Epsilon-Equilibria of Perturbed Games*

Matthew O. Jackson, Tomas Rodriguez-Barraquer, and Xu Tan†

May 2010, This Draft: February, 2011

Abstract

We prove that for any equilibrium of a (Bayesian) game, and any sequence of perturbations of that game, there exists a corresponding sequence of ex-ante ε-equilibria converging to the given equilibrium of the original game. We strengthen the conclusion to show that the approaching equilibria are interim equilibria (ε-best responses for almost all types) if beliefs in the perturbed games converge in a strong-enough sense to the limit beliefs. These results imply that equilibrium selection arguments that are based on perturbations to a game are not robust to slight perturbations in best reply behavior (or to underlying preferences). This applies to many standard equilibrium selections, including Selten's (1975) definition of trembling hand perfect equilibrium, Rubinstein's (1989) analysis of the electronic mail game, and Carlsson and van Damme's (1993) global games analysis, among others.

Keywords: epsilon-equilibrium, epsilon-Nash equilibrium, electronic mail game, global games, Bayesian games, trembling hand perfection, Nash equilibrium, lower hemi-continuity

JEL Classification Codes: C72, D82

1 Introduction

As many games have multiple equilibria, and some may seem more "natural" than others, game theorists have examined a variety of arguments that refine the set of equilibria. A

*We thank Attila Ambrus, Ben Golub, Willemien Kets, Stephen Morris, and Bob Wilson for helpful comments, and Ehud Kalai, and advisory editor, and three anonymous referees for suggestions that improved the content and presentation of the paper.

†All three authors are at the Department of Economics, Stanford University, Stanford, California 94305-6072 USA. Jackson is also an external faculty member at the Santa Fe Institute and a member of CIFAR. Emails: jacksonm@stanford.edu, trodrig@stanford.edu, and xutan@stanford.edu.

primary technique for refinement, as used in the seminal equilibrium refinement of trembling hand perfection, is to consider some sequence of approximations of the game and then only consider equilibria of the limit game that are the limit of a sequence of equilibria in the perturbed games.[1] A motivation for such an approach is to look for equilibria that are "robust" to slight variations in the specification of the game. The point that we make here is that these sorts of results in the literature are *all* dependent on the fact that the games were perturbed slightly in terms of the actions that players can take, the information that the players might have, or the underlying uncertainty about what the types of the players and or state might be, but that the underlying preferences (as functions of actions, types and states) are not perturbed. If one allows a slight perturbation in best reply behavior, then all equilibria of the original game can be approximated by sequences of approaching (approximate) equilibria. As the modeler may not have precise knowledge of preferences, allowing for such perturbations to best replies is often a natural form of robustness. In fact, one can generally interpret ε-best replies as *exact* best replies to some perturbed preferences.[2] Thus, our main result shows that allowing for slight perturbations in preferences along the sequence, along with the perturbations in games, neutralizes any selective power that the sequences have.

Specifically, we prove that for any equilibrium of a continuous (Bayesian) game, and any sequence of perturbations of that game, there exists a corresponding sequence of ex-ante ε-equilibria converging to the given equilibrium of the original game, with the ε converging to zero. Thus, selection arguments that have been made in the literature, such as those in Selten's (1975) definition of trembling hand perfect equilibrium, Rubinstein's (1989) analysis of the electronic mail game, and Carlsson and van Damme's (1993) global games analysis, are not robust to slight perturbations in best reply behavior (or thus to corresponding perturbations in underlying preferences). In a sense, our result can be thought of as a sort of lower hemi-continuity result, where one loosens the exactness of the equilibria that approach the limit point.

We also prove that if the supports of the type distributions of the perturbations approach the support of the limit game and a stronger notion of belief convergence holds, then the

[1]There are different ways to think of trembling hand perfection. An easy one, following Selten's original definition, is to think of the perturbed games being one which restricts strategies to place minimal weight on all actions.

[2]There are a variety of interpretations of ε-equilibria, beyond the simple one of players only approximating their best responses. As just pointed out, one is that the modeler may have incorrectly specified some of the preferences and types of the players. Another interpretation is that players themselves have misinterpreted some of their payoffs, or are somehow not completely informed of their payoffs. Yet another is that the players optimize subject to some additional costs of fully discovering their type or fully tailoring their strategy to their types (see the discussion in Section 3.3). Such equilibria have proven useful in a variety of contexts (e.g., Kalai (2004)).

same result holds with the stronger notion of *interim* equilibrium, so that almost all types are best responding, rather than just a measure approaching one. The difference between our interim and ex-ante results cannot be avoided and sheds some light on aspects of some perturbation arguments. In particular, some perturbation arguments, such as global games arguments, rely on introducing small probabilities of types who are very different from any type in the original game.

The contribution of our work is three-fold. First, although there are some previous papers that provide lower hemi-continuity sorts of results in some specific contexts, they do not make the point that refinements of equilibria are not robust to perturbations of best response behavior. Second, previous works on ε-equilibria are limited to countable settings and the substantial technical advance that we provide is dealing with continuum settings, which allows us to cover many important applications with our main theorem. Third, the relation between our ex-ante and interim results point out how some perturbation arguments depend on the introduction of new types that are "far" from the original game, and to what extent they are robust to perturbations in best replies.

In particular, two prominent previous studies of ε-equilibria that relate to our analysis are Radner (1980) and Fudenberg and Levine (1986). Radner examined finitely repeated oligopoly games and showed that 'cooperative' behavior could be sustained in such games as ε-equilibria of finite horizon games with long enough horizons. If one views those games as approximating the infinite horizon game, then one can view Radner's result as showing that the cooperative equilibrium that is sustainable in the infinitely repeated game can be approximated by a limit of ε-equilibria of a sequence of approaching games. While our results appear superficially to be limited to finite horizon games, the results in Fudenberg and Levine (1986) show how with appropriate topological definitions, one can view finite horizon games as approximating the infinite horizon ones, and so a Radner result is subsumed in our approach. Moreover, Fudenberg and Levine (1986) show that equilibria of the (complete information) infinite horizon repeated games can be approximated by ε-equilibria of the finite horizon games. In doing this they prove that an equilibrium of a complete information game can be approximated by a sequence of ε-equilibria of approaching games where there are successively looser restrictions on the strategy spaces in a carefully defined way. That result is a special case of our main result, as our results also cover other approximations of strategies. However, the even more substantial distinction is that our results apply to a general class of Bayesian games, and allow for perturbations in uncertainty. On a technical level, the treatment of Bayesian games is the substantial complication in proving our main result, and where much of the effort of the proof is directed especially with the added continuum structure.[3] But most importantly, beyond the treatment of a general class of games of

[3]Engl(1995) extends Fudenberg and Levine (1986) from a complete information to an incomplete infor-

incomplete information, the main point and motivation of our paper is much different that those of Radner (1980) and Fudenberg and Levine (1986), as they were motivated to reconcile the seeming discontinuity between finitely repeated and infinitely repeated games. The main motivation of our paper is completely different and is to shed light on the robustness of various refinements of equilibria that are based on perturbations of a game. That requires proving theorems in Bayesian settings as some of the most prominent of such refinements include changes in the information structure.

Another paper considering ε-equilibria is Monderer and Samet (1989), who examine loosenings of common knowledge and show that with sufficient commonality in players' beliefs about the payoff structure of the game, one can find an ε-equilibrium close to any Nash equilibrium of the original game. Their analysis thus also covers Rubinstein's email game, and so the point that ε-equilibria revive the desired equilibrium in Rubinstein's setting is not new to our paper. However, the analysis of Monderer and Samet (1989) centers on underlying complete information games and (countable) uncertainty about such games, and on commonality of beliefs, and so otherwise is quite orthogonal to ours (and see the concluding remarks for additional discussion).

Our paper is also related to that of Kajii and Morris (1998). Their focus was on defining what it means for beliefs to be 'strategically close' to each other in terms of having nearby ε-equilibria. As such, they provide a variation of a lower hemi-continuity result for ε-Bayesian equilibria. Our results move beyond those of Kajii and Morris on several dimensions. First, they consider games with countable sets of types, which simplifies the analysis dramatically, as all of the technical difficulties that we face here deal with continuity arguments that are associated with the continuum. This is not a mere technicality, since it is necessary to deal with continua to cover applications like global games. Second, they only consider variations in beliefs, and not in the underlying action spaces and payoff functions, which is also not a mere technicality since to cover applications like trembling hand perfection one needs to allow for variations in action spaces. Third, their motivation is quite different, and our main point about the sensitivity of equilibrium selections to ε-best responses is new and not implied by their results (and in fact their results do not cover any of the applications that we discuss here).

Finally, we remark that our results and their implications are quite different from, and complementary to, those of Weinstein and Yildiz (2007). Weinstein and Yildiz show how variations of beliefs can lead to any selection among rationalizable strategies, which then implies that techniques used in global games may not be robust to the specification of the uncertainty. Here, our results apply to (Bayesian) Nash equilibria, rather than rationalizability, and work with loosening best response behavior rather than adjusting (higher order)

mation case, but still only allowing countable uncertainty.

beliefs, and thus the results have no overlap in conclusions, proof techniques, or intuitions. Moreover, our results include a different set of applications - including things like trembling hand perfection that are not covered by rationalizability arguments.

2 Bayesian Games and Equilibria

We begin with some basic definitions.

A Bayesian game is a collection $\langle N, (A_i)_{i\in N}, (T_i)_{i\in N}, S, P, (u_i)_{i\in N}\rangle$, which is as follows

- $N = \{1, \ldots, n\}$ is a finite set of players,
- A_i is the action space of player i, which is a compact metric space for each $i \in N$,
- T_i is the type space of player i, which is a compact metric space for each $i \in N$,
- S is a 'state' space, which is a compact metric space,
- P is a probability measure on $T \times S$, and
- $u_i : A \times T \times S \to [0, 1]$ is a continuous payoff function for each i.

A and T denote $A = A_1 \times \cdots \times A_n$ and $T = T_1 \times \cdots \times T_n$.

Note that this formulation allows for mixed strategies by allowing A_i to be a set of randomized strategies over some pure strategy set.

A *strategy* for player i is a (Borel) measurable function $\sigma_i : T_i \to A_i$. Let $\sigma(t) = (\sigma_1(t_1), \ldots, \sigma_n(t_n))$.

Given a profile of strategies $\sigma = (\sigma_1, \ldots, \sigma_n)$, let $V_i(\sigma, t_i)$ be the expected utility of agent i with type t_i and let $U_i(\sigma)$ be the ex ante expected utility:

$$V_i(\sigma, t_i) = E_P[u_i(\sigma(t), t, s)|t_i]$$

$$U_i(\sigma) = E_P[u_i(\sigma(t), t, s)]$$

where E_P denotes any version of the conditional expectation with respect to P relative to the Borel σ-algebra on T_i.

A profile of strategies $\sigma = (\sigma_1, \ldots, \sigma_n)$ is an interim ε-*equilibrium* of the game $\langle N, (A_i), (T_i), S, P, (u_i)\rangle$ for some $\varepsilon \geq 0$ if for all i and strategies σ_i' of player i:

$$V_i(\sigma, t_i) \geq V_i(\sigma_i', \sigma_{-i}, t_i) - \varepsilon$$

for P-almost all types t_i.

