Theoretical Economics 16 (2021), 1055–1093 1555-7561/20211055

Strict pure strategy Nash equilibria in large finite-player games

Guilherme Carmona
School of Economics, University of Surrey

Konrad Podczeck
Institut für Volkswirtschaftslehre, Universität Wien

In the context of anonymous games (i.e., games where the payoff of a player is, apart from his/her own action, determined by the distribution of the actions made by the other players), we present a model in which, generically (in a precise sense), finite-player games have strict pure strategy Nash equilibria if the number of agents is large. A key feature of our model is that payoff functions have differentiability properties. A consequence of our existence result is that, in our model, equilibrium distributions of non-atomic games are asymptotically implementable by pure strategy Nash equilibria of large finite-player games.

Keywords. Large games, pure strategy, Nash equilibrium, generic property.

JEL classification. C72.

1. Introduction

It is now well established in economics to address problems of strategic interaction among many negligible individuals by models of anonymous games. In such games, the impact on the payoff of a player by the actions chosen by the other players factors through the distribution of these actions.[1] A particular and convenient class of models of anonymous games is formed by games with a "continuum of players" or, more precisely, "non-atomic games" (henceforth, *continuum games* for short). In continuum games, there is no longer a distinction between the distribution of the actions chosen by all players and the distribution of the actions chosen by all but one player, so that, concerning equilibrium existence, these games are rather easy to analyze. In fact, there are several results on existence of pure strategy Nash equilibrium for such games, the pioneering ones provided by Schmeidler (1973) and Mas-Colell (1984). In these results, no linear structure is imposed on players' action sets. After all, of course, continuum games are idealizations of situations with a large but finite number of players, and in this respect the following questions naturally arise:

Guilherme Carmona: g.carmona@surrey.ac.uk
Konrad Podczeck: konrad.podczeck@univie.ac.at

Thanks to Daniel García, Michael Greinecker, and Klaus Ritzberger for helpful comments and suggestions.

[1] We follow Kalai (2004) in defining the notion of anonymous game this way. Actually, Kalai (2004) speaks of semi-anonymous games to indicate that in the incomplete information setting he considers, a player's identity is encoded in his type. Since we do not consider incomplete information, the prefix "semi" has been dropped. We note that in Khan and Sun (1999) the notion of anonymous game is reserved for games with a continuum of players specified solely by distributions of players' characteristics.

Available at https://econtheory.org. https://doi.org/10.3982/TE3967

Q1. To what extent do pure strategy Nash equilibrium existence results for continuum games carry over to pure strategy Nash equilibrium existence results for games with a large but finite number of players?

Q2. Are equilibria of continuum games artifacts of having a continuum of players or are they realizable as "limits" of pure strategy Nash equilibria of large finite-player games? That is, given a continuum game and an equilibrium of this game, are there "close" large finite-player games having "similar" pure strategy Nash equilibria?

In other words, are continuum games and equilibria of such games good idealizations of situations with a large but finite number of players?

The "to what extent" clause is incorporated in the first question because it is well known that finite-player games may fail to have pure strategy Nash equilibria (unless action sets are convex and quasiconcavity assumptions on payoff functions are made). In fact, this may be the case for a fixed distribution of payoff functions, regardless of the number of players in such games (see the example after the statement of Theorem 1 in Section 3.5). Thus, the best one can hope for in regard to Q1 is to get positive results in terms of genericity analysis. As made clear by the literature on competitive equilibrium in exchange economies, a suitable and convenient setting for genericity analysis is a setting where agents' characteristics have differentiability properties. In this paper, we develop such a setting for anonymous games, so that there is a generic set of continuum games such that finite-player games forming a sequence with an increasing number of players and a "limit" in this set have pure strategy Nash equilibria if the number of players is large enough.

Of course, a differentiability assumption on payoff functions necessarily implies that the domains of these functions are subsets of a linear space (for which we take a Euclidean space). In particular, we have a linear structure on the action sets of players. This is a restriction compared with the analysis in Mas-Colell (1984), where the action sets of players (which in Mas-Colell (1984) are the same for all players in a game) can be any compact metric space.[2]

Based on our existence result, we deal with Q2. We show that, in our model, any equilibrium of any continuum game is asymptotically implementable in the sense that there exists a sequence of finite-player games, with an increasing number of players, and a corresponding sequence of strict pure strategy Nash equilibria such that the given continuum game and its equilibrium arise as limit (in an appropriate sense).[3] This bears

[2]In fact, action sets in our model have nonempty interior in some Euclidean space. We remark in this regard that pure strategy Nash equilibria with actions belonging to the interior of action sets do not have a mixed strategy interpretation unless payoff function are linear in the own actions of players. Models of games with action sets having nonempty interior in Euclidean spaces arise in several applications. We mention models of Cournot competition where firms can vary outputs continuously and models of auctions where bids can vary in products of intervals.

[3]This notion of asymptotic implementation is more specific than that in Khan and Sun (1996), where asymptotic implementation signifies, in the sense of nonstandard analysis, a "transfer" of any results for games with a hyperfinite Loeb spaces of players to games with a finite number of players (see Loeb and Wolff 2015, p. 355).

some surprise: It is well known that an equilibrium correspondence need not be lower hemicontinuous, which suggests that the set of limits of strict equilibria of *all* sequences of finite-player games "converging" to a given continuum game could still form a proper subset of the equilibrium set of the limit continuum game. As our result shows, this intuition is incorrect.[4]

The above result is important for applications that assume a continuum of players since, by considering the set of all equilibria of a game with a continuum of players, one can rest assured that none of these equilibria is a spurious implication of the (tractable but unrealistic) assumption of an infinite population. The intuition for this result is that any equilibrium of a continuum game can be made strict in a nearby continuum game and that any strict equilibrium of a continuum game is robust, i.e., it yields a strict equilibrium for all finite-player games close to it.

As mentioned above, an important aspect of our analysis is that the equilibria that we obtain for large finite-player games are strict, i.e., best reply sets in an equilibrium are singletons; thus, there is no issue of specifying certain actions in the best reply sets as equilibrium actions. This makes our results immune against an objection as formulated by Mas-Colell (1977), which addresses general equilibrium theory, but applies to game theoretic contexts as well. To quote Mas-Colell (1977), "[i]mportant as those results are, the notion of equilibrium they deal with has some unattractive features. In particular, knowledge by the consumers of the equilibrium price system (plus the preference maximization hypothesis) does not determine the equilibrium; one needs, in addition, a possibly very careful specification of each consumer's commodity bundles. This makes the equilibrium a 'decentralized' one only in some weak sense."

