheta) \text{ s.t. } w \equiv \gamma \right\} \supseteq W,
\]

where the second inclusion is due to the condition in Eq. (10). Conclude that \(W\) is self-generating with respect to the APS operator in the uncertain repeated game normalized by \(p\). As a result, every element of \(W\) can be attained in an XPE.

8.2 Appendix B

Fix \(p \in \prod_{\theta \in \Theta} \text{co}\{u(a|\theta) | a \in A(\theta)\}\). In what follows we will use a convention that \(\emptyset + x = \emptyset\) for every \(x \in \mathbb{R}^n\).

To formally prove the proposition, we will use the auxiliary operator \((G, x) \mapsto (\mathcal{I}(G, x), \mathcal{X}(G, x))\), where \(G\) is a bounded subset of \(\mathbb{R}^n\) and \(x \in \mathbb{R}^n\), defined by

\[
\mathcal{I}(G, x) := \bigcap_{\vartheta \in \Theta} T^\vartheta_p G + \delta x, \quad \mathcal{X}(G, x) := \begin{cases} \bigvee_{\vartheta \in \Theta} T^\vartheta_p G + \delta x & \text{if } \bigcap_{\vartheta \in \Theta} T^\vartheta_p G \neq \emptyset, \\ \bigvee_{\vartheta \in \Theta} \bigvee_{a(\theta) \in A(\theta)} (u(a|\theta) - p(\theta)) & \text{if } \bigcap_{\vartheta \in \Theta} T^\vartheta_p G \neq \emptyset. \end{cases}
\]

Define also the partial order \(\succeq\) over pairs \((G, x)\) as follows:

\[(G, x) \succeq (G', x') \text{ if } G + x \supseteq G' + x' \text{ and } x \leq x'.\]

The following two lemmas establish important properties of the auxiliary operator that will be used later in the proof of Proposition 1.

**Lemma 6.** The operator \((\mathcal{I}, \mathcal{X})\) is monotone in \(\succeq\), that is

\[
(\mathcal{I}(G, x), \mathcal{X}(G, x)) \succeq (\mathcal{I}(G', x'), \mathcal{X}(G', x')) \text{ if } (G, x) \succeq (G', x').
\]

**Proof.** The claim is trivially true when \(\mathcal{I}(G', x') = \emptyset\), so assume the opposite. Observe that for each \(\theta \in \Theta\),

\[
T^\vartheta_p G \supseteq T^\vartheta_p (G + x - x) \supseteq T^\vartheta_p (G) + \delta (x - x),
\]

Electronic copy available at: https://ssrn.com/abstract=4222516
where the first inclusion is due to monotonicity of $T^p_\theta$ (in the set inclusion sense) and the second is due to $x - \bar{x} \geq 0$. Taking “∩” and “∪” “∧” in the above equation, we obtain

$$\bigcap_{\theta \in \Theta} T^p_\theta G \supseteq \bigcap_{\theta \in \Theta} T^p_\theta G + \delta(\bar{x} - x) \quad \text{and} \quad \bigvee_{\theta \in \Theta} T^p_\theta G \subseteq \bigvee_{\theta \in \Theta} T^p_\theta G + \delta(\bar{x} - x).$$

Rearrange to see the claim. \(\square\)

**Lemma 7.** Set $x^0 := \bigvee_{\theta \in \Theta} \bigwedge_{a \in A(\theta)} \{u(a|\theta) - p(\theta)\}$ and $G^0 := \bigcap_{\theta \in \Theta} co\{u(a|\theta) - p(\theta) | a \in A(\theta)\} - x^0$. Then, 

$$(G^0, x^0) \geq (\mathcal{G}(G^0, x^0), \mathcal{X}(G^0, x^0)) \text{ and } (G^0, x^0) \geq (G, x) \text{ if } (\mathcal{G}(G, x), \mathcal{X}(G, x)) \geq (G, x).$$