A profile of strategies $\sigma = (\sigma_1, \ldots, \sigma_n)$ is an ex-ante ε-*equilibrium* of the game $\langle N, (A_i), (T_i), S, P, (u_i) \rangle$ for some $\varepsilon \geq 0$ if

$$U_i(\sigma) \geq U_i(\sigma_i', \sigma_{-i}) - \varepsilon$$

for all i and all strategies σ_i' of player i.

Interim equilibrium and *ex-ante equilibrium* are the corresponding special cases where $\varepsilon = 0$.

Below, we provide theorems relating to both of these definitions of equilibrium. The result on the stronger notion of interim equilibrium requires a stronger definition of convergence, and the difference between these notions clarifies some aspects of arguments that are used in the literature. For instance, the trembling-hand example discussed below is a special case of the stronger result since it has complete information, while global games results do not satisfy the stronger convergence requirement and are only covered by the ex-ante result. Regardless of that stronger result, the ex-ante notion of equilibrium is still an important and interesting one. For instance, if players formulate rules of thumb that do very well (are almost best responses) in most situations, then the ex-ante notion applies.

3 Applications: Trembling Hand Perfect Equilibria, Commitment in Games, the Email Game, and Global Games

Before stating the main theorems, we illustrate some of their implications with applications to four classic papers that have used perturbations of a game in different ways to make selections of equilibria.

3.1 Trembling Hand Perfect Equilibria

We begin with the canonical equilibrium refinement. Here our main point is particularly easy to see and so some of the basic intuition is conveyed, before we deal with the complications introduced with incomplete information.

Consider a finite game in normal form $\Gamma^0 = \langle N, (A_i^0), (u_i^0) \rangle$ (so here states and types are degenerate),[4] where

$$A_i^0 = \{a_i^0 \in [0,1]^{m_i} \mid \sum_k a_{ik}^0 = 1\}$$

[4]One can easily define corresponding concepts for general Bayesian games, and it is clear that our main theorem still directly covers those cases.

is a set of possibly mixed strategies over m_i pure actions, and u_i^0 are standardly defined (expected) payoff functions. Let

$$A_i^r = \{a_i^r \in [\eta_{ik}^r, 1]^{m_i} \mid \sum_k a_{ik}^r = 1\},$$

and let $\Gamma^r = \langle N, (A_i^r), (u_i^r) \rangle$ be the game where each player i is restricted to place at least weight η_{ik}^r on action k; and such that η_{ik}^r converges to 0 as $r \to \infty$ for each i and k, but where $u_i^r = u_i^0$.

A Nash equilibrium a^0 of Γ^0 is a *trembling-hand perfect equilibrium*, as defined by Selten (1975), if there exists a sequence of perturbed games Γ^r (for some associated sequence of minimal trembles on actions η_{ik}^r converging to 0) and an associated sequence of equilibria a^r of Γ^r that converge to a^0.[5]

We define a corresponding concept allowing for slight perturbations in the best response function, but with the stronger requirement that things work for all perturbations.

We say that Nash equilibrium a^0 of the game Γ^0 is a *trembling*-hand perfect equilibrium* if *for every* sequence of perturbed games Γ^r (for some associated sequence of minimal trembles on actions η_{ik}^r converging to 0) there exists a sequence of (interim) ε^r-equilibria a^r of Γ^r that converge to a^0, such that $\varepsilon^r \to 0$.

We remark that the above definition requires the equilibrium be approached by ε^r regardless of the perturbation, which is stronger than the analogy to trembling hand perfect equilibrium which only requires things hold for some perturbation. Thus, our comment here also holds for proper equilibrium as defined by Myerson (1978), which requires specific perturbations.

PROPOSITION **1** *All of the Nash equilibria of any finite normal form game are trembling*-hand perfect equilibria.*

Thus, while trembling hand perfection can be a very discerning refinement, if we perturb best response correspondences along with perturbing the game, the refinement power is completely eliminated.[6]

Proposition 1 is a corollary of Theorem 3, but again we can see a fairly easy direct proof. Begin with a^0 and define a^r by choosing

$$a_i^r \in \text{argmin}_{\widehat{a}_i^r \in A_i^r} |\widehat{a}_i^r - a_i^0|.$$

[5] We omit the obvious definition of convergence, but complete details are provided in the next section.

[6] The set of trembling hand perfect equilibria have a close relationship to the set of undominated Nash equilibria, and in fact coincide for two-person games. If one loosens domination to only say that an action is dominated if some other action gives at least an ε improvement regardless of the actions of other players, then we would obtain a similar conclusion: all Nash equilibria are undominated* equilibria, where domination is defined in this weaker sense.

It then follows easily that if a^0 is an equilibrium, then a_i^r must be a ε^r-best response to a_{-i}^r (noting that expected utility is continuous in mixed strategies), where ε^r is proportional to $\max_{ik}(\eta_{ik}^r)$.

To see how this works in the context of a simple example, consider the two-by-two game pictured in Figure 1.

	L	R
U	1, 1	0, 0
D	0, 0	0, 0

Figure 1: There are two equilibria: (U,L) and (D,R). Only (U,L) is a trembling hand perfect equilibrium.

In this game, there are two Nash equilibria, (U,L) and (D,R), and only one trembling hand perfect equilibrium (U,L). If we require that either player to place weight at least $\frac{1}{r}$ on each strategy then the unique best response is for the players to play U, L.

Note however, that if, for instance, the column player plays L with weight $\frac{1}{r}$ and R with the remaining weight, then the row player only loses $\frac{1}{r}$ in expected payoff by playing D rather than U. Thus, (D,R) is still an $\frac{1}{r}$-equilibrium of the game where players are forced to mix their strategies.

In fact, if along with forcing the players to mix their strategies, we allow for a slight perturbation in payoffs, then (D,R) is an exact equilibrium for the sequence of perturbed games. For example, if the payoffs are slightly perturbed as in Figure 2, then (D,R) is in fact a strict equilibrium of the corresponding perturbed games.

The point of this is example is *not* that we find (D, R) to be a compelling equilibrium of this game. The point is that in order to rule out this equilibrium, one has to rely on arguments that involve more than simply perturbing the game if one believes that along with possible trembles in the strategies come possible misperceptions of the payoffs by either the modeler or the players. In that case, (D, R) becomes a strict equilibrium, and ruling out strict equilibria requires other sorts of approaches. Approaches such as stochastic stability (e.g.,

	L	R
U	1, 1	0, 0
D	0, 0	$\frac{1}{r-2}$, $\frac{1}{r-2}$

Figure 2: Perturbed Payoffs: Now D is a strict best reply even when $\frac{1}{r}$ weight is played on L.

Kandori, Mailath and Rob (1993), Young (1993)), or other sorts of dynamic/evolutionary or reasoning arguments (e.g., Harsanyi and Selten (1988)), can lead to a prediction of (U,L) even with perturbations of the payoffs, but these rely on very different sorts of machinery and are not subject to the same criticism that we are making of trembling hand perfection.

3.2 Commitment in Games

Bagwell (1995) studies the robustness of predictions regarding the ability of a party to commit to a course of action by examining what happens under small perturbations to the observability of their actions. An example that illustrates his results is as follows.[7]

A seller decides whether to produce a low quality version (*Low*) or a high quality version (*High*) of a product. A buyer observes the quality, and then decides whether to buy (*Buy*) or not buy (*Not*) the product at some fixed price. The payoffs are described in Figure 3.

In the limit game Γ^0, the seller moves first, and the buyer decides whether to buy after she perfectly observes the quality. In the unique subgame perfect equilibrium of Γ^0 the seller produces the high quality version and the buyer buys if the quality is high and does not buy if the quality is low. The fact that the quality of the product can be observed by the buyer before making his decision allows the seller to commit to producing the high quality version. Bagwell studies the robustness of the commitment prediction by examining perturbations of this sequential game in which the buyer has to make the decision after an imperfect observation of the product's quality.

[7]We thank Ehud Kalai for suggesting that we use this example.

	Buy	*Not*
High	1, 1	0, 0
Low	2, -1	0, 0

Figure 3: Payoffs as a function of the actual quality and the decision of the buyer

Specifically consider the game Γ^r in which the seller chooses the quality and the buyer decides whether to buy or not to buy based upon a noisy signal s, taking values (H) or (L) and which is consistent with the true quality of the product with probability $1 - \frac{1}{r}$,

$$p(s = H|High) = 1 - \frac{1}{r} = p(s = L|Low).$$

In terms of mapping this into our setting, there is a state of nature which can be thought of as "correct" or "not correct," which determines whether the type of the buyer matches the actual quality or not, and types can be thought of as the signals. We can then list a buyer's strategy as a function of the signal.

Bagwell's interesting conclusion can be seen by noting two key features of these games with imperfect observability which hold regardless of how small the amount of noise is[8]. The first one is that the only pure strategy Nash equilibrium of Γ^r involves the buyer never buying the product and the seller producing the low quality version[9]. The second one is that in the only equilibrium in which the seller produces the high quality version with high probability, the buyer buys for sure upon observing the H signal and buys with probability approximately $\frac{1}{2}$ upon observing signal L[10]. So no matter how large r is, there are no equilibria which are *close* to the unique subgame perfect equilibrium of Γ^0. In this sense, it seems that the ability of the seller to commit to producing the high quality version disappears even with

[8] That is, how large r is.

[9] It has a family of mixed strategy Nash equilibria in which the seller produces the high quality version with probability at most $\frac{1}{r}$ and the buyer never buys.

[10] Note that since in Γ^r, all information sets are reached on the equilibrium path, none of the standard refinements of the concept of Nash equilibrium apply.

the slightest noise in the observations made by the buyer prior to making his decision. Nevertheless if we allow for a corresponding small perturbation in the players' best response behavior then this commitment ability is restored.

Claim 1 *For any equilibrium σ^0 of the limit game (Γ^0), there is a sequence σ^r of interim ε^r-equilibria of the corresponding games Γ^r that converge to σ^0, where $\varepsilon^r \to 0$. That is, each σ^r is an ε^r-equilibrium with the property that the strategy of player 2 is ε^r-optimal conditional on each possible signal that he may observe.*

Consider the strategy profile $\sigma^r = (\sigma_s^r, \sigma_b^r)$, according to which the seller chooses high quality with probability $1 - \frac{1}{r}$ and low quality with probability $\frac{1}{r}$; the buyer chooses to buy when seeing the high-quality signal and doesn't buy when seeing the low-quality signal. For large enough r and given the buyer's strategy, producing the high quality version is the seller's best choice. However this optimal strategy only beats σ_s^r by $\frac{1}{r}$, and therefore σ_s^r is $\frac{1}{r}$-optimal. The buyer's strategy is in fact optimal because conditional on observing the high quality signal, the probability that the product is of high quality is high enough for him to prefer to buy, and conditional on the low-quality signal, the item has an even probability of being high-quality and low-quality so he is indifferent between buying and not buying. Thus σ^r is an interim r-equilibrium, and it is easy to see that σ^r converges to the unique subgame perfect equilibrium of Γ^0 σ^0 as $r \to \infty$.

3.3 ε-Equilibria in Rubinstein's (1989) Electronic-Mail Game

Let us now turn to a different example, where a different sort of perturbation is made to the games, but our claim will still apply (by a slightly different argument).