To note some details of our model, action sets are compact subsets of some Euclidean space, with dense interior, and for any player, the externality, i.e., the channel through which his/her payoff is affected by the actions of the other players, is given, as in Araújo et al. (2000), Balder (2002), Rauh (2003), or Yu and Zhu (2005), by finitely many summary statistics (e.g., the first noncentral moments) of the distribution of these actions. From the viewpoint of applications, this is not a big restriction; in fact, in many applications of anonymous games, e.g., Cournot oligopoly games, it is just the mean action of the other players that determines a player's payoff in addition to his/her own action. Payoff functions in our model are twice continuously differentiable. The main costs of our results, compared with standard game theory, are an assumption that implies that the best replies of a player against the distributions of the actions of the other players are always in the interior of his/her action set. This assumption is needed to be in a position in which differentiability assumptions on payoff functions can be conveniently exploited. On the set of players' characteristics (i.e., payoff functions), we define a suitable topology; because the actual definition requires some technical constructions, we refer to Section 3.4 and here say only that this topology is defined in terms of graphs of payoff functions to accommodate the fact that action sets may differ across players. A continuum game is specified as a Borel probability measure with compact support on

[4]This does not imply lower hemicontinuity, which means in the above context that the set of limits of strict equilibria of *any* sequence of finite-player games converging to a given continuum game equals the equilibrium set of the limit continuum game.

the space of players' characteristics. The compact support condition requires that players' characteristics in a continuum game are not too dissimilar. Equilibria of continuum games are described by equilibrium distributions, as in Mas-Colell (1984).

"Generic" in the set of continuum games is formally expressed as open and dense in the topology that treats two such games as close if they are close in the narrow topology and if their supports are close in the Hausdorff metric topology. "Close" for two continuum games in the former topology means that they involve similar players' characteristics with similar frequencies; the extra requirement of being close in the latter topology means that they are close only if they involve similar players' characteristics. In the notion of a generic set of continuum games, "open" means stability against perturbations, and "dense" means that every continuum game can be approximated by continuum games belonging to the generic set.

We remark that the generic set of continuum games we identify in the proof of our main result (Theorem 1) is defined intrinsically in the sense that no reference to the particular problem of equilibrium existence in large finite-player games is made. Roughly, this set consists of those continuum games ν that have an equilibrium distribution such that the corresponding externality (which is the same for all players in a continuum game) has a neighborhood on which (a) the correspondence that sends externalities to the externalities determined by the best replies of the players with characteristics in the support of ν can be identified with a differentiable function, and (b) at each point, the derivative of this function minus the identity matrix has maximal rank.

The organization of the paper is as follows. In the next section, we mention some of the related literature. In Section 3, the model is set up and the results are stated. Detailed proofs can be found in Section 5, after Section 4 provides an outline of the proof of the main result. In the Appendix, some auxiliary lemmata, which combine some more or less well known mathematical facts, are stated and proved.[5]

2. Related literature

Results related to ours can be found in Rashid (1983), Khan and Sun (1999), Kalai (2004), Carmona and Podczeck (2009), and Carmona and Podczeck (2012), who have established that sufficiently large finite-player games have pure strategy approximate equilibria.[6] Relative to that literature, the contribution of the present paper consists in presenting a setting of games that allows us to drop the "approximate" qualifier generically (see Theorems 1 and 3 below). This is important because approximate equilibria are not always appealing;[7] thus, by dispensing with this qualification, this is something we do

[5]A longer version of this paper (Carmona and Podczeck 2019) contains additional material, including a somewhat broader literature review, an application of our results to Cournot oligopoly, and some discussion of games with essentially ordinally nonequivalent payoff functions.

[6]Here *pure strategy approximate equilibrium* means a strategy profile such that for some numbers $\epsilon > 0$ and $0 \leq \eta < 1$, players who make up a fraction of at least $1 - \eta$ cannot deviate so that payoffs increase more than ϵ (in Kalai (2004) and some of the results of Carmona and Podczeck (2009) the number η is zero), and *sufficiently large* means that these numbers can be chosen arbitrarily small if one takes the number of players large enough.

[7]Consider a finite-player game where each player can choose one number in the interval $[0, 1]$ and each player's payoff is the average choice of the entire population (including his own choice). If this game is

not have to worry about in our results. In fact, we obtain strict equilibria (i.e., players have unique best replies) generically, so in an equilibrium there is no indeterminacy regarding which actions players will choose. For these reasons, our analysis provides an important supplement to the above-mentioned results.

Asymptotic implementation of equilibria of continuum games, as defined in the Introduction, is considered in Housman (1988) and Carmona and Podczeck (2020a). The corresponding results in those papers are stated in terms of approximate equilibria of finite-player games. As above, we can drop the "approximate" qualifier. In fact, in our model, every equilibrium of every continuum game, not just an equilibrium of a generic continuum game, can be asymptotically implemented in terms of strict pure strategy Nash equilibria of finite-player games (see Theorems 2 and 3).

In our companion paper (Carmona and Podczeck (2020b)), we deal with the same questions as in this paper, but without differentiability assumptions on payoff functions. There, for the case of finite action spaces, we obtain results that are analogous to those in the present paper. In contrast, beyond the finite actions case, our results in Carmona and Podczeck (2020b) rely, directly or indirectly, on the use of "highly concentrated" mixed strategies; here, we show that mixed strategies can be completely dispensed with in our setting.

In the present paper, the goal is to allow players' actions to vary continuously and at the same time to obtain strict pure strategy Nash equilibria for large finite-player games. For this, we adopt a setting in which payoff functions have differentiability properties. The main idea of our analysis is that, with such payoff functions, there is a generic set of continuum games that have equilibria for which the following statements are true: (i) at the equilibrium actions, payoff functions are locally strictly concave in the own actions of players, so that best reply sets are singletons locally; (ii) this implies that best reply sets are singletons globally, (iii) this property remains true in the product of some neighborhood of a game belonging to the generic set and some neighborhood of the externality arising in the equilibrium under consideration; (iv) we can use (iii) to get the existence of strict pure strategy Nash equilibrium along sequences of finite-player games with a limit in the generic set of continuum games if the number of players is large enough.[8]

A similar idea underlies the analysis of Mas-Colell (1977, Theorem 2), who considers nonconvex differentiable preferences in a general equilibrium framework of large economies. Concerning (i) above, see also Trockel (1984, p. 10). For an analysis of large economies where agents' preferences are convex, see Dierker (1975).

