The Truth is in the Eye of the Beholder: or Equilibrium in Beliefs and Rational Learning in Games

Yaw Nyarko[1]
New York University
This version: September 18, 1998

ABSTRACT. Games with incomplete information or randomness in the moves of others typically have many decision-theoretically equivalent formulations of the type space. These different formulations correspond to different ways of encoding the realizations of randomizations in the type of a player. Solution concepts, assumptions or paradoxes in games should be independent of the formulation of the game used. I refer to this axiom as TIGER, for *"Type Independence among Games which are equivalently Re-formulated."* Results like convergence of beliefs to a Nash equilibrium (e.g., Jordan (1995)), obey TIGER. On the other hand, I show that results on Bayesian Learning and convergence of true play to Nash equilibrium (e.g., Kalai and Lehrer (1993a)) violate TIGER. Similarly many of the paradoxes in the learning literature (e.g., on the possibility of having optimization and prediction at the same time - Nachbar (1997)) disappear when we require TIGER to hold. The message is that in games with incomplete information, (i) the appropriate solution concept is Nash equilibrium of beliefs rather than "true" strategies, and (ii) the type-space formulation is important for other solution concepts. As regards (ii), we note that there is one formulation of the type space under which the Kalai and Lehrer (1993) assumptions always imply that type space is countable.
JEL Classification Numbers: C70, C73, D81, D82, D83, D84.
Keywords: Bayesian learning, Bayesian Games, types, repeated games

Mailing Address: New York Univ. , Dept. of Economics, 269 Mercer St. 7th fl., NYC, NY 10003.

[1]I am very grateful to the NSF and the C.V. Starr Center at New York University for their generosity. I thank Professors L. Blume, D. Easley, E. Kalai, J. Nachbar, A. Schotter and the seminar participants at Yale, NYU and Univ. of Minnesota University for helpful comments. The usual disclaimer applies.

1. Introduction

Games with incomplete information or randomness in the moves of others typically have many decision-theoretically equivalent formulations of the type space. For example, suppose you believe that your opponent is going to choose one of two actions with equal probability. Your beliefs could be "formulated" in any one of the following two ways: (i) you could believe that your opponent is randomizing with equal probability; or (ii) you could assign equal probability to your opponent being one of two types - the type who chooses the first action (non-randomly) or the type who chooses the second action. To players within the game, the different formulations of the game change neither their beliefs nor their play. These different formulations are equivalent in terms of "decision-theory" since under each beliefs about others and payoffs are unaffected. This paper will formalize this notion of equivalence of type-reformulations, and study an axiom which requires solution concepts in games to be independent of the formulation of the game used. We then use this formalization and the axiom to study some of the results in the rational learning literature.

Among game theorists there are many who argue forcefully that people do not really use mixed strategies - instead mixed strategies represent players' conjectures. Other game theorists do not mind the use of mixed strategies. These differing perspectives are really arguments over the appropriate type-formulation of the game: in the former there are many types of players each choosing a pure strategy; in the latter there may be only one type of player - one who chooses a mixed strategy.

Simple questions like "is the game in a Nash equilibrium?" or "do players learn to play a Nash equilibrium" will be impossible to answer since the answer will depend on the type-formulation under consideration. This is because there is no unambiguous notion of what constitutes the "truth." Instead it will be in the eye of the beholder (actually, the game theorist). The true type of a player is ambiguous because that player's type can encode different amounts of the outcomes of randomizations. How would one go about determining what players are "truly" doing? Since the different formulations of true player are not decision-theoretically relevant, they may be hard to test in a "laboratory" in the same manner one could possibly elicit beliefs[1].

For all of the above reasons, one may want to require solution concepts, assumptions or paradoxes in games to be independent of the formulation of the game used. This independence enables

---

[1]Note well that what we are claiming here has nothing to do with the Allais or other framing paradoxes. The Allais paradox shows that two different "frames" or ways of presenting information can lead to different behavior. Here we are considering different frames in the mind of a player that lead to the same behavior.

one to avoid taking a stand on mixed versus pure strategies, or on decision-theoretically irrelevant issues. I call this axiom TIGER, for "Type-Independence among Games which are Equivalently Re-formulated." The message of this paper will be the following: (i) if one is interested in satisfying the axiom TIGER, then one has to move away from concepts of equilibrium in strategies and move instead toward concepts of equilibrium in beliefs or conjectures; and (ii) even if one is not interested in satisfying the axiom TIGER per se, the type-space formulation is a very important and often ignored part of the specification of incomplete information games.

A review of the recent literature on Bayesian learning in repeated games highlights this issue. In the rational learning literature there are two broad approaches. The papers of Jordan (1991, '95) and later Nyarko (1994, '97b) where there is convergence of beliefs to a Nash or subjective equilibrium take the first approach. The vast majority of the literature, however, follows the second approach and includes Kalai and Lehrer (1993a) (henceforth KL93), Lehrer-Smorodinsky (1997), Nachbar (1997), Sandroni (1995a,b) and many others. The second group of papers is concerned with the convergence of true strategies to an equilibrium.

The conclusions of papers by Jordan and Nyarko, which provide results on convergence of beliefs to an equilibrium, all obey TIGER. On the other hand, to get conditions for convergence to an equilibrium, the KL93-type papers impose conditions which must hold for each true vector of types of players. Because the concept of truth is ambiguous, so too is whether the KL93 type assumptions hold for a given game. The KL93 assumptions may hold for one type-formulation of the game, yet fail in another. This is despite the fact that the formulations are decision-theoretically equivalent. On the positive side, the strength of the KL93 result is the following: if the KL93 assumption holds for a given formulation of the type-space, there will be convergence to an equilibrium whose meaning is determined by that given formulation of the type-space. I argue simply that the type-space formulation is ambiguous, and in particular that two reasonable game theorists can argue about what is the correct formulation for a given game.

This ambiguity in what constitutes the "truth" is also an issue in Nachbar (1997). Nachbar shows that in some games a paradox may arise from an inherent inconsistency between prediction of the true play and optimization. As in KL93, Nachbar's definitions (in particular his definition of prediction) requires a specification of the true strategies; it may therefore hold in one game but fail in another equivalent one. Hence the paradox in Nachbar (1997) violates the axiom TIGER. Furthermore, as we shall show, if one goes from prediction of true strategies to prediction of beliefs, TIGER will hold. All this illustrates our first message: to satisfy TIGER one should use notions of equilibrium of beliefs rather than of true strategies.

Regardless of your stand on whether the axiom TIGER should hold and equilibrium in beliefs is appropriate, my second message is that the type-space formulation is important[2]. In Section 6 below I show that for the KL93-type results to hold, the type-space formulation must be sufficiently coarse - meaning that a type should not encode too much information. Specifically, I show that for any collection of equivalent games which are linearly ordered in terms of their type-refinement - from coarsest to finest - there will exist a critical game such that all coarser games obey the KL93 type assumptions and all finer ones will violate it. Similar monotonicity results hold for the Nachbar paradox.

Related to the above, I prove a result which has been the subject of some controversy in interpreting the KL93 results. Many researchers believe that the KL93 assumptions imply the countability of the set of types - indeed, this has already appeared in the published literature! Even when a player has only two actions in each period, the set of possible infinite-horizon plays is uncountable. A restriction to countably many possible types of play is therefore a strong one. By using the framework of type-formulations, I am now able to answer the question as regards countability and the KL93 assumptions. The KL93 assumptions by themselves do not immediately imply countable types. Everything, however, depends upon the type-formulation. Suppose that we are in what we refer to as the "comprehensive" formulation of the game, where each player-type chooses a different pure strategy. Then for the KL93 assumption to hold in this formulation, the set of types must be countable.

The paper proceeds as follows: In Section 2, I provide a leading example illustrating all the issues discussed in this paper. My call for the use of equilibrium in beliefs is not new in the literature - although I believe the arguments presented here for the use of equilibrium in beliefs are novel. Section 3 comments on this literature. That section also points out the issues discussed here, as regards a certain arbitrariness in the specification of the type space, also arises in the standard definitions of a Bayesian Nash equilibrium. Section 4 contains basic notation, while Section 5 discusses type-reformulations of a game and presents the axiom TIGER. Section 6 discusses the rational learning literature. Concluding remarks are presented in Section 7. Section 8 is Appendix A and contains examples where we compute critical type formulations for the KL93 assumptions that we mentioned earlier. Appendix B in Section 9, contains all the major proofs.

2. An Example to Illustrate Everything

---

[2]For related work on type-space representations and Bayesian learning, see. Jackson et. al. (1997).

Example 2.1: Consider the following "matching pennies" stage game:

		Player B LEFT	Player B RIGHT
Player A	TOP	1,-1	-1,1
	BOTTOM	-1,1	1,-1

We shall describe two different formulations of the game, differing only in the specification of the type-space and the behavior strategies chosen. To fix the main ideas we will first discuss the one-period version of the game. We will then go to the infinite-horizon model and show that the same conclusions are obtained there, even with "learning."

Consider the following two type-formulations for the one-period model:

f1: Player A (resp. B) chooses a behavior strategy which selects actions TOP and BOTTOM (resp. LEFT and RIGHT) with equal probability. Each player knows the behavior strategy being used by the other.

f2: Let τ^A be a realization from a coin-tossing experiment where an outcome from {HEADS,TAILS} is chosen with equal probability. Hence τ^A is an element of {HEADS,TAILS}. Let τ^B be another realization from another coin-tossing experiment where an outcome from {HEADS,TAILS} is chosen with equal probability, but which is independent of the coin from which τ^A was obtained. At date 0 Player A is told of τ^A (and is not informed about τ^B) and player B is told of τ^B (and is not informed about τ^A). We may consider τ^A to be player A's "type" and τ^B to be player B's type. Suppose that each player knows how the types are drawn. Consider the following play of the game: Player A chooses TOP or BOTTOM according to as τ^A is HEADS or TAILS. Similarly, Player B chooses LEFT or RIGHT according to as τ^B is HEADS or TAILS. (A nicer story is to think of a player as being one of a pair of twins. Each twin has a birthmark which says either "HEAD" or "TAILS." The twin's birthmark is her "type." Each twin chooses an action as a function of her type or birthmark as described above. Player B is similarly one of a pair of twins. A game is an encounter between one twin of A and one twin of B.)

In f1 each player has only one possible type. In f2 each player is one of two possible types.

In f2 all we have done is to use "types" to encode the outcomes of randomizations. Alternatively, we may think of types in f2 as being used to purify the mixed (or randomized) actions in f1. Note the following:

a. Each player's belief about her opponent will assign probability 1/2 to each of the opponent's actions being played. This is true for each player-type and regardless of which of the two formulations, f1 or f2, is used. Hence, regardless of the formulation, these *beliefs of players form a Nash equilibrium.*

b. In formulation f1 it is immediate that players' true actions form a Nash equilibrium - there is only one vector of types, with each player-type mixing at each date with equal probability. In formulation f2, each player-type is choosing a pure action. Since the matching pennies game does not have a Nash equilibrium in pure strategies, the vector of actions of any vector of player-types does NOT constitute a Nash equilibrium. In particular, *in f1 "true" play is a Nash equilibrium, while in f2 it is not!* The answer to the question "are players' actions a Nash equilibrium?" cannot therefore be answered unambiguously unless a statement is made as to the formulation of the type-space being used.

c. To all intents and purposes, the two formulations represent the same "game." In particular they are decision-theoretically equivalent (and this will be made precise later). Suppose these players are happily playing in the manner described above. The players are indifferent and oblivious to the names that will be assigned to them, i.e., whether they are one of one or one of two possible types. Even though they are content, game theorists, who may disagree as to the type-formulation, may have lifelong fights as to whether or not they are playing a Nash equilibrium in strategies! The notion of what precisely is a type is completely in the mind of the modeler; players may not even be thinking in terms of types, only actions.

d. It should be clear that f1′ and f2′ are not the only formulations of the type-space to the above game. Define f3′ to be the situation where a fair coin is tossed, if it is HEADS the player plays first action with probability 2/3 and the second with probability 1/3; and if it is TAILS the player reverses the probabilities of the two actions. This is yet another decision-theoretically equivalent formulation of the game.

One may be tempted to conclude that in the infinite-horizon model all these problems disappear because of some sort of "learning." They do not. We illustrate this below and provide additional remarks pertaining to the rational learning literature. Consider the following formulations, analogous to f1 and f2:

F1: Player A (resp. B) chooses a behavior strategy which picks actions TOP and BOTTOM (resp. LEFT and RIGHT) with equal probability in <u>each period</u>, independently of the past. Each player knows the behavior strategy being used by the other.

F2: Let τ^A be a realization from infinitely many independent and identical coin-tossing experiments where an outcome from {HEADS,TAILS} is chosen with equal probability. Hence τ^A is an element of $\{\text{HEADS,TAILS}\}^\infty$. Let τ^B be another realization from an i.i.d sequence of coin-tosses, $\{\text{HEADS, TAILS}\}^\infty$, which is independent of the sequence from which τ^A was obtained. At date 0 Player A is told of the entire sequence τ^A (and is not informed about τ^B) and player B is told of τ^B (and is not informed about τ^A). We may consider τ^A to be player A's "type" and τ^B to be player B's type. Suppose that each player knows how the types are drawn. Consider the following play of the game: at date n Player A looks at the n-th coordinate of her sequence of coin-tosses - if it is a HEADS she plays her first action, TOP and if it is a TAILS she plays her second action, BOTTOM. Similarly, if the n-th element of τ^B is HEADS player B plays the action LEFT at date n and otherwise she plays action RIGHT. Suppose further that each player knows that the other is choosing actions via this rule.