**Proof.** Take $g = \frac{1 - \delta}{\delta} d(a|\theta)$ in the definition of $T^p_\theta \mathbb{R}^n$ to obtain 

$$\bigwedge_{\theta \in \Theta} T^p_\theta G^0 \geq \bigwedge_{\theta \in \Theta} T^p_\theta \mathbb{R}^n = (1 - \delta)(x(\theta) - p(\theta)) \quad \forall \theta \in \Theta,$$

which implies 

$$\mathcal{X}(G^0, x^0) = \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} T^p_\theta G^0 + \delta x^0 \geq x^0 = \bigvee_{\theta \in \Theta} \bigwedge_{a \in A(\theta)} \{u(a|\theta) - p(\theta)\}.$$

The first relationship in the statement of the lemma then follows from Eq. (30) and the following observations:

$$\mathcal{G}(G^0, x^0) \cup \mathcal{X}(G^0, x^0) = \bigcap_{\theta \in \Theta} T^p_\theta G^0 + \delta x^0 \subseteq \bigcap_{\theta \in \Theta} T^p_\theta (G^0 + x^0) \subseteq G^0 + x^0.$$

The first inclusion is due to $x^0 \leq 0$, thus $T^p_\theta (G^0) + \delta x^0 \subseteq T^p_\theta (G^0 + x^0)$ for all $\theta \in \Theta$. The second follows from the definition of the operator $T^p_\theta$ and the fact that $G^0$ is bounded.

We now show the second relationship in the statement of the lemma. By assumption about the pair $(G, x)$, 

$$\mathcal{X}(G, x) = \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} T^p_\theta G + \delta x \leq x.$$

The bound in Eq. (30) also applies to $G$, thus $x \geq x^0$. Finally, note that for each $\theta \in \Theta$, 

$$G + x \subseteq T^p_\theta G + \delta x \subseteq \text{co}\{(1 - \delta)(u(a|\theta) - p(\theta)) + \delta(G + x) | a \in A(\theta)\},$$

which establishes $G + x \subseteq G^0 + x^0$ and concludes the proof. \(\square\)

We now have all the ingredients to complete the proof of Proposition 1.

**Proof of Proposition 1.** Define $(G^k, x^k)_{k=1}^\infty$ by repeatedly applying the auxiliary operator $(\mathcal{G}, \mathcal{X})$ to $(G^0, x^0)$, which is defined in Lemma 7, and note that the sequence $(G^k)_{k=1}^\infty$ is exactly as described in
the statement of the proposition.

Lemma 6 and the first part of Lemma 7 jointly imply that \((G^k, x^k)\) is non-increasing in \(\succeq\). Since \((x^k)\) is non-decreasing in \(\succeq\) and bounded, it is convergent. Let \(x(p)\) be its limit point, and define \(G(p) := \bigcap_{k=1}^{\infty} (G^k + x^k - x(p))\), which might be empty. Clearly, \((G^k)\) converges to \(G(p)\) in the Hausdorff distance \(d_H\) with our convention \(d_H(\emptyset, \emptyset) = \emptyset\).

We now show that \((G(p), x(p))\) is the largest fixed point of the auxiliary operator \((\mathcal{G}, \mathcal{X})\). First, the fact that it is a fixed point holds trivially when \(G(p) = \emptyset\). Otherwise, it is due to the Knaster-Tarski (or Bourbaki-Witt) theorem as \(\{(G^k, x^k)\}_{k=0}^{\infty} \cup \{(G(p), x(p))\}\) is a complete chain. The second part of Lemma 7 implies that any fixed point is lower than \((G^0, x^0)\), thus \((G(p), x(p))\) is the largest fixed point.