Apart from this similarity, the arguments in our proofs and in those of Mas-Colell (1977) differ to a large extent. For instance, in Mas-Colell (1977) the existence of equilibria in large finite economies close to generic continuum economies can be established

played by n players, then the pure strategy profile where each player chooses 0 is an ε-equilibrium with $\varepsilon = 1/n$; however, for each player, 1 is a strictly dominant strategy.

[8]We can only be sketchy here. We just note that even though we do not assume action sets to be convex, (i) makes sense because we may assume them to have dense interior, and open sets in a Euclidean space include convex neighborhoods of each of their points.

by just applying an implicit function theorem to a (locally defined) excess demand function depending on prices and distributions of preferences (actually, Mas-Colell (1977) refers to Dierker (1975) at this point). In our setting, we cannot proceed this way, because, in finite-player games, different players may face different summary statistics of the actions of the respective other players, so that the dimension of domain of the problem increases with the number of players; as a consequence, we need a fixed-point result to get pure strategy Nash equilibria. We also remark that in Mas-Colell (1977) there is no counterpart of the treatment of asymptotic implementability of any equilibrium of any continuum game.

As already pointed out in the Introduction, Schmeidler (1973) and Mas-Colell (1984) established the existence of pure strategy Nash equilibrium for games with a continuum of players. We contribute to this literature by establishing the generic existence of pure strategy strict equilibrium for sufficiently large finite-player games.

Schmeidler (1973) is part of a literature where, different from the distributional approach adopted in the present paper to describe continuum games and their equilibria, an explicitly specified atomless measure space of players and a measurable map from this space to some space of players' characteristics are taken as given, and equilibria are described directly by strategy profiles (see also, e.g., Balder (2002) or the more recent treatment by Khan et al. (2017)). For the purpose of our paper, the distinction between the individualistic and the distributional approach is not important for two reasons. First, the focus of our asymptotic generic existence result is on finite-player games; continuum games matter only as a "reference point" and their equilibrium distributions provide all the information needed. Second, our asymptotic implementation result can be translated to the individualistic context using standard arguments: Given an explicit atomless measure space of players and a map from this space to a space of players' characteristics, the joint distribution of this map and an equilibrium strategy profile is an equilibrium distribution to which we can apply our asymptotic implementation result.

It bears some repetition to emphasize that our asymptotic implementation result provides a robust version for the equilibrium existence results of Schmeidler (1973) and Mas-Colell (1984). Regarding the latter result, it provides a sense in which equilibrium distributions of continuum games can be regarded as involving pure strategies, as they are shown to be limits of strict pure strategy equilibria of large finite-player games.

3. The model and the results

3.1 *General notation and terminology*

If X is any metric space, we write ρ_H for the Hausdorff metric on the set of nonempty compact subsets of X.[9] Recall that on the set of nonempty compact subsets of a metric space X, the topology defined from the Hausdorff metric depends only on the topology of X, not on the particular metric.

[9]That is, $\rho_H(A, B) = \max\{\max_{a\in A} d(a, B), \max_{b\in B} d(b, A)\}$ for any two nonempty compact sets $A, B \subseteq X$, writing d for the metric of X.

For a subset A of a topological space X, $\operatorname{int} A$ denotes the interior of A, ∂A denotes the boundary of A, $\operatorname{cl} A$ denotes the closure of A, and $A \setminus B$ denotes set-theoretic subtraction. If $A \subseteq \mathbb{R}^n$, then $\operatorname{co} A$ denotes the convex hull of A.

If μ is a Borel measure on a separable metric space X, we write $\operatorname{supp}(\mu)$ for the support of μ, i.e., the smallest closed subset of X with full measure. (Recall that every Borel measure on a separable metric space has a support.) If μ is a Borel measure on a product $X \times Y$ of metric spaces, μ_X and μ_Y denote the marginal measures on X and Y, respectively.

Euclidean spaces are regarded as being equipped with the Euclidean norm. For any point x in such a space and for any number $r > 0$, we write $B(x, r)$ for the open ball of center x and radius r, and write $\bar{B}(x, r)$ for the closed ball of center x and radius r.

Let $X \subseteq \mathbb{R}^k$ be such that $\operatorname{int} X$ is dense in X (which is true, in particular, if X is convex and $\operatorname{int} X \neq \varnothing$). We say that a function $f \colon X \to \mathbb{R}^\ell$ is continuously differentiable if there is an open $\tilde{X} \subseteq \mathbb{R}^k$ including X such that f can be extended to a function $\tilde{f} \colon \tilde{X} \to \mathbb{R}^\ell$ that is continuously differentiable in the usual sense; the derivatives of f at non-interior points of X are defined to be those of $\tilde{f}$ (note that these derivatives do not depend on the particular choice of the extension $\tilde{f}$ if $\operatorname{int} X$ is dense in X.) In this case, we write $Df(x)$ for the derivative of f at $x \in X$ and write Df for the map $x \mapsto Df(x)$. If Df happens again to be continuously differentiable in the above sense, we say that f is twice continuously differentiable; in this case, $D^2 f(x)$ stands for $DDf(x)$, and $D^2 f$ stands for the map $x \mapsto D^2 f(x)$. If f is defined on a product $X \times Y$, where Y is any set and X is as above, then $D_x f(x, y)$ means the derivative of $f(\cdot, y)$ at $x \in X$ if $f(\cdot, y)$ is continuously differentiable; if $f(\cdot, y)$ is twice continuously differentiable, $D_x^2 f(x, y)$ stands for $D_x D_x f(x, y)$.

3.2 *Actions and externality*

We consider games with a large number of players (as a particular case, with a continuum of players), where the payoff of a player is determined by his/her own action and an externality that is given by finitely many summary statistics of the distribution of the actions of the other players, as in Balder (2002) or Rauh (2003), and payoff functions have differentiability properties. We start setting up the model by fixing an ambient space so as to get suitable sets for the domains of payoff functions.

There is a universe A of possible actions, which is a nonempty and compact subset of $\mathbb{R}^n$, with dense interior. Action sets of games are included in A and also have dense interior.