Rubinstein's (1989) electronic mail game is one where with complete information (and common knowledge) of a state of nature, two players are able to coordinate on an equilibrium that leads to a socially efficient outcome that is unambiguously good for both of them. However, if their communication is imperfect, and the state is no longer common knowledge, then that equilibrium disappears. One implication of our theorem is that even though that equilibrium disappears, it is still an ε-equilibrium. In fact, in this case the ε is proportional to the noise in the communication.

The game is formally described as follows. There are two states of nature x and y and two players who choose an action in $A_i = \{X, Y\}$. The states of nature correspond to the payoff matrices in Figures 4 and 5. State y occurs with probability $p < \frac{1}{2}$ and x occurs with probability $1-p$, and the state is privately revealed to player 1 at the beginning of the game. The players are connected by an automatic e-mail exchange system that works as follows:

- An initial message is sent by player 1's terminal if and only if the state of nature is y.

- Each terminal automatically replies to any message it receives.
- Any given message fails to get from the emitting terminal to the receiving terminal with a probability $0 < \gamma < 1$.

After the state of nature is revealed to player 1, players learn how many messages $t_i \in \{0, 1, 2....\}$ were sent by their respective machines, and then the players simultaneously choose action X or action Y, and their payoffs are determined according to the state of nature as shown in Figures 4 and 5, where $L > M$ and δ is taken to be a small number.[11]

	X	Y
X	M, M	0, $-L$
Y	$-L$, 0	$-\delta$, $-\delta$

Figure 4: The Game G_x

There is a unique equilibrium of the above game for any $\gamma > 0$ which is for both players to choose action X regardless of their types. The key insight, is that type $t_1 = 1$ realizes that she sent a message to the other player, but did not get one in return. There is a better than 1/2 chance that player 2 never received 1's initial message and so will then find it optimal to play X.[12] Given these odds, player 1 of type $t_1 = 1$ should choose X. The game unravels from there, as similar reasoning then applies to player 2 of type $t_2 = 1$, and then to player 1 of type $t_1 = 2$, and so forth.

The intriguing aspect of this game, is that when the state is really y, the players' types are likely to be quite large, so with high probability, they will each know the state, and know

[11] Rubinstein's (1989) game has $\delta = 0$. Our formulation results in a unique equilibrium, rather than simply a unique equilibrium where X is played in state x. If $\delta = 0$ there is also an equilibrium where both players always play Y. With any $\delta > 0$, that is no longer an equilibrium.

[12] It is a dominant strategy for player 1 to choose X in state x. This implies that when $t_2 = 0$, player 2 places at least weight 1/2 on the state being x and player 1 playing X and then has a unique best reply to play X as well.

	X	Y
X	0, 0	0, $-L$
Y	$-L$, 0	M, M

Figure 5: The Game G_y

that the other knows the state, and so forth, to a high level. However, the γ-wedge is enough to break down full common knowledge, and that is all that is needed to eliminate the quite natural equilibrium where both players choose X when the state is x and Y when the state is y.

When we look at the limiting game, where $\gamma = 0$, which intuitively corresponds to a game where both players are publicly told the state and it is common knowledge, then there is another quite natural and Pareto optimal equilibrium where players play X when the payoff-state is x and Y when the state is y. Thus, any slight introduction of noise that breaks down the common knowledge of the state in the way described above, eliminates the higher payoff equilibrium, and results in a very different play of the game.

Our point is that if in addition to introducing a slight noise in beliefs one also allows for slight deviations from full optimization, then the original equilibria are approximately recaptured.

Claim **2** *For any sequence of $\gamma \to 0$ (the probability of the message not getting across), there is a corresponding sequence of ex ante ε^γ-equilibria, σ^γ, of the corresponding email games, where $\varepsilon^\gamma \to 0$, and which converge to having players play X when the payoff-state is x and Y when the state is y.*

Claim 2 shows that for slight perturbations in the information away from the complete information game, if one also allows for slight perturbations in best reply behavior (in an ex ante sense), then one recovers all equilibria including the Pareto optimal equilibrium that disappeared in the perturbations that Rubinstein considered.

Claim 2 follows as a corollary to Theorem 3.[13] In this case, it is also easy to see a direct proof of the claim. Consider the strategies $\sigma_i^*(0) = X$, $\sigma_i^*(t_i) = Y$ $\forall t_i \geq 1$ and all $\gamma > 0$. For any $\gamma > 0$ (the probability of the message not getting across), $\sigma^* = (\sigma_1^*, \sigma_2^*)$ constitute an ε-equilibrium of the game where ε is proportional to γ. To see this, note that it is a strictly dominant strategy for $t_1 = 0$ to play X. It then follows easily that the unique best response for $t_2 = 0$ is to play X. Next, under the prescribed strategy $\sigma^* = (\sigma_1^*, \sigma_2^*)$, it is a unique best response for any $t_i > 1$ to play Y, since they are sure that the other player has $t_i \geq 1$ and will thus play Y. The only remaining types are $t_i = 1$, and the only type who is not best replying is $t_1 = 1$, which is a type with a probability proportional to γ.

It is worth noting that the above-specified ε^γ-equilibrium of the game Pareto dominates the exact equilibrium of the perturbed game (for small γ). It is only a single low-probability type of player 1 who is not optimizing. Player 1 is better off ignoring some of his or her information: if player 1 only observes the true state, but not the number of messages received (t_1 is unknown to player 1 but the state is known), then playing a strategy corresponding to what each player has heard the state to be is an equilibrium of the game. In fact, if player 1 could commit not to pay attention to the messages received - all players would strictly benefit. This also provides an interesting interpretation of the ε-equilibrium. If player 1 has to incur some cost to observe t_1 (presuming the player sees the state $\{x, y\}$), then it would be optimal for player 1 to choose not to pay to observe the type. This is an example where a ε-equilibrium that involves low probabilities of some types acting substantially suboptimally, can be understood by having players face costs of discovering information or a cost of fully tailoring strategies to information (in addition to other justifications that are based on players or the modeler incorrectly perceiving payoffs).

Note that this example illustrates why the ε-equilibrium needs to be an ex-ante notion for this setting. The statement is not true from an interim perspective: type $t_1 = 1$ always has a strict best reply (with a fixed gap) of playing X. Proposition 2 echoes a finding of Monderer and Samet (1989) who argue that the distinction between ex-ante and interim equilibrium is due to a lack of common belief (a weakening of common knowledge) for some of the types. We discuss the relation to their results more fully in the concluding remarks.

3.4 Global Games

Next, let us consider the class of global games as first studied by Carlsson and van Damme (1993).

[13] Reset the type spaces to be $[0, 2]$, and then remap so that each $t_i > 0$ becomes $1 + 1/t_i$ and keep $t_i = 0$ to be 0, so that all type spaces can then be taken to be $T_i = [0, 2]$. Then let the Γ^0 game be such that with probability p the state is y and both players are of type 1, and with probability $1 - p$ the state is x and both players are of type 0.

A "global game" $\Gamma^r = \{N, S, (T_i), (A_i), (u_i), (P_i^r)\}$ is defined as follows, following the structure from Frankel, Morris, and Pauzner (2003). A state $s \in S \subset I\!R$ is drawn according to a continuous density ϕ with connected and compact[14] support. Each player i observes a signal $t_i^r = s + r\eta_i$, where η_i is distributed according to an atomless density f_i with support in the interval $[-1/2, 1/2]$, and η_i is independent of η_j for all $i \neq j$. The action set of player i, $A_i \subseteq [0, 1]$, is a closed, countable union of closed intervals and points, and contains 0 and 1. If player i chooses action $a_i \in A_i$, his or her payoff is $u_i(a_i, a_{-i}, s)$.

Frankel, Morris, and Pauzner (2003), in generalizing the earlier results of Carlsson and van Damme (1993), make a series of assumptions about the structure of the payoffs of the game, and also about the uncertainty structure. Under those assumptions, they show that even though the limit game (where $r = 0$) may have multiple equilibria, there is a unique equilibrium along the sequence. Thus, the approximation by a series of equilibria of slightly noisy versions of the game "selects" a unique equilibrium of the limit game. Here, we show that even without any structure on the game, we regain all of the equilibria of the limit game as the limit of ε-equilibria along the sequence. So, no selection can be made if one allows for perturbations to the best replies as well as the uncertainty in the game.

PROPOSITION **2** *For any sequence of "global games"* Γ^r*, and any equilibrium* σ^0 *of the limit game (*Γ^0*), there is a sequence* σ^r *of ex-ante* ε^r*-equilibria of the corresponding global games* Γ^r *that converge to* σ^0*, where* $\varepsilon^\gamma \to 0$.

It is important to note that this result holds only for the ex-ante notion of equilibrium and not the interim. The global games approach relies on introducing small measures of types that have very different preferences from the types in the original game, and then using those far-away types to anchor the strategies of the equilibrium. If one requires a stronger definition of convergence in the game, then the global games conclusions could not be obtained, or if one allows that perhaps the far-away types' strategies are mispecified.

In contrast to the earlier propositions, Proposition 2 is more difficult to prove directly, but it does follow as a direct corollary of Theorem 3. Thus, we now move on to the main theorem.

4 The Theorems

Before stating the main theorems, we provide some definitions that are useful in defining convergence in a general setting of Bayesian games.

[14]Frankel, Morris, and Pauzner (2003) do not assume compact support, but their results could be recast in such a setting with an appropriate compactification. We can prove Proposition 2 without compactness (via a series of successive approximations and repeated applications of our main theorem), but then it is no longer a direct corollary of Theorem 3.

4.1 Sequences of Games

For $r = 0, 1, \ldots,$ let $\Gamma^r = \langle N^r, (A_i^r), (T_i^r), S^r, P^r, (u_i^r) \rangle$ be a Bayesian game. All the games in question have the same set of players $N^r = N = \{1, \ldots, n\}$, and satisfy the following conditions: The action spaces A_i^r lie in a compact subset A_i of a fixed locally convex linear metric space,[15] the closed type spaces T_i^r lie in a fixed compact metric space T_i, and the closed state spaces S^r lie in a fixed compact metric space S, and the P^r are Borel probability measures on $T \times S$, with support in $T^r \times S^r$. All u_i^r are defined on $A \times T \times S$.

4.2 Convergence of a Sequence of Games

We say that the sequence of games $\{\Gamma^r\}$ *converges to* Γ^0 if

[1] $S^r \to S^0$, and for each i: $A_i^r \to A_i^0$ and $T_i^r \to T_i^0$; all in the Hausdorff metric,

[2] the u^r are a uniformly equicontinuous sequence of functions.[16,17]

[3] convergence of beliefs:

(a) $P^r \to P^0$ in the sense of weak convergence of measures (as measures on $T \times S$),

(b) For each r and i there exists a P^r-measure one set, $B_i^r \subset T_i^r$, such that: for any continuous function $f : T \times S \to [0, 1]$ and for every $\varepsilon > 0$ there exists δ and $\underline{r}$ such that if r and r' are larger than $\underline{r}$ (including the limit $r = 0$) and $t_i^r \in B_i^r$ and $t_i^{r'} \in B_i^{r'}$ are such that $d(t_i^r, t_i^{r'}) < \delta$ then $|E_{P^r}[f|t_i^r] - E_{P^{r'}}[f|t_i^{r'}]| < \varepsilon$.