Just as in the one-shot game we have the following: **(i)** Regardless of the formulation, the *beliefs of players (of each type) form a Nash equilibrium.* **(ii)** In formulation F1 players' true behavior strategies form a Nash equilibrium. In formulation F2, each player-type is choosing a pure strategy, the true strategies (or limit points of continuation true strategies) do not constitute a Nash equilibrium for the (zero discount factor) infinitely repeated matching pennies game. **(iii)** The two formulations are decision-theoretically equivalent as we mentioned in (c) above. **(iv)** There are a zillions of other equivalent type-formulations of the game, as in (d). In the infinite-horizon model, there are further items to be noted:

(v) *Absolute Continuity Conditions:* It is easy to show that the ex ante absolute continuity conditions required to get the Jordan and Nyarko convergence of beliefs results will hold here, regardless of which formulation, F1 or F2, is used. Indeed, both formulations obey the common prior assumption. In Section 6 we show, as this example illustrates, that the Jordan and Nyarko assumptions and conclusions obey our axiom TIGER. The absolute continuity conditions required in KL93 on the other hand are in ex post terms (i.e., they must hold for each vector of types). Formulation F1, with only one vector of types, can be shown to obey the KL93 assumptions (and hence the KL93 conclusion on convergence of true play). Formulation F2, with a continuum of possible types, can be shown to violate the KL93 assumptions and conclusions (beliefs assign probabilities (1/2,1/2) to each action at each date, while each player-type chooses a pure strategy). Hence the KL93 assumptions and conclusions violate our axiom TIGER.

(vi) *On Nachbar (1997):* Nachbar (1997) has pointed out that sometimes prediction of the "truth" and optimization may be in conflict. The repeated "matching pennies" game above is an example covered by the Nachbar paper (indeed it is the leading example used in that paper). His conclusion, however, depends critically upon the formulation of the game used. Optimization holds in both formulations F1 and F2 of the game. Nachbar's definition of prediction, however, only occurs in formulation F1. Hence the conflict between optimization and prediction of the truth occurs under formulation F2 but does not occur under formulation of F1. In particular, under formulation F1 the paradox disappears completely. This shows that the Nachbar paradox violates the axiom TIGER. We will show in Section 6 that there if instead we insist on prediction of beliefs then both optimization and prediction will be easily obtained, and further they will obey TIGER.

3. Some more Comments and Related Literature

It has often been suggested to me that the issues relating to the ambiguity in the definition of a type can somehow be resolved by appealing to the Harsanyi (1967,68) concept of a Bayesian Nash equilibrium (henceforth BNE). On the contrary, the problems discussed here are also relevant in the definition and the use of the BNE concept. Harsanyi's definition a Bayesian-Nash equilibrium presupposes the existence, for each player i, of a type space T_i and a mapping $\varphi_i:T_i \to F_i$ from that player's type space T_i to that player's strategy space F_i. The vector of these mappings, $\{\varphi_j\}_j$ for all the players j are assumed, in the Harsanyi's definition, to be known by each player, and given this knowledge each player maximizes expected utility. Harsanyi, however, does not give too much guidance as to what the type space should be. On one extreme a type could specify only a player's payoff function (this is related to the notion of a sparse type used in this paper). On the other extreme a type could specify an individual's payoff function and behavior strategy (and beliefs about others' payoffs and strategies, and beliefs about beliefs, etc). (This is related to the concept of a comprehensive type used in this paper.) Under the former notion of a type, the concept of a BNE provides tight predictions on the game (e.g., if the attribute vector is common knowledge the BNE becomes a Nash equilibrium). Under the latter notion of a type, the concept of BNE implies nothing other than expected payoff maximization given beliefs (so that, e.g., if the payoff matrix is common knowledge the only prediction from a BNE is that players are not using strictly dominated actions). (See Nyarko (1993) for details.) The main point of this paper is that care should be taken whenever using a model where "types" are used to model imperfect information. Precisely the same caution is required with the Harsanyi BNE!

This paper is also related to the argument over mixed versus pure strategies. Many have argued that players do not really use mixed strategies. Instead they choose certain, non-random actions at each date. Mixed strategies then become representations of the beliefs of players. This argument has been made by Harsanyi (1973), Aumann (1987) and many others who study the "decision-theoretic" approaches to game theory.[3] This argument has also been eloquently made by Binmore (1991, p.286). In this paper we provide yet another justification for the interpretation of mixed strategies as beliefs: the need for the consistency across type formulations, formalized in our main axiom TIGER.

Our work is related to the main theorems of Aumann and Brandenburger (1995), who provide epistemic conditions for ***beliefs*** or conjectures to be Nash equilibria. It should be clear that even under the hypotheses of their main theorems, it is possible for the actual play of the players not to be a Nash equilibrium even though the beliefs are. This distinction is captured in formulations F2 as opposed to F1 of Example 2.1. One could interpret the Aumann and Brandenburger (1995) paper as giving epistemic conditions for why we should study Nash equilibrium of beliefs as opposed to actual play. As suggested in this paper, working with ***beliefs*** as opposed to actual play enables one to achieve the

[3]See Aumann (1987) and Aumann-Bradenburger (1995) for further comments and references.

consistency formalized in the axiom TIGER.

4. The Repeated Game Structure

I is the *finite* set of players. The set A_i represents the *finite* set of actions available to player i at each date n=1,2,... , and $A \equiv \prod_{i \varepsilon I} A_i$. $\boldsymbol{H}^N \equiv AxAx..xA$ (N-times) is the set of histories of length N; $\boldsymbol{H}^0$ is the singleton set consisting of the null history, which we denote by h^0; $\boldsymbol{H} \equiv \cup_{n=0}^{\infty} \boldsymbol{H}^n$ is the set of all finite histories; $Z \equiv \prod_{n=1}^{\infty} A$ is the set of infinite histories or *play paths*. The projection of z ε Z onto the period n coordinate is denoted by z_n while the projection onto the coordinates of periods 1 through n is denoted by z(n). For any i in I, $z_{i,n}$ and $z_i(n)$ are the i-th coordinates of z_n and z(n) respectively. Perfect recall is assumed: when player i is choosing her date n+1 action she will know the date n history z(n).

Given any metric space X, we let $\mathcal{P}(X)$ denote the set of probability measures defined over (Borel) subsets of X. Unless otherwise stated the set $\mathcal{P}(X)$ will be endowed with the weak topology. Given any fεF, ν(f) ε $\mathcal{P}(Z)$ denotes the probability distribution over *Z* induced by f. The set of *behavior strategies* for player i is the set $F_i \equiv \{f_i:\boldsymbol{H} \rightarrow \mathcal{P}(A_i)\}$ and $F \equiv \prod_{i \varepsilon I} F_i$. The space F is endowed with the topology of weak convergence. A *pure strategy* for i is any behavior strategy which takes values on the vertices of $\mathcal{P}(A_i)$.

Each player i in I has an *attribute vector* which is some element θ_i of the set Θ_i, assumed to be a *compact* subset of a finite dimensional Euclidean space. $u_i:\Theta_i xA \rightarrow \mathbb{R}$ is player i's within-period *continuous and bounded* utility function which depends upon her attribute vector, θ_i, as well as the vector of actions, a ε A, chosen by the players. Each player i knows her own attribute vector θ_i but does not necessarily know those of other players, θ_j for $j \neq i$. Player i has a discount factor which is a known *continuous* function, $\delta_i:\Theta_i \rightarrow [0,1)$, of the player i's attribute vector. Specifically, $\Theta_i = \Theta_i^{\#} x[0,1)$, where $\Theta_i^{\#} \subseteq \mathbb{R}^{Ai}$ represents the set of stage game payoff vectors and where $\delta_i:\Theta_i \rightarrow [0,1)$ is the projection of Θ_i onto its second coordinate, and represents the discount factor. Any set of this form will be called an attribute vector space for i. Define $U_i:\Theta_i xZ \rightarrow \mathbb{R}$ by $U_i(\theta_i,z) \equiv \sum_{n=1}^{\infty} [\delta_i(\theta_i)]^{n-1} u_i(\theta_i,z_n)$ and $V_i:\theta_i xF \rightarrow \mathbb{R}$ by $V_i(\theta_i,f) \equiv \int_Z U_i(\theta_i,z)d\nu(f)$.

Each player i is characterized by a type, τ_i, which specifies, among other things, that player's attribute vector, θ_i. A *type-space for i* is any set $T_i = \Theta_i x T_i^{\#}$ where Θ_i is an attibute vector space and $T_i^{\#}$ is a complete and separable metric space. (This definition allows $T_i = \Theta_i$.) We let $\theta_i(\tau_i)$ denote the attribute vector of player i of type τ_i, so that $\theta_i(.)$ is the projection of T_i onto Θ_i. A type-space will be any $T \equiv \prod_{i\varepsilon I} T_i$, where T_i is a type-space for i.

If X is a cartesian product X=YxZ and $\eta \, \varepsilon \, \mathcal{P}(X)$, we denote by $\mathrm{Marg}_Y \, \eta$ the marginal of η on Y. Define[4]

$$\Omega \equiv T \times F \times Z \quad . \tag{1}$$

We shall say that a probability measure $\mu \varepsilon \mathcal{P}(\Omega)$ <u>*respects ν*</u> if $\mathrm{Marg}_Z \, \mu(.|\tau,f) = \nu(f)$ for μ-a.e. $(\tau,f)\varepsilon TxF$. An <u>*ex ante subjective belief*</u> for a player i is any probability μ_i over Ω which respects ν. To see how an ex ante subjective belief may be obtained, we note here that it constructed from the following three building blocks:

(i) *(Beliefs about others)* Each player-type τ_i has some belief over the types and behavior strategies of others. This defines $\mathrm{Marg}_{T_{-i} x F_{-i}} \, \mu_i(.|\tau_i)$.

(ii) *(Own Behavior Strategies)* Each such player-type τ_i chooses some behavior strategy $f_i(\tau_i)$. This defines $\mathrm{Marg}_{F_i} \, \mu_i(.|\tau_i)$, which assigns point mass to $f_i(\tau_i)$.

(iii) *(Ex ante distribution of types)* Each player i has an ex ante distribution over the set of own-types in T_i. This defines $\mathrm{Marg}_{T_i} \, \mu_i$.

The product of the measures in parts (i) and (ii) produces a measure over $T_{-i}xF$ conditional on τ_i. Combining this with (iii) results in a measure over TxF. Using the measure ν then results in a measure over Ω. This is player i's ex ante subjective belief. Part (iii) is required primarily for the measure-theoretic technicalities. One may object to the requirement in (i) on the grounds that players should be expected to have beliefs over the behavior strategies of others and not over the types of others. This is a valid objection. We discuss this in Section 5, and argue that this objection supports the use of our main axiom, TIGER.

If η_k is a probability measure on a complete and separable metric space X_k,for k=1,2,...K, we let $\otimes_{k=1}^K \eta_k$ denote the product of the measures $\{\eta_k\}_{k=1}^K$ on the cartesian product $\prod_{k=1}^K X_k$. We impose the following assumptions on μ_i:

Assumption 4.1

---

[4]The cartesian product of metric spaces will always be endowed with the product topology.

a. (*independence of strategies*) $\text{Marg}_F\, \mu_i(.|\tau) = \prod_{j\varepsilon I} \text{Marg}_{F_j}\, \mu_i(.|\tau_j)$ for μ_i a.e. $\tau=\{\tau_j\}_{j\varepsilon I}$.

b. (*independence of types*) $\text{Marg}_T\, \mu_i = \prod_{j\varepsilon I} [\text{Marg}_{T_j} \mu_i]$.

Assumption 4.1 says that player i believes that (a) conditional on players' types, players choose behavior strategies independently and (b) that players' types are drawn independently. The set of beliefs of players that we allow is therefore the set

$$B(\Omega) = \{\mu \,\varepsilon\, \mathcal{P}(\Omega) | \mu \text{ respects } \nu \text{ and satisfies Assumption 4.1 (with } \mu=\mu_i)\}. \qquad \textbf{(2)}$$

Given any sets $\{X_i\}_{i\varepsilon I}$ and $\{Y_i\}_{i\varepsilon I}$, and any functions f_i:$X_i \rightarrow Y_i$ for iεI, we define $X \equiv \prod_{i\varepsilon I} X_i$, $X_{-i} \equiv \prod_{j\neq i} X_j$, and $f_{-i}(x_{-i}) \equiv \prod_{j\neq i} f_j(x_j)$. For each i$\varepsilon$I, define the equivalence relation, $\sim^{\mathbf{k}}$, on F_i as follows: $\forall f_i$ and $f_i' \,\varepsilon\, F_i$, $f_i \sim^{\mathbf{k}} f_i'$ if $\forall f_{-i} \,\varepsilon F_{-i}$, $\nu(f_i,f_{-i}) = \nu(f_i',f_{-i})$. Let $F_i^{\sim \mathbf{k}}$ denote the set of equivalence classes of $\sim^{\mathbf{k}}$. From Kuhn's (1953) Theorem (see Aumann (1964) for the infinite-horizon case) $\exists \kappa_i$:$\mathcal{P}(F_i) \rightarrow F_i^{\sim \mathbf{k}}$ such that for any (behavior strategy mixture) $\varphi_i \,\varepsilon \mathcal{P}(F_i)$, $\forall$ f_i $\varepsilon\kappa_i(\varphi_i)$ and $\forall f_{-i} \varepsilon F_{-i}$, the probability distribution on Z induced by φ_i and f_{-i} is equal to $\nu(f_i,f_{-i})$. The behavior strategy $f_i \varepsilon \kappa_i(\varphi_i)$ is said to be ***realization equivalent*** to the mixed strategy φ_i. We refer to f_i as the ***KSR*** (for "Kuhn Strategic Representation") of φ_i. We shall write $f_i = \kappa_i(\varphi_i)$ when we mean that "f_i is any member of the equivalence class of $\kappa_i(\varphi_i)$."