Summary statistics of a distribution τ_A of actions in A are given by the formulae $\int g_1(a)\, \mathrm{d}\tau_A(a), \dots, \int g_m(a)\, \mathrm{d}\tau_A(a)$, where g_i, $i = 1, \dots, m$, is a continuously differentiable function from A to $\mathbb{R}$ (given independently of τ_A). We write g for the vector $(g_1, \dots, g_m)$ and write $\int g(a)\, \mathrm{d}\tau_A(a)$ for the vector $\big(\int g_1(a)\, \mathrm{d}\tau_A(a), \dots, \int g_m(a)\, \mathrm{d}\tau_A(a)\big)$. A natural example is obtained by setting

$$g(a) = \big(a_{(1)}, a_{(1)}^2 \dots a_{(1)}^k, a_{(2)}, a_{(2)}^2 \dots a_{(2)}^k, \dots, a_{(n)}, \dots, a_{(n)}^k\big)$$

for each $a \in A$, where superscripts are exponents and the subscript (h) means the hth coordinate of a, $h = 1, \dots, n$; in this case, $m = kn$ and $\int g(a)\, \mathrm{d}\tau_A(a)$ is the vector of the first k noncentral moments of the coordinate distributions determined by τ_A; see Rauh (2003). A special case of this example is given if $m = n$ and g is the restriction to A of the identity on $\mathbb{R}^n$, so that $\int g(a)\,\mathrm{d}\tau_A(a)$ is the "mean action" corresponding to the distribution τ_A on A.

For any player and any distribution τ_A on A induced by the actions of the other players in a game, the externality is $e(\tau_A) = \int g(a)\,\mathrm{d}\tau_A(a)$. Set

$$E = \left\{ \int g\,\mathrm{d}\tau_A \colon \tau_A \text{ is a probability measure on } A \right\}.$$

Note that E is a convex and compact subset of $\mathbb{R}^m$, and that E is just the convex hull of the compact set $g(A)$. Every point of E can arise as an externality for a player in some continuum game. Thus, the set E can be seen as the universe of possible externalities.

In finite-player games, the set of externalities an individual player could actually face is of the form $\sum_{j=1}^{l} g(A_j)/l$. To ensure that in games with sufficiently many players such sets have dense interior, i.e., are appropriate for differentiability assumptions on payoff functions, we make the following assumption on the map g: Whenever O is an open set in $\mathbb{R}^n$ with $O \subseteq A$, then $g(O)$ affinely spans $\mathbb{R}^m$ (in other words, the smallest affine subspace in $\mathbb{R}^m$ including $g(O)$ is $\mathbb{R}^m$ itself); see Lemma 9 in the Appendix for the desired conclusion. This assumption simply imposes some kind of homogeneity property on the map g. We remark that the assumption is satisfied if g is such that $\int g\,\mathrm{d}\tau_A$ is, as in the example above, the vector of the first k noncentral moments of the coordinate distributions determined by τ_A; see Lemma 10 in the Appendix.[10]

3.3 *Payoff functions*

A payoff function is a real-valued function u with $\operatorname{dom} u = A_u \times E_u$, where A_u and E_u are subsets of the actions universe A and the externalities universe E, respectively. The set A_u is the action set of a player with payoff function u. (We thus specify actions sets of players by components of the domains of payoff functions; this is for notational efficiency.) The set E_u is referred to as the externalities factor in $\operatorname{dom} u$.

This formalization of payoff functions is convenient because it gives an easy way to set up a space of payoff functions so that in actual games, the domain of the payoff function of any player does not depend on action profiles that this player cannot observe. For example, in a finite-player game, the set of externalities relevant for the payoff of an individual player is of the form $\sum_{j=1}^{l} g(A_j)/l$, where the A_js are the action sets of the other players, so this set is the natural choice of the externalities factor in the domain of the payoff function of this player. (Also see Remark 2 below.)

We write $\varphi(u, e)$ for the best reply set of a player with payoff function u if he faces $e \in E_u$ as externality. Thus,

$$\varphi(u, e) = \left\{ a \in A_u \colon u(a, e) = \max_{a' \in A_u} u(a', e) \right\}.$$

[10]For an example where this assumption fails, put $A = [0, 1]$ and let $g\colon A \to \mathbb{R}$ be a continuously differentiable function with a nonzero derivative at some point of A, but such that the derivative vanishes on some nondegenerate subinterval of A.

We assume that for a payoff function u, and the associated sets A_u and E_u, the following conditions are satisfied:

(U1) The sets A_u and E_u are nonempty compact subsets of A and E, respectively, such that both A_u and E_u have dense interior.

(U2) The function u is twice continuously differentiable.

(U3) If $(a, e) \in \partial A_u \times E_u$, then there is an $a' \in A_u$ such that $u(a', e) > u(a, e)$.

Note that (U1) and (U2) imply that $\varphi(u, e)$ is nonempty for each $e \in E_u$, and that (U3) implies that $\varphi(u, e) \subseteq \operatorname{int} A_u$ for each $e \in E_u$.

3.4 *Space of payoff functions*

The set of payoff functions is denoted by $\mathcal{U}$. In Section 5.1, we show that there is a unique metrizable topology on $\mathcal{U}$ such that a sequence $\langle u_k \rangle$ in $\mathcal{U}$ converges to some $u \in \mathcal{U}$ if and only if the following statements hold:

(a) We have $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$ and $\rho_H(\partial A_u, \partial A_{u_k}) \to 0$.

(b) If $(a, e) \in \operatorname{dom} u$ and $(a_k, e_k) \in \operatorname{dom} u_k$, $k \in \mathbb{N}$, are such that $(a_k, e_k) \to (a, e)$, then $u_k(a_k, e_k) \to u(a, e)$, $Du_k(a_k, e_k) \to Du(e, a)$, and $D^2 u_k(a_k, e_k) \to D^2 u(e, a)$.[11]

In the rest of the paper, $\mathcal{U}$ is regarded as being equipped with this topology, and for definiteness we fix any metric ρ that induces this topology, so that $\mathcal{U}$ can be regarded as a metric space if necessary. None of our results depends on any specific metric that induces the topology of $\mathcal{U}$; in particular, none of them depends on the choice of the metric ρ.

