$\emptyset$ MARKS THE SPOT

MAXWELL B. STINCHCOMBE

Abstract. A population game consists of a non-atomic, finitely additive probability space of agents, a set of actions, and, for each agent, a utility function that depends continuously on their action and the population distribution of actions. If the probability is not countably additive, then approximate equilibria may not exist. Existence failures are due to the positive mass of agents that can only be represented as elements of the empty set in the original model. These mislaid agents can be characterized using nonstandard analysis or compactification-based representations of the distribution of utility functions. Restoring the missing agents yields equilibrium existence and the finite approximability of equilibria.

1. Introduction

A finite, positive, non-atomic measure space can model a large population of "agents." If the measure is non-atomic, then no individual agent's choice of action can change population aggregates. With agents' utility depending on their own choice and population aggregates, this class of models provides an extremely powerful tool for the study of population-wide optimizing behavior.

A probability space $(T, \mathcal{T}, \mu)$ in which μ fails to be countably additive is one that, following Uhl (1984) "was unfortunate enough to have been cheated on its domain." The "cheated" aspect is the positive mass that has no representation, but connotations of the words "unfortunate" and "cheated" can be too negative. There are many cases in which such probabilities are a useful device, allowing one to capture limit phenomena more easily. However, in other cases, the mass that is missing a representation is quite "unfortunate," it can mean that the model is mis-leading.

This paper studies non-atomic population games when the measure space of agents, $(T, \mathcal{T}, \mu)$, is finitely additive but may not be countably additive. This is another case in which the results can be misleading. For some games, the set of agent characteristics that have no representation can, seemingly, lead to lack of even approximate equilibria. Restoring the mislaid agents remedies this and related problems.

1.1. **Non-atomic Population Games.** A non-atomic population game is specified by three objects: a non-atomic probability space of agents or "types," $(T, \mathcal{T}, \mu)$; a compact metric space of actions, A, that each $t \in T$ chooses from; and for each

Date: February 22, 2019. Many thanks to Ali Khan, Lei Qiao, Kali Rath and Yeneng Sun for sharing with me early versions of their paper "The Necessity of Countable Additivity in Modeling Infinitely Many Agents." The present paper would not have been possible without their pellucid examples of population games, and discussions with Ali Khan and Kali Rath were especially helpful. Also, many thanks to Bob Anderson for discussions of nonstandard models in economics, and to Bill Zame for lessons on compactifications.

$t \in T$, there is a utility function, $\mathcal{G}(t)$, that is jointly continuous in own choices and the population distribution of choices. If type t chooses an action according to $a(t)(\cdot)$, a probability on A, then the population distribution of choices, $\nu_a \in \Delta(A)$ is defined by $\nu_a(B) = \int a(t)(B)\, d\mu(t)$, B a Borel measurable subset of A.

Let $C(A \times \Delta(A))$ denote the set of continuous functions on $A \times \Delta(A)$ with the supnorm topology and associated Borel σ-field. Assume that $t \mapsto \mathcal{G}(t) \in C(A \times \Delta(A))$ is measurable. A measurable $t \mapsto a(t)$ is an **ϵ-equilibrium** if

$$\mu(\{t : \mathcal{G}(t)(a(t), \nu_a) \geq \max_{b \in A} \mathcal{G}(t)(b, \nu_a) - \epsilon\}) \geq 1 - \epsilon, \tag{1}$$

and is an **equilibrium** if it is a 0-equilibrium.

There is no loss in assuming, *as we do from here on*, that $\|\mathcal{G}(t)(\cdot,\cdot)\|_\infty \leq +1$ for all t. Letting $\mathbb{U} = \mathbb{U}(A \times \Delta(A))$ denote the unit ball in $C(A \times \Delta(A))$, this is the assumption that $\mathcal{G} : T \to \mathbb{U}$.

The induced distribution on $\mathbb{U}$ is denoted $p = \mathcal{G}(\mu)$ and defined by $p(E) = \mu(\mathcal{G}^{-1}(E))$. With the sup norm topology, $\mathbb{U}$ is a complete separable metric space (csm). A probability q on a complete separable metric space (M, d) is **tight** if for every $\epsilon > 0$, there exists a compact K such that $q(K) > 1 - \epsilon$. Strictly weaker than tightness is **neighborhood-tightness** (n-tightness), for every $\epsilon > 0$, there exists a compact K such that that for every $\delta > 0$, $q(K^\delta) > 1 - \epsilon$ where K^δ is the δ-neighborhood of K. Strictly stronger than tightness is **conditional tightness** (c-tightness), which requires that for every measurable E with $q(E) > 0$, the conditional probability $q(\cdot|E)$ is a tight probability on the metric space (E, d).

Results here, combined with small variants of examples in Khan et al. (2016), given in §2, show the following.

- If p is c-tight, then equilibria exist and the approximate equilibrium correspondence is continuous.
- If p is n-tight, then ϵ-equilibria exist for every $\epsilon > 0$, but equilibria might not exist.
- If p is not n-tight, then ϵ-equilibria may not exist for a range of ϵ.

The existence problems arise because a game model failing c-tightness has a positive mass set of agents with utility functions that seem to be elements of the empty set. This paper gives two methods for finding these mislaid agents and utility functions and restoring them to the game, and both methods lead to equilibrium existence and the continuity of the approximate equilibrium correspondence. The empty set that needs filling arises from failures of countable additivity.

1.2. **The Empty Set Marks the Spot.** A finitely additive probability p is countably additive if for every $E_n \downarrow \emptyset$, $p(E_n) \downarrow 0$, and it is **purely finitely additive** if there exists a sequences of measurable sets $E_n \downarrow \emptyset$ with $p(E_n) \equiv 1$. Limits strictly between 0 and 1 are possible. From Yosida and Hewitt (1952), any finitely additive p can be decomposed into a purely finitely additive part, q^{pfa}, and a countably additive part, q^{ca},

$$p = \delta q^{pfa} + (1 - \delta) q^{ca}. \tag{2}$$

The δ in the composition is always unique, and is equal to the supremum of the set of possible values of $\lim_n p(E_n)$ when $E_n \downarrow \emptyset$.

Kingman (1967) calls the δ the **discrepancy** of the probability. He shows that the lack of representation for the positive mass in the sets $E_n \downarrow \emptyset$ leads to e.g. Poisson jump processes or Brownian motions with polynomial time paths. The mass that has no representation in the space of polynomials is the mass that should be assigned to time paths with jumps or with unbounded variation over finite intervals.

For population games, if the discrepancy of $p = \mathcal{G}(\mu)$, the population distribution of utility functions, is strictly positive, then it gives the mass of the agents that are missing representations for their utility functions. Their utility functions have been mislaid, and 'should' belong to $\cap_n E_n$, but $E_n \downarrow \emptyset$ prevents this. The agents too have been mislaid because $\mathcal{G}^{-1}(E_n) \downarrow \emptyset$ although $\mu(\mathcal{G}^{-1}(E_n)) \geq \delta > 0$.

1.3. **Representing the Mislaid Agents and Utility Functions.** This paper presents two methods to recover, from μ and $p = \mathcal{G}(\mu)$, representations of the mislaid utility functions and agents.

- The first method replaces the spaces involved in specifying the game by their counterparts in Robinson's non-standard model for real and functional analysis. The sets $E_n \downarrow \emptyset$ correspond to sets ${}^*E_n$, which are internal subsets of the non-standard version of the unit ball, ${}^*\mathbb{U}$, of utility functions. Nested sequences of internal sets have a finite intersection property — $\cap_n {}^*E_n = \emptyset$ if and only if ${}^*E_N = \emptyset$ for some finite N. As each E_N is non-empty, the mislaid utility functions can be found in the non-empty set $\cap_n {}^*E_n$, and the mislaid agents in the inverse image of this set in ${}^*T$.
- The second involves imbedding the space of utility functions, $\mathbb{U}$, as a dense subset of a compact Hausdorff space, $\widehat{\mathbb{U}}$, so that p has a unique countably additive extension, $\widehat{p}$ that extends p in the sense that $\widehat{p}(\widehat{E}_n) = p(E_n)$. Being countably additive probability on a compact space, there is a nested collection of non-empty compact $K_n \subset \widehat{E}_n$ that "nearly fill" the $\widehat{E}_n$. The nested sequence of compact sets, $\{\widehat{K}_n : n \in \mathbb{N}\}$, has the finite intersection property, hence has non-empty intersection, implying that $\cap_n \widehat{E}_n$ is not empty. The mislaid utility functions belong to $\cap_n \widehat{E}_n$, and the mislaid agents to $\widehat{T}$, the appropriate compactification of T.

Restoring the mislaid agents and their utility functions to the model yields countably additive games. After restoration, these game have equilibria and a continuous approximate equilibrium correspondence. With either method, the basic properties of the mislaid agents and utility functions can be recovered from the information in μ, the population measure, and p, the induced distribution of utility functions. The most basic properties of the utility functions depend on the tightness properties of p. For the nonstandard analysis extensions: if p is n-tight, then the new utility functions are nearstandard in ${}^*\mathbb{U}$; and if p is not n-tight, then a positive mass of agents have utility functions that are not close to any element of $\mathbb{U}$. For the compactifications: if p is n-tight, then with probability 1, the utility functions belong to $\mathbb{U}$ after it has been imbedded in $\widehat{\mathbb{U}}$; and if p is not n-tight, then a positive mass of agent utility functions belong to the complement of $\mathbb{U}$ in $\widehat{\mathbb{U}}$.

1.4. **Outline.** The next section covers, in decreasing order of stringency, the c-tightness, tightness, and n-tightness of finitely additive probabilities on complete separable metric spaces (such as $\mathbb{U}$). Because c-tightness is equivalent to countable additivity, games that have c-tight population measures have equilibria and the limits of ϵ-equilibria are equilibria. We will see that n-tight games have ϵ-equilibria for each $\epsilon > 0$, and examples will show that such games may not have have equilibria. Other examples will show that games that fail n-tightness may, or may not, have ϵ-equilibria. Following this, §3 defines the nonstandard versions of population games and gives the corresponding equilibrium existence and finite approximability results. By taking the appropriate quotient spaces, §4 develops a useful representation of the requisite compactification of the space of utility functions and gives the existence and finite approximability results for equilibria in the compactified games. The final section discusses the relations between this work and a literature in which finitely additive probabilities, as a tool, range from immensely useful to completely misleading. The applications include maximization and noncooperative equilibrium models, stochastic process theory, betting/investing, social choice theory, Pareto optimality in intergenerational equity, and learning foundations for multiple prior models. The appendix contains the proofs as well as synopses of finitely additive probabilities and nonstandard analysis.

Throughout, probabilities are assumed to be finitely additive, but they may fail to be countably additive.

2. Tightness and Population Games

The space of agent characteristics, $\mathbb{U}$, is the closed unit ball in the separable Banach space $C(A \times \Delta(A))$, hence is a csm (complete separable metric space). The two variants of tightness, c-tightness and n-tightness (mnemomic for conditional-tightness and neighborhood-tightness respectively) are defined for a general csm (M, d) with Borel σ-field $\mathcal{M}$. The examples are given in the unit ball of $C([0,1])$. This space is both well-studied and relevant to present concerns — if A is finite, then $C(A \times \Delta(A))$ can be represented as a finite disjoint union of the sets $\{a\} \times C([0,1])$, $a \in A$.

2.1. Definitions and Characterizations.

Tightness is a crucial property of countably additive probabilities, and it is always satisfied for countably additive probabilities on a csm (this is Ulam's theorem, e.g. Billingsley (1968, Thm. 1.4, p. 10)). For finitely additive probabilities, there is a stronger and a weaker version of tightness. For any non-empty subset $F \subset M$ and any $\delta > 0$, F^δ denotes the δ-neighborhood of F, $F^\delta = \{x \in M : d(x, F) < \delta\}$.

Definition 1. *A finitely additive Borel probability q on a csm (M, d) is*

(a) ***conditionally tight (c-tight)*** *if for every measurable $E \subset M$ and for every $\epsilon > 0$, there exists a compact $K \subset E$ such that $q(K) \geq (1-\epsilon)q(E)$, it is*
(b) ***tight*** *if for every $\epsilon > 0$, there exists a compact $K \subset M$ such that $q(K) > (1-\epsilon)$, and it is*
(c) ***neighborhood tight (n-tight)*** *if for every $\epsilon > 0$, there exists a compact K such that for all $\delta > 0$, $q(K^\delta) > (1-\epsilon)$.*

Countable additivity is equivalent to c-tightness.[1]

Lemma 1. *A finitely additive Borel probability q on M is c-tight if and only if it is countably additive.*

[1]All proofs are in Appendix B. Christensen (1971) systematically studies properties of finitely additive Borel probabilities that guarantee countable additivity, showing that they are countably additive if they "do not behave very irregularly on the closed sets." The result closest to the present study of finitely additive Borel probabilities is his Theorem 5.