Most of the conditions are straightforward, simply requiring convergence of corresponding action spaces, type spaces, payoffs, and probability measures. The requirement [2] that preferences are uniformly equicontinuous guarantees that preferences can be compared across games. In fact in most applications the preferences are the same across games and it is only the information and uncertainty structure, states, perhaps actions spaces that differ along the sequence. In such a case, [2] is satisfied as long as preferences are continuous, given the compactness of the domains.

[15] It is thus sufficient to have a compact subset of a normed vector space that is a metric space.

[16] We say that a sequence of functions $f^r : X \to Y$ indexed by r (possibly including a limit element indexed by $r = 0$), where X and Y are compact metric spaces, is uniformly equicontinuous if f^r is uniformly continuous for every r and for every $\varepsilon > 0$ there is a $\delta > 0$ and an index $\overline{r}$ such that $d(f^r(x), f^{r'}(x')) < \varepsilon$ for every pair $r, r' \geq \overline{r}$ (including $r = 0$) and $x \in X$ and $x' \in X$ such that $d(x, x') < \delta$.

[17] When working with product spaces, we take d to be the ∞-product metric, So, for instance, $d((a_i, t_i, s), (a_i', t_i', s')) = \max\{d_{A_i}(a_i, a_i'), d_{T_i}(t_i, t_i'), d_S(s, s')\}$, and then $d((a, t, s), (a', t', s')) = \max_i\{d((a_i, t_i, s), (a_i', t_i', s'))\}$. Note that uniformity is implied in compact spaces, but we state it to make clear the implications that are used in the proof.

The only condition that might be unexpected is [3b]. Some condition like this is needed, as weak convergence of the underlying type distributions is not sufficient to guarantee that agents' interim beliefs in the sequence of games converge to the beliefs in the limit game.[18] An example of a sequence of underlying distributions that converge weakly to a limit but such that the corresponding sequence of interim beliefs is completely different from the limit can be seen in Example 1 in Milgrom and Weber (1985).[19] They exhibit a game where players' types are perfectly correlated and revealing of a state all along a sequence, but not at all revealing of a state in the limit under weak convergence. Such sequences need to be ruled out in order to establish any sort of equilibrium convergence.

Condition [3b] simply requires convergence of beliefs in the weakest way required for the proof technique of our main theorem, and is satisfied in the relevant applications. For instance, if the type space is finite, there exists some δ such that the distance between any two types are larger than δ. Thus [3b] is just the convergence of conditional expectation upon the same type, which is true by weak convergence. Moreover, if the type space is countable and there exist some δ such that the distance between any two types are larger than δ, [3b] applies as well. Another sufficient condition is independent conditional distributions such that $E_{P^r}[f|t_i^r] = E_{P^r}[f|t_i^{r'}]$ for all P^r, f and all types, and [3b] is implied by the same argument as in the finite case. More generally, the condition is a continuity condition in terms of how conditional beliefs vary with types, and is satisfied in many standard applications (such as some sequences of affiliated values auctions that have the same supports on types as the uncertainty varies).

4.3 Interim ε-Equilibrium

We begin with results on the stronger notion of interim ε-equilibrium, which requires approximate best responses for all types of all players.[20]

Before stating the theorems on interim ε-equilibrium, let us discuss why it is important to allow for such equilibria in order to have a robust analysis, rather than simply focusing on exact equilibria. There are at least three reasons that justify examining robustness with

[18]For more on the necessity of such a condition, see Kajii and Morris (1998).

[19]That example has one player and an uncertain state, so that $S = \{0, 1\}$ and $T = T^1 = [0, 1]$. P^r is such that $P^r[s = 0, t_1 = \frac{2j}{2r}] = \frac{1}{2r}$ and $P^r[s = 1, t_1 = \frac{2j-1}{2r}] = \frac{1}{2r}$ for every $j \in \{1, \ldots, r\}$. Thus, all along the sequence player 1's type exactly reveals the state. Note that P^r weakly converges to P^0 such that $P^0[s = 0, t_1 \leq t] = P^0[s = 1, t_1 \leq t] = t/2$ for $t \in [0, 1]$ and so in the limit the player's type reveals nothing about the state. If the player's utility is state dependent, then equilibria of the sequence of games and the limit game might be completely different from each other as the information is not converging in an appropriate sense.

[20]Of course, formal statements are up to sets of measure zero given that all things are only defined up to versions of corresponding probability measures.

regards to interim ε-equilibrium:

- The modeler may have (slightly) mispecified or mismeasured the payoffs of the players.
- The players may make slight errors or may slightly misperceive their payoffs or have slight noise in their beliefs or perceptions of payoffs.
- Complexity or other costs may prevent players from exact optimization, and it may suffice for them to approximately optimize.

Given this perspective, if one relies on refinements based on exact equilibrium, when those refinements are not robust with respect to approximate equilibrium, then one may be mistaken in predictions.

THEOREM **1** *Let $\{\Gamma^r\}$ be a sequence of (Bayesian) games converging to Γ^0 such that the supports of the P^rs converge to the support of P^0 in the Hausdorff metric. For any continuous equilibrium*[21] *σ^0 of the limit game Γ^0, there exists a sequence of interim ε^r-equilibria σ^r of the corresponding games $\{\Gamma^r\}$ that converge uniformly to σ^0*[22] *such that $\varepsilon^r \to 0$.*

This theorem requires that the support of P^r be close to the support of P^0 when r is large enough. This is an important condition for our interim results. As we saw in the email game above, the support of P^0 on the type space was $t \in \{0, +\infty\}$ while the support of P^r was $t \in \{0, 1, 2, ..., +\infty\}$ for all $r > 0$. For agents with types $(t = 1)$, which are far from the support of P^0, we cannot find a type close to $t = 1$ in $t \in \{0, +\infty\}$ such that we cannot find the corresponding strategy in σ^0 to assign to $t = 1$ and make sure it is close to best response.

The proof of this theorem is a variation of the proof of our theorem about ex ante ε-equilibrium. That proof already implies that given the identical supports on types and identical action spaces then the actual original strategies would be interim ε-equilibria. The only non-obvious portion is the approximation of strategies when type and action spaces do not coincide. This is taken care of by first matching a type in the support of game Γ^r to a nearest type in the support of Γ^0, being careful about measurability, and then choosing its corresponding strategy to be as close as possible to that of its matched type (the details appear in the proof of Theorem 3), subject to any constraints imposed by differences between A^r and A^0. Then continuity of beliefs and utility functions guarantee that these strategies are approximate best responses for all types.

[21] If the support of P^0 is finite, then any equilibrium σ^0 is continuous. Thus, this applies to any setting where the limit game has a finite support of types, including oth the trembling-hand example and the commitment-games example.

[22] In our proof we construct $\sigma^r : T \to A^r$ so that the σ^r are defined for all types in T, even those outside of T^r, and also mapping within the range A^r, and such that if we restrict the domain of σ^r to be T^r then these are then strategies in Γ^r.

This theorem shows that an interim approximation result holds generally for continuous equilibria when the approximating games have type spaces close to the limit game. If there are some discontinuity points in the equilibrium strategies, slight belief changes in Γ^r can have significant effects since there can be types facing agents on one side of some discontinuity point who face agents on the other side of the discontinuity point with very different behaviors under a slight belief change. Thus, in order to extend the theorem to address discontinuous equilibria, one has to strengthen the conditions on the convergence of beliefs. In order to deal with this issue, a stronger form of convergence of beliefs is needed, and we consider the following conditions that correspond to conditions that Milgrom and Weber (1985) used to study existence of equilibria and upper-hemi continuity:

- Absolute Continuity: let P_i^r be the marginal distribution of P^r with respect to T_i and P_S^r be the marginal on S. P^r is absolutely continuous with respect to the product measure $\hat{P}^0 = P_S^0 \times P_1^0 \times \cdots \times P_n^0$, with corresponding Radon-Nykodym derivative by f^r, for every r (including 0).
- Uniform Convergence: The $\{f^r\}$ converge uniformly to f^0 on every compact subset of $T \times S$, and f^0 is continuous almost everywhere with respect to $\hat{P}^0$.

Note that the first condition requires that the measures associated with the perturbed games be absolutely continuous with respect to the product measure of the limit game.

With this stronger notion of convergence of beliefs in hand, we can show that the interim approximation results hold for all equilibria, including discontinuous ones.

THEOREM **2** *Let $\{\Gamma^r\}$ be a sequence of (Bayesian) games converging to Γ^0 satisfying absolute continuity and uniform convergence. For any equilibrium σ^0 of the limit game Γ^0 there exists a sequence of interim ε^r-equilibria σ^r of the corresponding games $\{\Gamma^r\}$ that converge in probability to σ^0 relative to P^0 and such that $\varepsilon^r \to 0$.*

These theorems imply that all equilibria of limiting games may be approached by approximate equilibria of any sequence of nearby games, provided the type spaces are sufficiently close in well-defined senses. Thus, refinements relying on perturbations of exact equilibria are not robust to perturbations of payoffs or best reply behavior, unless they rely on perturbations of the type spaces that introduce types that are quite distant from the original ones. Thus, these theorems also shed light on the difference between interim and ex-ante ε-equilibrium notions, as to satisfy the interim notion, we require the stronger notion of convergence. We next discuss more drastic perturbations, which then must rely on ex-ante equilibrium, in order to capture applications to Rubinstein's email game, to global games, and related settings.

4.4 Ex-Ante ε-Equilibrium

Before stating our theorem on ex-ante ε-equilibrium, we provide additional justifications regarding the importance and interest of this as an equilibrium notion. In justifying the notion of interim approximate equilibria, we mentioned three possible justifications: slight misspecifications or errors by the modeler, slight errors or noise in beliefs or perceptions by the players, or slight costs of full optimization. When we move to a notion of *ex-ante* ε-equilibrium, we now allow some vanishing set of types of players not to be even approximately optimizing. How does one justify an equilibrium notion that allows some subset of players not even to approximately best respond? Here there are several answers. First, as in the justification of interim equilibrium, it could be that the modeler has slightly misspecified the game. Any equilibrium refinement relying on perturbed games that escapes the interim approximations we discussed above must then be relying critically on some small set of types who are quite different from any types in the limit game. If those types are eliminated, then any equilibrium in the limit can be approximated. Thus, it is on at most a vanishing fraction of types on which any perturbation-based refinement must depend and it will then be non-robust to a misspecification of the types in that way. Second, from the vast literature on game theoretical experiments, one often sees, in fact, non-trivial fractions of players who fail to even approximately best respond, having some more basic misunderstanding of the game. Here we show that that fraction need only be tiny to allow approximation of any equilibrium of a limiting game. Third, players may in fact optimize from an ex-ante point of view and then choose behaviors that pay attention to most possible situations, but are not fully optimal in some vanishing fraction of situations.

In order to state the theorem, we need one additional definition regarding the form of convergence that we establish. Given compact metric spaces X and $X^0 \subset X$, we say that a sequence of corresponding functions $f^r : X \to Y$ for some compact metric space Y *converges** to $f^0 : X^0 \to Y$ relative to a probability measure P on X, denoted $f^r \to_P^* f^0$, if f^r is continuous for each r and $d(f^r(x), f^0(x)) < \varepsilon(r)$ for a set of P-measure at least $1-\varepsilon(r)$ where $\varepsilon(r) \to 0$.