Evidently on subsets of $\mathcal{U}$ consisting of functions with a common domain, the topology of $\mathcal{U}$ is just the topology of C^2-uniform convergence. We also note that $\mathcal{U}$ is a separable topological space (see Lemma 2 in Section 5.1).

Concerning the condition $\rho_H(\partial A_u, \partial A_{u_k}) \to 0$ in (a) above, this condition is needed in addition to the condition $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$ to have a notion of closeness of action sets such that an interior point of an action set is also an interior point of nearby action sets. This is central for our results, but unfortunately is not implied by the condition $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$ alone unless action sets are assumed to be convex. In fact, for convex action sets, $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$ implies $\rho_H(\partial A_u, \partial A_{u_k}) \to 0$ (see Wills 2007). Of course, in scenarios where all players have the same action set (as considered in Theorem 3 below), the condition $\rho_H(\partial A_u, \partial A_{u_k}) \to 0$ can be dropped from (a).

3.5 *Finite-player games*

Recall that m is the dimension of the ambient Euclidean space of the externalities universe E. We consider finite-player games given by pairs (I, G), where I is a finite set of players, with $\#(I) \geq \max\{2, m + 1\}$, and G is a map from I to $\mathcal{U}$ such that

[11] Concerning (a), note that by remarks in Section 3.1, convergence for the Hausdorff metric on the family of nonempty compact subsets of a Euclidean space X is topological, i.e., does not depend on the particular metric that induces the topology of X.

$E_i = \sum_{j \in I \setminus \{i\}} g(A_j)/((\#I)-1)$ for each $i \in I$, writing E_i for $E_{G(i)}$ and A_j for $A_{G(j)}$. Note that for a player i in a finite-player game (I, G), any distribution of actions chosen by the other players is of the form $\sum_{j \in I \setminus \{i\}} \delta_{a_j}/((\#I)-1)$, where δ_{a_j} denotes Dirac measure at point a_j in the action set A_j of $j \in I \setminus \{i\}$. Thus, the equality $E_i = \sum_{j \in I \setminus \{i\}} g(A_j)/((\#I)-1)$ means that the externalities factor in the domain of the payoff function of a player i is exactly the set of externalities the player could actually face in the game (I, G). If $\#(I) > m$, then by Lemma 9, $\operatorname{int} \sum_{j \in I \setminus \{i\}} g(A_j)/((\#I)-1)$ is dense in $\sum_{j \in I \setminus \{i\}} g(A_j)/((\#I)-1)$, so the conditions $E_i = \sum_{j \in I \setminus \{i\}} g(A_j)/((\#I)-1)$ are consistent with (U1) in the assumptions on payoff functions and, therefore, are consistent with the definition of $\mathcal{U}$. Because the focus of our paper is on large games, there is no problem with a condition based on imagining that a game has sufficiently many players.

A strategy profile in a finite-player game (I, G) is a map $f: I \to A$ such that $f(i) \in A_i$ for each $i \in I$. Given any strategy profile f, we write $e_{f,i}$ for the externality faced by player i; that is, $e_{f,i} = \sum_{j \in I \setminus \{i\}} g(f(j))/(\#(I)-1)$ or, in other words, $e_{f,i} = \int g(a)\, \mathrm{d}\tau_{A,f,i}(a)$, where $\tau_{A,f,i}$ is the distribution of the actions chosen by the players $j \in I \setminus \{i\}$. Thus, for any Borel set $B \subseteq A$,

$$\tau_{A,f,i}(B) = \#\big(\{j \in I \setminus \{i\} : f(j) \in B\}\big)/\big(\#(I)-1\big).$$

A strategy profile $f: I \to A$ is a pure strategy Nash equilibrium if $f(i) \in \varphi(G(i), e_{f,i})$ for each $i \in I$. A pure strategy Nash equilibrium is called strict if $\#(\varphi(G(i), e_{f,i})) = 1$ for each $i \in I$.

Every finite-player game (I, G) defines a distribution on $\mathcal{U}$, i.e., a distribution of payoff functions. We write ν_G for such a distribution. Thus, for any Borel set B in $\mathcal{U}$, $\nu_G(B) = \#(\{i \in I : G(i) \in B\})/\#(I)$.

3.6 *Continuum games*

Recall from Section 3.4 that $\mathcal{U}$ can be regarded as a metric space. Let $\mathcal{M}$ be the set of all Borel probability measures on $\mathcal{U}$ with compact support. (By Lemma 2 in Section 5.1, $\mathcal{U}$ is separable, so any Borel measure on $\mathcal{U}$ has a support.) We regard $\mathcal{M}$ as being given the topology such that $\nu_n \to \nu$ if both $\nu_n \to \nu$ in the narrow topology[12] and $\rho_H(\mathrm{supp}(\nu_n), \mathrm{supp}(\nu)) \to 0$, i.e., $\mathrm{supp}(\nu_n) \to \mathrm{supp}(\nu)$ in the Hausdorff metric topology. Note that for any finite-player game, ν_G as defined in the previous section belongs to $\mathcal{M}$. Given $\nu \in \mathcal{M}$, let

$$E(\nu) = \left\{ \int g\, \mathrm{d}\tau_A : \tau \text{ is a probability measure on } \mathcal{U} \times A \text{ such that } \tau_{\mathcal{U}} = \nu \text{ and } (u, a) \in \mathrm{supp}(\tau) \text{ implies } a \in A_u \right\}.$$

[12]Recall that the narrow topology on the set of Borel measures on a metrizable topological space is the topology of pointwise convergence on the bounded continuous functions defined on this space, evaluation being given by integration.

Following Mas-Colell (1984), we define a continuum game as a distribution on the space of players' characteristics. We add the assumption that the support of the distribution of players' characteristics in a continuum game is compact, i.e., that players' characteristics in a continuum game are not too dissimilar. With a continuum of players, every player is negligible in the strict mathematical sense, so that there is no distinction between the distribution of the actions chosen by all players and the distribution of the actions chosen by all but one player. Thus, because in any game, the externalities factor in the domains of payoff functions must be equal to the set of externalities players could actually face, in our model, a continuum game is an element $\nu \in \mathcal{M}$ such that $E_u = E(\nu)$ for each $u \in \operatorname{supp}(\nu)$. By Lemma 5(a), requiring these equalities is consistent with the definition of $\mathcal{U}$. We write $\mathcal{G}$ for the set of continuum games and give $\mathcal{G}$ the subspace topology induced by the topology of $\mathcal{M}$ (see the Introduction for the meaning of this topology).