Being at Fortet-Mourier distance 0 from some countably additive probability is equivalent n-tightness. For finitely additive p and q, define the pseudo-metric

$$\beta_M(p,q) = \sup\left\{\left|\int f\,dp - \int f\,dq\right| : \|f\|_{Lip} \leq 1\right\} \tag{3}$$

where $\|f\|_{Lip} := \sup_{a \in A} |f(a)| + \sup_{a \neq b} |f(a) - f(b)|/d(a,b)$ is the Lipschitz norm of the function f. From Fortet and Mourier (1953), $\beta_M(\cdot,\cdot)$ is a metric, not merely a pseudo-metric, on the class of countably additive probabilities. Further, $\beta_M(\cdot,\cdot)$ metrizes the weak* topology for countably additive probabilities on csm's — $\beta_M(p_n, p) \to 0$ if and only if $\int_M f\,dp_n \to \int_M f\,dp$ for every continuous $f : M \to \mathbb{R}$ (Dudley, 1989, Theorems 11.3.1 and 11.3.3).

Lemma 2. *A finitely additive Borel probability q on M is n-tight if and only if there exists a countably additive q^{ca} such that $\beta_M(q, q^{ca}) = 0$.*

2.2. **Examples.** It is clear that c-tightness implies tightness and that tightness implies n-tightness. Neither of the implication reverses, and n-tightness can fail. The first and the third examples use the probability space $(T, \mathcal{T}, \mu)$ where $T = \mathbb{R}_{++}$, $\mathcal{T}$ is the (usual) Borel σ-field, and μ is a non-atomic, purely finitely additive probability on $\mathcal{T}$ with $\mu([r,\infty)) = 1$ for all $r \in \mathbb{R}$.

Example 1 (Tight but not c-tight)**.** *Let $f \neq g$ be two functions in the unit ball $C([0,1])$. For each $t \in T$, define $\mathcal{G}(t) = f + \frac{1}{t}g$ and let $p = \mathcal{G}(\mu)$ denote the induced measure on $C([0,1])$. For each $n \in \mathbb{N}$, $p(E_n) = 1$ where $E_n := \{f + rg : 0 < r < \frac{1}{n}\}$. Because $E_n \downarrow \emptyset$, p is purely finitely additive, and by Lemma 1, not c-tight. However, $p(K) = 1$ where K is the compact set $\{f + rg : 0 \leq r \leq 1\}$.*

The next probability assigns mass 0 to every compact set, but assigns mass 1 to every open neighborhood of the origin in $C([0,1])$, hence satisfies n-tightness. Recall that convergence in the weak* topology for finitely additive probabilities is defined by $p^\alpha(E) \to p(E)$ for every measurable E, equivalently $\int g\,dp^\alpha \to \int g\,dp$ for every bounded measurable g. By Alaoglu's theorem (e.g. Corbae et al. (2009, Thm. 10.7.1)), the set of finitely additive Borel probabilities on $C([0,1])$ are compact in the weak* topology.

Example 2 (N-tight but not tight)**.** *It can be shown[2]that for every $n \in \mathbb{N}$, there exists a finitely additive p_n with $p_n(\frac{1}{n} \cdot \mathbb{U}) = 1$ and $p_n(K) = 0$ for all compact $K \subset \mathbb{U}$. Let p be a weak* accumulation point of the p_n. For each $\delta > 0$, $p(B_\delta(\{0\}) = \lim_n p_n(B_\delta(\{0\}) = 1$ and for each compact K, $p(K) = \lim_n p_n(K) = 0$.*

In the next example, the probability assigns mass 0 to the δ-ball around every compact set in $C([0,1])$, $\delta < 1/4$.

Example 3 (Not n-tight)**.** *For each $t \in T$, define $\mathcal{G}(t)(\cdot)$ by*

$$\mathcal{G}(t)(x) = \begin{cases} 1 & \text{if } x \leq \frac{1}{2}, \\ 1 - t(x - \frac{1}{2}) & \text{if } \frac{1}{2} \leq x \leq \frac{1}{2} + \frac{2}{t}, \text{ and} \\ -1 & \text{if } \frac{1}{2} + \frac{2}{t} \leq x. \end{cases} \tag{4}$$

For each t, $\|\mathcal{G}\| = 1$ so that $p = \mathcal{G}(\mu)$ is a finitely additive probability on the unit ball. The maximal absolute value of the slope of $\mathcal{G}(t)$ is t. Therefore, for any $r \in \mathbb{R}$, a μ-mass 1 set of agents have slope greater than r.

---

[2]See Lemma 7 in Appendix B1.

Pick arbitrary compact $K \subset C([0,1])$. By the Arzelà-Ascoli theorem (e.g. Corbae et al. (2009, Theorem 6.2.61)*), $K \subset C([0,1])$ must be uniformly equicontinuous. Since a μ-mass 1 set of agents have range from -1 to $+1$ with a slope having absolute value greater than r, it can be shown*[3]*that $p(K^{1/4}) = 0$.*

2.3. **Failures of Tightness in Population Games.** A non-atomic population game, Γ, is specified by the tuple $((T, \mathcal{T}, \mu), \mathbb{U}, \mathcal{G})$. Here, $(T, \mathcal{T}, \mu)$ is a finitely additive, non-atomic probability space. $\mathbb{U}$ is the unit ball on $C(A \times \Delta(A))$ where A is a compact metric space, and $\Delta(A)$ is the set of countably additive Borel probabilities on A metrized by $\beta_A(\rho, \rho') = \sup\{|\int_A f\, d\rho - \int_A f,\, d\rho'| : \|f\|_{Lip} \leq 1\}$. $\mathbb{U}$ is endowed with the supnorm topology and the corresponding Borel σ-field, and $\mathcal{G}$ is a measurable mapping from T to $\mathbb{U}$. The population distribution of utility functions, $\mathcal{G}(\mu)$, is denoted p.

Definition 2. *A game Γ is **c-tight/tight/n-tight** as p is c-tight/tight/n-tight.*

By Lemma 1, c-tightness is equivalent to countable additivity. Therefore, c-tight games have equilibria and a continuous approximate equilibrium correspondence.

- Game 1 shows that games that are n-tight but not c-tight may fail to have equilibria even if they have ϵ-equilibria for all $\epsilon > 0$.
- Game 2 shows that games that not n-tight may not have ϵ-equilibria for a range of ϵ.
- Game 3 shows that games that are not n-tight may have equilibria.

The games use the following probability space to model the agents: $T = \mathbb{R}_{++}$; $\mathcal{T}$ is the (usual) Borel σ-field; and μ is a non-atomic, purely finitely additive probability on $\mathcal{T}$ with $\mu([r, \infty)) = 1$ for all $r \in \mathbb{R}$. For the first two games, the action space is $A = \{0, 1\}$, $\Delta(A) = [0, 1]$, and the utility function $\mathcal{G}(t)$ depends on a_t and $\nu \in [0, 1]$ interpreted as the mass of agents in T that play $a = 1$.

2.3.1. *N-tight but not C-tight.* The following example, from Khan et al. (2016), is tight but not c-tight. It has ϵ-equilibria for all $\epsilon > 0$ but has no equilibrium. Directly after the example is a sketch of how to modify the game to be n-tight but not tight.

Game 1. *If $\mathcal{G}(t)$ is the utility function $a \cdot (\frac{1}{t} - \nu)$, the game has no equilibrium, but it does have ϵ-equilibria for every $\epsilon > 0$.*

- For the no-equilibrium argument, suppose first that $\nu_a > 0$ is the population distribution of an equilibrium $a(\cdot)$. The action 1 is one of the best responses for $t \in T$ only if $(\frac{1}{t} - \nu_a) \geq 0$, that is only if $t \leq 1/\nu_a$. By assumption, $\mu((0, 1/\nu_a]) = 0$, so, if the proportion of the population choosing 1 is strictly positive, then it must be 0. On the other hand, if $\nu_a = 0$, then for each $t \in T$, $\frac{1}{t} > \nu_a$, so every t should (apparently) play the action 1, making $\nu_a = 1$.

[3]Let $t_n \uparrow \infty$ in T. If $d(\mathcal{G}(t_n), K) \leq \frac{1}{4}$ for infinitely many n, then, taking a subsequence if necessary, there exists a sequence $f_n \in K$ such that $\|f_n - \mathcal{G}(t_n)\| \leq \frac{1}{4}$. Taking another subsequence if necessary, $f_n \to f$ for some f in the compact set K, and $\|f - \mathcal{G}(t_n)\| \leq \frac{1}{4}$ for each t_n in the subsequence. But $\|f - \mathcal{G}(t_n)\| \leq \frac{1}{4}$ implies that $f(\frac{1}{2}) \geq \frac{3}{4}$ because $\mathcal{G}(t_n)(\frac{1}{2}) \equiv 1$. Because f is continuous, there exists some $\delta > 0$ such that for each x in the interval $(\frac{1}{2} - \delta, \frac{1}{2} + \delta))$, $|f(\frac{1}{2}) - f(x)| < \frac{1}{4}$. For all n with $\frac{1}{t_n} < \delta$, $\|f - \mathcal{G}(t_n)\| \geq \frac{1}{2}$. Since $\mu(\{t : \frac{1}{t} < \delta\}) = 1$, $p(K^{\frac{1}{4}}) = 0$.

- For the ϵ-equilibrium existence, let $a(t) = 1_E(t)$ for E with $\mu(E) = \epsilon$, $\epsilon > 0$. Each t in a full measure subset of E^c is strictly best responding, while each $t \in E$ is losing at most ϵ of their possible utility.

In Example 2, the probability p puts mass 0 on all compact sets K and mass 1 on all neighborhoods of 0. One can change the construction to yield a probability that also puts mass 0 on the compacts while putting mass 1 on the intersection of each $B_\delta(0)$ and the strictly positive orthant in $\mathbb{U}$. Let $t \mapsto \mathcal{H}(t)$ be a random point in the strictly positive orthant of $\mathbb{U}$ having distribution q and set the utility function equal to $a \cdot (\mathcal{H}(t) - \nu)$. The analysis of the ϵ-equilibria and the non-existence of an equilibrium is essentially unchanged, and we now have a game that is n-tight but fails to be tight.

2.3.2. *Not N-tight.* Unlike the previous game, the following, also due to Khan et al. (2016), has no approximate equilibria. The induced distribution on utility functions is taken from Example 3.

Game 2. *$\mathcal{G}(t)$ is the utility function $a \cdot u(t, \nu)$ where*

$$u(t,\nu) = \begin{cases} 1 & \text{if } \nu \leq \frac{1}{2}, \\ 1 - t(\nu - \frac{1}{2}) & \text{if } \frac{1}{2} \leq \nu \leq \frac{1}{2} + \frac{2}{t}, \text{ and} \\ -1 & \text{if } \frac{1}{2} + \frac{2}{t} \leq \nu. \end{cases} \tag{5}$$

Note that the utility function of type t has a section between $\frac{1}{2}$ and $\frac{1}{2} + \frac{2}{t}$ with the absolute value of its slope equal to t. This means that the average maximal slope, $\int_T t \, d\mu(t)$ is undefined (or equal to $-\infty$ if one prefers that set of conventions).

The approximate best response sets in Game 2 have the following properties:

- for $\nu \leq \frac{1}{2}$, every t has a best response of 1, and they lose utility of 1 by playing 0, so that ϵ-best responses must put mass at least $1 - \epsilon$ on $a = 1$; and
- for $\nu > \frac{1}{2}$, we have $\mu(\{t : \frac{1}{2} + \frac{2}{t} < \nu\}) = \mu(\{t : t > 2(\nu - \frac{1}{2})\}) = 1$, so that a mass 1 set of players loses utility of 1 by playing $a = 1$, and ϵ-best responses must put mass at least $1 - \epsilon$ on $a = 0$.

The implications for ϵ-equilibria follow.

- If $\nu_a \leq \frac{1}{2}$ is an ϵ-equilibrium distribution, then $\int a(t)(\{1\}) \, d\mu(t) \geq (1 - \epsilon)^2$ because at least $1 - \epsilon$ of the population must put mass at least $1 - \epsilon$ on 1. Thus, if $\nu_a \leq \frac{1}{2}$, then $\nu_a \geq (1 - \epsilon)^2$, which cannot be satisfied for $\epsilon < 1 - 1/\sqrt{2}$.
- In the same fashion, if $\nu_a > \frac{1}{2}$ is the population play in an ϵ-equilibrium $a(\cdot)$, then $\nu_a \leq \epsilon(1 - \epsilon)$, which cannot be satisfied for any ϵ.

2.3.3. *Not N-Tight with Equilibria.* It the following example, the population measure fails n-tightness, but the equilibrium analysis can, with the exception of some interpretational difficulty, proceed without impediment.

Game 3. *The type space, $(T, \mathcal{T}, \mu)$, is unchanged, but now the action space is $A = [\pi, 2\pi]$, and agent t's utility when $\nu_a \in \Delta(A)$ is the population distribution is*

$$\mathcal{G}(t) = \sin(t \cdot a(t)) \left[1 + \textstyle\int_A x \, d\nu_a(x)\right]. \tag{6}$$

- The population measure puts no mass on δ-neighborhoods of any compact set: the slope of type t's utility function ranges from $-t$ to positive t as the function cycles between +1 to −1; for any $r \in \mathbb{R}$, $\mu([r, \infty)) = 1$, so the slope must be, with probability 1, larger than any $r \in \mathbb{R}$.