Fix any subjective belief $\mu_i \,\varepsilon\, B(\Omega)$ of player i. Define

$$f_j^i = \kappa_i(\text{Marg}_{F_j}\, \mu_i(.)) \;\forall i,j \quad \text{and} \quad f^i \equiv \{f_j^i\}_{j\varepsilon I}. \qquad \textbf{(3)}$$

Under the independence assumption 4.1, each type of player i (μ_i -a.e) has the same belief about player $j\neq i$ equal to the marginal of $\mu_i(.)$ on F_j. Hence for $i\neq j$, f_j^i is the KSR of the belief of each type of player i about player j's behavior strategy. Next define

$$f^i(\tau_i) \equiv \{f_i(\tau_i), f^i_{-i}\} . \qquad \textbf{(4)}$$

The only difference between f^i in **(3)** and $f^i(\tau_i)$ in **(4)** is in the i-th coordinate - the latter conditions on the type τ_i while the former does not. The behavior strategy profile $f^i(\tau_i)$ is the KSR of player-type τ_i's beliefs. The behavior strategy vector f^i is the KSR of someone who, like player i, has beliefs given by

μ_i , but who, unlike player i, does not observe the own-type τ_i. (To illustrate the difference consider Example 2.1. Let $\underline{f}_{A,0.5}$ and $\underline{f}_{B,0.5}$ denote the strategies chosen by A and B in formulation F1 - randomize at each date with equal probability - and define $\underline{f}_{0.5}=(\underline{f}_{A,0.5}, \underline{f}_{B,0.5})$. In both formulations F1 and F2, $f^i = \underline{f}_{0.5}$ for all i. In formulation F1, which has only one vector of types, $f^i(\tau_i) \equiv \underline{f}_{0.5}\ \forall i$; in formulation F2, $f^i(\tau_i)=\{f_i(\tau_i), \underline{f}_{j,0.5}\}$ where $j\neq i$ and $f_i(\tau_i)$ is the pure strategy determined by τ_i ε {HEADS,TAILS}$^\infty$ as described in example 2.1.)

Fix a collection of ex ante beliefs $\{\mu_i\}_{i\varepsilon I}$ and let $\{f_i(\tau_i)\}_{i\varepsilon I}$ denote the behavior strategies chosen by the vector of player-types $\tau=\{\tau_i\}_{i\varepsilon I}$. We now define a distribution μ^* ε B(Ω), *the outside observer's belief induced by* $\{\mu_i\}_{i\varepsilon I}$. In the literature this is often defined as the "true" distribution, and is the measure with respect to which theorems are proved. The measure μ^* is defined to be the unique element of B(Ω) such that $\mathrm{Marg}_{TxF}\ \mu^* = \otimes_{i\varepsilon I}\ \mathrm{Marg}_{T_i x F_i}\ \mu_i$. In particular, μ^* is the ex ante belief of an outside observer who knows how players choose behavior strategies as a function of their types, and for each i in I has the same ex ante beliefs about each player i's own-types as player i herself. Define f^* to be the KSR of μ^* *not conditioning on types* (and observe that f^* below is the same as f^i in **(3)** above except that μ^* replaces μ_i):

$$f_j^* \equiv \kappa_i(\mathrm{Marg}_{F_j}\ \mu^*)\ \forall j\varepsilon I \qquad \text{and} \qquad f^* \equiv \{f_j^*\}_{j\varepsilon I}. \qquad \textbf{(5)}$$

Of course the KSR of μ^* conditioning on types $\tau=\{\tau_i\}_{i\varepsilon I}$ is precisely the vector $\{f_i(\tau_i)\}_{i\varepsilon I}$ of behavior strategies chosen by those player-types.

5. Equivalent Type-Reformulations

5.1. Consider as fixed the following: the set of players I, the attribute vector space, Θ, and the space of actions A (and hence the spaces F and *Z*). The set **G** below is the set of all "games" g which can be constructed *given the primitives* {I, Θ,A}:

$$\mathbf{G} \equiv \Big\{ g= \langle T, \{\mu_i\}_{i\varepsilon I}, \mu^* \rangle \;|\; T \text{ is a type-space; } \{\mu_i\}_{i\varepsilon I}\ \varepsilon \prod_{i\varepsilon I} B(\Omega) \text{ where } \Omega \equiv TxFxZ\text{; and } \mu^* \text{ is the outside observer belief induced by } \{\mu_i\}_{i\varepsilon I} \Big\}. \qquad \textbf{(6)}$$

Definition 5.1 (Equivalent type-Refinements). Fix any pair of tuples $g=\langle T,\{\mu_i\}_{i\varepsilon I},\mu^*\rangle$ and $g'=\langle T',\{\mu'_i\}_{i\varepsilon I},\mu^{*\prime}\rangle$ in **G**. We say that g′ is an equivalent type-refinement of g (or g is an equivalent type-coarsening of g′) and we write $g' \geq g$, if there exist sets $\{\Gamma_i\}_{i\varepsilon I}$ (each a complete and separable metric space) such that $\forall i\varepsilon I$, $\tau_i \varepsilon T_i$, and $\gamma_i \varepsilon \Gamma_i$,

i. (*type-refinement*) $T'_i = T_i x \Gamma_i$;

ii. (*same own payoffs*) $\theta_i(\tau_i') = \theta_i(\tau_i)$ when $\tau_i' = (\tau_i, \gamma_i)$;

iii. (*same beliefs about others*) $\text{Marg}_{F_{-i}}\ \mu_i'(.|\tau_i' = (\tau_i, \gamma_i)) \equiv \text{Marg}_{F_{-i}}\ \mu_i(.|\tau_i)$;

iv. (*same expected own play*) $\text{Marg}_{F_i}\ \mu_i'(.|\tau_i) \equiv \text{Marg}_{F_i}\ \mu_i(.|\tau_i)$; and

v. (*same distribution of types*) $\text{Marg}_{T_i}\ \mu_i' = \text{Marg}_{T_i}\ \mu_i$.

Definition 5.2: $g\varepsilon$ **G** is an <u>*equivalent type-reformulation*</u> of $g'\varepsilon$ **G** if there exists a $\hat{g}\ \varepsilon$ **G** such that either (i) $\hat{g}$ is an equivalent type-refinement of both g and g′ or (ii) $\hat{g}$ is an equivalent type-coarsening of both g and g′. We then say that g and g′ are <u>*equivalent*</u>, and we write $g \sim g'$.

The equivalences in (iii) and (iv) of Definition 5.1 are in the sense of "Kuhn equivalences" defined in Section 4. If $g' \geq g$, a type τ_i in g is sub-divided into other types, with generic member (τ_i, γ_i). Parts (ii) and (iii) require that the new type have the same attribute vector and belief about others as the original ones. This implies that if each player is optimizing then conditional on τ_i , each player-type (τ_i, γ_i) is indifferent between her own play and the play of player-type (τ_i, γ_i'). Part (iv) requires that after "integrating out" the γ_i's we obtain the same original play. This shows that the role of the type-refinement is to encode in the new type some information that may be used to pick a realization from what was originally a randomization over actions. In summary, if g and g′ are two equivalent games, player-types (τ_i, γ_i) and τ_i will have the same payoffs and same beliefs and the games g and g′ only differ in the fact that the γ_i's are used in encoding the outcomes of realizations of randomizations. I therefore interpret the games g and g′ as "decision-theoretically" equivalent.

The notion of equivalent type-refinements results in the partial ordering, $\geq$, on **G**. A minimal "refinement", defined as "sparse" below, is one where a type is the same as an attribute vector, so that the type has absolutely no realizations of randomizations encoded in it. A "maximal" refinement on the other hand, defined as "comprehensive" below, would require each player to do all the randomizations at date 0 and encode them entirely in his/her type. In Example 2.1, formulation F1 is a sparse formulation, while

F2 is a comprehensive formulation.

Definition 5.3: A $g= \langle T,\{\mu_i\}_{i\varepsilon I},\mu^*\rangle$ in **G** is a _sparse_ formulation if $T_i = \Theta_i\ \forall i\varepsilon I$, and is a _comprehensive_ formulation if each (or μ^*-almost every) player-type in g is choosing a pure strategy.

One can think of integrating out the realizations of any randomizations which are encoded in a type. For example, suppose that there are two possible types of Player A, (θ_A, γ') and (θ_A, γ''),each with the same attribute vector θ_A and each occurring with probability 1/2. This can be collapsed into one type by integrating out the γ's in the following manner: Consider there being only one type, called θ_A , and suppose that this player-type chooses the (KSR of the) mixed strategy which assigns probability 1/2 to the behavior strategy chosen by player-type (θ_A, γ') and probability 1/2 to that chosen by player-type (θ_A, γ') . This operation shows how to construct a sparse formulation for any given game. Next, one can think of an operation going the other way where one encodes in a type the realizations of all randomizations. This will produce from any game g an equivalent game which is comprehensive. We therefore obtain the following:

Proposition 5.1: Fix any game g in **G**. $\exists$ $\underline{g}$ ε**G** and $\overline{g}$ ε**G** such that $\underline{g} \sim g \sim \overline{g}$, $\underline{g}$ is sparse and $\overline{g}$ is comprehensive.
Proof: The proof of this and all other main results appear in Appendix B.■

We now state our axiom TIGER. A _property_ is any statement pertaining to games which, in any given game, may be true or false. In particular, a property is a binary relation b:**G** → {0,1}, which assigns a value "true" or "false" to each game. All our numbered definitions in Section 6 below implicitly define a property for games. An "assumption", "conclusion" or "paradox" is a property of games, which may or may not hold in any particular game. The axiom TIGER below will be a requirement on properties of games:

Axiom of Type-Independence among Games which are Equivalently Re-formulated , TIGER: *A property of games obeys the axiom TIGER if for each game g ε* ***G*** *the property holds for g if and only if it also holds for all other games g′ε* ***G*** *which are equivalent to g.*

The axiom TIGER is a requirement not only among games which are ordered by $\succeq$, but rather among all games which are equivalent. The following, however, is immediate:

Proposition 5.2: Let b be a property of games and suppose that b is true in game g ε **G** if and only if it is true for all other games g′ ε **G** for which either g $\succeq$ g′ or g′ $\succeq$ g. Then b satisfies TIGER.

5.2. Beliefs over the Types of Others?

We now return to an issue brought up in Section 4. One may argue that it is not appropriate to model players as having beliefs over the types of others. All that should be important is a player's beliefs about the strategies that will be used by other players. So, suppose that each player i is not characterized by a μ_i in $B(\Omega)$, but rather by the following three components: (i) beliefs about the strategies of others, $\bar{f}_{-i}^{i} \varepsilon \mathcal{P}(F_{-i})$, common to all types of player i; (ii) the behavior strategy $f_i(\tau_i)$ chosen by each player type $\tau_i \varepsilon T_i$; (iii) the ex ante distribution over the possible own-types of player i in T_i. The above three components do not specify player i's beliefs about the types of other players τ_{-i} in T_{-i}: there are many joint distributions over $T_{-i}xF_{-i}$ for which the KSR of the marginal on F_{-i} is $\bar{f}_{-i}^{i}$. Given two games $g=<T,\{\mu_i\}_{i\varepsilon I},\mu^*>$ and $g'=<T,\{\mu'_i\}_{i\varepsilon I},\mu^{*\prime}>$ in **G** on the same type space , let us say that g and g′ are strategically equivalent and write $g \sim^{\mathbf{S}} g'$ if they share the same components (i)-(iii) above - that is, if for all i in I, (i) $Marg_{F_{-i}}\ \mu_i = Marg_{F_{-i}}\ \mu_i'$; (ii) $Marg_{F_i}\ \mu_i(.|\tau_i) = Marg_{F_i}\ \mu_i'(.|\tau_i)$ for all τ_i (μ_i-a.e); and (iii) $Marg_{T_i}\mu_i = Marg_{T_i}\mu_i'$. Both g and g′ are in some sense valid representations of the interaction between the players. After all, why should it be important for player i to be correct in specifying player j's *type*? After we have specified i's beliefs about j's behavior strategy, knowledge by i of j's type should be "decision-theoretically" irrelevant in some sense. This sense is captured by the equivalences of Definition 5.1 above. In particular, since Definition 5.1 does not make reference to i's beliefs about j's type space for j≠i, it is easy to see that whenever $g \sim^{\mathbf{S}} g'$ then $g \sim g'$ in the sense of Definition 5.1[5]. (The converse of course is immediate.) This therefore provides further justification for the use of the axiom TIGER: two games which are strategically equivalent should be considered "equivalent" so TIGER must be imposed on any property we will use for such games!