This is essentially a standard convergence notion, namely convergence in probability, except that we additionally demand that the approaching functions be continuous, which makes it a stronger notion of convergence. Nevertheless, we will show that for any equilibrium of the limit game, there exists a sequence of ε-equilibria converging it in this sense.[23]

THEOREM **3** *Let* $\{\Gamma^r\}$ *be a sequence of (Bayesian) games converging to* Γ^0*. For any equilibrium* σ^0 *of the limit game* Γ^0*, there exists a sequence of ex ante* ε^r*-equilibria* σ^r *of the*

[23] As will be clear in the proof, if the action spaces coincide then the convergence will be pointwise and so even stronger (in particular, step 3 of the proof of Theorem 3 is no longer necessary).

corresponding games $\{\Gamma^r\}$ *such that* $\sigma^r \to^* \sigma^0$ *relative to* P^0 [24] *and* $\varepsilon^r \to 0$.

The proof of the theorem follows a fairly simple intuition, but has substantial complications due to the fact that the games may have different type and strategy spaces, and that we allow for arbitrary distributions over types. These are important elements of generality to cover many applications. So, let us describe the basic intuition for the proof in the case where all of the type and strategy spaces are the same and the distributions over types all have the same finite support. We then outline the more general proof technique and refer the reader to Appendix for details.

In the case where all of the spaces are the same and the type distributions all have the same finite support, then the original equilibrium strategies form a sequence of ε^r-equilibria of the sequence of games. This is straightforward. If this were not true, then for some player and type of that player that has positive measure in the limit, there would be a sequence of deviations for that type that would all result in a benefit of more than some fixed amount $\gamma > 0$. Given the finite support, any strategy can be taken to be a continuous function, so then weak convergence of the beliefs implies that the limit of the deviations would result in a better reply in the limit, which contradicts the fact that the original strategy is an equilibrium.

The complications for the general proof come on several fronts. First, without a finite support on types, the original strategies (as well as deviations from them) might not be continuous functions. In that case, weak convergence of the measures of types does not guarantee any convergence of beliefs. In fact, it alone is not sufficient for the theorem in the general case, and that is the role of condition [3b]. Thus, the proof needs to work with approximations to the original strategies. We approximate the original strategies by a sequence of continuous functions that converge to the original strategies in the sense of $\to^*$. This involves using a theorem by Aldaz (1996) on continuous approximations, as well the Tietze extension theorem, in order to make sure that the convergence is strong enough for our purposes (Lemma 1). We then show that these approximations form a sequence of ε-equilibria of the original game due to the uniform continuity of the utility functions in the original game and the $\to^*$ convergence of our approximations to the original strategies. The second complication comes from then showing that these continuous approximations of the original strategies are approximate equilibria of the approaching games. The difficulty is that optimal deviations from these strategies need not have nice convergence properties, as the set of strategies itself is not compact; and we need to work with those in order to bound the potential gain from deviations along the sequence of converging games. We address this by a careful and fine partitioning of the space of types in a way such that the boundaries of the partition have negligible measure (Lemma 3) and then argue that any optimal deviation can

[24] See footnote 22.

be approximated by step functions on those partitions. The set of step functions on these carefully constructed partitions is compact. We then show that convergence of beliefs holds for such sequences of step functions (and here [3b] comes into play, together with Lemma 2), given that the non-deviating strategies are continuous and that the boundaries of the partitions have been chosen to have negligible measure. Next, we face the third complication, which is that the action and type spaces of the sequence of games might not be the same, and so these ε-equilibria of the original game cannot be directly taken to even be feasible in the approximating games. Thus, we need to approximate these by strategies in the converging games. In order to do that, we need to appeal first to the Kuratowski-Ryll-Nardzewski Selection Theorem to show that we can find a measurable map for large enough r which is an approximation, and then again apply the same machinery as above to get a continuous approximation. We then verify that these remain ε^r-equilibria, for appropriate choices of ε^r.

4.5 Correlated Equilibria and Upper Hemi-Continuity

Our next remark concerns correlated equilibria, as defined by Aumann (1974). Theorem 3 is stated for Bayesian (Nash) equilibria. Given that a correlated equilibrium can be written as a Bayesian equilibrium where the uncertainty encodes the correlation device, our results also cover correlated equilibria. We state the remark for the case of a complete information game, which then covers both the interim and ex-ante notions of equilibrium. The remark extends to Bayesian games as well, but then there are various potential definitions one can work with in terms of defining correlated equilibrium and we simply refer the reader to Forges (1993) for details on defining correlated equilibria in Bayesian settings for an appropriate notion in a strategic form setting; and then one can correspondingly define ε-correlated equilibria by extending her definition to only require approximate best responses to extend the remark.

REMARK **1** *Consider a finite game in normal form $\Gamma^0 = \langle N, (A_i^0), (u_i^0) \rangle$ where A_i^0 is the set of mixed strategies over a finite number of pure actions. For any correlated equilibrium of Γ^0 and a sequence of finite games in normal form $\Gamma^r \to \Gamma^0$, there is a corresponding sequence of ε^r-correlated equilibria[25] of the games Γ^r converging* to the correlated equilibrium of Γ^0.*

Theorem 3 establishes a sort of lower hemi-continuity result for Bayesian equilibrium, when we allow the approaching sequence to be ε^r-equilibria. An upper hemi-continuity result also holds, as one would expect from standard upper hemi-continuity results on Bayesian equilibria (e.g., see Milgrom and Weber (1985) and Jackson, Simon, Swinkels and Zame (2004)). We state such a result as a remark as its proof is more standard, and the necessary

[25]We omit the formal definition, as it is the obvious extension of correlated equilibria to have players ε-best respond given the information they have from the correlating device and the strategies of other players.

modifications of previous results to work with the general form of convergence of games that we allow here can be easily adapted from the proof of Theorem 3. We state this for the weaker concept of ex-ante equilibrium, since that then provides a stronger result since the limiting equilibrium with be both interim and ex-ante.

REMARK **2** *If σ^r is a sequence of ex-ante ε^r-equilibria of Γ^r such that $\sigma^r \to^* \sigma^0$ and $\varepsilon^r \to 0$, then σ^0 is an equilibrium of the limit game Γ^0.*

A sketch of the proof of the remark is as follows. Suppose to the contrary that σ^0 is not an equilibrium of the limit game Γ^0. Then there exists i, $\varepsilon > 0$ and $\overline{\sigma}_i^0$ such that

$$\varepsilon + E_{P^0}(u_i^0(\sigma^0)) < E_{P^0}(u_i^0(\overline{\sigma}_i^0, \sigma_{-i}^0)).$$

Then using some of the arguments in the proof of Theorem 3, we can find a (continuous) approximation of $\overline{\sigma}_i^0$, denoted $\widetilde{\sigma}_i^r$, such that

$$\varepsilon/2 + E_{P^r}(u_i^r(\sigma^r)) < E_{P^r}(u_i^r(\widetilde{\sigma}_i^r, \sigma_{-i}^r)).$$

However, σ^r is a sequence of ε^r equilibria of Γ^r, and so when r is large enough

$$\varepsilon/4 + E_{P^r}(u_i^r(\sigma^r)) \geq E_{P^r}(u_i^r(\overline{\sigma}_i^r, \sigma_{-i}^r))$$

for any possible strategy $\overline{\sigma}_i^r$, which contradicts the previous inequality.

5 Concluding Remarks

5.1 Continuity of Beliefs and Common-p Beliefs

Our analysis has shown that under suitable convergence and continuity conditions, any equilibrium of an underlying game can be approximated by ex ante ε-equilibria of an approaching sequence of games. We also provided a result for interim ε-equilibria under stronger conditions on the underlying sequences of perturbations.

Our approach is quite different from that of Monderer and Samet (1989), who prove a result with a similar structure when the approximating sequence is with countable uncertainty and satisfies a common p-belief condition. This raises a question regarding the relationship between the continuity that we work with here and the commonality of p-beliefs studied by Monderer and Samet.

Effectively, what is needed to get our approximation result is that players' beliefs about the uncertainty they face in the perturbed games not differ too much from what they face in the limit. Players face two forms of related uncertainty: one about what their payoffs in the game are as a function of the actions ("payoff uncertainty") and the other about what other players are likely to do ("strategic uncertainty"). A common p-belief condition (for very high p) is sufficient to ensure that both of these forms of uncertainty become negligible in the limit. Our continuity condition is a more direct requirement that ensures that beliefs (both payoff and strategic) in the perturbed games are similar to those in the limit. The advantage of working with a continuity condition is that it is a more primitive and more general mathematical condition, which allows us to expand the scope of the analysis quite substantially to cover settings where the underlying games are Bayesian and where we can allow for continua of types. Nonetheless, it would still be interesting to explore which sorts of common p-belief and other epistemic conditions are sufficient for continuity in the uncertainty beyond the countable-perturbations setting that was studied by Monderer and Samet (1989)[26].

5.2 Extensive Form Games and Rationalizability

Our approach has covered Bayesian games and has focused on equilibrium notions. Other potential applications of our reasoning include rationalizability notions, as well as extensive form games.

In fact, extensive form games can already be analyzed by applying our analysis since, for instance, our results already cover the associated normal form game and its trembling hand perfect equilibria*. This then implies that applying our analysis to the extensive form leads one to recover all Nash equilibria of the extensive form, even when one applies perfection arguments that would usually lead to sequential rationality. To understand why the admittance of ε-best responses leads one to recover all Nash equilibria rather than some variation of perfection in extensive-form settings, it is again easy to see that with only slight trembles, a change in any given player's off-the-equilibrium path behavior will only lead to slight changes in payoffs.[27] Thus, such trembles do not have the usual refining power, and all Nash equilibria are recovered, even in the extensive form. Despite our results being directly applied to extensive forms, there are also some interesting questions that arise concerning what the appropriate ε-notion of sequential rationality should be in doing so. Does one simply require ε-best responses from the perspective of the overall payoffs of the game (in

[26]Kajii and Morris (1997) provide a definition of common p-beliefs in a continuum setting, which could potentially provide a foundation for some analogs to Monderer and Samet's (1989) results with richer uncertainty.

[27]This applies whether one works with an associated agent normal form following Selten (1975) (where each information set is treated as a separate player), or with a standard formulation where players' fully contingent strategies are the actions in the strategic form of the game (e.g., von Neumann and Morgenstern (1947)).

which case our results here can be applied as just discussed) or does one require that players' payoffs be ε-best responses restricting attention to the payoffs *conditional on already being in some subgame*?[28] If it is the latter, then details of the extensive form need to be taken into account.

In terms of rationalizability, one can similarly loosen exact best responses to some beliefs to instead require ε-best replies. For instance, such an approach would eliminate much of the refinement power of concepts like perfect and cautious rationalizability that refine rationalizability concepts (see Bernheim (1984) and Pearce (1984)) in similar ways that we have seen here.

6 References

Aldaz (1996) "On the approximation of measurable functions by continuous functions," *Journal Rendiconti del Circolo Matematico di Palermo*, 45:2, 289 - 302.

Aumann, R. (1974) "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, 1, 67 - 96.

Bagwell, K. (1995) "Commitment and Observability in Games," *Games and Economic Behavior*, 8, 271 - 280.

Bernheim, D. (1984) "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007 - 1028.

Carlsson H. and E. van Damme (1993) "Global Games and Equilibrium Selection," *Econometrica*, 61:5, 989 - 1018.

Dugundji, J. (1951) "An extension of Tietze's theorem," *Pacific Journal of Mathematics* 1, 353 - 367.