- For any $t \in \mathbb{R}$, the strict best responses, denoted $a^*(t)$, are independent of what other agents choose. They consist of all solutions to $\sin(t \cdot a) = 1$ for $\pi \leq a \leq 2\pi$.
- Equilibrium play involves each t playing some $a \in a^*(t)$ and receiving the utility $1 + \int x \, d\nu_a(x)$. For different selections from $a^*(\cdot)$, the integral $\int x \, d\nu_a(x)$ can be anywhere in $[\pi, 2\pi]$.

The interpretational difficulty is that when t is infinite, the utility function cycles between +1 and −1 infinitely often as t's choice runs from π to 2π. Here however, the game is easy to analyze because the measurable selections from $a^*(\cdot)$ are also viable strategies for the expanded versions of the game.

3. Nonstandard Analysis Versions Population Games

A non-atomic population game, Γ, is specified by $((T, \mathcal{T}, \mu), \mathbb{U}, \mathcal{G})$ as above. This section works replaces the spaces defining Γ by larger versions of the spaces that appear in nonstandard analysis.

3.1. **The Nonstandard Analysis Setting.** We work in a κ-saturated extension of a superstructure containing, as a bounded element, the measure space T and the compact set of actions A and κ is a cardinal larger than the class of all subsets of the game Γ. For E a bounded element of the superstructure, ${}^*E$, read "star E," denotes its nonstandard version.

The most relevant implication of κ-saturation is the existence of **exhaustive** ***-finite** (or **exhaustive hyperfinite**) sets: for non-empty $X \subset \Gamma$, $\mathcal{P}_F(X)$ denotes the finite subsets of X; κ-saturation guarantees that there exists an $X^f \in {}^*\mathcal{P}_F$ with $x \in X_f$ for each $x \in X$. It will be convenient to replace T and $\mathcal{T}$ by exhaustive hyperfinite version of themselves below.

For any space X with topology τ_X in the superstructure, and for any $x \in X$, let $m(x) \subset {}^*X$ denote the **monad** of x, defined as $\bigcap_{G \in \tau_X; x \in G} {}^*G$. The standard part relation is defined by $x = \mathrm{st}(y)$ if $y \in m(x)$. A point y in ${}^*X$ is **nearstandard** if $y \in m(x)$ for some $x \in X$, and the class of nearstandard points is denote $\mathrm{ns}({}^*X)$. If the topology τ_X is Hausdorff, then $y \mapsto \mathrm{st}(y)$ is a function on the set of nearstandard points in ${}^*X$.

For the space $\mathbb{R}$ with the usual (Euclidean) topology, the standard part of $y \in \mathrm{ns}({}^*\mathbb{R})$ is denoted ${}^\circ y$, and when $|y| < \epsilon$ for every standard $\epsilon > 0$ in $\mathbb{R}$ (not ${}^*\mathbb{R}$), we write $y \simeq 0$ for "y is infinitesimal." For nearstandard y, $|y - {}^\circ y| \simeq 0$.

For an internal function $g : {}^*X \to {}^*[-1, +1]$, $\mathrm{st_V}(g)$, is the **standard value version of** g defined by $x \mapsto {}^\circ g(x)$. Thus, $\mathrm{st_V}({}^*\mathbb{U})$ is the set of functions $(a, \nu) \mapsto {}^\circ f(a, \nu)$ where $f \in {}^*\mathbb{U}$, and for $t \in {}^*T$, $\mathrm{st_V}({}^*\mathcal{G})(t)$ is the function $(a, \nu) \mapsto {}^\circ{}^*\mathcal{G}(t)(a, \nu)$.

The set of finitely additive probabilities on $\mathcal{T}$, $\Delta(\mathcal{T})$, is given the weak* topology, which is defined by the property that a net $\mu_\alpha \to \mu$ if and only if $\int g \, d\mu_\alpha \to \int g \, d\mu$ for all $\mathcal{T}$-measurable $g : T \to [-1, +1]$. Because $\Delta(\mathcal{T})$ is a compact Hausdorff space in the weak* topology, every $\mu' \in {}^*\Delta(\mathcal{T})$ is nearstandard to a unique $\mu \in \Delta(\mathcal{T})$. Explicitly, for $\mu' \in {}^*\Delta(\mathcal{T})$, there exists a unique $\mu \in \Delta(\mathcal{T})$ such that for all $\mathcal{T}$-measurable $g : T \to [-1, +1]$, $\int g \, d\mu = {}^\circ \int {}^*g \, d\mu'$.

3.2. **Games with Nonstandard Pieces.** There is a focal nonstandard version of the games Γ, while hyperfinite versions of the game are also useful. The focal nonstandard versions of the games are denoted Γ^{NS}, they expand each part of Γ to its nonstandard version (by "putting *'s in from of everything"), but, and this is crucial, they take the standard part of the value of the probability and the utility functions so that they take values in $[0,1]$ and $[-1,+1]$ respectively.

Definition 3. *For every game Γ, the game Γ^{NS} is defined as*

$$\Gamma^{NS} = (({}^*T, \sigma({}^*\mathcal{T}), {}^{\circ*}\mu), \mathrm{st}_{\mathrm{V}}({}^*\mathbb{U}), \mathrm{st}_{\mathrm{V}}({}^*\mathcal{G})). \tag{7}$$

In Γ^{NS}: set of player coalitions is given by $\sigma({}^*\mathcal{T})$, the smallest σ-field containing ${}^*\mathcal{T}$; the population measure is given by the countably additive (Loeb) probability ${}^{\circ*}\mu : \sigma(\mathcal{T}) \to [0,1]$; the set of possible utility functions, $\mathrm{st}_{\mathrm{V}}({}^*\mathbb{U})$, is the set of functions $\mathrm{st}_{\mathrm{V}}(u)$, $u \in {}^*\mathbb{U}$; the strategies are the ${}^*\mathcal{T}$-measurable, functions from ${}^*T$ to ${}^*\Delta(A)$; and the value of characteristics mapping, $\mathrm{st}_{\mathrm{V}}({}^*\mathcal{G})$ at any $t \in {}^*T$ is the function $\mathrm{st}_{\mathrm{V}}({}^*\mathcal{G}(t))$. This belongs to $\mathrm{st}_{\mathrm{V}}({}^*\mathbb{U})$, and maps ${}^*(A \times \Delta(A))$ to $[-1,+1]$.

The role of n-tightness can be seen in the following consequence of Lemma 2.

Lemma 3. *If $p = \mathcal{G}(\mu)$ is an n-tight probability on $\mathbb{U}$, then ${}^{\circ*}p(\mathrm{ns}({}^*\mathbb{U})) = 1$.*

3.3. **Restored, Nonstandard Games.** Restoring mislaid agents and their utility functions delivers equilibrium existence to Games 1, 2, and provides interpretation for the equilibria of Game 3.

For Game 1, the existence of ϵ-equilibria for all $\epsilon > 0$ leads, by overspill, to ϵ-equilibria for $\epsilon \simeq 0$, and by the definition of $\mathrm{st}_{\mathrm{V}}(\cdot)$ and ${}^{\circ*}\mu$, these are equilibria.

Recall that the type space for all three games is $(T, \mathcal{T}, \mu) = (\mathbb{R}_+, \mathcal{B}, \mu)$ where μ is non-atomic and satisfies $\mu([r, +\infty)) = 1$ for all $r \in \mathbb{R}$. Because the finitely additive probabilities are a compact Hausdorff space in the weak* topology, μ is the weak* standard part of ${}^*\mu$, and indeed many other $\mu' \in {}^*\Delta^{fa}(\mathcal{B})$.

NS-Game 1. *If $\mathcal{G}(t) = a \cdot (\frac{1}{t} - \nu)$, then for a ${}^{\circ*}\mu$-mass 1 set of agents, the utility function $\mathrm{st}_{\mathrm{V}}({}^*\mathcal{G}(t)) = a \cdot (-\nu)$ because the nearstandard $t \in {}^*T$ are a null set. In this game, the equilibria involve all but a null set of $t \in T$ playing $a_t = 0$.*

Game 2 had no approximate equilibria, and the population distribution of utility functions put no mass on the neighorhood of any compact set. The mislaid utility functions are therefore not nearstandard, and the equilibrium joint distribution of actions and utilities depends on the cdf (cumulative distribution function) of ${}^*\mu$ on ${}^*\mathbb{R}$. This is an important point, details of μ, as they appear in ${}^*\mu$, matter for the analysis of the equilibrium.

NS-Game 2. *Suppose that $\mathcal{G}(t)$ is the utility function $a \cdot u(t, \nu)$ where*

$$u(t,\nu) = \begin{cases} 1 & \text{if } \nu \leq \frac{1}{2}, \\ 1 - t(\nu - \frac{1}{2}) & \text{if } \frac{1}{2} \leq \nu \leq \frac{1}{2} + \frac{2}{t}, \text{ and} \\ -1 & \text{if } \frac{1}{2} + \frac{2}{t} \leq \nu. \end{cases} \tag{8}$$

Let $F_{{}^*\mu}(\cdot)$ denote the cdf of ${}^*\mu$. Let $t_c \in {}^*\mathbb{R}_{++}$ solve $F_{{}^*\mu}(t_c) = \frac{1}{2} + \frac{1}{t_c}$. The equilibria involve — up to ${}^{\circ*}\mu$-null sets of agents — all $t < t_c$ playing $a_t = 1$, and the remaining t playing $a_t = 0$. The distribution of utility for those playing $a_t = 1$ depends on the cdf, e.g. if μ is the standard part of the uniform distribution on $[0, N]$, N infinite, then $t_c \simeq N/2 + 2$, and in equilibrium, an infinitesimal more than $\frac{1}{2}$ of the population is playing $a = 1$ and receiving utility uniformly distributed on $[0,1]$, while the remainder of the population is playing $a = 0$ and receiving utility equal to 0.

In the previous games, the action space A was finite, hence equal to ${}^*A$. In the next game, ${}^*A$ contains every $a \in A$, as well as nonstandard points.

NS-Game 3. *Suppose that the action space is $A = [\pi, 2\pi]$, and $\mathcal{G}(t)$ is*

$$\mathcal{G}(t) = \sin(t \cdot a(t)) \left[1 + \int_A x \, d\nu_a(x)\right] \tag{9}$$

when $\nu_a \in {}^*\Delta(A)$ is the population distribution. For a ${}^{\circ *}\mu$-mass 1 set of agents, $\mathcal{G}(t)$ is the infinitely variable utility function $a \mapsto \kappa \cdot \sin(t \cdot a)$ where $\kappa \geq 1$ is independent of t's choice. The Pareto dominant equilibria involve every such agent playing the largest solution to $\max_{a \in {}^*A} \sin(t \cdot a)$, and in this equilibrium, a mass 1 set of agents receive utility $1 + 2\pi$.

3.4. **Equilibrium Existence and Approximation.** Games models are "well-behaved' if equilibria exist and the approximate equilibrium outcome correspondence is continuous. For this, we need the outcome mapping and finite approximation to games. Taking the outcome mapping first, a strategy $t \mapsto \sigma_t$ in Γ^{NS} induces a joint distribution on the space of actions and utilities $({}^*A) \times [-1, +1]$. Denote by $\mathbb{O}(\sigma)$ the distribution on the compact space $A \times [-1, +1]$ induced by the mapping $(a, r) \mapsto (\mathrm{st}(a), r)$.

To discuss finite approximations and the approximate equilibrium correspondence, let $\mathbb{F}$ denote the set of finite, $\mathcal{T}$-measurable partitions of T. Partially order $\mathbb{F}$ by $f \succsim f'$ if f refines f' and define an $f \in {}^*\mathbb{F}$ to be **infinitely fine** if $f \succsim f'$ for each $f' \in \mathbb{F}$. By κ-saturation, ${}^*\mathbb{F}$ contains infinitely fine f's. For any positive $\delta > 0$ and $u \in \mathbb{U}$, the inverse image of the δ-ball around u, $\mathcal{G}^{-1}(B_\delta(u))$, belongs to $\mathcal{T}$. Therefore, any infinite f has the property that the diameter of ${}^*\mathcal{G}(B)$ is infinitesimal for any $B \in f$.

For any such f, and define the hyperfinite approximation to Γ^{NS} by

$$\Gamma^f = (({}^*T, \mathcal{T}(f), {}^*\mu), {}^*\mathbb{U}, {}^*\mathcal{G}(f)) \tag{10}$$

where $\mathcal{T}(f)$ is the smallest internal field of sets containing f, and $\mathcal{G}^f(t) = E^\mu(\mathcal{G}|B_t)$ if $t \in B_t \in f$ and ${}^*\mu(B_t) > 0$, and $\mathcal{G}^f(t) = u_t$ for some $u_t \in \mathcal{G}(B_t)$ if $t \in B_t \in f$ and ${}^*\mu(B_t) = 0$. In this game, utilities take values in the nonstandard set of numbers, ${}^*[-1, +1]$, rather than in the set $[-1, +1]$.

Recall that $\beta(\cdot, \cdot)$ is the Fortet-Mourier distance between probabilities.