---

[5]Formally, define $\hat{\mathbf{T}} = Tx\Gamma$ with Γ a trivial singleton set and construct $\hat{g}$ in **G** from g by extending g onto $\hat{\mathbf{T}}$ in the obvious manner. It is then immediate that $\hat{g} \succeq g$ and $\hat{g} \succeq g'$, so $g \sim g'$.

6. Rational Learning in Games

In most of the rational learning literature, there are three basic parts: the first assumes that players are optimizing; the second part imposes an absolute continuity assumption over beliefs; the third part then shows that these assumptions imply convergence to some sort of equilibrium. We now consider the three parts formally. The message of this section will be the following: (i) for the conclusions of the rational learning literature to satisfy TIGER, we will have to insist on statements pertaining to equilibrium in beliefs as opposed to equilibrium in strategies; and (ii) for many of the assumptions and conclusions which violate TIGER, they hold if the type space is sufficiently coarse and are violated whenever the type space is sufficient refined (or vice versa).

6.1. Optimization

Definition 6.1: (*optimization*) The game $g=<T,\{\mu_i\}_{i\varepsilon I}, \mu^* > \varepsilon$ **G** satisfies optimization if $\mu_i(M_i \cup M_i^0)=1$ $\forall i \varepsilon I$, where

$$M_i \equiv \{(\tau,f,z)\varepsilon\Omega: \delta_i(\theta_i(\tau_i))>0 \text{ and } f_i \text{ maximizes } \int_{F_{-i}} V_i(\theta_i(\tau_i), f_{-i}', \cdot)d\mu_i(.|\tau_i)\} \text{ and}$$
$$M_i^0 \equiv \{(\tau,f,z)\varepsilon\Omega: \delta_i(\theta_i(\tau_i))=0 \text{ and } \forall n,\ z_{i,n+1} \text{ maximizes } \int_{A_{-i}} u_i(\theta_i(\tau_i), z_{-i,n+1}, \cdot)d\mu_i(.|z(n),\tau_i)\}.$$

where the integration in the definition of M_i is over $f_{-i}' \varepsilon F_{-i}$.

Definition 6.1 requires each player i to be maximizing her subjective expected discounted sum of utilities with μ_i probability one. Whenever the discount factor is equal to zero (i.e., on the set M_i^0 above) player i will be required to maximize her expected utility at each date. Under equivalent type-reformulations, players' beliefs and payoff functions do not change, so the following should be immediate:

Proposition 6.1: Definition 6.1, optimization, obeys TIGER.

6.2. The Absolute Continuity Assumptions. Given any two probability measures μ' and μ'' on some measure space X, μ' is absolutely continuous with respect to μ'' if for all measurable subsets $D\subseteq X$, $\mu'(D)>0$ implies that $\mu''(D)>0$. We then write $\mu' \ll \mu''$. μ' and μ'' are mutually absolutely continuous if $\mu' \ll \mu''$ and $\mu'' \ll \mu'$. Fix a game $g=<T, \{\mu_i\}_{i\varepsilon I}, \mu^* > \varepsilon$ **G**, and consider the following definitions:

Definition 6.2 ***(CPA)*** **:** g obeys the *Common Prior Assumption (CPA)* if and only if $\mu_i = \mu^* \ \forall i \,\varepsilon I$.

Definition 6.2* ***(CPA*)*** **:** g obeys the *Common Prior Assumption (CPA*)* if and only if $\forall i \,\varepsilon I$, $\text{Marg}_{F}\ \mu^* = \text{Marg}_{F}\ \mu_i$, $\forall i \varepsilon I$.

Definition 6.3 ***(GGH)*****:** g obeys the *generalized Harsanyi consistency condition* (GGH), (or ex ante absolute continuity) if and only if $\text{Marg}_{T_i \times Z}\ \mu^* \ll \text{Marg}_{T_i \times Z}\ \mu_i$, $\forall i \varepsilon I$.

Definition 6.4 ***(KL-T)*****:** g obeys *(KL-T)* (or ex post absolute continuity) if and only if $\mu^*(T_{KL\text{-}T})=1$ where $T_{KL\text{-}T} \equiv \cap_{i \varepsilon I}\{\tau=(\tau_i,\tau_{-i})\varepsilon T \,|\, \text{Marg}_{Z}\ \mu^*(.|\tau) \ll \text{Marg}_{Z}\ \mu_i(.|\tau_i)\}$.

The learning results of Jordan (1991,95) use the common prior assumption, *(CPA)*. Condition (GGH)[6] is used in the learning results of Nyarko (1994 and 1997b). Assumption *(KL-T)* is the natural extension of the KL93 assumption to the model with many types, and delivers the KL93 conclusions for μ^*-almost every τ. It is easy to see that the common prior assumption implies but is not implied by condition *(GGH)*. Indeed, *(CPA)* is strictly stronger than *(GGH)*. Further, condition *(KL-T)* implies but is not implied by condition *(GGH)* (for a proof that *(KL-T)* implies *(GGH)* see Nyarko (1997b)). Example 2.1 (formulation F2) shows that *(GGH)* can be true while *(KL-T)* fails, so *(KL-T)* is strictly stronger than *(GGH)*.

The common prior assumption *(CPA)* (as opposed to *(CPA*)*) violates TIGER. The problem is that *(CPA)* requires each player to get correct the mapping from each other player's type to that player's behavior strategy. For example fix a game g with two types of each player and suppose that g obeys *(CPA)*. Define another game exactly the same as the first except that in the new game Player A mis-labels the types of B . In particular, if $\bar{\tau}_B$ and $\hat{\tau}_B$ are the two types of player B and $f(\bar{\tau}_B)$ and $f(\hat{\tau}_B)$ are the behavior strategies they use, then in game g′ Player A incorrectly believes player-type $\bar{\tau}_B$ is using $f(\hat{\tau}_B)$ and player-type $\hat{\tau}_B$ is using $f(\bar{\tau}_B)$. It should be clear that the re-labeling is "harmless" and that g and g′ are equivalent games. (Indeed, this was precisely the discussion at the end of Section 5.) The game g,

[6] Elsewhere (Nyarko (1997a)) I have defined Condition (GH) "for generalized Harsanyi common prior assumption" to be where μ_i and μ_j are mutually absolutely continuous with respect to each other. (GGH) above generalizes this latter condition (hence the name "GGH") by first requiring merely absolute (and not mutual absolute) continuity, and this with respect to marginals.

however, obeys *(CPA)* while g′ violates it. Hence (*CPA)* violates TIGER. For this reason *(CPA)* is really not a good assumption to place on a game. Assumption *(CPA*)* gets around this problem: *(CPA*)* obeys TIGER and implies the Jordan convergence results which were originally proved with CPA (see Propositions 6.2 and 6.5 below) .

If one distribution is absolutely continuous with respect to another distribution over a product space, then the same will be true of their marginal distributions. Coarsening a type-space is similar to the operation of going from a joint to a marginal distribution. It is therefore not surprising that if condition *(GGH)* (resp. *(KL-T)*) holds in a game and we coarsen the type-formulation, condition *(GGH)* (resp. *(KL-T)*) will continue to hold. The difficulty lies in going from one type-formulation to a finer one. Indeed, Example 2.1 shows that *(KL-T)* may hold in one game (as in F1) but may be violated when the type-formulation is refined (as in F2)- so *(KL-T)* violates TIGER. This does not happen with *(GGH)*, however, and in particular *(GGH)* obeys TIGER. This is because player i's beliefs about others are independent of i's type, and i's type is the only type used in the statement of *(GGH)* - in particular, under *(GGH)* player-type τ_i only has to be "correct" (in terms of absolute continuity) about the average play of others; and beliefs about average play do not change with i's type or with the type formulation. Similar reasoning shows why assumption *(CPA*)* obeys TIGER. Condition *(KL-T)* on the other hand conditions on the true <u>vector</u> of types, which will change as the types, and hence type-formulations, change. In particular, under *(KL-T)* player-type τ_i has to be "correct" (in terms of absolute continuity) about the true play of others, and this true play will change with the type formulation. The ambiguity of the notion of what is a true type leads to the failure of *(KL-T)* to satisfy TIGER. Formally we have the following two results:

Proposition 6.2: (a) *(CPA*)* obeys TIGER; **(b)** condition *(GGH)* obeys TIGER; **(c)** condition *(KL-T)* violates TIGER; and **(d)** *(CPA)* violates TIGER.

Proposition 6.3 (Monotonicity in *(KL-T)*): Fix any g in **G** and suppose that g obeys *(KL-T)*. Then all equivalent type-coarsenings of g also obey *(KL-T).*

Proposition 6.3 implies that given any chain of equivalent games linearly ordered by $\geq$ from coarsest to most refined[7], there will exist a critical game such that all coarser ones obey *(KL-T)* and all finer ones violate it (or else either all obey or all violate *(KL-T))*. Examples 8.1 and 8.2 in Appendix A compute such critical games in two different chains of games. Example 8.1 provides a chain of games g_1, g_2, ..., g_∞ where g_k for $k<\infty$ has a countable type-space and g_∞ an uncountable type-space. The game g_k will satisfy *(KL-T)* for $k<\infty$ and will violate it for $k=\infty$. This suggests a connection between condition *(KL-T)* and the countability of the type-space. Indeed, we have the following:

Proposition 6.4 *((KL-T) and the Countability of the Type-space)***:** Fix any game $g=<T,\{\mu_i\}_{i\varepsilon I},\mu^*>$ in **G** with at least two players. Suppose g obeys *(KL-T)* and is comprehensive. Then the set of plays is countable: i.e., there exists a ***countable*** subset Φ of *Z* such that the event $\{\omega=(\tau,f,z) \text{ in } \Omega \mid z\ \varepsilon\ \Phi\}$ has μ^* and μ_i probability one $\forall$i in I. If, in addition, each vector of types results in a different play path, then the type-space must be countable.

In F2 of Ex. 2.1 each player-type chooses a pure strategy at each date. Proposition 6.4 therefore implies that (KL-T) cannot hold for that example. When the type-formulation is not comprehensive, so each type does not choose a pure strategy at each date, it is possible to have a model with uncountably many types which satisfy the Kalai and Lehrer assumption[8]. Proposition 5.1, however, then shows that the "game" can always be equivalently reformulated so that in the new game each player-type chooses a pure strategy. In that case, as Example 2.1 indicates, the notion of exactly what a type is becomes ambiguous, as is whether or not assumption (KL-T) holds. We note that this result relies critically on the independence assumptions used.[9]

---

[7]By "linearly ordered by ...." we mean that we can index the games by k in [0,1] such that $k\geq k'$ implies $g_k \geq g_{k'}$ (or vice versa).

[8]A very nice example to this effect was independently provided to me by R. Smorodinsky (1995) while this paper was undergoing revision.

[9]Suppose that there are two players each with a type in [0,1], but that the set of possible vectors of types is the diagonal on the unit square. Suppose each player-type chooses a different pure strategy, and beliefs are such that conditional on a player observing her own type, she knows perfectly the type of her opponent (which is equal in value to her's). Then KL-T holds with an uncountable set of types, a violation of Proposition 6.4.

6.3. The Convergence Results of Jordan, KL93 and Nyarko. The norm $\|.\|$ denotes the total variation norm on $\mathcal{P}(Z)$; i.e., given p, q$\varepsilon\mathcal{P}(Z)$, $\|p-q\| \equiv \text{Sup}_E |p(E)-q(E)|$, where the supremum is over Borel measurable subsets E of Z. Fix any attribute vector $\theta \equiv \{\theta_i\}_{i\varepsilon I}\, \varepsilon\Theta$. Define $\forall i\varepsilon I$,

$$N_i(\theta_i) \equiv \{f=\prod_{j\varepsilon I} f_j \varepsilon F\text{: } f_i \;\varepsilon\; \text{argmax } V_i(\theta_i,f_{-i},\cdot)\} \quad \text{and} \quad N(\theta) \equiv \cap_{i\varepsilon I} N_i(\theta_i)\;; \tag{7}$$

$$ND(\theta) \equiv \{\nu' \;\varepsilon\mathcal{P}(Z)\text{: } \exists\; f\varepsilon N(\theta) \text{ with } \nu' = \nu(f)\}. \tag{8}$$

$$SE_\varepsilon(\theta) \equiv \{f=\prod_{j\varepsilon I} f_j\varepsilon F\text{: } \forall i\varepsilon I,\ \exists (f_j^i)_{1\le j\le n}\; \varepsilon\; F \text{ with } f_i^i = f_i \text{ such that}$$

$$\text{(i) } f_i \;\varepsilon \text{ argmax } V_i(\theta_i, f^i_{-i}, \cdot) \text{ and (ii) } \|\nu(f) - \nu(f^i)\| \le \varepsilon\} \text{ and } SE(\theta) \equiv SE_0(\theta). \tag{9}$$

N(θ) is the set of Nash equilibrium behavior strategy profiles for the complete information game with attribute vector θ. ND(θ) is the set of all distributions of play that can be generated by some Nash equilibrium behavior strategy profile. $SE_\varepsilon(\theta)$ (resp. SE(θ)) is the set of subjective ε-equilibria (resp. subjective equilibria). One can show that ND(θ) is equal to the set of all distributions induced by some f ε SE(θ). (The definitions on subjective equilibria are taken from Kalai-Lehrer (1993b). See also Battigali et. al. (1988) and (1992).)