Pure strategy Nash equilibria of continuum games are specified in our model in terms of equilibrium distributions, as in Mas-Colell (1984). In our notation, an equilibrium distribution of a continuum game $\nu \in \mathcal{G}$ is a Borel probability measure τ on $\mathcal{U} \times A$ such that $\tau_{\mathcal{U}} = \nu$ and $\operatorname{supp}(\tau) \subseteq \{(u, a) \in \mathcal{U} \times A\colon a \in \varphi(u, e(\tau_A))\}$. By Mas-Colell (1984), every continuum game $\nu \in \mathcal{G}$ has an equilibrium distribution.

Remark 1. Given $\nu \in \mathcal{G}$, there are plenty of sequences $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$ in $\mathcal{M}$. Indeed, in Lemma 7 in Section 5.1, we show, based on the law of large numbers and the Shapley–Folkman theorem, that such sequences do exist. Of course, given such a sequence, there are uncountably many (ordinally nonequivalent) modifications of the countably many payoff functions involved such that the resulting sequences of finite-player games still converge to ν.

Remark 2. In our setup, whether a sequence of finite-player games converges to some continuum game depends only on players' payoffs at actions in their own actions sets and on the externalities they can potentially observe in an actual game. This would not be the case had we taken $\mathcal{U}$ to be a space of functions defined on the product of the entire actions universe and the entire externalities universe.

3.7 *Results*

Our first result gives precision to the idea that, generically, pure strategy Nash equilibrium existence results for continuum games carry over to large finite-player games in a setting with differentiable payoff functions. In the context of this result, the compact support condition on the distributions of players' characteristics in continuum games means that along sequences of finite-player games, players' characteristics must not become too dissimilar if the number of players increases toward infinity.

Theorem 1. *There is an open and dense subset $\mathcal{G}^*$ of $\mathcal{G}$ such that whenever $\nu \in \mathcal{G}^*$, and $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$, there is an $N \in \mathbb{N}$ such that (I_n, G_n) has a strict pure strategy Nash equilibrium if $\#(I_n) \geq N$.*

The idea of the proof is first to identify an open and dense set $\mathcal{G}^* \subseteq \mathcal{G}$ such that if $\nu \in \mathcal{G}^*$, then ν has an equilibrium distribution τ such that for some neighborhood V of $\operatorname{supp}(\nu)$ and some neighborhood W of $e(\tau_A)$, best replies on $V \times W$ are given by a (continuously differentiable) function. Given such a ν, if $\langle (I_n, G_n)\rangle$ is a sequence of finite-player games such that $\nu_{G_n} \to \nu$, then $\operatorname{supp}(\nu_{G_n})$ must be in V eventually. In addition, at this point, the idea is to get strict pure strategy Nash equilibria for the games (I_n, G_n) if n is large enough by a fixed-point argument, using the fact that best replies on $V \times W$ are given by a function, but taking care of the fact that in finite-player games, different players may face different externalities. A more detailed overview of the proof of Theorem 1 is provided in Section 4.

The following example shows that in the context of Theorem 1, we cannot do better than to obtain a result for generic distributions of players' characteristics, regardless of the number of players, i.e., of the size of I.

Example. Let $A = [-1/2, 3/2]$ and let $v\colon A \to \mathbb{R}_+$ be a twice continuously differentiable function, with $Dv(1/2) = D^2v(1/2) = v(1/2) = 0$, assuming a global maximum, equal to 1, exactly at the points 0 and 1. Let $g\colon A \to \mathbb{R}$ be the restriction to A of the identity on $\mathbb{R}$. Then $E = [-1/2, 3/2]$, and for each $f\colon I \to A$ and each $i \in I$, the externality $e_{f,i} \in E$ faced by i is the mean of the actions of the players different from i. Let $\#(I)$ be even, with $\#(I) \geq 4$. Partition I into sets H and J of equal size. For $i \in H$, the payoff function is $u_H\colon A \times E \to \mathbb{R}$ defined by setting

$$u_H(a, e) = v(a)(3/2 - e) \quad \text{if } a < 1/2 \quad \text{and} \quad u_H(a, e) = v(a) \quad \text{if } a \geq 1/2,$$

and for $i \in J$, the payoff function is $u_J\colon A \times E \to \mathbb{R}$ defined by setting

$$u_J(a, e) = v(a)(e + 1/2) \quad \text{if } a < 1/2 \quad \text{and} \quad u_J(a, e) = v(a) \quad \text{if } a \geq 1/2.$$

Note that for all $i \in I$ and all values of $e_{f,i}$, the best reply sets are included in $\{0, 1\}$, and that if $f\colon I \to A$ is a strategy profile such that $f(i) \in \{0, 1\}$ for all $i \in I$, then $e_{f,i} = \#\{j \in I \setminus \{i\}\colon a_j = 1\}/\#(I - 1)$ for each $i \in I$.

Now suppose $f\colon I \to A$ is a pure strategy Nash equilibrium. Consider $i, i' \in H$ and suppose $f(i) = 0$ and $f(i') = 1$. Then, by optimal choice of actions, $e_{f,i} \leq 1/2$ and $e_{f,i'} \geq 1/2$. However, calculating frequencies, we see that $f(i) = 0$ and $f(i') = 1$ together imply that $e_{f,i} > e_{f,i'}$, and from this contradiction it follows that all members of H must choose the same action, say 0. But then, because $\#(I)$ is even, $e_{f,i} < 1/2$ for all members i of J, so they all must play 1, by optimal choice of actions. This, however, means that $e_{f,i} > 1/2$ for all members of H, again because $\#(I)$ is even, so their optimal actions are also equal to 1, and this contradiction shows that no pure strategy Nash equilibrium exists.

In regard to this example, note that the set of games that have pure strategy Nash equilibria is closed in the set of all games with the given number of players. Thus, this example actually provides a counterexample to the idea that it could be possible in our model to define, for any finite number of players, a generic set of payoff functions such that every game with the given finite number of players and payoff functions in this set had a pure strategy Nash equilibrium.

The arguments in the proof of Theorem 1 can be used to show that in our model, every equilibrium distribution of every continuum game is the limit of some sequence of finite-player games and corresponding strict pure strategy Nash equilibria, in the sense of the following definition.