Theorem A. *Every Γ^{NS} has an equilibrium. Further,*

1. *if $f \in {}^*\mathbb{F}$ is infinitely fine, $\epsilon > 0$ is infinitesimal, and a^f is an ϵ-equilibrium of Γ^f, then a^f is an equilibrium of Γ^{NS},*
2. *if a^* is an equilibrium of Γ^{NS} and $f \in {}^*\mathbb{F}$ is infinitely fine, then $E^\mu(a^*|f)$ is an ϵ-equilibrium of Γ^f for some infinitesimal $\epsilon > 0$,*
3. *the approximate equilibrium outcome correspondence, $(f, \epsilon) \mapsto \mathbb{O}(Eq^\epsilon(\Gamma^f))$, is continuous as $\epsilon \downarrow 0$ and f becomes infinitely fine,*
4. *if $p = \mathcal{G}(\mu)$ is c-tight, then there exists a one-to-one onto mapping between the outcomes of the equilibria of Γ and the outcomes of equilibria in Γ^{NS}, and*
5. *if $p = \mathcal{G}(\mu)$ is n-tight and $\hat{a}$ is an equilibrium of Γ^{NS}, then for all $\epsilon > 0$, there exists an ϵ-equilibrium, a^ϵ of Γ with $\beta(\mathbb{O}(\hat{a}), \mathbb{O}(a^\epsilon)) < \epsilon$.*

Comments. The first two parts use conditional probabilities for hyperfinite spaces as developed in Anderson (1982). Equilibrium existence follows from the first part and transfer of the statement that equilibria exist in finite population games. The continuity of the approximate equilibrium correspondence is a direct application of the conversion of nonstandard results into statements about ultrafilter limits, e.g. Robinson (1964, Thm. 5.1). The last part implies that n-tight games always have ϵ-equilibria.

4. Compact Imbeddings and Population Games

The previous section found the missing pieces of the games using nonstandard extensions. This section finds the missing pieces via imbeddings into subsets of compact spaces. The results here are direct corollaries of the previous section because the compact spaces into which we are imbedding the original game can be expressed as quotient spaces of their nonstandard versions.[4]

If (H, τ) is a compact Hausdorff space and X is a non-empty set, then any one-to-one $\varphi : X \to H$ is a **compact imbedding**. If X has a topology and φ is a homeomorphism, then the closure of $\varphi(X)$ in H provides a **compactification**, that is, a homeomorphism between the original space and a dense subset of a compact Hausdorff space. Endowing X with the topology $\varphi^{-1}(\tau)$ turns a compact imbedding into a compactification.

This section begins with the the general quotient space construction that we use, then gives the properties of the two spaces $(\widehat{T}, \widehat{\mathcal{T}}, \widehat{\mu})$ and $\widehat{\mathbb{U}}$. The extension, $\widehat{\mathcal{G}}$ of $\mathcal{G}$ follows, and the properties of the game $\widehat{\Gamma}$ follow from the previous section.

4.1. Generating Compact Imbeddings.

The following construction will be used repeatedly.

Definition 4. *For X a non-empty internal set and $\mathcal{F}$ as a set of internal functions from X to ${}^*[-1, +1]$, define $x \sim_{\mathcal{F}} y$ if $\mathrm{st}_{\mathrm{V}}(f)(x) = \mathrm{st}_{\mathrm{V}}(f)(y)$ for each $f \in \mathcal{F}$, define $X' = X/\sim$ as the quotient space of equivalence classes, and define $\mathbb{K}(X; \mathcal{F})$ as the topological space $(X', \tau_{\mathcal{F}})$ where $\tau_{\mathcal{F}}$ is the smallest topology on X making each function $\mathrm{st}_{\mathrm{V}}(f) : X \to [-1, +1]$, $f \in \mathcal{F}$, continuous.*

Comments. The sets X' and X are identical if $\mathcal{F}$ **separates points**, that is, if $x \neq y$ in X if and only if $\mathrm{st}_{\mathrm{V}}(f)(x) \neq \mathrm{st}_{\mathrm{V}}(f)(y)$ for some $f \in \mathcal{F}$. The set $\mathcal{F}$ need not be internal and will, in one instance, contain functions that are *not* of the form ${}^*f$ for some $f : X \to \mathbb{R}$.

The following result is, basically, a restatement of the parts of Anderson (1982, Theorem 4.3.3) needed here. For completeness, a proof is included in the appendix.

Lemma 4. $\mathbb{K}(X; \mathcal{F})$ *is a compact Hausdorff space, and the supnorm closure of the smallest algebra containing the functions* $\mathrm{st}_{\mathrm{V}}(\mathcal{F})$ *is the set of continuous functions.*

Two classic compactifications demonstrate what is involved.

- If $X = {}^*\mathbb{N}$ and $\mathcal{F}$ is the set of functions ${}^*g$ where g is a function from $\mathbb{N}$ to $[-1, +1]$, then $\mathbb{K}(X; \mathcal{F})$ is homeomorphic to $\beta\mathbb{N}$, the Stone-Čech compactification $\mathbb{N}$.
- If $(T, \mathcal{T})$ is a non-empty set and a σ-field of subsets that separates points, $X = {}^*T$ and $\mathcal{F}$ is the set of functions ${}^*g$, g a $\mathcal{T}$-measurable function from T to $[-1, +1]$, then $\mathbb{K}(X; \mathcal{F})$ is homeomorphic to the Stone space for $(T, \mathcal{T})$.

The Stone space has a central role to play.

[4] Robert Anderson, in private communication, has explained the utility of nonstandard versions of a space by the observation that they contain, simultaneously, compactifications with respect to all the properties one could want. This section provides one more instance of this general pattern.

4.2. **The Compact Imbedding of** $(T, \mathcal{T}, \mu)$**.** Define $\widehat{T} = \mathbb{K}({}^*T; \mathcal{F})$ where $\mathcal{F}$ is

$$\{\mathrm{st_V}({}^*g) : g \text{ is a measurable function from } T \text{ to } [-1, +1]\}.$$

It is immediate that $T \subset \widehat{T}$, and, from the construction, that every measurable $g : T \to [-1, +1]$ has a unique continuous extension $\widehat{g} : \widehat{T} \to [-1, +1]$.

4.2.1. *More Continuous Extensions.* Every measurable g taking values in a compact Hausdorff space, H, not just in $[-1, +1]$, also has a unique continuous extension. This is useful for population games when the measurable function is $\mathcal{G}$ and the unit ball, $\mathbb{U}$, has been compactified. The unique continuous extension is given in steps. First, homeomorphically imbed H in $\times_{f \in \mathbb{U}(H)} I_f$ where $\mathbb{U}(H)$ is the unit ball in $C(H)$ and $I_f = f(H) \subset [-1, +1]$ is the corresponding, necessarily compact range of f. Second, each a measurable $g : T \to H$ becomes, after the homeomorphic imbedding of H, a collection, g_f, indexed by the $f \in \mathbb{U}(H)$, of $[-1, +1]$-valued measurable functions. Each of these has a unique continuous extension, $\widehat{g}_f$, to $\widehat{T}$, and the mapping $\widehat{g}$ is the homeomorphic image of $(\widehat{g}_f)_{f \in \mathbb{U}(H)}$.

4.2.2. *Extensions of Probabilities.* For each $E \in \mathcal{T}$, the function 1_E is measurable and takes values in $\{0, 1\}$, hence its unique continuous extension, denoted h_E, takes values in this two-point set. Let $\widehat{E}$ denote the set of $t \in \widehat{T}$ for which $h_E(t) = 1$, and define $\widehat{\mathcal{T}}$ to be the smallest σ-field containing all of the $\widehat{E}$.

The probability μ has a unique continuous extension to $\widehat{\mathcal{T}}$. Let $M_b(T)$ denote the bounded measurable functions on T and define the positive, continuous, normalized, linear functional $L_\mu : M_b(T) \to \mathbb{R}$ by $L_\mu(g) = \int g \, d\mu$. $C(\widehat{T})$ is the set of extensions, $\widehat{g}$, of the $g \in M_b$. Thus, $\widehat{L}(\widehat{g}) := L_\mu(g)$ defines a positive, continuous, normalized, linear functional on $C(\widehat{T})$. By the Riesz representation theorem, $\widehat{L}$ can be uniquely represented as an integral against a countably additive probability, denoted $\widehat{\mu}$, on $\widehat{\mathcal{T}}$.

4.3. **The Compactification of the Unit Ball.** $\mathbb{U}$ is the unit ball in $C(M)$, where M is the compact metric space, $A \times \Delta(A)$. Define $\widehat{\mathbb{U}}$ as $\mathbb{K}({}^*\mathbb{U}; \Psi)$ where Ψ is the set of functions ${}^*\psi$ with $\psi : \mathbb{U} \to [-1, +1]$ a Lipschitz function.

4.3.1. *Visualizing the Compactification.* The next result prepares for the result that makes these two spaces easy to visualize: $\widehat{M}$ is ${}^*M$, and $\widehat{\mathbb{U}}$ is the unit ball in the set of continuous functions on $\widehat{M}$.

Lemma 5. *The sup norm topology on* $\mathbb{U}$ *is the smallest topology making every* $\psi \in \Psi$ *continuous.*

Using this Lemma, the following tells us that for $M = [0, 1]$, $\widehat{M} = {}^*[0, 1]$ and $C(\widehat{M})$ is the set of functions $\mathrm{st_V}(f)$ where $f \in {}^*C([0, 1])$ and $\|f\| \leq B$ for some finite B. In particular, $C(\widehat{M})$ contains e.g. the infinitely steep functions in NS-Game 2.

Proposition 1. $\widehat{\mathbb{U}}$ *is the unit ball in* $C(\widehat{M})$.

4.3.2. *Visualizing the Expanded Spaces of Actions.* We have already seen, in Game 2, that when $A = \{0, 1\}$ and $\Delta(A) = [0, 1]$, we need to expand $\Delta(A)$ to capture the equilibrium phenomena. The expansion did not involve expanding the two-point space of actions nor the set purely finitely additive distributions on $\{0, 1\}$ — a finitely additive probability on a finite set is vacuously countably additive. What was needed was ${}^{*}\mu((0,t_c]) = \frac{1}{2} + \epsilon$ where ϵ is a strictly positive infinitesimal and t_c is the nonstandard cut-off type indifferent between the actions $a = 0$ and $a = 1$. Proposition 1 shows that there *is* a need to expand the space of actions A when it is not finite — from Proposition 1, the utility functions are continuous functions on a domain (much) larger than $A \times \Delta(A)$, and that domain is what the agents must be choosing from.

The classical compactifications arise when $\mathcal{F}$ is a set of standard functions, but for general analyses of population games, the requisite set of utility functions is ${}^{*}\mathbb{U}$, which contains functions that are not of the form ${}^{*}f$ for $f \in \mathbb{U}$. These nonstandard functions include the infinitely steep functions required for the analysis and the compactification, $\widehat{M}$, of $M = A \times \Delta(A)$ covers this possibility.

Suppose, for the purposes of illustration, that the space of actions $A = [0,1]$. If $\mathcal{F}$ is a set of standard functions, then the quotient map ${}^{*}[0,1] \mapsto \mathbb{K}({}^{*}[0,1]; \mathcal{F})$ is many to one. By contrast, if $U([0,1])$ is the unit ball in $C([0,1])$, then the quotient map ${}^{*}[0,1] \mapsto \mathbb{K}({}^{*}[0,1]; {}^{*}U([0,1]))$ is one-to-one and onto.[5]

4.4. **Mislaid Pieces and Compactifications.** A game is specified by a measure space of types, the space of actions and population distributions over actions, a set of utility functions, and an assignment of types to utility functions. Using the pieces defined above, the compactified version of a game Γ is

$$\widehat{\Gamma} = ((\widehat{T}, \widehat{\mathcal{T}}, \widehat{\mu}), \widehat{\mathbb{U}}, \widehat{\mathcal{G}}). \tag{11}$$

These compact imbeddings/compactifications deliver the following.

Theorem B. *Each $\widehat{\Gamma}$ has an equilibrium. Further, if the characteristics measure, $p = \mathcal{G}(\mu)$,*

1. *is c-tight, then there exists a one-to-one onto mapping between the outcomes of the equilibria of Γ and the outcomes of equilibria in $\widehat{\Gamma}$, and*
2. *if it is n-tight and $\hat{a}$ is an equilibrium of $\widehat{\Gamma}$, then for all $\epsilon > 0$, there exists an ϵ-equilibrium, a^{ϵ} of Γ with $\beta(\mathbb{O}(\hat{a}), \mathbb{O}(a^{\epsilon})) < \epsilon$.*

5. Discussion and Interpretation

Khan et al. (2016) argue that Games 1 and 2 show that countable additivity "is a necessity" for nonatomic population models, I argue that this is only part of the story of finitely additive population models. The other part of the story starts with the observation that finitely additive probabilities are the traces of countably additive probabilities on larger spaces. The traces contain enough information about the mislaid agents and their utility functions that representations can be given. With the representations in place, the games become a useful tool to study population wide maximization behavior.

These models on larger spaces are countably additive, and, as Khan et al. (2016) show, countable additivity is essential for them to be well-behaved. But the resultant models on the larger spaces are a new, and potentially very interesting, class of models, and it is the finitely additive models that deliver the new class. §5.1 discusses a wide range of models used in economics and related fields in which the larger spaces represent mislaid pieces, pieces that are crucial to the use and understanding of the models. §5.2 discusses several settings in which purely finitely

[5] This follows from transfer of following statement, "$x \neq y \in [0,1]$ if and only if there exists $u \in U([0,1])$ such that $|u(x) - u(y)| = 1$."

additive probabilities have been a very productive tool. §5.3 returns to discuss this new class of models for large population games with these perspectives in mind.