Note that \(x(p) = \frac{\bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} G(p)}{1-\delta}\), and define \((\Gamma(\theta|p))_{\theta \in \Theta}\) as in the statement of the proposition, that is
\[
\Gamma(\theta|p) := T^p_\theta G(p) + \delta x(p).
\]
By construction, since \((G(p), x(p))\) is a fixed point of the auxiliary operator,
\[
G(p) = \bigcap_{\theta \in \Theta} \Gamma(p|\theta) - \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(p|\theta) \quad \text{and} \quad x(p) = \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(p|\theta).
\]
Let \((\Gamma(\theta))_{\theta \in \Theta}\) be a family of sets satisfying the conditions of Proposition 1, and define
\[
G := \bigcap_{\theta \in \Theta} \Gamma(\theta) - \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(\theta) \quad \text{and} \quad x = \bigvee_{\theta \in \Theta} \bigwedge_{\theta \in \Theta} \Gamma(\theta).
\]
By Eq. (10),
\[
(\mathcal{G}(G, x), \mathcal{X}(G, x)) \succeq (G, x).
\]
As a result, \((G(p), x(p))\) isn’t lower than \((G, x)\) in \(\succeq\), thanks to Lemma 7. Finally, by Lemma 6 and Eq. (10), for every \(\theta \in \Theta\),
\[
\Gamma(\theta|p) = T^p_\theta G(p) + \delta x(p) \supseteq T^p_\theta G + \delta x \subseteq \Gamma(\theta),
\]
which completes the proof.

\[\square\]

8.3 Appendix C

Suppose that the set of XPE is non-empty so that \(\pi\) in Eq. (21) is well-defined. The following lemma makes several important observations, which will be used later in the proof of the proposition, about this function.

**Lemma 8.** Let \(\pi\) be defined as in Eq. (21). Then, \(\pi \succeq 0\), non-decreasing and linearly homogeneous.
Furthermore, it satisfies $\pi \leq \mathcal{D}(\pi)$, where $\mathcal{D}$ is given by Eq. (23), and

$$\pi \leq \pi^0 := \min_{\theta \in \Theta} \max_{a \in A(\theta)} \lambda \cdot (u(a|\theta) - r(\theta)).$$

Proof. Monotonicity and linear homogeneity follow directly from the definition of $\pi$. As for non-negativity, since every element of $E^r(\delta)$ is non-negative, $\pi$ is non-negative.

We now show that the auxiliary function satisfies $\pi \leq \mathcal{D}(\pi)$. By Lemma 1, $E^r(\delta)$ satisfies the APS recursion in the uncertain repeated game normalized by $r_S$. Since every element of $E^r(\delta)$ is non-negative, and, as shown by Lemma 5, only action profiles in $(S(\theta))_{\theta \in \Theta}$ can ever arise in an XPE,

$$E^r(\delta) \subseteq \prod_{\theta \in \Theta} \co \left\{ w \in \mathbb{R}^{n \times \Theta^w} | \exists a \in S(\theta), \exists v \in E^r(\delta) \text{ s.t.} \right\}$$

$$w = (1 - \delta)(u(a|\theta) - r_S(\theta)) + \delta v, \frac{1-\delta}{\delta} d(a|\theta) \leq \bigwedge_{e \in \Theta^w} v(e). \quad (31)$$

The set of discounted XPE payoffs $E^r(\delta)$ might have a complex geometry. We now relax it be looking at its projections onto individual environments. Specifically, for each $e \in \Theta^w$, denote by $E^r_e(\delta)$ the set of discounted XPE payoffs in this environment. Clearly, $E^r(\delta)$ belongs to the hypercube $\prod_{e \in \Theta^w} E^r_e(\delta)$. Thus, the object on the right hand side of Eq. (31) can be made larger by replacing $E^r(\delta)$ with the corresponding hypercube:

$$E^r(\delta) \subseteq \prod_{\theta \in \Theta} \co \left\{ w \in \mathbb{R}^{n \times \Theta^w} | \exists a \in S(\theta), \exists v \in \prod_{e \in \Theta^w} E^r_e(\delta) \text{ s.t.} \right\}$$

$$w = (1 - \delta)(u(a|\theta) - r_S(\theta)) + \delta v, \frac{1-\delta}{\delta} d(a|\theta) \leq \bigwedge_{e \in \Theta^w} v(e). \quad (32)$$