Engl (1995) "Lower Hemicontinuity of the Nash Equilibrium Correspondence," *Games and Economic Behavior* 9, 151 - 160.

Forges, F. (1993) "Five legitimate definitions of correlated equilibrium in games with incomplete information," *Theory and Decision* 35:3, 277 - 310.

Frankel, D.M., S. Morris, and A. Pauzner (2003) "Equilibrium selection in global games with strategic complementarities," *Journal of Economic Theory*, 108:1, 1 - 44.

[28] To see the difference, note that some plays can be rationalized from an overall payoff perspective because players anticipate a low probability of reaching some subgames. However, if one requires near best replies *conditional* on being in some subgame, then that can have very different implications.

Fudenberg, D., and D. Levine (1986) “Limit Games and Limit Equilibria,” *Journal of Economic Theory*, 38:2, 261 - 279.

Harsanyi, J.C. (1967-1968) “Games with Incomplete Information Played by ”Bayesian” Players, I-III.” *Management Science*, 14:3, Theory Series, 159 - 182.

Harsanyi, J., Selten, R., (1988) *A general theory of equilibrium selection in games*, MIT Press, Cambridge.

Jackson, M.O., Simon, L.K., J.M. Swinkels, and W.R. Zame (2002) “Communication and Equilibrium in Discontinuous Games of Incomplete Information,” *Econometrica*, 70:5, 1711 - 1740.

Kajii, A. and S. Morris (1997) “Common p-Belief: The General Case,” *Games and Economic Behavior*, 18, 73 - 82.

Kajii, A. and S. Morris (1998) “Payoff Continuity in Incomplete Information Games,” *Journal of Economic Theory*, 82, 267 - 276.

Kalai, E. (2004) “Large Robust Games,” *Econometrica*, 72:6, 1631 - 1665.

Kandori, M., Mailath, G., Rob, R., (1993) “Learning, mutation, and long run equilibria in games,” *Econometrica*, 61, 29 - 56.

Milgrom, P.R. and R.J Weber (1985) “Distributional Strategies for Games with Incomplete Information,” *Mathematics of Operations Research*, 10:4, 619 - 632.

Monderer, D. and D. Samet (1989) “Approximating Common Knowledge with Common Beliefs,” *Games and Economic Behavior*, 1, 170-190.

Myerson, R.B. (1978) “Refinements of the Nash equilibrium concept,” *International Journal of Game Theory*, 15: 133-154.

Pearce, D. (1984) “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica*, 52, 1029 - 1050.

Radner, R. (1980) “Collusive behavior in oligopolies with long but finite lives,” *Journal of Economic Theory*, 22, 136 - 156.

Rubinstein, A. (1989) “The Electronic Mail Game: Strategic Behavior under ‘Almost Common Knowledge”.” *American Economic Review*, 385 - 391.

Selten, R. (1975) “A reexamination of the perfectness concept for equilibrium points in extensive games,” *International Journal of Game Theory*, 4, 25 - 55.

von Neumann, J. and O Morgenstern (1947) *Theory of games and economic behavior*, Princeton NJ: Princeton University Press.

Weinstein, J. and M. Yildiz (2007) "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica* 75:2, 365 - 400.

Young, H.P. (1993) "The evolution of conventions," *Econometrica*, 61, 57 - 84.

Appendix

The following lemmas are useful in the proofs.

LEMMA **1** *Given any $\delta > 0$, any i, any measurable strategy σ_i defined on a closed subset of T_i and any Borel probability measure P on that set, there exists a continuous σ_i' defined on all of T_i such that $d(\sigma_i'(t_i), \sigma_i(t_i)) < \delta$ for all t_i in a set of P-measure of at least $1-\delta$. If a second probability measure P' is also present, then $d(\sigma_i'(t_i), \sigma_i(t_i)) < \delta$ for all t_i in a set that has both P' and P-measure of at least $1-\delta$.*

Proof of Lemma 1: By Corollary 2.4 in Aldaz (1996) [noting that a compact subset of a locally convex metric space is a Frechet space, and that any Borel measure on a metric space is completed], there exists a sequence of continuous functions σ_i^k defined on the domain of σ_i converging pointwise to σ_i, P-almost everywhere. By a variation of the Tietze Extension Theorem[29] there exists an extension of σ_i^k which is continuous on all of T_i and is equal to σ_i^k on its domain, and so without loss of generality we take σ_i^k to be continuous on all of T_i. By Egorov's theorem, for any $\delta > 0$, we can find a set of P-measure at least $1-\delta$ such that the pointwise convergence of the σ_i^k to σ_i is uniform, and applying Egorov's Theorem a second time leads to the second conclusion. ∎

LEMMA **2** *Consider a sequence of (possibly discontinuous) functions $f^k : X \to [0,1]$ where X is a compact metric space, such that $|f^k(x) - f^*(x)| < \varepsilon^k$ for a set of P-measure at least $1-\varepsilon^k$ where $\varepsilon^k \to 0$ and P is a probability measure on X. Then $E_P[f^k] \to E_P[f^*]$.*

Proof of Lemma 2: Since there is a sequence $\varepsilon^k \to 0$ such that $|f^k(x) - f^*(x)| < \varepsilon(k)$ for a P-measure set of at least $1-\varepsilon^k$, it follows that

$$E_P[|f^k - f^*|] < \varepsilon^k(1-\varepsilon^k) + (1)\varepsilon^k,$$

and the right hand side goes to 0 with ε^k. Thus, $E_P[f^k] \to E_P[f^*]$. ∎

[29]This follows, since A_i is a locally convex linear compact metric space, by a theorem due to Dugundji (1951). Thank you to Ben Golub for pointing us to this version of the theorem.

LEMMA **3** *For every i and k there exists a finite partition of T_i denoted $\mathcal{C}_i^k$ such that the diameter of every $C_i \in \mathcal{C}_i^k$ is less than $1/k$ and $\sum_{C_i \in \mathcal{C}_i^k} P^0(\partial C_i) < 1/k$, where ∂C denotes the boundary of C.*

Proof of Lemma 3: First, we cover T_i with balls $B(t_i, k) = \{t_i' \in T_i | d(t_i, t_i') < 1/(4k)\}$. T_i is compact, so we can find a finite number of balls covering T_i; and we denote the finite set of the centers of the balls by T_i^c. Note that for any ϕ, $\frac{1}{2} < \phi < 1$, the balls $B(t_i, \phi k)$ for $t_i \in T_i^c$ still cover T_i and each have diameter less than $1/k$. Order the points in T_i^c $t_i^1, t_i^2, \ldots, t_i^m$ where $m = |T_i^c|$. Find $\frac{1}{2} < \phi^1 < 1$ such that $P^0(\partial B(t_i^1, \phi^1 k)) < 1/(m^2 k)$. This is possible since the intersection of the $\partial B(t_i^1, \phi k)$s is empty for any set of distinct ϕs such that $\frac{1}{2} < \phi < 1$, and so at most a finite number, $m^2 k$, of the boundaries can have measure of at least $1/(m^2 k)$. Similarly, find ϕ^j for each t_i^j in T_i^c such that $P^0(\partial B(t_j^1, \phi^j k)) < 1/(m^2 k)$. The balls $B(t_i^j, \phi^j k)$ form a finite covering of the set T_i such that the sum of their boundaries has a total measure of less than $1/mk$. Now, form a partition of T_i as follows. Let C_i^1 be $Cl(B(t_i^1, \phi^1 k))$, where Cl denotes closure. Inductively, let C_i^j be $Cl(B(t_i^j, \phi^j k))$ less the points in $\cup_{j' < j} C_i^{j'}$. Note that the boundary of the elements of the partition is the union of parts of the boundaries of the collection of balls $B(t_i^j, \phi^j k)$, and so $P^0(\partial C_i^j) < m\frac{1}{m^2 k} = \frac{1}{mk}$. Therefore $\sum_{C_i \in \mathcal{C}_i^k} P^0(\partial C_i) < 1/k$. Also, the diameter of each element of the partition is less than $1/k$ by construction. ∎

LEMMA **4** *σ is an interim ε-equilibrium relative to some P if and only if for every i and every positive measure Borel set $B_i \subset T_i$*

$$\int_{S \times B_i \times T_{-i}} u_i(\sigma, t, s) dP \geq \int_{S \times B_i \times T_{-i}} u_i(\sigma_{-i}, a_i, t, s) dP - \varepsilon P(B_i)$$

for all $a_i \in A_i$.

Proof of Lemma 4: If $\sigma = (\sigma_1, \ldots, \sigma_n)$ is an interim ε-*equilibrium* then for all i and actions a_i of player i:

$$V_i(\sigma, t_i) \geq V_i(a_i, \sigma_{-i}, t_i) - \varepsilon$$

for P-almost all types t_i.

Thus, for any positive measure B_i

$$\int_{B_i} V_i(\sigma, t_i) dP \geq \int_{B_i} V_i(a_i, \sigma_{-i}, t_i) dP - \varepsilon P(B_i).$$

Thus, by the definition of conditional expectations,

$$\int_{B_i \times S \times T_{-i}} u_i(\sigma, t, s) dP \geq \int_{B_i \times S \times T_{-i}} u_i(\sigma_{-i}, a_i, t, s) dP - \varepsilon P(B_i).$$

To see the converse, suppose that σ is not an interim ε-equilibrium, so that there exists σ_i' and a positive measure of types B_i' such that

$$V_i(\sigma, t_i) < V_i(\sigma_i', \sigma_{-i}, t_i) - \varepsilon$$

for all $t_i \in B_i'$. This implies that we can find some positive measure of types B_i'' such that

$$V_i(\sigma, t_i) < V_i(\sigma_i', \sigma_{-i}, t_i) - \varepsilon'$$

for all $t_i \in B_i''$ where $\varepsilon' > \varepsilon$. Then, by the uniform continuity of the utility (and the compactness of the domain), we can find a single a_i such that for a positive measure of types $B_i \subset B_i''$

$$V_i(\sigma, t_i) < V_i(a_i, \sigma_{-i}, t_i) - \varepsilon$$

for all $t_i \in B_i$. Thus,

$$\int_{B_i} V_i(\sigma, t_i) dP < \int_{B_i} V_i(\sigma_i', a_{-i}, t_i) dP - \varepsilon P(B_i).$$

Thus, by the definition of conditional expectations,

$$\int_{B_i \times S \times T_{-i}} u_i(\sigma, t, s) dP < \int_{B_i \times S \times T_{-i}} u_i(\sigma_{-i}, a_i, t, s) dP - \varepsilon P(B_i).$$

The claim follows. ∎

Corollary 1 *Suppose that P is absolutely continuous with respect to $\widehat{P}$, with Radon-Nykodym derivative f. Then σ is an interim ε-equilibrium relative to P if and only if for every i, a_i and P_i-almost every t_i:*

$$\int_{S \times T_{-i}} u_i(\sigma, t_{-i}, t_i, s) f(s, t_i, t_{-i}) dP_S(s) \times_{j \neq i} dP_j(t_j) \geq$$

$$\int_{S \times T_{-i}} u_i(\sigma_{-i}, a_i, t_{-i}, t_i, s) f(s, t_i, t_{-i}) dP_S(s) \times_{j \neq i} dP_j(t_j) - \varepsilon.$$

We begin our proofs with the ex-ante theorem, as some of its steps are useful in proving the interim theorems.