5.1. **Mislaid Pieces can Mislead.** Finitely additive probabilities mislay pieces of models in a wide range of contexts. The following examples demonstrate many ways in which the lack of representations of parts of the model can be misleading, or can render sensible analysis difficult, even impossible.

- If representations of maximizing behavior do not exist in the model, then the study of the determinant of changes in maximizing behavior is not possible.
- If representations of equilibria in normal form games do not to exist in the model, then the study of their dependence on e.g. the utility functions of the agents is not possible.
- If representations of time paths of stochastic processes with jumps do not exist in the model, then the study of e.g. queueing systems becomes impossible.
- If representations of the states at which a bettor/investor is glad to have taken a bet/invested fail to exist, then expected utility maximizing agents can, seemingly, be money pumped, and the inclusion of such agents in models of markets renders interpretations of any equilibria that do exist quite problematic.
- If representations of the dictators in Arrow's impossibility theorem cannot be found in the population model, then one has circumvented the force and meaning of the "no dictatorship" condition in social choice theory.
- If representations of generations receiving 0 transfers do not exist in the models used to study intergenerational equity, then sensible interpretation of Pareto optimality becomes impossible.

These will be treated in turn, with more details for the first topic as the subsequent ones have strong parallels.

5.1.1. *Maximization.* Fix a non-empty, infinite set X with non-trivial σ-field $\mathcal{X}$. The finitely additive probabilities on $\mathcal{X}$, Δ^{fa}, are a convex subset of the dual space of the bounded $\mathcal{X}$-measurable functions on X. The weak* topology on Δ^{fa} is defined as the smallest making the mappings $\eta \mapsto \int f\, d\eta$ continuous for each bounded measurable f. With this topology, the convex set Δ^{fa} is compact, and its extreme points are "point-masses," that is, the probabilities that satisfy $\nu(E) = 0$ or $\nu(E) = 1$ for all $E \in \mathcal{T}$. Compactness and continuity are at the core of useful models of maximizing behavior.

Suppose that $f : X \to \mathbb{R}$ is bounded and measurable, but does not achieve its supremum, r_f, that is, there is no $x \in X$ such that $f(x) = r_f$. Despite the non-existence of a solution to $\max_{x\in X} f(x)$, the problem

$$\max_{\nu\in\Delta^{fa}} \int f(x)\, d\nu(x) \tag{12}$$

does have a solution, and its value is r_f. Since (12) involves maximizing a continuous linear function over a compact convex set, at least one of the solutions is an extreme point, that is, a point mass. However, this point mass has no support in the set X.

For the purposes of more explicitly demonstrating how this paper's two techniques for finding and representing the mislaid pieces of models work, take X to be $[0,1]$, $\mathcal{X}$ to be the Borel σ-field, and set $f(x) = x$ if $x < 1$ and $f(1) = 0$ so that $r_f = 1$. The first approach replaces X by ${}^*X$ and replaces the utility function being maximized, $f(\cdot)$, by $u(x) := \text{st}_V({}^*f)(x)$. For any x in ${}^*X$ having standard part less than 1, $u(x)$ is equal to ${}^\circ x$, the standard part of x. A point $x \in {}^*X$ solves the nonstandard version of the problem if and only if it belongs to $\cap_{n\in\mathbb{N}} E_n$ where $E_n := {}^*f^{-1}([r_f - \frac{1}{n}, r_f]) = {}^*\{x \in X : r_f - \frac{1}{n} \leq f(x) \leq r_f\}$. This non-empty set of points consists of the numbers $1 - \epsilon$ where $\epsilon > 0$ is infinitesimal. The solution set is the set of nonstandard numbers in ${}^*[0,1]$ that are strictly larger than any real number in $[0,1)$ but strictly smaller than 1.

The second approach imbeds $[0,1]$ as a dense subset of a compact, Hausdorff space $\widehat{[0,1]}$ having two dual properties: every bounded measurable function, g, on $[0,1]$ has a unique continuous extension, $\widehat{g}$, to $\widehat{[0,1]}$; and every finitely additive probability, ν, has a unique countably additive extension, $\widehat{\nu}$, to the Borel σ-field on $\widehat{[0,1]}$. The continuous function $\widehat{f}$ achieves its maximum on the compact set $\widehat{[0,1]}$. The set of maxima belong to $\cap_{n\in\mathbb{N}} \widehat{E}_n$ where $\widehat{E}_n$ is the closure, in the compact space $\widehat{X}$, of the set $E_n := f^{-1}([r_f - \frac{1}{n}, r_f])$. These points are the standard part of the $1-\epsilon$ solutions from the previous problem, but now the standard part is taken in the compact, Hausdorff topology on $\widehat{[0,1]}$. These are, again, representations of points to the right of every real number in $[0,1)$ but to the left of 1, points that do not exist in the usual model of $[0,1]$, but which do exist in both of the extensions given here.

5.1.2. *Noncooperative Games.* Harris et al. (2005) study normal form games with a finite set of players, I, infinite sets of strategies X_i, and bounded utility functions $u_i : \times_{j\in I} X_i \to \mathbb{R}$. For each $i \in I$, let η_i be a finitely additive mixed strategy on the class of all subsets of X_i, and let $\eta = (\eta_i)_{i\in I}$. The authors characterize, from several points of view, the class of games in which the existence of finitely additive solutions to each agent's maximization problem

$$\max_{\nu_i \in \Delta^{fa}(X_i)} \int u_i(x)\, d(\eta \setminus \nu_i)(x) \tag{13}$$

combine to yields a non-empty set of mixed strategy equilibria (where $(\eta \backslash \nu_i)$ denotes the vector η with ν_i in the i'th position). For this, it is necessary and sufficient that Fubini's theorem holds for all vectors of finitely additive probabilities (i.e. that the order of integration not matter). Stinchcombe (2005) treats the complementary class of normal form games.

For the class of games having utility functions for which Fubini's theorem holds for all vectors of finitely additive mixed strategies, Harris et al. (2005) show that the equilibrium set is non-empty, that it varies upper-hemicontinuously in the agents' utility functions, and that the approximate equilibrium correspondence is continuous. One can arrive at these conclusions by the following steps: replace each X_i by an exhaustive, hyperfinite X_i^f; replace the utility functions by $\text{st}_V({}^*u_i)$; transfer the Nash (1950) existence theorem for finite games to find an equilibrium for the hyperfinite game; and take the standard part of the equilibria in the smallest topology on the X_i such that the $u_i(\cdot)$ are jointly continuous. From Harris et al. (2005), Fubini's theorem holds if and only if this is possible, and the quotient relation involved takes the ${}^*X_i$ to $\widehat{X}_i$ that are compact metric spaces.

5.1.3. *Stochastic Process Theory.* Kingman (1967) shows that there are purely finitely additive probabilities, p, on the set $\mathcal{P}$ of polynomial time paths on $[0,\infty)$ that have the same finite dimensional distributions as e.g. a Poisson process. In more detail, let $\mathcal{X}^\circ$ denote the smallest field making measurable the coordinate projections $f \mapsto f(t)$, $t \in [0,\infty)$, $f \in \mathbb{P}$ and containing $\{\mathbf{0}\}$ where $\mathbf{0}$ denotes the polynomial identically equal to 0. Restricted to $\mathcal{X}^\circ$, a distribution μ specifies the finite dimensional distributions (fidi's) of a stochastic process. The fidi's on a Poisson process lead to a well-defined μ on the field $\mathcal{X}^\circ$ of subsets of the polynomials. This means that, sampling at points in the time line $[0,\infty)$, we are looking at a non-constant pure jump process, but the space of paths is the polynomials.

To show that μ is pfa, define

$$E_n = \{f \in \mathbb{P} : f \neq \mathbf{0},\ f(k/2^n) \in \{0,1,2,\ldots\},\ k = 1,\ldots,n\cdot 2^n\}. \tag{14}$$

It is immediate that E_n is a non-empty element of $\mathcal{X}^\circ$, properties of polynomials imply that $E_n \downarrow \emptyset$. However, $\mu(E_n) \equiv 1$.

Both the nonstandard and the compactification approaches give representations for this. In ${}^*\mathcal{P}$, the probability distribution ${}^\circ{}^*p$ puts mass 1 on *-polynomials that rise from $n-1$ to n at over nearstandard, infinitesimal intervals $(\tau_n - \epsilon, \tau_n + \epsilon) \subset {}^*[0,\infty)$ and are within ϵ of n between $\tau_n + \epsilon$ and $\tau_{n+1} - \epsilon$, where the τ_n's are the jump times of a Poisson process. The compactification of e.g. the unit ball in $\widehat{\mathcal{P}}$ can be constructed just as the compactification of $\mathbb{U}(A \times \Delta(A))$. In the resulting space, the domain of the continuous, here polynomial, functions is expanded so that step size 1 jumps can be realized by infinitely steep climbs over infinitesimal intervals. When infinitesimal time intervals are not observable, this reduces to the usual space of rcll step functions for point processes.

5.1.4. *Expected Utility Maximization.* An expected utility maximizer with the von Neumann-Morgenstern utility function $u(\cdot)$ has wealth w, and the option to pay an amount R to take a bet X. For any countably additive probability p, the law of iterated expectations tells us $E\,u(w+X)$ satisfies, for any partition $\{A_n : n \in \mathbb{N}\}$ of the state space,

$$E\,u(w+X) = \textstyle\sum_n E\,(u(w+X)|A_n)\cdot p(A_n). \tag{15}$$

Suppose that decision maker would take the bet, that $E\,u(w+X) > u(w-R)$.

When X has variability larger than R, a partition can be arranged with the property that for some A_n, $E\,(u(w+X)|A_n) < E\,u(X-R)$, which, by the balance equation (15), implies that for others, denoted A_m, $E\,(u(w+X)|A_m) > E\,u(X-R)$. Learning that the true state belongs to an A_n means that the decision maker has learned that they regret taking the bet, but on the upside, learning that the true state belongs to an A_m means that the decision maker has learned that they are happy to have taken the bet.

With finitely additive probabilities, the situation is much different, the balance or "conglomerability" condition in (15) can fail. Dubins (1975) shows that when p is nonatomic and purely finitely additive, as the priors in Savage (1972) are, there exists a countable partition $\{A_n : n \in \mathbb{N}\}$ with the property that $E\,(u(w+X)|A_n) < E\,u(w+X)$ for each $n \in \mathbb{N}$. This delivers, seemingly, the possibility of money pumps: suppose that the decision maker has willingly paid R to take the bet X because $E\,u(w+X) > u(R)$; reveal which A_n in the partition $\{A_n : n \in \mathbb{N}\}$ has happened; arrange matters so that conditional on each A_n, $E\,(u(w+X)|A_n) < u(w-R)$; at this point, after the realization of each element in a partition of the state space, the decision maker is willing to pay a positive amount to get out of the bet.

What makes the seeming paradox go is that $\sum_n p(A_n) = 1 - \delta$ for a strictly positive δ. Defining $E_n = \cup_{i \geq n} A_i$ yields a sequence $E_n \downarrow \emptyset$ with $p(E_n) \geq \delta$. The mislaid states, the ones that should belong to $\cap_n E_n$, are the ones in which the decision maker is happy to have taken the bet. Stinchcombe (1997) details how both nonstandard versions of the spaces and compactifications of the spaces can be used to construct isomorphic versions of sets of decision problems that contain representations of the mislaid states. The isomorphic versions avoid money pumps, and the other paradoxes, that arise in the finitely additive theory of choice under uncertainty.

5.1.5. *Social Choice Theory.* Fishburn (1970) shows that when the domain is an infinite set of individuals, all of the assumptions in Arrow's impossibility theorem can be satisfied by taking, as the class of decisive coalitions, sets in any maximal free ultrafilter on the space of individuals. It is easy to check that $\mathfrak{F}$ is a maximal free ultrafilter if and only if $\eta(E) := 1_{\mathfrak{F}}(E)$ is a purely finitely additive point mass. Kirman and Sondermann (1972) interpret the support points for such an η as an "invisible dictator," one that is not present in the original space of individuals. The compactification of the measure space of agents used above parallels the Armstrong (1980) interpretation of invisible dictators as " 'agents' in the Stone space of the quotient algebra, in which context social welfare functions induce continuous preference profiles."

Here the mislaid agents are the dictators that Arrow's axiom is meant to exclude. Without a representation for the dictators, one comes to an Arrovian *possibility* theorem for large population models. This is misleading — the ultrafilter social choice rule in Fishburn (1970) picks a dictator, but one without a representation. When the mislaid agents have a representation, the conclusion is the opposite, that Arrow's impossibility theorem also holds for large population models.

5.1.6. *Intergenerational Equity.* In studies of intergenerational equity, patient preferences are operationalized as immunity to a range of perturbations of the generations. Khan and Stinchcombe (2018) show that this requires the tangents to patient social welfare functions be representable as integrals against purely finitely additive measures on the set of generations. This in turn allows for seeming violations of the Pareto principle.