Finally, we allow sunspots to depend on the future stage games disconnecting the recursion across environments. Mechanically, we drop $\bigwedge_{e \in \Theta^w}$ in Eq. (32) thereby relaxing its constraint:

$$E^r(\delta) \subseteq \prod_{(\theta,e) \in \Theta \times \Theta^w} \co \left\{ w \in \mathbb{R}^{n} | \exists a \in S(\theta), \exists v \in E^r_e(\delta) \text{ s.t.} \right\}$$

$$w = (1 - \delta)(u(a|\theta) - r_S(\theta)) + \delta v, \frac{1-\delta}{\delta} d(a|\theta) \leq v. \quad (33)$$

Note that for each fixed environment, the recursion on the right-hand side of Eq. (33) is now finite-dimensional but the result of this recursion are aggregated across environments as a hypercube.

We now have the necessary ingredients to establish $\pi \leq \mathcal{D}(\pi)$. The idea is to apply certain duality arguments to the bound in Eq. (33). Without further ado, observe that for each $\lambda \in \mathbb{R}_g^x$, the function
\[ \pi(\lambda) = \min_{\theta \in \Theta} \inf_{a \in \Theta^w} \sup_{\varepsilon \in \mathbb{E}^q(\delta)} \lambda \cdot v(\varepsilon) \leq 0 \]  
\leq \min_{\theta \in \Theta} \inf_{a \in \Theta^w} \inf_{(a, \varepsilon) \in \mathbb{S}(\theta) \times \mathbb{E}^q(\delta)} \sup_{\mu \in \mathbb{R}_+^*} (1 - \delta)(\lambda \cdot u(a|\theta) - \lambda \cdot r_\delta(\theta) - \mu \cdot d(a|\theta)) + \delta(\lambda + \mu) \cdot v \leq 
= \min_{\theta \in \Theta} \max_{a \in \mathbb{S}(\pi|\theta)} \left(1 - \delta\right) \left(\lambda \cdot u(a|\theta) - \lambda \cdot r_\delta(\theta) - \mu \cdot d(a|\theta)\right) + \delta \pi(\lambda + \mu) . \]

The first line in Eq. (34) is just a definition of \( \pi \). The second line is due to the relationship in Eq. (33) where \( \delta \mu \) is the Lagrange multiplier attached to the inequality constraint on the right hand side of this equation. The third line uses weak duality to swap \( \inf_{\mu \in \mathbb{R}_+^*} \) and \( \sup_{(a, \varepsilon) \in \mathbb{S}(\theta) \times \mathbb{E}^q(\delta)} \). Note that at this step \( \inf_{\mu \in \mathbb{R}_+^*} \) and \( \max_{a \in \mathbb{S}(\theta)} \) can be interchanged because their optimal choices are unrelated. The fourth line takes into an account and explicitly substitutes for \( \pi \), which gives us exactly \( \mathcal{D}(\pi)(\lambda) \).

The only part of the lemma left is the upper bound on \( \pi \). Relax \( \mathcal{D}(\pi)(\lambda) \) by setting \( \mu = 0 \) and allowing all action profiles to obtain

\[ \pi(\lambda) \leq \mathcal{D}(\pi)(\lambda) \leq \max_{a \in \mathbb{A}(\theta)} (1 - \delta)(u(a|\theta) - \pi(\theta)) + \delta \pi(\lambda) = (1 - \delta)\pi^0(\lambda) + \delta \pi(\lambda) . \]

which gives \( \pi \leq \pi^0 \).

Lemma 8 suggests to seek for the largest function satisfying \( q \leq \mathcal{D}(q) \) in the set of non-negative, non-decreasing, linearly homogeneous functions \( q \) such that \( \pi \leq q \leq \pi^0 \). Let \( Q \) be the set of such functions. If we found the largest function \( \pi \in Q \), then we would be able to bound \( \pi \) from above by \( \pi \).