Definition. Let $\nu \in \mathcal{G}$ be a continuum game and let τ be an equilibrium distribution for ν. A sequence $\langle (I_n, G_n)\rangle_{n\in\mathbb{N}}$ of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$ is said to *asymptotically implement* (ν, τ) if for all n larger than some $N \in \mathbb{N}$, (I_n, G_n) has a strict pure strategy Nash equilibrium f_n such that the sequence of distributions of the maps $G_n \times f_n$ converges to τ narrowly. We say that (ν, τ) is *asymptotically implementable* if it can be asymptotically implemented by some sequence $\langle (I_n, G_n)\rangle_{n\in\mathbb{N}}$ with $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$.

Theorem 2. *Every (ν, τ), where $\nu \in \mathcal{G}$ is a continuum game and τ is an equilibrium distribution for ν, is asymptotically implementable by a sequence $\langle (I_n, G_n)\rangle_{n\in\mathbb{N}}$ of finite-player games such that $\nu_{G_n} \in \mathcal{G}$ for each n.*

We emphasize that every continuum game can be taken in Theorem 2, not just a game from a generic set. Thus, Theorem 2 shows that in our model, no equilibrium distribution of any continuum game is an artifact of having continuum many players.

The idea of the proof is to show first that any continuum game $\nu \in \mathcal{G}$ and any equilibrium distribution τ of ν can be approximated by ν's belonging to the generic set $\mathcal{G}^*$ from Theorem 1, together with equilibrium distributions that witness this. We then extend the arguments of the proof of Theorem 1 and show that, given any $\nu' \in \mathcal{G}^*$ and any equilibrium distribution τ' that witnesses this, along every sequence $\langle (I_n, G_n)\rangle$ of finite-player games such that $\nu_{G_n} \to \nu'$, there are strict pure strategy Nash equilibria f_n if n is large enough such that the distributions of the maps $G_n \times f_n$ converge to τ' narrowly. Coupling these arguments with an argument (based on the law of large numbers) that shows that sequences of finite player games converging to a given continuum game do exist, Theorem 2 follows by an obvious diagonal argument.

We note that the approximating sequence of finite player games is, in general, different from a finite-player truncation of the given continuum game, i.e., it is not generally the case that $\operatorname{supp}(\nu_{G_n}) \subseteq \operatorname{supp}(\nu)$. However, if $\langle (\hat{I}_n, \hat{G}_n)\rangle_{n\in\mathbb{N}}$ is a sequence of finite-player games such that $\#(\hat{I}_n) \to \infty$, $\nu_{\hat{G}_n} \to \nu$ and $\operatorname{supp}(\nu_{\hat{G}_n}) \subseteq \operatorname{supp}(\nu)$, then $(\hat{I}_n, \hat{G}_n)$ and (I_n, G_n) are arbitrarily close when n is large.

In much of the literature on continuum games it is assumed, following Schmeidler (1973) and Mas-Colell (1984), that all players have the same action set. In our model, the set of continuum games in which all players have the same action set is not open in the set of all continuum games, so Theorems 1 and 2 do not apply to this special case. We address this case with another theorem. To this end, it is convenient to settle some additional notation.

Let $\mathcal{C}$ be the set of all closed subsets C of the actions universe A such that C has dense interior. For each $C \in \mathcal{C}$, let $S_C = \{u \in \mathcal{U} \colon A_u = C\}$. Note that S_C is closed in $\mathcal{U}$ for each $C \in \mathcal{C}$. For each $C \in \mathcal{C}$, let $\mathcal{M}_C = \{\nu \in \mathcal{M} \colon \operatorname{supp}(\nu) \subseteq S_C\}$ and $\mathcal{G}_C = \mathcal{G} \cap \mathcal{M}_C$.

THEOREM 3. *Given any $C \in \mathcal{C}$, there is a relatively open dense subset $\mathcal{G}_C^*$ of $\mathcal{G}_C$ such that whenever $\nu \in \mathcal{G}_C^*$ and $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$, then there is an $N \in \mathbb{N}$ such that (I_n, G_n) has a strict pure strategy Nash equilibrium if $\#(I_n) \geq N$. Moreover, every (ν, τ), where $\nu \in \mathcal{G}_C$ is a continuum game and τ is an equilibrium distribution for ν, is asymptotically implementable by a sequence $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ of finite-player games such that $\nu_{G_n} \in \mathcal{M}_C$ for each n.*

The ideas of the proof are the same as those of the proofs of Theorems 1 and 2, modulo that concerning asymptotic implementation of equilibrium distributions of continuum games, one has to take care to choose finite-player games in which all players have the correct action set.

4. Overview of the proof of the main result

In this section we sketch the argument for our main result, Theorem 1, in the special case in which all players have the same action set $A = [0, 1]$ and the externality map g is the identity on A, i.e., $g = \mathrm{id}_A$ (which is actually a special case of Theorem 3).

In this special case, regardless of the number of players in a game, in particular, regardless of whether the set of players is finite or a continuum, the externality set of any player is equal to $[0, 1]$, so that $[0, 1]$ can be taken to be the universe E of externalities. Thus, the space $\mathcal{U}$ of possible payoff functions is the space of twice continuously differentiable functions on $A \times E \equiv [0, 1] \times [0, 1]$, and the topology of $\mathcal{U}$ is just the standard topology of C^2-uniform convergence. Note also for the sequel that the choice of g to be id_A entails that the boundary assumption (U3) on payoff functions implies that equilibrium externalities are always in the interior of E.

We build up to the generic set $\mathcal{G}^*$ by considering first the subset $\mathcal{G}_1$ of $\mathcal{G}$ that consists of those ν such that for some equilibrium distribution τ of ν, the following conditions hold:

(i) We have $\#(\varphi(u, e(\tau_A))) = 1$ for each $u \in \mathrm{supp}(\nu)$.

(ii) For each $u \in \mathrm{supp}(\nu)$, $D_a^2 u(a_u, e(\tau_A)) < 0$, where a_u is the unique element of $\varphi(u, e(\tau_A))$.