A probability, μ, is purely finitely additive if and only there exists a strictly positive function g such that $\int g\, d\mu = 0$ (see Lemma 6 in §A1 below for this). The seeming violations of the Pareto principle arise when an allocation X is indifferent to $X + g$ where g is strictly positive for every generation but $\int g\, d\mu = 0$. For nonatomic population models, the existence of a reallocation that increases the utility of a null set of agents does not violate Pareto optimality. The missing points in the domain have been mislaid in the full μ-measure set of generations at which a nonnegative function integrating to 0 takes the value 0. Again, one can represent the elements of the full mass set of mislaid generations either with nonstandard versions of the space of generations or with the appropriate compactification.

5.2. **The Virtues of Finite Additivity.** Purely finitely additive probabilities have been spectacularly useful tools in several areas precisely because they do mislay pieces of the model, and they have been modestly useful in other contexts, e.g. providing learning foundations for a subset of multiple prior models of choice.

5.2.1. *Ultraproducts.* The essential construction in both Robinson's nonstandard analysis and the ultraproduct approach to Banach spaces uses a purely finitely additive point mass, η, on the set of all subsets of $\mathbb{N}$. Define two sequences $(x_m)_{m\in\mathbb{N}}$ and $(y_m)_{m\in\mathbb{N}}$ in a Banach space $\mathbb{V}$ to be equivalent if $\mu(\{m \in \mathbb{N} : x_m = y_m\}) = 1$. Then define ${}^*\mathbb{V}$, the nonstandard extension of $\mathbb{V}$, as the set of equivalence classes of all sequences in $\mathbb{V}$.[6] For a recent survey of the uses of ultraproducts across a broad range of mathematics (e.g. stochastic differential equations, ergodic theory, Ramsey theory, Banach densities on $\mathbb{N}$), see Keisler (2010).

For ultrafilter and nonstandard analysis constructions, the equivalence class structure directly captures asymptotic properties of sequences precisely because μ is purely finitely additive and $\mathbb{N}$ lacks points at which one must represent the behavior of the sequence "at infinity" as a point in $\mathbb{V}$. This allows many analyses of infinite dimensional spaces to proceed using tools and intuitions from finite dimensional spaces. In general equilibrium theory, it allows the treatment of nonatomic population models of exchange economies to proceed using tools and intuitions from finite population exchange economies (see Anderson (1992) for a survey).

5.2.2. *Multiple Priors.* One way to provide a foundation for the multiple prior theory of choice under ambiguity is to suppose that an independent and identically distributed sequence of random variables, X_n, has been observed in $(0, 1]$, and that the probability that $X_n \in (a, b]$, $p((a, b])$, has been learned for each $a < b$. Take as the set of priors, $\Pi(p)$, the set of all finitely additive probabilities consistent with p on the field of sets generated by the $(a, b]$. These are countably additive probabilities on the Stone space for $(0, 1]$. The set $\Pi(p)$ has cardinality as larger than the class of all subsets of $\mathbb{R}$. The missing points that support these priors are 'hidden' as purely finitely additive point masses on the unit interval.

The sets of priors $\Pi = \Pi(p)$ has a definite, albeit limited, use for multiple prior models of choice. If K is a compact metric space of consequences and choices in the presence of ambiguity/multiple priors are modeled as choices between measurable functions from the state space to K, then the objects of interest are the sets of induced distributions $f(\Pi) = \{f(\mu) : \mu \in \Pi\}$, f a measurable function. A "face" in $\Delta(K)$ is a set of distributions $\Delta(F)$, F a closed subset of K. Stinchcombe (2016) shows that the class of sets of induced distributions always contains the faces, and if p is non-atomic, then it is exactly the set of convex combinations of faces. Some interesting decision problems in the presence of ambiguity are of this form, but most are not.

5.3. **Population Games, with Mislaid Pieces Restored.** For nonatomic population games, failures of countable additivity on the space of agents involve a lack of representation for a positive mass of agent characteristics. This lack of representation can make the models misleading. Restoring the mislaid agents to the model delivers equilibrium existence and continuity of the approximate equilibrium correspondence. At issue is how complicated the restoration must be to "fix" nonatomic population games.

[6] An index set larger than the integers is sometimes desirable, see the discussion of saturation, p. 24 in §A2 below.

5.3.1. *The Complications.* For n-tight nonatomic population games, Theorem B shows that the compactification approach allows a one-to-one onto mapping between equilibrium outcomes in the original game and the game with added points. For non-atomic population games failing n-tightness, more complicated constructions are required.

There is one essential reason for the complications — when the induced distribution on characteristics/utility functions, $\mathcal{G}(\mu)$, fails to be n-tight, representing the mislaid pieces of the game requires adding extra utility functions. But the extra utility functions are functions on a larger domain, part of that domain is the choices that the agents make, so that A must be expanded, and this in turn requires expanding the set of population distributions of agent choices, $\Delta(A)$.

The most direct way to represent all of the requisite types and utility functions is to use nonstandard extensions of the space of types, $(T, \mathcal{T}, \mu)$, the unit ball in the space of continuous utility functions, $\mathbb{U}(A \times \Delta(A))$, and the space of distributions induced on $\mathbb{U}$ by the characteristics mapping $\mathcal{G}$, and this is the content of the results in §3.4. Including the requisite types and utility functions by compactification requires a new class of compact spaces, studied in §4.

Proposition 1 contains the essential construction for §4. It gives the new, compact Hausdorff space, $\widehat{M}$, that is the domain for the new unit ball of continuous utility functions, $\widehat{\mathbb{U}}$. With this space in hand, it is easy to recognize when the game needs no extra points — when it is n-tight — and this is the content of Theorem B.

5.3.2. *The Potential.* The failures of countable additivity that matter most for population games are the failures of n-tightness of the induced distribution on characteristics. In these cases, it is necessary to add agents and utility functions to analyze the equilibria of the game. Filling in the missing pieces has perhaps been an uncomfortably elaborate process. After all, one could argue that the original formulation of the game contained all that the modeler thought relevant, and if this means that there are no equilibria or there are no representations of the limits of ϵ-equilibria, then so be it. There are two counter arguments. First, the use of purely finitely additive probabilities has, here, as in many other instances, mislaid major parts of the model, and analysis of a model with pieces mislaid can be very misleading. Second, and this points to possibilities for future research, finitely additive probabilities are the traces of countably additive probabilities on larger spaces, and these larger spaces provide a new class of models. These new models may prove to be a more apt tool for the study of population wide optimization.

References

R. M. Anderson. Star-finite representations of measure spaces. *Trans. Amer. Math. Soc.*, 271(2):667–687, 1982.

R. M. Anderson. The core in perfectly competitive economies. *Handbook of game theory with economic applications*, 1:413–457, 1992.

T. E. Armstrong. Arrow's theorem with restricted coalition algebras. *Journal of Mathematical Economics*, 7(1):55–75, 1980.

P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York-London-Sydney, 1968.

J. P. R. Christensen. Borel structures and a topological zero-one law. *Math. Scand.*, 29:245–255 (1972), 1971.

D. Corbae, M. B. Stinchcombe, and J. Zeman. *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton University Press, Princeton, NJ, 2009.

L. E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *Annals of Probability*, 3:89–99, 1975.

R. M. Dudley. *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

P. C. Fishburn. Arrow's impossibility theorem: Concise proof and infinite voters. *Journal of Economic Theory*, 2(1):103–106, 1970.

R. Fortet and E. Mourier. Convergence de la répartition empirique vers la répartition théorique. *Ann. Sci. Ecole Norm. Sup. (3)*, 70:267–285, 1953.

C. J. Harris, M. B. Stinchcombe, and W. R. Zame. Nearly compact and continuous normal form games: characterizations and equilibrium existence. *Games and Economic Behavior*, 50(2):208–224, 2005.

A. E. Hurd and P. A. Loeb. *An Introduction to Nonstandard Real Analysis*, volume 118 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1985.

H. J. Keisler. The ultraproduct construction. In *Ultrafilters Across Mathematics*, volume 530 of *Contemp. Math.*, pages 163–179. Amer. Math. Soc., Providence, RI, 2010.

M. A. Khan, L. Qiao, K. P. Rath, and Y. Sun. The necessity of countable additivity in modeling infinitely many agents. *Working paper, Department of Economics, Johns Hopkins University*, 2016.

U. Khan and M. B. Stinchcombe. Planning for the long run: Programming with patient, Pareto responsive preferences. *Journal of Economic Theory*, 176:444 – 478, 2018.

J. F. C. Kingman. Additive set functions and the theory of probability. *Proc. Cambridge Philos. Soc.*, 63:767–775, 1967.

A. P. Kirman and D. Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972.

T. Lindstrøm. An invitation to nonstandard analysis. In *Nonstandard Analysis and its Applications*, volume 10 of *London Math. Soc. Stud. Texts*, pages 1–105. Cambridge Univ. Press, Cambridge, 1988.

P. A. Loeb. A nonstandard representation of measurable spaces and L_∞. *Bulletin of the American Mathematical Society*, 77:540–544, 1971.

A. Mas-Colell. On a theorem of Schmeidler. *J. Math. Econom.*, 13(3):201–206, 1984.

J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

A. Robinson. On generalized limits and linear functionals. *Pacific J. Math.*, 14: 269–283, 1964.

L. J. Savage. *The Foundations of Statistics*. Dover Publications Inc., New York, revised edition, 1972.

M. B. Stinchcombe. Countably additive subjective probabilities. *Review of Economic Studies*, 64(1):125–146, 1997.

M. B. Stinchcombe. The gap between probability and prevalence: Loneliness in vector spaces. *Proceedings of the American Mathematical Society*, 129(2):451–457, 2001.

M. B. Stinchcombe. Nash equilibrium and generalized integration for infinite normal form games. *Games and Economic Behavior*, 50(2):332–365, 2005.
M. B. Stinchcombe. Objective and subjective foundations for multiple priors. *Journal of Economic Theory*, 165:263–291, 2016.
J. J. Uhl. Review of Rao and Rao (1983). *Bulletin of the London Mathematical Society*, 16(4):431–432, 1984.
K. Yosida and E. Hewitt. Finitely additive measures. *Transactions of the American Mathematical Society*, 72(1):46–66, 1952.

Appendix

The first two parts of the appendix, A.1 and A.2, contain a synopsis of the mathematics of finitely additive probabilities relevant to this paper, as well as a synopsis of the properties of nonstandard extensions. The second two parts, B.1 and B.2, contain the proofs, first of the tightness results, and then the results on the compactifications and equilibrium existence and approximability.

A. Synopses

A1. **Synopsis: Finitely Additive Probabilities.** Fix a non-empty set X and a σ-field of subsets, $\mathcal{X}$. $M_b(X)$ denotes the Banach space of bounded $\mathcal{X}$-measurable functions with the supnorm.

Basics

The finitely additive probabilities on $(X, \mathcal{X})$, denoted $\Delta^{fa}(X)$, are functions $\mu : \mathcal{X} \to [0, 1]$ satisfying $\mu(X) = 1$ and $\mu(E \cup F) = \mu(E) + \mu(F)$ for disjoint $E, F \in \mathcal{X}$. Δ^{fa} is a subset of the dual space of the Banach space $M_b(X)$, and is given the weak* topology, defined by $\mu_\alpha \to \mu$ if and only if $\int g\, d\mu_\alpha \to \int g\, d\mu$ for every $g \in M_b(X)$.

A probability μ is **countably additive** if $\mu(\cup_n E_n) = \sum_n \mu(E_n)$ for each disjoint collection $\{E_n : n \in \mathbb{N}\} \subset \mathcal{X}$, equivalently, if $\mu(E_n) \downarrow 0$ for every nested collection $\{E_n : n \in \mathbb{N}\} \subset \mathcal{X}$ with $E_n \downarrow \emptyset$. The probability is **purely finitely additive** if there exists a sequence of measurable sets, $E_n \downarrow \emptyset$, with $\mu(E_n) \equiv 1$.

Every finitely additive probability μ has a unique decomposition as $\mu = \delta\mu^{ca} + (1 - \delta)\mu^{fa}$ where $0 \le \delta \le 1$ is unique, μ^{ca} is countably additive, and μ^{fa} is purely finitely additive (Yosida and Hewitt, 1952). Further, δ is the supremum of $\lim_n \mu(E_n)$ where the supremum is taken over all $E_n \downarrow \emptyset$.

A Characterization

The following result, though simple, is often quite useful for understanding the peculiarities of purely finitely additive probabilities.

Lemma 6. *The following conditions are equivalent for a probability μ.*

(a) *μ is purely finitely additive.*
(b) *There is a countable partition $\{F_n : n \in \mathbb{N}\} \subset \mathcal{X}$ with $\mu(F_n) \equiv 0$.*
(c) *There exists a strictly positive $g \in M_b(X)$ with $\int g\, d\mu = 0$.*

Proof. (a) $\Rightarrow$ (b). Suppose that $E_n \downarrow \emptyset$ and $\mu(E_n) \equiv 1$. Set $E_0 = X$ and define $F_n = E_{n-1} \setminus E_n$. The F_n are disjoint, and because $P(E_n) = 1$, $P(F_n) = 0$.