The following lemma confirms that \( \mathcal{D} \) is a well-defined operator, establishes its monotonicity and the fact that it preserves continuity. The last point isn’t required for \( \pi \) to be defined by our iterative procedure of Proposition 2; however, it helps in computations of \( \pi \). We include the result on continuity for completeness, and we shall discuss later how to numerically approximate \( \pi \).

**Lemma 9.** \( \mathcal{D} \) is an order-preserving map from \( Q \) to itself. Furthermore, if the set of static Nash equilibria is non-empty in every individual stage game, then for every continuous \( q \in Q \), there exists a number \( m \in \mathbb{R}_+ \) such that for all \( \lambda \in \mathbb{R}_+^* \) with \( \|\lambda\| = 1 \),

\[ \mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \min_{\mu \in \mathbb{R}_+^*} \max_{a \in \mathbb{S}(q|\theta)} \left(1 - \delta\right) \left(\lambda \cdot u(a|\theta) - \lambda \cdot r_\delta(\theta) - \mu \cdot d(a|\theta)\right) + \delta q(\lambda + \mu) . \]

As a result, \( \mathcal{D}(q) \) is continuous when \( q \) is continuous.
Proof. Note that for each $\lambda \in \mathbb{R}_+^n$,

$$
\mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \inf_{\mu \in \mathbb{R}_+^n} \max_{a \in S(q|\theta)} \left( (1 - \delta) \lambda \cdot \left( r(a|\theta) - \int_{\bar{a} \in S(q|\theta)} r(\bar{a}|\theta) \right) + \delta \left( q(\lambda + \mu) - \frac{1 - \delta}{\delta} (\lambda + \mu) \cdot d(a|\theta) \right) \right),
$$

where the term in the brackets is non-negative by the definition of $S(q|\theta)$. It follows that the value on the right hand side of Eq. (35) is bounded from below by a non-negative number, thus the function $\mathcal{D}(q)$ is well-defined.

Clearly, $\mathcal{D}(q)$ is linearly homogeneous. We now establish that it is non-decreasing. To begin note that the first term in Eq. (35) is non-decreasing in $\lambda$, because $r(a|\theta) - \int_{\bar{a} \in S(q|\theta)} r(\bar{a}|\theta) \geq 0$. As for the second term, observe that

$$
\inf_{\mu \in \mathbb{R}_+^n} q(\lambda + \mu) - \frac{1 - \delta}{\delta} (\lambda + \mu) \cdot d(a|\theta) = \inf_{\mu \in \mathbb{R}_+^n : \mu \geq \lambda} q(\mu) - \frac{1 - \delta}{\delta} \mu \cdot d(a|\theta) = \lambda \cdot u(a|\theta) - \int_{\bar{a} \in S(q|\theta)} r(\bar{a}|\theta) - \mu \cdot d(a|\theta) + \delta q(\lambda + \mu).
$$

is non-decreasing in $\lambda$, because the expression on the right becomes more constrained when $\lambda$ gets larger. Conclude that $\mathcal{D}(q)$ is a non-decreasing function.

It is easy to see that $\mathcal{D}$ is order-preserving, i.e., $\mathcal{D}(q) \leq \mathcal{D}(\tilde{q})$ whenever $q \leq \tilde{q}$. Then, monotonicity of $\mathcal{D}$ and Lemma 8 imply

$$
\mathcal{D}(q) \geq \mathcal{D}(\pi) \geq \pi.
$$

Similarly, the reader can verify that $\mathcal{D}(\pi^0) \geq \pi^0$, thus

$$
\mathcal{D}(q) \leq \mathcal{D}(\pi^0) \leq \pi^0,
$$

which verifies that $\mathcal{D}(q)$ is a well-defined element of $Q$.