Proof of Theorem 3:

Suppose that $\{\Gamma^r\}$ is a sequence of Bayesian games converging to Γ^0. Let σ^0 be an equilibrium of the limit game Γ^0. We prove the theorem in the following three steps.

The first step finds a sequence of continuous strategies mapping all of T_i into A^0, denoted σ^δ, approximating σ^0 that are ex-ante ε^δ equilibria of the game Γ^0.

The second step then shows that for each $\delta > 0$ these are also ex-ante $2\varepsilon^\delta$ equilibria of the games Γ^r for large enough r, excepting for the fact that these strategies do not map into A^r.

The third step approximates these σ^δ by strategies mapping into A^r and shows that they are ex-ante $3\varepsilon^\delta$ of the games Γ^r for large enough r.

Step 1: For every $\delta > 0$ there exists an approximation of σ^0, which we denote σ^δ, such that:

- $\sigma_i^\delta : T_i \to A^0$ and σ^δ is continuous,
- $\sigma^\delta \to^* \sigma^0$ (as $\delta \to 0$), and
- the σ^δ are ex-ante ε^δ equilibria of the game Γ^0, where $\varepsilon^\delta \to 0$ (as $\delta \to 0$).

Step 2: For every $\delta > 0$ there exists $r(\delta)$ such that if $r \geq r(\delta)$ then

$$E_{P^r}[u_i^r(\sigma^\delta)] \geq E_{P^r}[u_i^r(\overline{\sigma}_i, \sigma_{-i}^\delta)] - 2\varepsilon^\delta$$

for any i and measurable $\overline{\sigma}_i : T_i \to A_i$.

Step 3: For every $\delta > 0$ and $r \geq r'(\delta)$ there exists a continuous $\sigma^{\delta,r}$ such that $\sigma_i^{\delta,r} : T_i \to A^r$, $\sigma^{\delta,r} \to^* \sigma^0$, and

$$E_{P^r}[u_i^r(\sigma^{\delta,r})] \geq E_{P^r}[u_i^r(\overline{\sigma}_i, \sigma_{-i}^{\delta,r})] - 3\varepsilon^\delta$$

for large enough r and any i and measurable $\overline{\sigma}_i : T_i \to A_i$.

Proof of Step 1:

By Lemma 1, for every $\delta > 0$ there exists an approximation of σ^0 denoted σ^δ, such that $\sigma_i^\delta : T_i \to A^0$ for each i and σ^δ is continuous and is such that $d(\sigma^\delta(t), \sigma^0(t)) \leq \delta$ for all t in a set of P^0-measure of at least $1 - \delta$.[30] We note that this implies that the σ^δ are such that $\sigma^\delta \to^* \sigma^0$ (as $\delta \to 0$).[31]

Thus, to conclude the proof of Step 1 it is sufficient to show that the σ^δ are ε^δ- equilibria of the game Γ^0 and where $\varepsilon^\delta \to 0$ (as $\delta \to 0$). In other words, it is sufficient to prove the claim that: $\forall \varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that $\forall \delta \leq \delta(\varepsilon)$ and any σ_i

$$E_{P^0}[u_i^0(\sigma^\delta)] \geq E_{P^0}[u_i^0(\sigma_i, \sigma_{-i}^\delta)] - \varepsilon. \tag{1}$$

[30] Note that although we stated Lemma 1 in the context of a given i, it also holds directly for a profile of strategies, simply dropping the i notation in the proof.

[31] This convergence is also pointwise convergence almost everywhere under P^0, since here we only need to use the result of Aldaz once to get the sequence of σ^δ.

To prove the claim above, we start with the fact that σ^0 is an equilibrium of Γ^0, such that for any σ_i,

$$E_{P^0}[u_i^0(\sigma^0)] \geq E_{P^0}[u_i^0(\sigma_i, \sigma_{-i}^0)]. \tag{2}$$

Given the uniform continuity of u_i^0, there exists h such that if $d(a, a') < 1/h$ then

$$|u_i^0(a, t, s) - u_i^0(a', t, s)| < \frac{\varepsilon}{4} \tag{3}$$

for all t, s.

Let $\delta(\varepsilon) = min(1/h, \varepsilon/4)$. Using the fact that $d(\sigma^\delta(t), \sigma^0(t)) \leq \delta$ for all t in a set of P^0-measure of at least $1 - \delta \leq 1 - \varepsilon/4$ combined with (3) implies that

$$|u_i^0(\sigma^\delta(t), t, s) - u_i^0(\sigma^0(t), t, s)| < \frac{\varepsilon}{4} \tag{4}$$

and also

$$|u_i^0(\sigma_i(t), \sigma_{-i}^\delta(t), t, s) - u_i^0(\sigma_i, \sigma_{-i}^0(t), t, s)| < \frac{\varepsilon}{4} \tag{5}$$

for a P^0-measure of at least $1 - \varepsilon/4$. Note that (4) and (5) imply that

$$E_{P^0}[u_i^0(\sigma^\delta)] - E_{P^0}[u_i^0(\sigma_i, \sigma_{-i}^\delta)] \geq \left(E_{P^0}[u_i^0(\sigma^0)] - E_{P^0}[u_i^0(\sigma_i, \sigma_{-i}^0)] - \frac{\varepsilon}{2}\right)\left(1 - \frac{\varepsilon}{4}\right) - \frac{\varepsilon}{4}.$$

By (2) the right hand side is at least $-\frac{\varepsilon}{2}\left(1 - \frac{\varepsilon}{4}\right) - \frac{\varepsilon}{4}$, establishing (1).

Proof of Step 2:

We show that for each $\delta > 0$ there exists $r(\delta)$ such that if $r \geq r(\delta)$ then

$$E_{P^r}[u_i^r(\sigma^\delta)] \geq E_{P^r}[u_i^r(\overline{\sigma}_i, \sigma_{-i}^\delta)] - 2\varepsilon^\delta \tag{6}$$

for any i and measurable $\overline{\sigma}_i : T_i \to A_i$.

Fix any δ and suppose that this is not true, so that there is some i and an infinite sequence of r with a $\overline{\sigma}_i^r$ violating the condition above, such that

$$E_{P^r}[u_i^r(\overline{\sigma}_i^r, \sigma_{-i}^\delta)] > E_{P^r}[u_i^r(\sigma^\delta)] + 2\varepsilon^\delta \tag{7}$$

Choose $k(\delta)$ and $h(\delta)$ as follows: $1/k(\delta) < \varepsilon^\delta/16$, $h(\delta) \leq \delta$, and find a finite subset $\mathcal{A}_i^{h(\delta)} \subset A_i^0$ such that for every every $a_i \in A_i^r$ (including the limit $r = 0$) there exists some $a_i' \in \mathcal{A}_i^{h(\delta)}$ such that $d(a_i, a_i') < 1/h(\delta)$ when r is large enough. Also by Lemma 3 find a finite partition $\mathcal{C}_i^{k(\delta)}$ of T_i such that the diameter of every $C_i \in \mathcal{C}_i^{k(\delta)}$ is less than $1/k(\delta)$ and $\sum_{C_i \in \mathcal{C}_i^{k(\delta)}} P^0(\partial C_i) < 1/k(\delta)$. In particular, given the uniform equicontinuity of the sequence of $\{u_i^r\}$, choose $k(\delta)$ and $h(\delta)$ such that if $d(t, t') < 1/k(\delta)$ and $d(a, a') < 1/h(\delta)$ then

$$|u_i^r(a, t, s) - u_i^r(a', t', s)| < \frac{\varepsilon^\delta}{16} \tag{8}$$

for all large enough r and all s.

In addition, by condition [3b] choose $k(\delta)$ so that if $d(t_i,t_i') < 1/k(\delta)$ then for every $a_i \in \mathcal{A}_i^{h(\delta)}$ and large enough r (and $\mathcal{A}_i^{h(\delta)}$ is finite, so this is possible having fixed $h(\delta)$ already):

$$|E_{P^r}[u_i^r(a_i,\sigma_{-i}^\delta)|t_i] - E_{P^r}[u_i^r(a_i,\sigma_{-i}^\delta)|t_i']| < \frac{\varepsilon^\delta}{16} \tag{9}$$

for P^r-almost every t_i, t_i' such that $d(t_i,t_i') < 1/k(\delta)$.

Let $\widetilde{\sigma}_i^r$ be a measurable selection[32] from

$$argmin_{a_i \in \mathcal{A}_i^{h(\delta)}} d\left(a_i, \overline{\sigma}_i^r(t_i)\right).$$

Then, by (8) and (7) it follows, since $d(\widetilde{\sigma}_i^r(t_i), \overline{\sigma}_i^r(t_i)) < 1/h(\delta)$ for each t_i, that

$$E_{P^r}[u_i^r(\tilde{\sigma}_i^r, \sigma_{-i}^\delta)] > E_{P^r}[u_i^r(\sigma^\delta)] + \frac{7}{4}\varepsilon^\delta \tag{10}$$

Next, note that by (9) it follows that

$$\max_{a_i \in \mathcal{A}_i^{h(\delta)}} E_{P^r}[u_i^r(a_i,\sigma_{-i}^\delta)|C_i] > E_{P^r}[u_i^r(\tilde{\sigma}_i^r,\sigma_{-i}^\delta)|C_i] - \frac{\varepsilon^\delta}{8} \tag{11}$$

for every $C_i \in \mathcal{C}_i^{k(\delta)}$. [33] Thus, define a_i^r to be a step function on the partition $\mathcal{C}_i^{k(\delta)}$ taking on values in $\mathcal{A}_i^{h(\delta)}$, and in particular taking on a value in $argmax_{a_i \in \mathcal{A}_i^{h(\delta)}} E_{P^r}[u_i^r(a_i,\sigma_{-i}^\delta)|C_i]$ such that

$$E_{P^r}[u_i^r(a_i^r,\sigma_{-i}^\delta)] > E_{P^r}[u_i^r(\sigma^\delta)] + \frac{3}{2}\varepsilon^\delta.$$

Given that the partition $\mathcal{C}_i^{k(\delta)}$ is finite, we can find a subsequence of large enough r such that the step function is the same for those r, denoted by a_i^*.

Applying Lemma 1, there exists a continuous function $\widehat{a}_i$ such that $|a_i^*(t_i) - \widehat{a}_i(t_i)| < 1/h(\delta)$ except on a set of measure $P^0(Set) < \frac{\varepsilon^\delta}{16}$. Let $Set' = Set \cup C$ where $C = \cup_{C_i \in \mathcal{C}_i^{k(\delta)}} \partial C_i$.

[32] Such a measurable selection exists given that $d\left(a_i, \overline{\sigma}_i^r(\cdot)\right)$ is measurable for each of the finite number of a_i's. We can then order the a_i's and whenever the argmin is multivalued, select the minimum under the ordering. Then note that the inverse image of $\widetilde{\sigma}_i^r$ at some a_i is then the intersection of sets of the form $\{t_i | d\left(a_i, \overline{\sigma}_i^r(t_i)\right) < d\left(a_i', \overline{\sigma}_i^r(t_i)\right)\}$ for a_i' below a_i in the ordering, and of the form $\{t_i | d\left(a_i, \overline{\sigma}_i^r(t_i)\right) \leq d\left(a_i', \overline{\sigma}_i^r(t_i)\right)\}$ for a_i' above a_i in the ordering. These are each measurable sets, and a finite intersection of measurable sets is measurable.