(b) $\Rightarrow$ (c). If $\{F_n : n \in \mathbb{N}\}$ is a countable partion with $\mu(F_n) \equiv 0$, define $g = \sum_n r_n 1_{F_n}$ for any strictly decreasing sequence $r_n \downarrow 0$. For any $\epsilon > 0$, pick $r_n < \epsilon$ and note that $\mu(\{g \leq r_n\}) = 1$, so that $0 \leq \int g \, d\mu < \epsilon$.

(c) $\Rightarrow$ (a). If g is strictly positive and $\int g \, d\mu = 0$, define $E_n = \{g \leq 1/n\}$. Strict positivity implies that $E_n \downarrow \emptyset$. If $\mu(E_n) = \gamma < 1$, then $\int g \, d\mu \geq (1-\gamma)/n$, which is strictly positive. □

Conglomerability

Finitely additive probabilities can fail to be **conglomerable**. From Dubins (1975), if μ is purely finitely additive and non-atomic, then for every measurable B with $0 < \mu(B) < 1$, there is a countable partition A_n of positive mass sets such that $\mu(B|A_n) < \mu(B)$ and another countable partition C_n such that $\mu(B|C_n) > \mu(B)$. If one believe that countable partitions captured all of the mass of purely finitely additive probabilities, one would conclude that models with such priors, e.g. Savage (1972), must suffer from money pumps if the model includes random variables that take on more than finitely many values.

Probabilities on Topological Spaces

Countably additive probabilities on topological spaces have special properties. The same is true for finitely additive probabilities. For the following, suppose that X is a topological space and $\mathcal{X}$ is its Borel σ-field.

- If X is a compact Hausdorff space, then for any finitely additive Borel probability μ, there exists a countably additive μ' such that $\int f \, d\mu = \int f \, d\mu'$ for every continuous $f : X \to \mathbb{R}$. That is, for continuous purposes, there is no difference between finitely additive and countably additive probabilities on compact spaces.
- If X is a csm, (complete separable metric space), then Ulam's theorem (Billingsley, 1968, Thm. 1.4, p. 10) tells us that every countably additive Borel probability μ is tight. Lemma 1 shows that c-tightness is equivalent to countable additivitity, while Lemma 2 shows that, for continuous purposes, there is no difference n-tight probabilities and countably additive probabilities.
- If X is a csm, then for any *purely* finitely additive Borel probability μ, any countably additive μ', and any $\epsilon > 0$, there is a measurable bounded function $g : X \to [0,1]$ such that $|\int g \, d\mu - \int g \, d\mu'| > 1 - \epsilon$. That is, for measurable purposes, there is a world of difference between finitely additive and countably additive probabilities.
- If X is a non-compact csm, then there exists a purely finitely additive μ such that for all $\epsilon > 0$ and all countably additive μ', there exists a Lipschitz continuous $f : X \to [0,1]$ with $|\int f \, d\mu - \int f \, d\mu'| > 1 - \epsilon$. Failures of n-tightness mean that even for continuous purposes, there can be a world of difference between finitely additive and countably additive probabilities.

A2. **Synopsis: Nonstandard Extensions.** This is a quick development of the essential ideas for nonstandard analysis. Lindstrøm (1988) is a more leisurely and detailed introduction to nonstandard analysis that builds directly on intuitions from sequences, while Hurd and Loeb (1985) gives a more axiomatic approach and covers, using the nonstandard analysis it develops, point-set topology and a good deal of functional analysis.

The Basic Construction

For an arbitrary non-empty set X, let $X^{\mathbb{N}}$ denote the set of sequences taking values in X, and fix a purely finitely additive point mass probability, η, on $\mathbb{N}$, that is, suppose that the probability η satisfies $\eta(E) = 0$ or $\eta(E) = 1$ for each $E \subset \mathbb{N}$.[7]

Define an equivalence relation in $X^{\mathbb{N}}$ by $(x_n) \sim (y_n)$ if $\eta(\{n \in \mathbb{N} : x_n = y_n\}) = 1$, let $\langle x_n \rangle$ denote the equivalence class of $(x_n) \in X^{\mathbb{N}}$. Define ${}^*X$ as $X^{\mathbb{N}}/\sim$, that is, as the set of all $\sim$-equivalence classes in $X^{\mathbb{N}}$. X is regarded as embedded in ${}^*X$ by the mapping $x \mapsto \langle x, x, x, x, \ldots \rangle$. If X is infinite, then any sequence (x_n) of distinct elements is not equivalent to any constant sequence. Thus, ${}^*X$ contains new elements if and only if X is infinite.

Examples

This construction can be used for arbitrary X, and examples help make clear the range of uses.[8]

(1) If $X = \mathbb{R}$, then ${}^*\mathbb{R}$ contains the equivalence classes of sequences $\epsilon_n \downarrow 0$. For any positive $r \in \mathbb{R}$, $\eta(\{n \in \mathbb{N} : 0 < \epsilon_n < r\}) = 1$. This means that $\epsilon := \langle \epsilon_n \rangle$ satisfies $0^* < \epsilon^* < r$ for any positive $r \in \mathbb{R}$. This is written $\epsilon \simeq 0$, and $1/\epsilon \simeq \infty$.

(2) Let $gr(f)$ denote the graph of a function from X to Y. The function ${}^*f : {}^*X \to {}^*Y$ is defined, for every $x := \langle x_n \rangle \in {}^*X$, by ${}^*f(x) = \langle f(x_n) \rangle$ in ${}^*Y$. Thus, if X and Y are metric spaces, then f is continuous at $x \in X$ if $[{}^*d_X(x', x) \simeq 0] \Rightarrow [d_Y({}^*f(x'), {}^*f(x)) \simeq 0]$, that is, infinitesimal movements in the domain yields infinitesimal movements in the range.

(3) A function $f : \mathbb{R} \to \mathbb{R}$ has derivative r at x° if for all non-zero $\epsilon \simeq 0$, $|\frac{f(x^\circ + \epsilon) - f(x^\circ)}{\epsilon} - r| \simeq 0$. Directional and multidimensional derivatives are essentially the same.

(4) If X is a metric space, then $x' \in {}^*X$ is **nearstandard** if $d(x', x) \simeq 0$ (where the notation "$d(x', x)$" is short for "${}^*d(x', x)$"). The **standard part** of a nearstandard x' is denoted $\text{st}(x')$, and when $X = \mathbb{R}$, it is also denote ${}^\circ x'$.

(5) If X is a compact metric space, then every $x' \in {}^*X$ is nearstandard: partition X into finitely many sets E_i having diameter less than $1/2^n$; $\eta(\{n \in \mathbb{N} : x_n \in E_i\}) = 1$ for only one of the $E_{n,i}$, denote it by $E_{n,i(n)}$; partition $E_{n,i(n)}$ into finitely many sets $E_{n+1,i}$ having diameter less than $1/2^{n+1}$; at most one of these sets, $E_{n+1,i(n+1)}$ has an η-mass 1 set of n with x_n in it; continuing inductively, the x_n belong to a sequence of sets having diameter less than $1/2^n$; in X, there is exactly one element in $\cap_n \text{closure}(E_n)$, and for every $\epsilon > 0$, $\mu(\{n \in \mathbb{N} : d(x_n, x) < \epsilon\}) = 1$, that is, $d(x, x') \simeq 0$.

(6) Let $\mathcal{P}_f([0,1])$ denote the class of finite subsets of $[0,1]$. The **hyperfinite subsets of** $[0,1]$ are ${}^*\mathcal{P}_f([0,1])$. Consider the set in ${}^*\mathcal{P}_f([0,1])$ given by $A := \langle A_n \rangle$ where $A_n = \{k/2^n : 0 \leq k \leq 2^n\}$. For any dyadic rational q, $\{n \in \mathbb{N} : q \in A_n\}$ has finite complement, hence $q \in \langle A_n \rangle$. The transfer principle (discussed below) tells us that A "behaves like" a finite set, so

[7] If $\mathcal{F}$ is a maximal free ultrafilter on $\mathbb{N}$, then setting $\eta(E) = 1_{\mathcal{F}}(E)$ delivers such a point mass. The Axiom of Choice, in its Hausdorff maximality principle form, implies that maximal free ultrafilters exists.

[8] For detailed introductions to this material see Hurd and Loeb (1985) or Lindstrøm (1988), for the bare essentials of the material used here, see Corbae et al. (2009, Ch. 11).

this is a set that is dense in $[0, 1]$ and yet allows for e.g. inductively moving from one point to the next.

(7) Fix a non-empty X, a σ-field of subsets $\mathcal{X}$, and a finitely additive probability $\mu : \mathcal{X} \to [0, 1]$. It is easy to verify that the function ${}^{\circ*}\mu : {}^*\mathcal{X} \to [0, 1]$ is a finitely additive probability on ${}^*\mathcal{X}$. It can be shown, e.g. Corbae et al. (2009, Thm. 11.2.4), that $\cap_{n\in\mathbb{N}} E_n = \emptyset$ in ${}^*\mathcal{X}$ if and only if there exists a finite N such that $\cap_{n\leq N} E_n = \emptyset$. This means that ${}^{\circ*}\mu$ is automatically countably additive on the field ${}^*\mathcal{X}$, hence by Carathéodory's extension theorem e.g. Corbae et al. (2009, Thm. 7.6.2), ${}^{\circ*}\mu$ has a unique, countably additive extension to $\sigma({}^*\mathcal{X})$, called the Loeb measure generated by μ.[9]

(8) When X is a topological space and τ_X is its topology, then the **monad** of $y \in X$ is the set $m(y) = \bigcap_{y\in G;G\in\tau_X} {}^*G$. When (X, τ_X) is a metric space, $m(y) = \{x \in {}^*X : d(x, y) \simeq 0\}$. In particular, $X = \Delta^{fa} = \Delta^{fa}(\mathcal{X})$ is compact with the weak* topology, so every $\mu' \in {}^*\Delta^{fa}$ is nearstandard (the argument for compact metric spaces generalizes), and the monad of a $\mu \in \Delta^{fa}$ is the set of $\mu' \in {}^*\Delta^{fa}$ with $|\int g\, d\mu - \int g\, \mu'| \simeq 0$ for all bounded measurable $g : X \to \mathbb{R}$.

Superstructures, Transfer, and Saturation

The unification of all of the sets ${}^*X$ happens in a **superstructure**. Let $\mathcal{P}(X)$ denote the class of all subsets of X. A superstructure on a set S is defined inductively: $V_0(S) := S$, $V_1(S) := \mathcal{P}(V_0(S)) \cup V_0(S)$, $V_{n+1} := \mathcal{P}(V_n(S)) \cup V_n(S)$, and $V(S) := \cup_n V_n(S)$. For example: if S contains a non-empty set X and $\mathbb{R}$, then $V_2(S)$ contains ordered pairs; $V_3(S)$ contains sets of ordered pairs, i.e. functions; and $V_4(S)$ contains subsets of the functions, e.g. the continuous or the measurable functions. For any object X in the superstructure $V(S)$, e.g. the set of continuous functions from $[a, b]$ to $\mathbb{R}$, one constructs ${}^*X$ as above, and $V({}^*S)$ is defined as the collection of all of these sets.

One of the most powerful tools for nonstandard analysis is the **transfer principle**, which says that logical statements that are true about objects in the superstructure $V(S)$ if and only if they are true in $V({}^*S)$ after "putting stars on everything." The simplest version of this reduces if-then statements to subset relations. For example, to express the idea that all compact and continuous normal form games have Nash equilibria, letting A denote the set of compact and continuous normal form games, and B the set of normal form games with Nash equilibria, the statement is $A \subset B$. It is clear from the "*" construction that $A \subset B$ if and only if ${}^*A \subset {}^*B$, and this is the canonical example of the transfer principle. To prove that all compact and continuous normal form games have equilibria, one transfers the statement that all finite games have equilibria to find that all *-finite games have equilibria, then show that the standard part of the mixed strategies are equilibria of the original game. This part of the arguments is where, typically, the hard work, if there is any, must be done.

The construction using a purely finitely additive η on the integers has the following **saturation** property: if $A \in V(S)$ is countable, i.e. has cardinality $\aleph_0$, then there exists an $A^f \in {}^*\mathcal{P}_f(A)$ such that for every $a \in A$, $a \in A^f$. Fix a cardinal number κ. It is possible to (carefully) construct point mass η's on larger index sets such

[9] In honor of its discoverer and initial developer, see Loeb (1971).

that for any $A \in V(S)$ with cardinality κ or smaller, there exists an $A^f \in {}^*\mathcal{P}_f(A)$ such that for every $a \in A$, $a \in A^f$. Such a $V({}^*S)$ is called **κ-saturated**.

B. Proofs

B1 contains the proofs of the characterization results for and examples about tightness, the equilibrium existence and continuity proofs are in B2.

B1. **Characterizations of Tightness and Related Material.** We now turn to the arguments for the three kinds of tightness introduced above.

Proof of Lemma 1. Suppose first that q is countably additive. Let $\mathcal{E} \subset \mathcal{K}$ denote the class of sets with $q(E) = \sup\{q(K) : K \subset E \text{ is compact}\}$. $\mathcal{E}$ contains the closed sets, is closed under complementation and countable unions, hence contains all of the measurable sets.