Given any history h ε ***H***, and any behavior strategy $f_i \,\varepsilon\, F_i$ for iεI, define the continuation strategy $f_{i,h}$ as follows: $\forall h' \,\varepsilon$ ***H***, $f_{i,h}(h') = f_i(hh')$ where hh′ is the concatenation of h and h′. Analogously, given $f=\{f_i\}_{i\varepsilon I}$ ε F and h ε ***H***, define $f_h \,\varepsilon$ F by $f_h(h') = f(hh')\ \forall h' \,\varepsilon$ ***H***. Given a sequence $\{x_n\}_{n=1}^\infty$ in some metric space X and a set $\chi \subseteq X$, write $x_n \to^c \chi$ if every cluster point of $\{x_n\}_{n=1}^\infty$ lies in the set χ.

Fix a game g in **G** and let f^i, $f(\tau)=\{f_i(\tau_i)\}_{i\varepsilon I}$ and f^* be as in eq.'s **(3) - (5)**. When the common prior assumption holds, either *(CPA)* or *(CPA*)*, $f^i = f^j = f^*$ $\forall$i and j. Define

$$\Omega_{Jordan} \equiv \{\omega=(\tau,f,z)\varepsilon\Omega\text{: } f^i = f^j = f^* \;\forall i,j \text{ and } f^*_{z(n)} \to^c N(\theta)\}. \tag{10}$$

The set Ω_{Jordan} is the set where the limit points of KSR's of beliefs of the future given the past (and not conditioning on types) lie in the set of Nash equilibrium strategies of the underlying complete information game of the realized attribute vector θ.

Next define

$$\Omega_{Nyarko} \equiv W \cap CD \cap C\,, \tag{11}$$

where $$W \equiv \{(\tau,f,z)\varepsilon\Omega\text{: } \lim_{N\to\infty} \|\nu(f^i_{z(N)}) - \nu(f^*_{z(N)})\| = 0 \;\forall i\varepsilon I\}; \tag{12}$$

$$CD \equiv \cap_{i\varepsilon I} CD_i \quad \text{with} \quad CD_i \equiv \{(\tau,f,z)\varepsilon\Omega: \nu(f^i_{z(N)}) \to^c ND(\theta)\} \quad \forall i\varepsilon I; \quad \text{and} \qquad \textbf{(13)}$$

$$C \equiv \cap_{i\varepsilon I} C_i \quad \text{with} \quad C_i \equiv \{(\tau,f,z)\varepsilon\Omega: f^i_{z(N)} \to^c N_i(\theta_i)\} \quad \forall i\varepsilon I . \qquad \textbf{(14)}$$

The set W is the event where each player i's' beliefs, μ_i, about the future, $\{z_{N+1},z_{N+2},...\}$, given the past, z(N), (and *not* conditioning on own-types) "merge" with those of μ^* (and hence with each other) as the date N tends to infinity. The set CD is the set where limit points of each player's beliefs about the future play conditional on the past (again, *not* conditioning on own-types) is the same as the play of some Nash equilibrium. The set C is the set where for each i ε I, cluster points of the continuation strategies of player i's KSR of beliefs not conditioning on own-types, f^i, lie in the set $N_i(\theta_i)$. In particular[10], if $f^\infty = (f_i^\infty, f_{-i}^\infty)\varepsilon F$ is such a cluster point then f_i^∞ is a best-response to f_{-i}^∞ for the player with attribute vector θ_i.

Finally, define

$$\Omega_{KL} \equiv \{\omega=(\tau,f,z)\varepsilon\Omega: \forall\varepsilon>0, \exists N=N(\varepsilon,\omega) \text{ such that } \forall n\geq N, f_{z(n)}(\tau) \text{ lies in } SE_\varepsilon(\theta)\}. \qquad \textbf{(15)}$$

The set Ω_{KL} is the set of sample paths where $\forall\varepsilon>0$ (the continuation of) the strategies of players eventually lie in the set of subjective ε-equilibria of the game with the realized attribute vector, θ.

The results in Proposition 6.5 below are the main results of Jordan (1995), Kalai and Lehrer (1993a), and Nyarko (1997b). The result we state for Jordan is a slight generalization of Jordan (1995), relaxing the assumption *(CPA)* he used to the weaker assumption *(CPA*)* - the proof appears in the Appendix B. Immediately following this, we record in Proposition 6.6 the fact that the Jordan and Nyarko results satisfy TIGER, while the Kalai and Lehrer result violates it (as is easily seen from formulations F1 versus F2 of Example 2.1).

Proposition 6.5: Fix any game g=<T, $\{\mu_i\}_{i\varepsilon I}$, μ^*> in **G**.

(a) **(Jordan):** Suppose g satisfies optimization (6.1) and *(CPA*)* . Then

$$\mu^*(\Omega_{Jordan}) = 1. \qquad \textbf{(16)}$$

---

[10] Note that $C \neq \cap_{i\varepsilon I}\{(\tau,f,z)\varepsilon\Omega: f^i_{z(N)} \to^c N(\theta)\}$. Instead, the former set contains the latter, usually strictly. In particular C is not the set where continuation of KSR of beliefs are Nash equilibria. The difference is the same as the difference between Nash and subjective Nash equilibria, and is due to the fact that on the set C players are allowed to have different (limit) beliefs about play off the equilibrium path.

(b) **(Nyarko):** Suppose g satisfies optimization (6.1) and *(GGH)*. Then

$$\mu^*(\Omega_{Nyarko}) = 1. \tag{17}$$

(c) **(Kalai and Lehrer):** Suppose g satisfies optimization (6.1) and *(KL-T)*. Then

$$\mu^*(\Omega_{KL})=1. \tag{18}$$

Proposition 6.6: (a) Eq. (16) obeys TIGER. **(b)** Eq. (17) obeys TIGER. **(c)** Eq. (18) violates TIGER.

The lemma below indicates an entire class of properties of games which will obey TIGER. The lemma is used in the proof of parts (a) and (b) of Proposition 6.6.

Lemma 6.1: Fix any sets $D_i \subseteq \Theta_i \times Z$ and $D^* \subseteq \Theta \times Z$ and any numbers k_i and k^* for i in I. Define the property b_i for i in I and the property b^* as follows: g= $<T, \{\mu_i\}_{i\varepsilon I}, \mu^* >$ in **G** satisfies b_i if and only if $\mu_i(D_i)=k_i$; and g satisfies b^* if and only if $\mu^*(D^*)=k^*$. Then b^* and each b_i satisfy TIGER.

In the monotonicity result below we use the definition of $SE_\varepsilon(\theta)$ of **(9)** where we insist that the f_j^i's of that definition are equal to the beliefs of player i about j. This monotonicity result is related to the work of Jackson[11], Kalai and Smorodinsky (1997).

Proposition 6.7 (Monotonicity): Fix any g in **G** and suppose that g obeys the KL93 conclusion in eq. (18). Then the conclusion also holds for all equivalent type-coarsenings of g.

6.4. On Nachbar (1997). The conclusion of Nachbar (1997) is that there is an inherent conflict between prediction and optimization when the strategy space is sufficiently rich. The definition of optimization is as in **(6.1)**. We now define prediction. Given any history h of length $\ell<\infty$ say, the cylinder set C(h) is the subset of all play paths z in *Z* whose first ℓ elements equal h. The definition below is an

---

[11]The Jackson et. al. paper was set in the context of a single agent inference problem. The connection with this paper however is immediate. The set Θ of their paper corresponds to the type space T here. What they refer to as a representation is very similar to what we call here an equivalent reformulation of a game. They introduce a concept of learnability, which, modulo technicalities, is Eq. (18). Proposition 6.7 then shows, loosely speaking, that if a game g in **G** is learnable (or specifically μ^*) then so too is any equivalent coarsening of g.

absolute continuity condition over cylinder sets.

Definition 6.5: The game g in **G** obeys *Ex Post Local Absolute Continuity (EPLAC)* if and only if for every h in ***H***, and μ^* -a.e. $\tau=\{\tau_i\}_{i\varepsilon I}$, $\nu(f(\tau))(C(h))>0$ implies $\nu(f^i(\tau_i))(C(h)) > 0$.

Given any $\varepsilon>0$ and any integer ℓ, and any f and f′ in F, f is said to *(ε,ℓ)-play like* f′ if for all histories h of length ℓ or less, $|\nu(f)(C(h)) - \nu(f')(C(h))| \leq \varepsilon$. Define

$$\Omega_P \equiv \left\{\omega=(\tau,f,z)\varepsilon\Omega: \forall\varepsilon>0, \forall i\varepsilon I, \forall \text{ integers } \ell, \exists N=N(\varepsilon,\ell,\omega) \text{ s.t. } \forall n\geq N, f_{z(n)}(\tau) \ \ (\varepsilon,\ell)\text{-plays like } f^i_{z(n)}(\tau_i)\right\}. \quad \textbf{(19)}$$

Definition 6.6: *(Nachbar Prediction)***:** The game g in **G** obeys Nachbar prediction if and only if (i) g satisfies EPLAC; and (ii)

$$\mu^*(\Omega_P) = 1 \, . \quad \textbf{(20)}$$

Definition 6.7: *(The Nachbar Paradox)*. The game g ε **G** satisfies the Nachbar Paradox if it is *not* the case that g obeys both optimization **(6.1)** and Nachbar Prediction.

Nachbar (1997) argued that in games where the strategy sets are sufficiently rich there is an inherent conflict between optimization and prediction in the sense of the definition above. As illustrated in Example 2.1, a sufficiently rich strategy set requires a sufficiently refined type-space. In particular, in that example under formulation F1 prediction and optimization hold, while under formulation F2, optimization holds but prediction fails. In particular F2, the more refined space, obeys the Nachbar paradox while F1 violates it. Among other things this implies the following:

Proposition 6.8: The Nachbar paradox violates TIGER.

We also have the following monotonicity result:

Proposition 6.9 (Monotonicity): Fix g and g′ in **G** with g′ ≥ g. **(a)** If the Nachbar paradox holds for g then it also holds for g′; and **(b)** if g′ violates the Nachbar paradox (i.e. if g′ satisfies optimization and prediction) then so too does g.

Proposition 6.9 above is proved with the aid of the following two lemmas which may be of interest in their own right. The first, Lemma 6.2 below, shows that EPLAC obeys TIGER, which is interesting when compared to the failure of (KL-T) to satisfy TIGER[12]:

Lemma 6.2: *(EPLAC)* satisfies TIGER.

Lemma 6.3: Fix g and g′ in **G** with $g' \geq g$. Then if the property in **(20)** holds in g′, it also holds in g.

The Nachbar paper asks whether ***true*** play can be predicted. We have argued that there may be problems with the notion of the "truth" here, and in particular that the Nachbar paradox violates TIGER. Instead of asking whether there is prediction of the truth, we could ask whether there is prediction of beliefs. Define the set Ω_{PB} as follows:

$$\Omega_{PB} \equiv \left\{ \omega=(\tau,f,z)\varepsilon\Omega: \forall\varepsilon>0, \forall \text{ integers } \ell, \exists N=N(\varepsilon,\ell,\omega) \text{ s.t. } \forall n\geq N,\ f^{*}_{z(n)}\ (\varepsilon,\ell)\text{-plays like } f^{i}_{z(n)},\ \forall i\varepsilon I \right\}. \quad \mathbf{(21)}$$

Observe that this is the same as the set where Nachbar's concept of prediction occurs (see Ω_P in **(19)**) except that we use beliefs $f^{*}_{z(n)}$ and $f^{i}_{z(n)}$ in place of true strategies and beliefs conditional on own-types, $f_{z(n)}(\tau)$ and $f^{i}_{z(n)}(\tau_i)$, respectively. Here it may be useful to stress again that $f^{i}_{z(n)}(\tau_i)$ and $f^{i}_{z(n)}$ represent the same beliefs over F_{-i}, the strategies of others; the only difference is as regards beliefs about own strategies - the latter conditions on own-types τ_i while the other does not. It is easy to see that on the

[12] Although *(EPLAC)*, ex post local absolute continuity of μ^* with respect to μ_i, obeys TIGER, it is easy to see that ex post local *mutual* absolute continuity - μ_i with respect to μ_j for all i and j - does not obey TIGER. Indeed consider Example 2.1. Under F1 all players' beliefs are mutually absolutely continuous with respect to each other (indeed the common prior assumption holds in that case). Consider now, however, F2 and fix a vector of types $\tau=(\tau_A, \tau_B)$. Player A of type τ_A assigns probability one to one particular date 1 action that she, Player A, will choose while player B assigns equal probability to both the actions of Player A. The players' beliefs, conditional on their types, are not locally mutually absolutely continuous in F2. Hence ex post local mutual absolute continuity does not satisfy TIGER.

set Ω_{PB}, for large enough n, $f^{i}_{z(n)}$ (ε,ℓ)-plays like $f^{j}_{z(n)}$, $\forall i,j \varepsilon I$ Hence ignoring beliefs about own play, over time players make approximately the same predictions about the future play. For example at any period n sufficiently large, players i and j will have approximately the same beliefs about the which action a third player k will choose in the next period. This is not exactly prediction of beliefs - indeed we have not modeled what i thinks about what j believes. It is however prediction beliefs in the sense that i and j will have the same beliefs about any third player k. Analogously to Definition 6.6 we therefore have:

Definition 6.8: *(Prediction of beliefs):* The game g in **G** obeys prediction of beliefs if and only if (i) g satisfies EPLAC; and (ii)

$$\mu^{*}(\Omega_{PB}) = 1 . \qquad \textbf{(22)}$$

From Lemmas 6.1 and 6.2 the following is immediate:

Proposition 6.10: Prediction of beliefs obeys TIGER.