It remains to establish the second part of the lemma. Suppose that static Nash equilibria exist in all stage games. Let $q \in Q$ be continuous, and note that $\mathcal{D}(q)$ can be re-written as follows:

$$
\mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \inf_{m \in \mathbb{R}_+^n} m \cdot f^1(\lambda|m|\theta),
$$

where for each $\theta \in \Theta$,

$$
f^1(\lambda|\theta) := \min_{\mu \in \mathbb{R}_+^n : \mu = 1} \max_{a \in S(q|\theta)} \left( (1 - \delta) \left( \lambda \cdot \left( u(a|\theta) - \int_{\bar{a} \in S(q|\theta)} r(\bar{a}|\theta) \right) - \mu \cdot d(a|\theta) \right) + \delta q(\lambda + \mu) \right).
$$

In other words, the inner minimization in the definition of $f^1$ is over directions $\mu$, which are required to have a unit norm, whereas the outer optimization is over values $m$ of this norm.
Since \( q \) is continuous, the theorem of the Maximum guarantees that \( f^1 \) is continuous, and

\[
f^1(\theta) = \min_{\mu \in \mathbb{R}^n_{>0} : \|\mu\| = 1} \max_{a \in S(q|\theta)} \delta \left( q(\mu) - \frac{1-\delta}{\delta} \mu \cdot d(a|\theta) \right)
= \min_{\mu \in \mathbb{R}^n_{>0} : \|\mu\| = 1} \delta q(\mu),
\]

where the second term drops out because each static Nash action profile is necessarily in \( S(q|\theta) \). If \( f^1(\theta) > 0 \), then there exists a number \( m^1(\theta) \) such that for every \( m \geq m^1(\theta) \),

\[
m \cdot f^1(\lambda/m|\theta) \geq \lim_{m \to 0} \tilde{m} \cdot f^1(\lambda/\tilde{m}|\theta) = \max_{a \in S(q|\theta)} (1-\delta) \lambda \cdot \left( \mu(a|\theta) - \int_{\tilde{a} \in S(q|\theta)} r(\tilde{a}|\theta) \right) + \delta q(\lambda).
\]

Suppose to the contrary that \( f^1(\theta) = 0 \). Let \( N^1 \) be the set of players such that there exists some \( \mu^i \in \mathbb{R}^n_{>0} \) such that \( \|\mu\| = 1, \mu^i > 0 \) and \( q(\mu^i) = 0 \). By definition of \( S(q|\theta) \), we must have \( d_i(a|\theta) = 0 \) for all \( a \in S(q|\theta) \) and every \( i \in N^1 \). Of course, if \( N^1 = N \), then every action profile in \( S(q|\theta) \) is a static Nash, thus Eq. (36) holds for all \( m > 0 \). So, set \( m^1(\theta) = 1 \).

Otherwise, since the function \( q \) is non-decreasing, any optimal \( \mu \) in the definition of \( \mathcal{D}(q) \) puts \( \mu_i = 0 \) for every \( i \in N^1 \); as a result, \( \mathcal{D}(q) \) can be re-written as follows:

\[
\mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \inf_{m \in \mathbb{R}^n_{>0}} m \cdot \tilde{f}^2(\lambda/m|\theta),
\]

where for each \( \theta \in \Theta \),

\[
f^2(\lambda|\theta) := \min_{\mu \in \mathbb{R}^n_{>0} : \|\mu\| = 1, \mu_i = 0 \forall i \in N^1 \} \max_{a \in S(q|\theta)} \lambda \cdot \left( u(a|\theta) - \int_{\tilde{a} \in S(q|\theta)} r(\tilde{a}|\theta) \right) - \mu \cdot d(a|\theta) + \delta q(\lambda + \mu),
\]

which is identical to \( f^1 \) but puts additional restrictions on \( \mu \). Apply the same argument as above, to either construct \( m^2(\theta) \) so that Eq. (36) holds for all \( m > m^2(\theta) \) or obtain some set \( N^2 \supseteq N^1 \), \( N^2 \neq N^1 \) and the function \( f^3 \) defined similarly to \( f^2 \) but with \( N^2 \) in the place of \( N^1 \). The second part of the lemma follows repeating this procedure (if necessary).