Suppose now that q is c-tight. It is sufficient to show that for all $E_n \downarrow \emptyset$ in $\mathcal{M}$, $q(E_n) \downarrow 0$. Suppose not, that is, suppose that for some $E_n \downarrow \emptyset$ in $\mathcal{M}$, $q(E_n) \downarrow \delta > 0$. Let $r_n \downarrow 0$ be a strictly decreasing sequence such that $\sum_n r_n q(E_n) < \delta/2$ (e.g. $r_n = \delta/2^{n+1}$). For each E_n, let K_n be a compact subset of E_n with $q(K_n) > (1-r_n)q(E_n)$. For any positive N, $q(E_N \setminus \cap_{n=1}^N K_n) \le \sum_{n=1}^N r_n q(E_n) < \delta/2$. Since $q(E_n) \ge \delta$, this implies that $q(\cap_{n=1}^N K_n) > \delta/2$ which is strict positive. Therefore, the collection $\{K_n : n \in \mathbb{N}\}$ has the finite intersection property. By compactness, $\cap_n K_n \neq \emptyset$, contradicting $\cap_n E_n = \emptyset$. □

Proof of Lemma 2. Note first that q is n-tight iff for every $\epsilon > 0$, there exists a compact K such that for all $\delta > 0$ and every continuous f satisfying $1_{K^\delta} \ge f \ge 1_K$, $\int f\,dq > (1-\epsilon)$. The proof uses this as an alternative definition of n-tightness.

If $\beta_M(q, q^{ca}) = 0$ for some countably additive q^{ca}, then for every bounded, Lipschitz continuous $f : M \to \mathbb{R}$, $\int f\,dq = \int f\,dq^{ca}$. Pick an $\epsilon > 0$ and pick K such that $q^{ca}(K) > (1-\epsilon)$. For any $\delta > 0$, consider the Lipschitz function $f(x) = \max\{0, 1 - \frac{1}{\delta}d(x,K)\}$. The function f satisfies both $1_{K^\delta} \ge f \ge 1_K$ and $\int f\,dq = \int f\,dq^{ca} \ge q^{ca}(K) > (1-\epsilon)$.

Fix a n-tight q, and for every bounded continuous $f : M \to \mathbb{R}$, define $L_q(f) = \int f\,dq$. L_q is a continuous linear function on $C_b(M)$, the bounded continuous function on M. In order that L_q be representable as an integral against a countably additive q^{ca}, it is necessary and sufficient that for every sequence $f_n \in C_b(M)$ satisfying $f_n(x) \ge 0$ and $f_n(x) \downarrow 0$ for all x, $L_q(f_n) \downarrow 0$ (this equivalence follows from Lebesgue's dominated convergence theorem and the Riesz representation theorem, e.g. Corbae et al. (2009, Theorems 7.5.6 and 9.8.2)).

Fix a non-negative sequence f_n in $C_b(M)$ with $f_n \downarrow 0$. Pick $\epsilon > 0$. We must show that for sufficiently large N, $L_q(f_N) < \epsilon$. Since f_1 is bounded, there is no loss in assuming that $1 \ge f_n(x) \ge 0$ for each x and each n. Pick a compact K such that for all $\delta > 0$ and every continuous g satisfying $1_{K^\delta} \ge g \ge 1_K$, $L_q(g) > (1-\epsilon/3)$. Because K is compact, we can pick N sufficiently large that f_N is uniformly less than $\epsilon/3$ on K (by Dini's theorem). Because f_N is continuous, for each $x \in K$, there exist $\delta_x > 0$ such that for all $y \in M$, if $d(x,y) < \delta_x$, we have $|f_N(x) - f_N(y)| < \epsilon/3]$. Because K is compact, there is a finite cover, $\{B(x_i, \delta_i) : i = 1, \ldots, I\}$, $x_i \in K$, of K by such open balls. Pick $\delta' > 0$ such that $K^{\delta'} \subset \cup_i B(x_i, \delta_i)$. By n-tightness, $q(M \setminus K^{\delta'}) < \epsilon/3$. Because $K^{\delta'} \subset \cup_i B(x_i, \delta)$, $f_N < 2\epsilon/3$ on $K^{\delta'}$. Combining, $L_q(f_N) < 2\epsilon/3 + \epsilon/3 = \epsilon$. □

Proof of Lemma 3. From Lemma 2, there is a (unique) countably additive q with $\beta(p, q) = 0$. Let ϵ_n be a decreasing of strictly positive numbers $\epsilon_n \downarrow 0$, let K_n be an increasing sequence of compact sets with $q(K_n) > 1 - \epsilon_n$, and define $F_n = K_n^{\epsilon_n}$. The set $F := \cup_N \cap_{n \geq N} {}^*F_n$ is a subset of $\text{ns}(\mathbb{U})$ and ${}^*p(F) \simeq 1$. □

The following shows that there are n-tight probabilities that put mass 1 on every δ-neighborhood of the origin but put mass 0 on any compact set.

Lemma 7. *For every $n \in \mathbb{N}$, there exists a finitely additive p_n with $p_n(\frac{1}{n} \cdot \mathbb{U}) = 1$ and $p_n(K) = 0$ for all compact $K \subset \mathbb{U}$.*

Proof. Let $\mathcal{K}$ denote the compact subsets of $\mathbb{U}$. For every finite $\mathcal{K}_F \subset \mathcal{K}$, let K_F denote the necessarily compact convex hull of the union of the compact subsets in $\mathcal{K}_F$. Because it is compact, the set $(\mathbb{R} \cdot K_F) \cap \mathbb{U}$ is a shy subset of $C([0, 1])$ (see e.g. Stinchcombe (2001) for this). That is, there exists a Borel probability η on $\mathbb{U}$ such that $\eta(f + (\mathbb{R} \cdot K_F) \cap \mathbb{U}) = 0$ for all $f \in C([0, 1])$. In particular, setting $f = 0$, we have $\eta(\mathbb{R} \cdot K_F) = 0$.

Let $\mathbb{P}(\mathcal{K}_F)$ denote the set of all finitely additive p such that $p((\mathbb{R} \cdot K_F) \cap \mathbb{U}) = 0$. This is a weak*-closed subset of $\Delta^{fa}(\mathbb{U})$, hence is compact. The class

$$\{\mathbb{P}(\mathcal{K}_F) : \mathcal{K}_F \text{ is a finite subset of } \mathcal{K}\} \tag{B1}$$

has the finite intersection property, hence has non-empty intersection. Any p in the intersection is a purely finitely additive probabilities satisfying $p(K) = 0$ for all compact $K \subset \mathbb{U}$.

Let p be any element of the intersection and for each $n \in \mathbb{N}$, let p_n be the image of p under the mapping $f \mapsto \frac{1}{n} f$. This guarantees that $p_n(\frac{1}{n} \cdot \mathbb{U}) = 1$. By construction, p_n also satisfies $p_n(K) = 0$ for all K. Finally, by the weak*-compactness of the set of finitely additive probabilities, $\{p_n : n \in \mathbb{N}\}$, has an accumulation point, and since $p_n(K) \equiv 0$ for each p_n, $p(K) = 0$ as well. □

B2. **Existence Results and Related Material.** We now turn to the equilibrium existence results, as well as the supporting material.

Proof of Theorem A. Pick infinite f that refines all finite partitions of T and an ϵ-equilibrium a^f for a positive infinitesimal ϵ (which exists by transfer of the statement that all finite population games have equilibria). By the definition of ϵ-equilibria, there exists an $E \in {}^*\mathcal{T}(f)$ with ${}^*\mu(E) > (1 - \epsilon)$ such that ${}^*\mathcal{G}(t)(a^f(t), \nu_{a^f}) > \max_{b \in {}^*A} {}^*\mathcal{G}(t)(b, \nu_{a^f}) - \epsilon$. The set E belongs to ${}^*\mathcal{T}$, hence belongs to $\sigma({}^*\mathcal{T})$ and ${}^{\circ *}\mu(E) = 1$. Since the sup norm diameter of ${}^*\mathcal{G}(f)(B)$ is infinitesimal for each $B \in f$, ${}^*\mathcal{G}(t)(a^f(t), \nu_{a^f}) \simeq \text{st}_{\text{V}}({}^*\mathcal{G}(t)(a^f(t), \nu_{a^f}))$, hence for each $t \in E$, $\text{st}_{\text{V}}(\mathcal{G}^*(t)(a^f(t), \nu_{a^f})) \geq \max_{b \in {}^*A} \text{st}_{\text{V}}(\mathcal{G}^*(t)(b, \nu_{a^f}))$ and a^f is an equilibrium for Γ^{NS}.

Now suppose that a^* is an equilibrium of Γ^{NS} and define $a^f = E^\mu(a^*|f)$. Let $B \in \sigma(\mathcal{T})$ denote the set

$$\{t \in {}^*T : \text{st}_{\text{V}}({}^*\mathcal{G}(t)(a^*(t), \nu_a)) \geq \max_{b \in {}^*A} \text{st}_{\text{V}}(\mathcal{G}^*(t)(b, \nu_a)). \tag{B2}$$

By the definition of equilibria, ${}^{\circ *}\mu(B) = 1$. Since the sup norm diameter of ${}^*\mathcal{G}(B)$ is infinitesimal for each $B \in f$,

$$ {}^{\circ *}\mu(\{t \in {}^*T : \text{st}_{\text{V}}({}^*\mathcal{G}(t)(a^f(t), \nu_{a^f})) < \max_{b \in {}^*A} \text{st}_{\text{V}}({}^*\mathcal{G}(t)(b, \nu_{a_f}))\}) = 0. \tag{B3}$$

By iterated expectation, $\nu_{a^f} = \nu_{a^*}$. Take any internal set E^f in ${}^*\mathcal{T}(f)$ such that ${}^{\circ *}\mu(E^f \Delta B^c) = 0$. Note that ${}^*\mu(E^f) \simeq 0$. Playing the strategy a^f on the complement of E^f delivers an ϵ-equilibrium for some $\epsilon \simeq 0$.

The continuity of the approximate equilibrium correspondence is the usual conversion of nonstandard results into statements about ultrafilter limits. For the last two parts, one works with the distributional formulation of games as developed in Mas-Colell (1984): any equilibrium outcome for Γ^{NS} induces a joint distribution over actions and types; if p is c-tight then the regular conditional probability for the joint distribution of an equilibrium; if p is n-tight, apply Lemma 2. □

Proof of Lemma 4. For each $f \in \mathcal{F}$, let $I_f = \{{}^\circ f(x) : x \in X\}$. Because $f(X)$ is an internal subset of ${}^*[-1,+1]$, I_f is compact. Give the compact product space, $I = \times_{f\in\mathcal{F}} I_f$, the product topology. For $x, y \in X$, define $x \sim y$ if ${}^\circ f(x) = {}^\circ f(y)$ for all $f \in \mathcal{F}$ and let $X' = X/\sim$ be the set of equivalence classes. The mapping $\varphi : X' \to I$ defined by $\varphi(x) = (f(x))_{f\in\mathcal{F}}$ from X' to I is, by definition, a homeomorphism. Since I is a compact Hausdorff space, X' is also compact and Hausdorff. Since X' is a quotient space of X, X is compact, and it is Hausdorff if and only if for each $x \neq y$ we have $x \not\sim_{\mathcal{F}} y$. □

Proof of Lemma 5. Every Lipschitz function is continuous. Therefore, for every open $G \subset \mathbb{R}$, every $\psi^{-1}(G)$ is open in the sup norm topology. In the other direction, for every $g \in \mathbb{U}$, the function $\psi_g(f) := \|f - g\|$ is Lipschitz and $\psi_g^{-1}(-\infty, \epsilon)$ is the ϵ-ball around g. □

Proof of Proposition 1 We have $x \neq y$ in M if and only if there exists an $f \in U(M)$ such that $f(x) = 0$ and $f(y) = 1$. By transfer, $x \neq y$ in ${}^*M$ if and only if there exists an $f \in {}^*U(M)$ such that $f(x) = 0$ and $f(y) = 1$. By Lemma 4, $\widehat{M} = {}^*M$ is compact and Hausdorff. To show that $\widehat{\mathbb{U}} = {}^*U(M)/\simeq$, note that by Lemma 5, $f \sim g$ if and only if ${}^*\psi(f) \simeq {}^*\psi(f)$ for all $\psi \in \Psi$. Finally, note that $\cup_{r>0}\, r \cdot \widehat{\mathbb{U}}$ is a sup norm closed vector algebra of continuous functions containing the constants and separating points in a compact Hausdorff space. By the Stone-Weierstrass theorem, it is therefore equal to the set of continuous functions. □

Proof of Theorem B. $\widehat{\Gamma}$ has an equilibrium because the standard part of any equilibrium for Γ^{NS} yields one. If q is countably additive, apply Mas-Colell (1984) to find an equilibrium for the distributional form of the game, then take a regular conditional probability to represent it. If q is tight, hence puts mass 1 on $\mathbb{U} \subset \widehat{\mathbb{U}}$, then the difference between Γ and $\widehat{\Gamma}$ is infinitesimal. The last part follows from Lemma 2, which shows that n-tight probabilities are at weak*-distance 0 from countably additive ones. □

Department of Economics, University of Texas at Austin