To give an indication of how prediction of beliefs can be obtained, we note that condition *(GGH)* implies prediction of beliefs[13]. (To see this apply Proposition 6.5b and observe that $W \subseteq \Omega_{PB}$). In particular, both formulations F1 and F2 of example 2.1 obey Prediction of Beliefs. When we move from strategies to beliefs, the Nachbar "paradox" disappears, and instead, under *(GGH)*, we very easily obtain both prediction and optimization!

7. Conclusion

Jordan (1996) states that a notable shortcoming of Bayesian learning models is that "convergence occurs at the level of expectations and not necessarily at the level of actual strategies." This paper shows that this should not be considered a shortcoming. Instead, if we want our results to be consistent in the sense of obeying TIGER, we can make statements *only* at the level of expectations or beliefs. Ambiguities in what constitutes the "truth" force us away from statements on true strategies and toward

[13]An alternate definition of prediction of beliefs would replace condition (ii) of Definition 6.8 with $\mu^{*}(W)=1$ where W is as in **(12)**. Since $W \subseteq \Omega_{PB}$ this results in a stronger notion of prediction of beliefs. This stronger definition also obeys TIGER, and also follows from condition *(GGH)*.

statements on expectations.

What is the correct type-formulation? There is no "correct" type-formulation. We advocate neither the sparse nor the comprehensive formulation. If you believe that people really do not mix, then you are advocating the comprehensive formulation. But then the Nachbar paradox holds (see formulation F2 of Example 2.1). On the other hand if you adopt the sparse formulation, you must deal with the arguments of those who insist that people do not mix. In Section 6 we showed that there is usually a critical type-formulation such that the KL93 assumptions and conclusions hold for all coarser formulations and fail to hold for all finer formulations. In that case you may define what is the correct formulation in terms of whether you want the KL93 conclusions to hold. In summary, if one is not willing to move to concepts like equilibrium in beliefs which satisfy TIGER, there is no obvious "correct" type-formulation. The best type-formulation, as with beauty, may lie in the eye of the beholder.

8. Appendix A: Examples

Example 8.1: *(A critical g for (KL-T))***:** Let I, A and Θ be as in the matching pennies game of Example 2.1. We now define for each k = 0,1,2,3,... ,∞, a game $g^k \varepsilon$ **G**. Define g^0 and g^∞ be formulations F1 and F2, respectively, of Example 2.1. For $1 \leq k < \infty$, define $T_i^k \equiv \{HEADS,TAILS\}^k$, and suppose that the types in T_i^k are generated via k independent tosses of a fair coin. The behavior strategy of player-type $\tau_i=\{\tau_{i,1}, \tau_{i,2}, ..., \tau_{i,k}\} \varepsilon T_i^k$ in game g^k is as follows: At date $n \leq k$, choose the first action (TOP for i=A and LEFT for i=B) if $\tau_{i,n}$=HEADS, and the second action otherwise; and at date n > k, randomize over both actions with equal probability. This defines the game g^k. It is easy to check that the collection of games $\{g^k\}_{k=1}^\infty$ are all equivalent with $g^\infty \geq g^{k+1} \geq g^k \geq g^0$ for all k. There is also obviously a weak-topology sense[14] in which the g^k's converge to g^∞. One can also check that the game g^k satisfies *(KL-T)* $\forall k < \infty$, and violates *(KL-T)* for k=∞. g^∞ in this case is a "critical" g, as mentioned in Section 6, for the given chain of games. ■

[14] To see this, define the game $\overline{g}^k$ to be the game g^k re-defined on the type space T_i= {HEADS, TAILS}$^\infty$ in the obvious manner: given any type τ_i in T_i in game $\overline{g}^k$, ignore the coordinates k+1, k+2,.... and proceed just as in game g^k . Next for each $\tau \varepsilon T$ and $k \leq \infty$, define ν_k to be equal to $\nu(f(\tau))$, the play induced by the vector of types τ in game g^k. Then it is easy to see that for each τ, $\nu_k(\tau)$ converges to $\nu_\infty(\tau)$ in the weak topology.

Example 8.2: *(A critical g for (KL-T))***:** Let I, A and Θ be as in the matching pennies game of Example 2.1. We now define for each k in $[0,\infty]$ a game $g^k \in$ **G**. Each of these games will be equivalent and will have exactly the same type-space. Further, there will exist a critical k (actually k=1) such that *(KL-T)* will hold for each k > 1 but will fail for each $k \leq 1$. The example will also be such that if k and k′ are close, then the games g^k and $g^{k'}$ will also be close in the weak-topology sense mentioned in Example 8.1.

Define for each k and each i in I, $T_i^k \equiv [0,1]^\infty$. The types in T_i^k are generated via the distribution over $[0,1]^\infty$ equal to the countably infinite product of uniform distributions over [0,1]. Define $\forall$ n=1,2,.... and k in $[0,\infty]$,

$$\Delta_{n,k} = (1/2)(1/(n+1)^{k/2}) . \qquad \textbf{(23)}$$

The behavior strategy of player i of type $\tau_i^k = \{\tau^k_{i,1}, \tau^k_{i,2}, \tau^k_{i,3}, ...\} \in T_i^k$ is defined as follows: the probability that she assigns to her first action (TOP for i=A and LEFT for i=B) at date n is

$$f_{i,n}(\tau_i) = \begin{cases} 1/2 + \Delta_{n,k} & \text{if } \tau_i^n \in [0,1/2) \\ 1/2 - \Delta_{n,k} & \text{if } \tau_i^n \in [1/2, 1] . \end{cases} \qquad \textbf{(24)}$$

The probability assigned to her second action at date n is therefore $1-f_{i,n}(\tau_i)$. We suppose that each player knows that this is how the types and behavior strategies are generated. Each player observes her own type but not the types of others. The probability that player i assigns to the event that player $j \neq i$ chooses her first action is $(1/2)[1/2 + \Delta_{n,k}] + (1/2)[1/2 - \Delta_{n,k}] = 1/2$. Each player is therefore indifferent between each of her two actions, so the behavior strategies just defined are best-responses to each other.

When $k=\infty$, $\Delta_{n,k}=0$ so each player-type is randomizing with probabilities 1/2 and 1/2 at each date. When k=0, $\Delta_{n,k}=1/2$ so each player is choosing a pure strategy, which depends on her type. Hence, the cases $k=\infty$ and k=0 correspond to formulations F1 and F2 respectively of Example 2.1. For each $k\neq 0$, each player-type's behavior strategy randomizes over her actions at each date with probabilities which converge, as the date $n \to \infty$, to the vector (1/2,1/2). The speed with which the probabilities converge to (1/2,1/2) is increasing in k. For *(KL-T)* to hold, this convergence must be sufficiently fast. It turns out that in the above example, the critical k for which this is true is k=1. In particular, when $k \leq 1$, the rate of convergence is so slow that *(KL-T)* fails to hold. To see this, define $\alpha_{i,n}$ to be the ratio of the true probability of player i choosing a given action at date n to the probability assigned by the beliefs of $j \neq i$ to that action (which, of course, is 1/2):

$$\alpha_{i,n} \;=\; \begin{cases} f_{i,n}(\tau_i)/(1/2) & \text{if the first action occurs at date n} \\ [1-f_{i,n}(\tau_i)]/(1/2) & \text{if the second action occurs at date n} \end{cases} \qquad \textbf{(25)}$$

Then $\sum_{n=1}^{\infty} (1-\alpha_{i,n})^2 = \sum_{n=1}^{\infty} (1-2f_{i,n}(\tau_i))^2 = 4\sum_{n=1}^{\infty} \Delta^2_{n,k} = \sum_{n=1}^{\infty} 1/(n+1)^k$. This sum is finite for k>1 and is infinite for $k \leq 1$. Applying (Shiryayev, Cor. 4 of p. 499), then shows that condition *(KL-T)* holds for k >1 and fails for $k \leq 1$. ■

9. Appendix B: Proofs

Proof of Proposition 5.1: Fix any $g = \langle T, \{\mu_i\}_{i\varepsilon I}, \mu^* \rangle$ in **G**. To obtain a $\underline{g}$ in G which is sparse and equivalent to g, define $\underline{T} = \Theta$ and define the behavior strategy of player-type $\underline{\tau}_i = \theta_i$ to be KSR of the marginal on F_i of $\mu_i(.|\theta_i)$. In particular, the sparse formulation, $\underline{g}$, is obtained by integrating out any randomization that was in the original types in g.

To obtain a g′ in G which is comprehensive and equivalent to g, one needs to perform at date -1 (i.e., before the game begins), all the possible future randomizations, and encode them in the type. The details of this are as follows: Define $\Gamma_i \equiv [0,1]^\infty$, $T_i' \equiv T_i x \Gamma_i$ and $T' = \prod_i T_i'$. Let Unif[0,1] be the uniform distribution over [0,1] and define $\xi_i \equiv \otimes_{k=1}^{\infty}$Unif[0,1]. We will later construct the behavior strategy, $f_i(\tau_i')$, of each player-type τ_i' in game g′. Define $1_{f(\tau')}$ (resp. $1_{f_i(\tau_i')}$) to be the probability measure on F (resp. on F_i) that assigns probability one to $f(\tau')$ (resp. $f_i(\tau_i')$). A unique $\mu^{*\prime} \varepsilon B(T'xFxZ)$ may be constructed with the following three components: (a) $\text{Marg}_{T'} \mu^{*\prime} = (\text{Marg}_T \mu^*) \otimes \prod_{i\varepsilon I} \xi_i$; (b) $\text{Marg}_F \mu^{*\prime}(.|\tau') = 1_{f(\tau')}$; and (c) $\text{Marg}_Z \mu^{*\prime}(.|\tau', f) = \nu(f)$. Similarly, a unique $\mu_i' \varepsilon B(T'xFxZ)$ may be constructed using the components (a) and (c) above, replacing μ^* with μ_i and $\mu^{*\prime}$ with μ_i', and replacing (b) with the requirement that $\text{Marg}_F \mu_i'(.|\tau'=(\tau,\gamma)) = (\text{Marg}_{F_{-i}} \mu_i(.|\tau)) \otimes 1_{f_i(\tau_i')}$.

We now construct the $\{f_i(\tau_i')\}_{i\varepsilon I}$ used above. Fix any i in I and $\tau_i \varepsilon T_i$. Since ***H*** is countable we can write $\boldsymbol{H} \equiv \{h^1, h^2, h^3, \ldots\}$. Then for any integer m, $f_i(\tau_i)(h^m)$ is a probability measure over A_i showing how player-type τ_i in game g chooses actions at history h^m . Since A_i is finite we may consider it an ordered set, $A_i = \{a_{i,\ell}\}_{\ell=1,2,\ldots,\#A_i}$, with $\#A_i$ the cardinality of A_i. Fix an m and let Supp $f_i(\tau_i)(h^m)$ be the support of $f_i(\tau_i)(h^m)$, similarly ordered. For each $h^m \varepsilon \boldsymbol{H}$, partition the unit interval [0,1] into distinct exhaustive sub-intervals with the ℓ-th sub-interval having lebesgue measure equal to the probability assigned by $f_i(\tau_i)(h^m)$ to the ℓ-th action in Supp $f_i(\tau_i)(h^m)$. Let $\mathcal{P}_i(\tau_i, h^m)$ denote this ordered set of sub-intervals. Each member of $\mathcal{P}_i(\tau_i,h^m)$ is associated with a unique action in A_i. For any $\tau_i' = (\tau_i,\gamma_{i,1},\gamma_{i,2},\gamma_{i,3},...)$ $\varepsilon T_i'$ define $f_i(\tau_i')$ by requiring it to choose at history h^m the action associated with the ℓ-th member of $\mathcal{P}_i(\tau_i,h^m)$, where ℓ is the unique integer such that $\gamma_{i,m}$ lies in the ℓ-th member of $\mathcal{P}_i(\tau_i,h^m)$.

We have completed the construction of $g'=<T',\{\mu'_i\}_{i\varepsilon I},\mu^{*\prime}>$. It should be obvious that it satisfies the conclusions of this proposition with $g'=\overline{g}$. ■

Proof of Proposition 6.1: Fix any two games g and g′ in **G** and suppose that $g'\succeq g$. Let Γ_i, τ_i and $\tau_i' = (\tau_i,\gamma_i)$ for $\gamma_i\ \varepsilon\ \Gamma_i$ be as in Definition 5.1. In particular, τ_i and τ_i' are two generic types in g and g′ respectively, which share the same attribute vector. From the independence assumption 4.1 these two player-types will also share the same beliefs about others. From condition 5.1(iv) the behavior strategy of player type τ_i in game g is equal to the KSR of the mixed strategy obtained by some randomization over the behavior strategies over player-types (τ_i,γ_i) for $\gamma_i\varepsilon\Gamma_i$. Standard arguments show that player-type τ_i 's behavior strategy is a best-response for her if and only if this is the case for all those of player-types (τ_i,γ_i) for $\gamma_i\ \varepsilon\,\Gamma_i$ (except possibly a set with zero probability). Hence optimization occurs in g if and only if it occurs in g′. An application of Proposition 5.2 therefore this proves the proposition. ■

Proof of Proposition 6.2: **(a)** This follows immediately from the fact that if $g \sim g'$ then the ex ante beliefs over F will be the same in both games g and g′.