[33] Let a_{t_i} be the best response for t_i from the actions in $\mathcal{A}_i^{h(\delta)}$. By (9), the difference of expected utilities for type t_i' between the action a_{t_i} and the action $a_{t_i'}$ is less than $\varepsilon^\delta/8$ when t_i and t_i' are in the same C_i, and the above inequality holds.

Then Lemma 3 implies that $P^0(Set') < \frac{\varepsilon^\delta}{8}$. Note that Set' is closed.[34] So when r is large enough on the subsequence, $P^r(Set') < \frac{3}{16}\varepsilon^\delta$. [35] Thus,

$$E_{P^r}[u_i^r(\widehat{a}_i, \sigma_{-i}^\delta)] > E_{P^r}[u_i^r(\sigma^\delta)] + \frac{5}{4}\varepsilon^\delta.$$

By weak convergence, it follows that

$$E_{P^0}[u_i^0(\widehat{a}_i, \sigma_{-i}^\delta)] > E_{P^0}[u_i^0(\sigma^\delta)] + \frac{9}{8}\varepsilon^\delta.$$

This contradicts the fact that σ^δ is a ε^δ equilibrium (relative to P^0) as shown in Step 1.

Proof of Step 3:

For each $\delta > 0$ and $r \geq r(\delta)$, we now adjust σ^δ slightly, so that it is a profile of strategies mapping into A^r.

For any $r \geq r(\delta)$, i, and $t_i \in T_i$, let

$$\Sigma_i^r(t_i) = \{a_i^r \in A_i^r | d(a_i^r, \sigma_i^\delta(t_i)) \leq 1/h(\delta)\},$$

where $h(\delta)$ is defined as in the proof of Step 2 as a function of δ. Since the sequence $\{A_i^r\}$ converges to A_i^0 in the Hausdorff metric, $\Sigma_i^r(t_i)$ is nonempty for all $t_i \in T_i$ and large enough r. Since $\Sigma_i^r(t_i)$ is a nonempty closed (by definition) subset of a compact space, it is compact. By the Kuratowski-Ryll-Nardzewski Selection Theorem, define $\widehat{\sigma}_i^{\delta,r}$ to be a measurable selection from $\Sigma_i^r(t_i)$. So $\widehat{\sigma}_i^{\delta,r}$ is a measurable mapping from T_i to A_i^r such that $d(\sigma_i^\delta(t_i), \widehat{\sigma}_i^{\delta,r}(t_i)) \leq 1/h$ for all $t_i \in T_i$. Thus by the uniform equicontinuity of the utility functions and (6) from Step 2,

$$E_{P^r}[u_i^r(\widehat{\sigma}^{\delta,r})] \geq E_{P^r}[u_i^r(\overline{\sigma}_i, \widehat{\sigma}_{-i}^{\delta,r})] - \frac{5}{2}\varepsilon^\delta$$

for large enough r and any possible $\overline{\sigma}_i$ mapping from T_i to A_i.

Again by Lemma 1, for any $\widehat{\sigma}_i^{\delta,r}$ there exists a continuous $\sigma_i^{\delta,r} : T_i \to A_i^r$ such that $d(\widehat{\sigma}_i^{\delta,r}(t_i), \sigma_i^{\delta,r}(t_i)) < 1/h(\delta)$ for all t_i in a set of P^r-measure of at least $1 - \varepsilon^\delta/8$ and also a set of P^0-measure of at least $1 - \varepsilon^\delta/8$. Then it follows that

$$E_{P^r}[u_i^r(\sigma^{\delta,r})] \geq E_{P^r}[u_i^r(\overline{\sigma}_i, \sigma_{-i}^{\delta,r})] - 3\varepsilon^\delta$$

for large enough r and any possible $\overline{\sigma}_i$ mapping from T_i to A. The proximity of the $\sigma_i^{\delta,r}$ to $\widehat{\sigma}_i^{\delta,r}$ on the P^0-measure of at least $1 - \varepsilon^\delta/8$, together with the proximity of $\widehat{\sigma}_i^{\delta,r}$ to σ_i^δ,

[34] Set' can be written as the union of $Set'' = \{t_i \notin C : |a_i^*(t_i) - \widehat{a}_i(t_i)| \geq 1/h(\delta)\}$ with C. C is clearly closed. Suppose that that there exists a sequence $t_i^{k(\delta)} \in Set'$ converging to some t_i^*. We show that then $t_i^* \in Set'' \cup C$. If t_i^* is a point of discontinuity of a_i^*, then $t_i^* \in C$ since a_i^* is a step function. If $t_i^* \notin C$ then it is a point of continuity of a_i^* and then $|a_i^*(t_i^*) - \widehat{a}_i(t_i^*)| \geq 1/h(\delta)$ since this is true along the sequence and the function is continuous at the limit point. But then it follows that $t_i^* \in Set''$.

[35] Otherwise, we can find a sequence of r with the measure greater than or equal to $\frac{3}{16}\varepsilon^\delta$. Given that Set' is closed, the limit measure should be larger than or equal to $\frac{3}{16}\varepsilon^\delta$ as well, which would be a contradiction.

and the fact that $\sigma^\delta \to^* \sigma$, implies that we can find a subsequence of δ and corresponding subsequence of $r > r'(\delta)$ for which $\sigma^{\delta,r} \to^* \sigma^0$ and these are $3\varepsilon^\delta$ equilibria where $3\varepsilon^\delta \to_\delta 0$, establishing the theorem. ∎

Proof of Theorem 1: Without loss of generality we take σ^0 to be a continuous mapping from all of T_i to A_i^0, using the variation of the Tietze Extension Theorem referenced in the proof of Lemma 1.

Since σ^0 is an equilibrium of Γ^0, it follows that for any i, for any $a_i \in A_i^0$, and P_i^0-almost any t_i^0:

$$E_{P^0}(u_i^0(\sigma^0)|t_i^0) \geq E_{P^0}(u_i^0(\sigma_{-i}^0, a_i)|t_i^0). \tag{12}$$

Since the A^rs converge to A^0, by the argument in Step 3 in the proof of Theorem 3, for any $\delta > 0$ there exists a large enough r^δ such that for all $r > r^\delta$ there exists a measurable $\sigma^r : T_i \to A_i^r$, such that for any $t_i \in T_i$, $d(\sigma^r(t_i), \sigma^0(t_i)) < \delta$. Moreover, then under the uniform equicontinuity of the utility functions, equicontinuity, for any $\gamma > 0$ there is a large enough r such that for any $a_i \in A_i$ and *every* t_i^r

$$|E_{P^r}(u_i^r(\sigma^r)|t_i^r) - E_{P^r}(u_i^r(\sigma_{-i}^r, a_i)|t_i^r)| \leq |E_{P^r}(u_i^r(\sigma^0)|t_i^r) - E_{P^r}(u_i^r(\sigma_{-i}^0, a_i)|t_i^r)| + \gamma. \tag{13}$$

To prove the theorem, it is enough to show that the σ^r's are interim ε^r equilibria with $\varepsilon^r \to 0$.

Suppose the contrary. Then there exists some ε such that for all large enough r the σ^r are not ε equilibria.

Thus, there is some i (given the finite set of agents, taking a subsequence if necessary) such that for each large enough r there is some a_i^r such that[36]

$$E_{P^r}(u_i^r(\sigma^r)|t_i^r) < E_{P^r}(u_i^r(\sigma_{-i}^r, a_i^r)|t_i^r) - \varepsilon/2$$

for a P^r-positive measure set of t_i^r.

It then follows, given the uniform equicontinuity of utility functions, and the uniform convergence of σ^r to σ^0 (see (13)), that for all large enough r

$$E_{P^r}(u_i^0(\sigma^0)|t_i^r) < E_{P^r}(u_i^0(\sigma_{-i}^0, a_i^r)|t_i^r) - \varepsilon/4 \tag{14}$$

for a P^r-positive measure set of t_i^r.

Take a fine enough finite grid of A_i^0 so that for every a_i there is an a_i^0 such that $|u_i^0(a_{-i}, a_i^0, t, s) - u_i(a_{-i}, a_i, t, s)| < \varepsilon/8$ for all a_{-i}, t, s. This can be done given the continuity of u on a compact domain, implying uniform continuity. Then mapping each a_i^r to such an a_i^0 on that grid, in (14) we can replace the a_i^rs with a sequence of corresponding a_i^0s

[36] The choice of a single a_i^r can be made via the continuity of the utility functions.

from that grid and the $\varepsilon/4$ by a $\varepsilon/8$. Then taking a subsequence, we can find some single a_i^0 such that for large enough r

$$E_{P^r}(u_i^0(\sigma^0)|t_i^r) < E_{P^r}(u_i^0(\sigma_{-i}^0, a_i^0)|t_i^r) - \varepsilon/8$$

for a P^r-positive measure set of t_i^r. However, this contradicts [3b], the convergence of the supports of the P^rs to P^0, and (12). Thus, the supposition was incorrect, establishing the result. ▮

Proof of Theorem 2: Recall that P^r is absolutely continuous with respect to the product measure $\hat{P}^0 = P_S^0 \times P_1^0 \times \cdots \times P_n^0$, with corresponding Radon-Nykodym derivative by f^r, for every r (including 0); and that $\{f^r\}_r$ converges uniformly to f^0 on all compact sets.

Let G be the support of P^0 (which is necessarily compact since a support is closed and a subset of a compact set). Since $P^0(G) = 1$ and $P^r(G) = 1$ for all r, we need only focus on the set G. Moreover, we have uniform convergence of f^r to f^0 on G.

Uniform convergence implies for any $\varepsilon > 0$, there exist a $r_1 > 0$ such that $|f^0 - f^r| < \varepsilon/2$ uniformly on G for all $r > r_1$. Thus,

$$|\int_{T_{-i}\times S} u_i^r(\sigma_{-i}(t_{-i}), a_i, s) f^r(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^0(t_j)$$

$$- \int_{T_{-i}\times S} u_i^r(\sigma_{-i}(t_{-i}), a_i, s) f^0(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^0(t_j)| < \varepsilon/2$$

for almost all types in G.

Also by the equicontinuity of the utility functions, there exist a $r_2 > 0$ such that $|u^0 - u^r| < \varepsilon/(2N)$ uniformly on G for all $r > r_2$ where N is the bound of the absolute value of f^0 on G. Then

$$|\int_{T_{-i}\times S} u_i^0(\sigma_{-i}(t_{-i}), a_i, s) f^0(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^0(t_j)$$

$$- \int_{T_{-i}\times S} u_i^r(\sigma_{-i}(t_{-i}), a_i, s) f^0(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^0(t_j)| < \varepsilon/2$$

for almost all types in G.

It follows that

$$|\int_{T_{-i}\times S} u_i^r(\sigma_{-i}(t_{-i}), a_i, s) f^r(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^r(t_j)$$

$$- \int_{T_{-i}\times S} u_i^0(\sigma_{-i}(t_{-i}), a_i, s) f^0(t_i, t_{-i}, s) dP_S^0(s) \times_{j\neq i} dP_j^0(t_j)| < \varepsilon$$

and so there cannot be an improving deviation bigger than ε for almost all types in G. Then by Corollary 1 the result follows. ▮