Finally, continuity of \( \mathcal{D}(q) \) directly follows from the theorem of the Maximum because the domains of choice variables are compact sets.

\[\square\]

We are now ready to complete the proof of Proposition 2.

**Proof of Proposition 2.** Construct the sequence \( (\pi^k)_{k=1}^\infty \) by repeatedly applying the operator \( \mathcal{D} \) to \( \pi^0 = \min_{\theta \in \Theta} \max_{a \in A(\theta)} \lambda \cdot (u(a|\theta) - r(\theta)) \).

Since \( \mathcal{D} : Q \to Q \) is an order-preserving function satisfying \( \mathcal{D}(\pi^0) \leq \pi^0 \), the sequence \( (\pi^k) \) is decreasing. This sequence admits an infimum in \( Q \), i.e., \( \pi(\lambda) := \lim_{k \to \infty} \pi^k(\lambda) \) for all \( \lambda \in \mathbb{R}^n_{>0} \), because the space of functions \( Q \) discussed in Lemma 9 is a complete lattice. The Knaster-Tarski theorem

Electronic copy available at: https://ssrn.com/abstract=4222516
ensures that $\bar{\pi}$ is a fixed point, that is $\bar{\pi} = \mathcal{D}(\bar{\pi})$, and that for every $q \in Q$ such that $q \leq \mathcal{D}(q)$ we have $q \leq \bar{\pi}$.

We end this section with a comment on a computation of the bound $\bar{\pi}$. If static Nash equilibria exist in all stage games, $\pi$ is necessarily continuous, thus it can be approximated on a finite grid over the portion of the unit ball that lies in $\mathbb{R}^n_{++}$. Continuity is due to the fact that $\pi^0$ is continuous, and each element of the sequence $(\pi^k)$ is continuous as well by the theorem of the Maximum. Then, the Dini’s theorem implies that $\pi$ as a monotone (pointwise) limit of continuous functions.

Note that the following representation, which is based on Eq. (35):

$$
\mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \inf_{\mu \in \mathbb{R}^n_{++}} \max_{\|\mu\| = 1} \left( (1 - \delta) \lambda \cdot \left( r(a|\theta) - \int_{a \in S(q|\theta)} r(\bar{a}|\theta) \right) + m \cdot \delta \left( q(\mu) - \frac{1 - \delta}{\delta} \mu \cdot d(a|\theta) \right) \right) \geq 0.
$$

Here for a fixed $\mu$ in the unit ball, which pins down a direction of the aggregate Lagrange multiplier, we first select the norm $m$ that minimizes the right-hand side and then take inf over $\mu$.

Since the term in the brackets is non-negative setting $m$ at the lowest value is optimal irrespective of how the maximizing action profile will change in response to it, thus

$$
\mathcal{D}(q)(\lambda) = \min_{\theta \in \Theta} \inf_{\mu \in \mathbb{R}^n_{++}} \max_{\|\mu\| = 1} \left( (1 - \delta) \lambda \cdot \left( r(a|\theta) - \int_{a \in S(q|\theta)} r(\bar{a}|\theta) \right) + m \cdot \delta \left( q(\mu) - \frac{1 - \delta}{\delta} \mu \cdot d(a|\theta) \right) \right) \geq 0.
$$

The representation derived in Eq. (37) is suitable for computations, because both $\lambda$ and $\mu$ here belong to a compact set.

So, fix some $\varepsilon > 0$ and let $\Lambda^\varepsilon$ be some finite grid over $\{ \lambda \in \mathbb{R}^n_{++} \| \lambda \| = 1, \lambda \geq \varepsilon \}$. By the second part of Lemma 9, for each $q \in Q$, the function $\mathcal{D}(q)$ can be well-approximated on the grid of $\Lambda^\varepsilon$ with $\mu \in \Lambda^\varepsilon$ and $\varepsilon > 0$ small enough.

References


Electronic copy available at: https://ssrn.com/abstract=4222516