(b) Fix any $g=<T,\{\mu_i\}_{i\varepsilon I},\mu^*>$ and $g'=<T',\{\mu_i'\}_{i\varepsilon I},\mu^{*\prime}>$ in **G** and suppose that $g'\succeq g$. Suppose g′ obeys *(GGH)*. By definition $T'=Tx\Gamma$ for some Γ. It is easy to verify that if one probability measure is absolutely continuous with respect to another on the cartesian product of two spaces then the same is true of their marginals. Hence if g′ obeys *(GGH)*, $\text{Marg}_{T_i x\, Z}\,\mu^* \ll \text{Marg}_{T_i x Z}\,\mu_i\ \forall i\varepsilon I$, so g obeys *(GGH).*

Next, suppose g obeys (GGH). Fix any i in I. Then, $\text{Marg}_Z\mu^* \ll \text{Marg}_Z\mu_i$. Since μ_i and μ_i' and also μ^* and $\mu^{*\prime}$ share the same marginal on *Z*, this implies that

$$\text{Marg}_Z\mu^{*\prime} \ll \text{Marg}_Z\mu_i' . \qquad \textbf{(26)}$$

Following each history each player chooses an action independently of the others. So, recalling the notation of eqn's **(4)** and **(5)**, for each date N history $h_N = (a_1,...,a_N)$,

$$\mu^{*\prime}(\{h_N\}\,|\,\tau_i') = \prod_{n=0}^{N-1} [f_i(\tau_i')(h_n)(a_{i,n+1})][f_{-i}^*(h_n)(a_{-i,n+1})] , \qquad \textbf{(27)}$$

$$\mu_i'(\{h_N\}|\tau_i') \quad = \quad \prod_{n=0}^{N-1} [f_i(\tau_i')(h_n)(a_{i,n+1})][f_{-i}^{\,i}(h_n)(a_{-i,n+1})] \,, \qquad \textbf{(28)}$$

$$\mu^{*\prime}(\{h_N\}) \quad = \quad \prod_{n=0}^{N-1} [f_i^{\,i}(h_n)(a_{i,n+1})][f_{-i}^{\,*}(h_n)(a_{-i,n+1})] \,, \text{ and} \qquad \textbf{(29)}$$

$$\mu_i'(\{h_N\}) \quad = \quad \prod_{n=0}^{N-1} [f_i^{\,i}(h_n)(a_{i,n+1})][f_{-i}^{\,i}(h_n)(a_{-i,n+1})] \,. \qquad \textbf{(30)}$$

Define

$$r_N(\tau_i') \equiv \frac{\mu^{*\prime}(\{h_N\}|\tau_i')}{\mu_i'(\{h_N\}|\tau_i')} \,, \quad \forall \tau_i' \varepsilon T_i' \quad \text{and} \quad r_N \equiv \frac{\mu^{*\prime}(\{h_N\})}{\mu_i'(\{h_N\})} \,, \qquad \textbf{(31)}$$

whenever the denominators of these expressions are positive, and define them to be equal to zero otherwise. Eq.'s **(27) - (30)** imply that $\forall N$, $h_N \varepsilon \boldsymbol{H}^N$, and $\tau_i' \varepsilon T_i'$ such that $\mu_i'(\{h_N\}) \neq 0$ and $\mu_i'(\{h_N\}|\tau_i') \neq 0$,

$$r_N = r_N(\tau_i'). \qquad \textbf{(32)}$$

It is easy to see that r_N is the Radon-Nikodym derivative of $\text{Marg}_Z\, \mu^{*\prime}$ with respect to $\text{Marg}_Z\, \mu_i'$ when the two measures are restricted to $\boldsymbol{H}^N$. Hence, using Shiryayev ((1984), Theorem 2, p.495), **(26)** implies that there exists an r_∞ such that $\lim_{N\to\infty} r_N = r_\infty$ which is finite with probability one with respect to $\mu^{*\prime}$ and μ_i'. From **(32)** this in turn implies that

$$\lim_{N\to\infty} r_N(\tau_i') = r_\infty \,. \qquad \textbf{(33)}$$

We will now argue that for each N and $\mu^{*\prime}$ a.e. τ_i',

$$\text{Marg}_{\boldsymbol{H}^N}\ \mu^{*\prime}(.|\tau_i') \ll \text{Marg}_{\boldsymbol{H}^N}\ \mu_i'(.|\tau_i'). \qquad \textbf{(34)}$$

To show this assume, on the contrary, that for some h_N and for a set of τ_i' 's with $\mu^{*\prime}$ positive probability, the following is true: (i) **(27)** is positive and (ii) **(28)** is zero. Noting that the products in **(27)** and **(28)** share some common terms, (i) implies that **(29)** is positive while (ii) implies that **(30)** is zero. This is a contradiction to **(26)**, which proves **(34)**.

It is also easy to see that $r_N(\tau_i')$ is the Radon-Nikodym derivative of $\text{Marg}_Z\, \mu^{*\prime}(.|\tau_i')$ with respect to $\text{Marg}_Z\, \mu_i'(.|\tau_i')$ when the two measures are restricted to $\boldsymbol{H}^N$. Eq.'s **(33)** and **(34)** and the Shiryayev result mentioned earlier then imply that $\text{Marg}_Z \mu^{*\prime}(.|\tau_i') \ll \text{Marg}_Z \mu_i'(.|\tau_i')$ for $\mu^{*\prime}$ a.e. τ_i'. Since by definition $\mu^{*\prime}$ and μ_i' have the same marginal on T_i', we conclude that $\text{Marg}_{T_i' x Z}\, \mu^{*\prime} \ll \text{Marg}_{T_i' x Z}\ \mu_i'$ So g′ obeys *(GGH)*. ■

Proof of Proposition 6.3: Fix any two games g= $\langle T,\{\mu_i\}_{i\varepsilon I},\mu^*\rangle$ and g′=$\langle T',\{\mu'_i\}_{i\varepsilon I},\mu^{*\prime}\rangle$ in **G** and suppose that $g' \succeq g$. Then, following the notation of Section 5, we may write $T'=Tx\Gamma$. Fix any i in I and define for each $(\tau,\gamma)=(\tau_i, \gamma_i, \tau_{-i}, \gamma_{-i})$, the following measures on Z: $P^*(.|\tau,\gamma) \equiv Marg_Z \mu^{*\prime}(.|\tau,\gamma)$ and $P_i(.|\tau,\gamma) \equiv Marg_Z \mu_i'(.|\tau_i,\gamma_i)$ (and note that the latter depends on (τ,γ) only through (τ_i,γ_i)). Suppose g' obeys *(KL-T)*. Then $P^*(.|\tau,\gamma) \ll P_i(.|\tau,\gamma)$ for μ^* almost every (τ,γ) in T'. We may integrate out the γ's conditional on τ, to conclude that

$$\int_\Gamma P^*(.|\tau,\gamma)d\mu^*(.|\tau) \ll \int_\Gamma P_i(.|\tau,\gamma)d\mu^*(.|\tau) \quad , \quad \text{for } \mu^* \text{ almost every } \tau . \qquad \textbf{(35)}$$

The left hand side of **(35)** is equal to $Marg_Z \mu^{*\prime}(.|\tau)$. From the definition of μ^*, the marginals of μ^* and of μ_i on $T_i x\Gamma_i$ are the same, so the right hand side of **(35)** is equal to $Marg_Z \mu_i'(.|\tau_i)$. Hence, $Marg_Z \mu^{*\prime}(.|\tau) \ll Marg_Z \mu_i'(.|\tau_i)$, so g obeys *(GGH)*.■

Proof of Proposition 6.4: We begin with the following claims:

Claim 1: A probability measure can assign positive probability to at most countably many distinct mutually disjoint sets.

Proof of Claim 1: Let P be a probability measure on a measure space $(\Omega_0,\Im)$. Fix an index set Q and let $\{\varphi_q\}_{q\varepsilon Q}$ be any collection of mutually disjoint measurable subsets of Ω_0 with $P(\varphi_q)>0\ \forall q\varepsilon Q$. Define for any integer k>0, $Q_k \equiv \{q\varepsilon Q | P(\varphi_q)>1/k\}$. Q_k cannot contain more than k distinct elements Q, for otherwise the total probability of the union of φ_q over $q\varepsilon Q_k$ would exceed 1. Since $Q = U_{k=1}^{\infty} Q_k$ this shows that Q must be countable. //

Claim 2: Let $\alpha:T_1 \rightarrow Z$ and $\beta:T_2 \rightarrow Z$ be two Borel measurable functions. Let $\pi = \pi_1 \otimes \pi_2$ be any product measure on T_1xT_2. Then there exists a countable set ψ such that if $J \equiv \{(\tau_1,\tau_2)\ \varepsilon\ T_1xT_2 \mid \alpha(\tau_1)=\beta(\tau_2)\notin\psi\}$, then $\pi(J)=0$.

Proof of Claim 2: Let α, β and π be as in the claim. Define $\psi \equiv \{z \text{ in } Z | \pi_2(\beta^{-1}(z))>0\}$. If $z \neq z'$ then $\beta^{-1}(z)$ and $\beta^{-1}(z')$ are disjoint. Hence from Claim 1, ψ is countable. With this ψ, let J then be as in the conclusion of this claim (Claim 2). Define $J(\tau_1) = \{\tau_2 \text{ in } T_2 \mid (\tau_1,\tau_2)\ \varepsilon\ J\}$. We proceed to show that

$$\pi_2(J(\tau_1)) = 0 \quad \forall\tau_1 \text{ in } T_1. \qquad \textbf{(36)}$$

So fix any τ_1 in T_1. Clearly **(36)** holds if $J(\tau_1)$ is empty. So suppose $J(\tau_1)$ is non-empty. Then $\alpha(\tau_1)\notin\psi$.

This implies first that $J(\tau_1) = \{\tau_2 | \alpha(\tau_1)=\beta(\tau_2)\} = \beta^{-1}(\alpha(\tau_1))$ and also that $\pi_2(\beta^{-1}(\alpha(\tau_1))=0$. Hence, eq. **(36)** again holds. Eq. **(36)** and the fact that π is a product measure in turn proves that $\pi(J)=0$. //

Proof of Proposition 6.4 (Cont'd): By assumption there are at last two players, 1 and 2 say. Let $\underset{\sim}{\mathbf{z}}:\Omega \to Z$ be the random variable on Ω which defines the play path z in each state ω in Ω. Fix any $\tau_1 \varepsilon T_1$ and $\tau_2 \varepsilon T_2$ and define $A(\tau_1) \equiv \{\bar{z}\, \varepsilon Z : \mu_1(\{\underset{\sim}{\mathbf{z}}=\bar{z}\} | \tau_1)>0\}$ and $C(\tau_2) \equiv \{\bar{z}\, \varepsilon Z : \mu_2(\{\underset{\sim}{\mathbf{z}}=\bar{z}\} | \tau_2)>0\}$. From Claim 1, $A(\tau_1)$ and $C(\tau_2)$ are both countable so we may write $A(\tau_1)=\{\alpha^1(\tau_1),\alpha^2(\tau_1),\alpha^3(\tau_1),...\}$ and $C(\tau_2)=\{\varsigma^1(\tau_2), \varsigma^2(\tau_2),\varsigma^3(\tau_2),...\}$. Next, define $\forall \tau_{-1}=\{\tau_j\}_{j\neq 1}$ in T_{-1} and $\forall$integers m, $\beta^m(\tau_{-1}) \equiv \varsigma^m(\tau_2)$ and $B(\tau_{-1}) \equiv \{\beta^1(\tau_{-1}), \beta^2(\tau_{-1}), \beta^3(\tau_{-1}),...\}$. One may check that we may order the points in the sets so that for each ℓ and m, α^ℓ and ς^m (and hence β^m) are measurable functions of their arguments. Fix any ℓ and m. Note that $\beta^m(\tau_{-1})$ depends on τ_{-1} only through τ_2. Apply Claim 2 with $\pi = (\mathrm{Marg}_{T_1} \mu^*) \otimes (\mathrm{Marg}_{T_2} \mu^*)$, $\alpha=\alpha^\ell$ and $\beta=\beta^m$, and let $\psi^{\ell,m} \subseteq Z$ be the countable set obtained from that claim. Define $\Psi = \mathrm{U}_{\ell,m=1,2,...}\, \psi^{\ell,m}$. Then Ψ is countable and

$$\mu^*(\{\tau=\{\tau_i\}_{i\varepsilon I} |\ \alpha^\ell(\tau_1) = \beta^m(\tau_{-1}) \notin \Psi \text{ for some } \ell \text{ and m}\}) = 0. \qquad \textbf{(37)}$$

By assumption the game is comprehensive. Hence $\forall\tau\ \varepsilon\ T$ there exists a unique play path $z(\tau)$ in Z resulting from that vector of types. Also $\mu^*(\{\underset{\sim}{\mathbf{z}}=z(\tau)\} | \tau) = 1$. So if $\tau=\{\tau_i\}_{i\varepsilon I}\, \varepsilon T_{KL\text{-}T}$, then $\mu_i(\{\underset{\sim}{\mathbf{z}}=z(\tau)\} | \tau_i) > 0\ \forall i\varepsilon I$, which in turn implies that $z(\tau)=\alpha^\ell(\tau_1)=\beta^m(\tau_{-1})$ for some ℓ and m. By assumption $\mu^*(T_{KL\text{-}T})=1$, so **(37)** implies that $\mu^*(\{\tau=\{\tau_i\}_{i\varepsilon I} |\ z(\tau) \notin \Psi\}) = 0$, from which the proposition follows. ■

Proof of Proposition 6.5: (a) Fix any $g=<T,\{\mu_i\}_{i\varepsilon I},\mu^*>$ in **G**. Suppose that g obeys (6.1) and *(CPA*)*. Define another game $g'=<T',\{\mu_i'\}_{i\varepsilon I}, \mu^{*\prime}>$ where we suppose each player's beliefs equal the "true" distribution μ^*: in particular, set $T_i'=T_i$, $\mu_i' = \mu^*\ \forall i$ in I, and $\mu^{*\prime}$ equal to the outside observer distribution induced by $\{\mu_i'\}_{i\varepsilon I}$. It should be clear that $\mu^* = \mu^{*\prime}$ and that the game g′ obeys the common prior assumption (CPA). Since g obeys *(CPA*)* and under the independence assumption 4.1 beliefs about others are independent of own-types, the belief of each player-type τ_i is the same in g as it is in g′. Hence, since g obeys optimization (6.1) so too does g′. The game g′ therefore satisfies the conditions of the original Jordan (1995) result, so the conclusion of Proposition 6.5(a) holds for g′. Since $\mu^* = \mu^{*\prime}$ and $f^* = f^{*\prime}$, that conclusion also holds in g, which is what we seek to prove. ■

Proof of Proposition 6.6: Any two equivalent games will have the same ex ante true play f^* and ex ante beliefs f^i for i in I. The proof of parts (a) and (b) therefore follows from Lemma 6.1.■

Proof of Lemma 6.1: Fix any $g = \langle T, \{\mu_i\}_{i \in I}, \mu^* \rangle$ and $g' = \langle T', \{\mu'_i\}_{i \in I}, \mu^{*\prime} \rangle$ in **G** and suppose that $g' \succeq g$. Then μ_i and μ_i' have the same marginal on Θ_i. From the independence assumption 4.1 and 5.1(iii) and (iv) we conclude that conditional on θ_i, μ_i and $\underline{\mu}_i$ have the same distribution over F and therefore over *Z*. Hence μ_i and $\underline{\mu}_i$ have the same marginal distribution on $\Theta_i x Z$. This implies $\mu_i(D) = \underline{\mu}_i(D)$ for any $D \subseteq \Theta_i \, x Z$, so g obeys b_i if and only if g' does. Apply Proposition 5.2 to conclude that b_i satisfies TIGER. Similar arguments imply that b^* obeys TIGER. ■

Proof of Proposition 6.7: Fix any $g = \langle T, \{\mu_i\}_{i \in I}, \mu^* \rangle$ and $g' = \langle T', \{\mu'_i\}_{i \in I}, \mu^{*\prime} \rangle$ in **G** and suppose that $g' \succeq g$. Then, following the notation of Section 5, we may write $T' = Tx\Gamma$. We concentrate on game g' for now. Fix any i in I, $\varepsilon > 0$ and date n. Define $\mathbf{\Omega}'(i,\varepsilon,n) = \{\omega \text{ in } \mathbf{\Omega}' \mid \| \nu(f_{z(n)}(\tau,\gamma)) - \nu(f^i_{z(n)}(\tau_i,\gamma_i)) \| \leq \varepsilon\}$. The inequality below follows from the fact that the left hand side is less than ε on $\mathbf{\Omega}'(i,\varepsilon,n)$, and, since ν is a probability, is less than one on $\mathbf{\Omega}'(i,\varepsilon,n)^c$:

$$\| \nu(f_{z(n)}(\tau,\gamma)) - \nu(f^i_{z(n)}(\tau_i,\gamma_i)) \| \leq \varepsilon + 1_{\mathbf{\Omega}'(i,\varepsilon,n)^c} \quad \textbf{(38)}$$

(where the superscript c denotes the "complement" of the set and 1_X is the indicator function on X). Hence upon integrating over γ with respect to $\mu^{*\prime}(.\mid\tau,z(n))$ and noting that $\mu^{*\prime}$ and μ_i' have the same marginal over Γ_i, we may conclude that

$$\left\| \int_\Gamma \nu(f_{z(n)}(\tau,\gamma)) d\mu^{*\prime}(.\mid\tau,z(n)) - \int_{\Gamma_i} \nu(f^i_{z(n)}(\tau_i,\gamma_i)) d\mu_i'(.\mid\tau_i,z(n)) \right\| \leq \varepsilon + \mu^{*\prime}(\mathbf{\Omega}'(i,\varepsilon,n)^c \mid \tau,z(n)). \quad \textbf{(39)}$$

Suppose that g' obeys the KL93 conclusion. Then $1 = \mu^{*\prime}(\mathbf{\Omega}'_{KL}) = \mu^{*\prime}(\cap_{i \in I} \cap_{\varepsilon>0} \cup_{N=1}^{\infty} \cap_{n \geq N} \mathbf{\Omega}'(i,\varepsilon,n))$. This implies that for each i and $\varepsilon > 0$, $\mu^{*\prime}(\mathbf{\Omega}'(i,\varepsilon,n)) \to 1$ as $n \to \infty$. So from **(39)** we conclude that

$$\mu^{*\prime}\left(\cap_{i \in I} \cap_{\varepsilon>0} \cup_{N=1}^{\infty} \cap_{n \geq N} \left\{ \left\| \int_\Gamma \nu(f_{z(n)}(\tau,\gamma)) d\mu^{*\prime}(.\mid\tau,z(n)) - \int_{\Gamma_i} \nu(f^i_{z(n)}(\tau_i,\gamma_i)) \, d\mu_i'(.\mid\tau_i,z(n)) \right\| \leq \varepsilon \right\}\right) = 1. \quad \textbf{(40)}$$

This is precisely the statement that g obeys the KL93 conclusion. ■

Proof of Proposition 6.9: **(a)** From Proposition 6.1 and Lemma 6.2 we know that optimization and EPLAC obey TIGER. The required monotonicity therefore follows from the monotonicity result of Lemma 6.3. **(b)** The statement in (b) is the counter-positive of (a). ■

Proof of Lemma 6.2: Fix any $g = \langle T, \{\mu_i\}_{i\varepsilon I}, \mu^* \rangle$ and $g' = \langle T', \{\mu'_i\}_{i\varepsilon I}, \mu^{*\prime} \rangle$ in **G** and suppose that $g' \geq g$. Let τ (respectively (τ,γ)) denote a generic type vector in g (resp. g′). We continue the proof in two steps, (a) and (b) below. Applying Proposition 5.2 then proves this lemma.

(a) *If g′ satisfies EPLAC then so too does g*: Suppose g′ satisfies EPLAC. Fix any $h \varepsilon \boldsymbol{H}$ and $\tau \varepsilon T$ and suppose that $\nu(f(\tau))(C(h))>0$. Then there exists a set of γ's with $\mu^{*\prime}$-positive probability such that $\nu(f(\tau,\gamma))(C(h))>0$. Since g′ satisfies EPLAC, this implies that $\nu(f^i(\tau_i, \gamma_i))(C(h))>0$, for a set of γ_i's with $\mu^{*\prime}$- positive probability. Integrating out the γ_i's implies that $\nu(f^i(\tau_i))(C(h)) > 0$. So g satisfies EPLAC.

(b) *If g satisfies EPLAC then so too does g′*: It is easy to see EPLAC holds in *<u>any</u>* game g′ if for that game for $\mu^{*\prime}$ almost every $(\tau,\gamma)=(\tau_i,\tau_{-i},\gamma_i,\gamma_{-i})$, for each n=0,1,2, ..., and for each history h of length n that occurs with $\mu^{*\prime}(.|\tau,\gamma)$ positive probability,

$$f(\tau,\gamma)(h)(\{z_{n+1}=\hat{a}\}) >0 \quad \text{implies} \quad f^i(\tau_i,\gamma_i)(h)(\{z_{n+1}=\hat{a}\}) > 0 \quad \forall \hat{a} \text{ in } A. \tag{41}$$

So fix any such n, h and $\hat{a}=(\hat{a}_i,\hat{a}_{-i}) \varepsilon A$. Suppose that the first inequality of **(41)** holds for a set of (τ,γ)'s with positive $\mu^{*\prime}$ probability. Since $f(\tau,\gamma)(h)(\{z_{n+1}=\hat{a}\}) = [f_i(\tau_i,\gamma_i)(h)(\{z_{i,n+1}=\hat{a}_i\})] . [f_{-i}(\tau_{-i},\gamma_{-i})(h)(\{z_{-i,n+1}=\hat{a}_{-i}\})]$, this implies that

$$f_i(\tau_i,\gamma_i)(h)(\{z_{i,n+1}=\hat{a}_i\}) >0 . \tag{42}$$

Next, integrating the first inequality of **(41)** over γ implies that $f(\tau)(h)(\{z_{n+1}=\hat{a}\}) > 0$, so if g satisfies EPLAC then $f^i(\tau_i)(h)(\{z_{n+1}=\hat{a}\}) > 0$, and in particular $f^i(\tau_i)(h)(\{z_{-i,n+1}=\hat{a}_{-i}\}) >0$. Since i's beliefs about j are independent of i's type, this implies that $f^i(\tau_i,\gamma_i)(h)(\{z_{-i,n+1}=\hat{a}_{-i}\}) >0$. Combining this with **(42)** implies the right hand side of **(41)** . So g′ obeys EPLAC.■

Proof of Lemma 6.3: The proof is almost identical to the proof of Proposition 6.7 so is omitted. ■

10. References

Aumann, R. (1964):"Mixed and Behavior Strategies in Infinite Extensive Games," Advances in Game Theory, Annals of Mathematical Studies, 5, 627-650.

_________ (1987):"Correlated Equilibrium as an Expression of Bayesian Rationality," Econometrica, 55, 1-18.

Aumann, R. and A. Brandenburger (1995):"Epistemic Conditions for Nash Equilibrium," Econometrica, 63(5), 1161-1180.

Battigalli, P. and D. Guaitoli (1988):"Conjectured Equilibrium and Rationalizability in a macroeconomic game with incomplete information," Bocconi University.

Battigalli, P., M. Gilli and C. Molinari (1992):"Learning and Convergence to Equilibrium in Repeated Strategic Interactions: An Introductory Survey," forthcoming in Ricerche Economiche 3-4.

Binmore, K. (1991): "Fun and Games," D.C. Heath, Lexington, Mass.

Harsanyi, J.C. (1967,1968): "Games with Incomplete Information Played by Bayesian Players," Parts I,II,III, Management Science, vol. 14, 3,5,7.

--------------(1973):"Games with Randomly Distributed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," pp. 1 - 23.

Jackson, M., E. Kalai and R. Smorodinsky (1997):"Patterns, Types, and Bayesian Learning," Manuscript, Carlifornia Institute of Technology; forthcoming Econometrica (1998).

Jordan, J. S. (1991): "Bayesian Learning in Normal Form Games," Games and Economic Behavior, 3, 60-81.

-------------- (1995): "Bayesian Learning In Repeated Games," Games and Economic Behavior, Vol. 9, No. 1, 1995, pp. 8-20.

-------------- (1996):"Bayesian Learning in Games: A NonBayesian Perspective," in "*The Dynamics of Norms,"* (eds.) C. Bicchieri, R. Jeffrey, and B. Skyrms, Cambridge University Press.

Kalai, E. and E. Lehrer (1993a):"Rational Learning Leads to Nash Equilibrium," Econometrica, 61, 1019-46.

--------------- (1993b):"Subjective Equilibrium in Repeated Games," Econometrica, 61, 1231-40.

Kuhn, H. (1953):"Extensive Form Games and the Problem of Information," Annals of Mathematics, Study No. 193-216.

Lehrer, E. and R. Smorodinsky (1997):"Repeated Large Games with Incomplete Information," Games and Economic Behavior, Vol. 18, No. 1, 1997, pp. 116-134.

Nachbar, J. (1997): "Prediction, Optimization, and Learning in Repeated Games," Econometrica, 65(2), 275-310.

Nyarko, Y. (1993):"The Types of a Bayesian Equilibrium" C.V. Starr Working Paper #93-36, New York University.

---------(1994):"Bayesian Learning in Leads to Correlated Equilibria in Normal Form games," Economic Theory, 4(6), 821-841.

---------(1997a): "Convergence in Economic Models with Bayesian Hierarchies of Beliefs," Journal of Economic Theory, 1997, 74(2), 266-296.

--------- (1997b):"Bayesian Learning Without Common Priors and Convergence to Nash Equilibrium," forthcoming, Economic Theory.

Sandroni,. A. (1995a): "Does Rational Learning Lead to Nash Equilibrium in Finitely Repeated Games?" Manuscript, Northwestern University.

--------- (1995b): "Necessary and Sufficient Conditions for Convergence to Nash Equilibrium: The almost Absolute Continuity Hypothesis." Manuscript, Northwestern University.

Shiryayev, A.N. (1984): "Probability," Springer-Verlag.

Smorodinsky, R. (1995): "The Absolute Continuity Hypothesis and Large Type Sets," Northwestern University.