Using the implicit function theorem and the choice of the topology on $\mathcal{U}$, we can show that the conditions imply that for each $\nu \in \mathcal{G}_1$, there are neighborhoods V of $\mathrm{supp}(\nu)$ in $\mathcal{U}$ and W of $e(\tau_A)$ in E such that, on $V \times W$, best replies of u against e can be described by a continuous map h such that $h(u, \cdot)$ is continuously differentiable. In particular, for each $e \in W$, $\int h(u, e)\, \mathrm{d}\nu(u)$ is the resulting externality when players reply against e optimally. Hence, the map $\xi \colon W \to \mathbb{R}^m$, defined by setting

$$\xi(e) = \int h(u, e)\, \mathrm{d}\nu(u) - e$$

for each $e \in W$, is such that its zeros are equilibrium externalities; for each equilibrium externality e, the corresponding equilibrium distribution is $\nu \circ (\mathrm{id}_{\mathcal{U}} \times h(\cdot, e))^{-1}$. Also, by Leibniz' rule, the map ξ is continuously differentiable.

Let $\mathcal{G}^*$ be the subset of $\mathcal{G}_1$ that consists of those ν such that for some equilibrium distribution τ of ν, the associated map ξ has a derivative $D\xi(e(\tau_A))$ different from zero. This is needed in the fixed-point argument sketched below. The set $\mathcal{G}^*$ is the generic set of Theorem 1. We next sketch the reason why $\mathcal{G}^*$ is open and dense in $\mathcal{G}$.

To see that $\mathcal{G}^*$ is open, fix $\nu \in \mathcal{G}^*$, and let V, W, and h be as above. By the choice of the topology on $\mathcal{G}$, there is a neighborhood U of ν in $\mathcal{G}$ such that $\operatorname{supp}(\nu') \subseteq V$ for all $\nu' \in U$. We can, therefore, define a map $\xi_U \colon U \times W \to \mathbb{R}^m$ by setting

$$\xi_U(\nu', e) = \int h(u, e)\,\mathrm{d}\nu'(u) - e$$

for each $\nu' \in U$ and $e \in W$, which has the property that $\xi_U(\nu', e) = 0$ implies that e is an equilibrium externality in the continuum game ν'. The defining properties of $\mathcal{G}^*$ imply that we can use a version of the implicit function theorem to obtain a continuous map $\nu' \mapsto e(\nu') \colon U \to W$ such that $\xi_U(\nu', e(\nu')) = 0$ for each $\nu' \in U$ (shrinking U, if necessary). Thus, $e(\nu')$ is an equilibrium externality of the continuum game ν' and the corresponding equilibrium distribution is, as above, $\tau' = \nu' \circ (\mathrm{id}_{\mathcal{U}} \times h(\cdot, e(\nu')))^{-1}$. Thus, any $\nu' \in U$ has an equilibrium distribution such that (i) in the definition of $\mathcal{G}_1$ holds. Condition (ii) is easy to verify as is the condition that $D_e\xi_U(\nu', e(\nu')) \neq 0$, the latter because $D_e\xi_U(\nu', e')$ depends continuously on (ν', e'). It thus follows that $\mathcal{G}^*$ is open.

Denseness of $\mathcal{G}^*$ is established by perturbing the payoff functions in any given continuum game. Such a perturbation is used first to show that $\mathcal{G}_1$ is dense in $\mathcal{G}$: Fix $\nu \in \mathcal{G}$ and let τ be an equilibrium distribution of ν. For each $(u, a) \in \operatorname{supp}(\tau)$, the action a is made to be a strict best reply for the payoff function u, roughly by adding a function ρ to u. Specifically, choose a twice continuously differentiable map $\rho \colon \mathbb{R} \to \mathbb{R}$ such that $\rho(0) = 0 > \rho(x)$ for all $x \in \mathbb{R} \setminus \{0\}$ and $D^2\rho(0) < 0$. For each $k \in \mathbb{N}$, $u \in \operatorname{supp}(\nu)$, and $a \in A$, define $u_{k,a} \colon A \times E \to \mathbb{R}$ by setting

$$u_{k,a}(a', e) = u(a', e) + \tfrac{1}{k+1}(\rho(a - a'))$$

for each $(a', e) \in A \times E$. If k is large enough, then all assumptions on payoff functions are satisfied by $u_{k,a}$, and $u_{k,a}$ is close to the corresponding u. We can then construct a continuum game by changing payoff functions u to $u_{k,a}$; formally, define a map $\lambda_k \colon \operatorname{supp}(\nu) \times A \to \mathcal{U}$ by setting $\lambda_k(u, a) = u_{k,a}$ for each $u \in \operatorname{supp}(\nu)$ and $a \in A$. Then τ_k defined as $\tau_k = \tau \circ (\lambda_k \times \mathrm{id}_A)^{-1}$ is an equilibrium distribution for $\nu_k = \tau \circ \lambda_k^{-1}$, witnessing that $\nu_k \in \mathcal{G}_1$. It follows that $\mathcal{G}_1$ is dense in $\mathcal{G}$.

For the denseness of $\mathcal{G}^*$, it now suffices to show that $\mathcal{G}^*$ is dense in $\mathcal{G}_1$. Fix $\nu \in \mathcal{G}_1$ and let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_1$. Set $\bar{e} = e(\tau_A)$. Roughly, the argument consists of constructing a family of a continuum of games ν_λ, indexed by $\lambda \in [0, 1)$, and corresponding equilibrium distributions τ_λ, both depending continuously on λ, such that the externality $e(\tau_{\lambda,A})$ is always $\bar{e}$ and the derivative $D\xi_\lambda(\bar{e})$ is zero for at most one value of λ, where ξ_λ is associated with τ_λ as above. This is done as follows. For each $u \in \operatorname{supp}(\nu)$ and each $\lambda \in [0, 1)$, define $u_\lambda \colon A \times E \to \mathbb{R}$ by setting

$$u_\lambda(a, e) = u(a, (1 - \lambda)\bar{e} + \lambda e)$$

for each $(a, e) \in A_u \times E(\nu)$. There is no difference between $u_\lambda(\cdot, \bar{e})$ and $u(\cdot, \bar{e})$; hence, changing the payoff function from u to u_λ yields a continuum game that belongs to $\mathcal{G}_1$. Indeed, define $\kappa_\lambda\colon \operatorname{supp}(\nu) \to \mathcal{U}$ by setting $\kappa_\lambda(u) = u_\lambda$. Set $\nu_\lambda = \nu \circ \kappa_\lambda^{-1}$ and $\tau_\lambda = \nu \circ (\kappa_\lambda \times h(\cdot, \bar{e}))^{-1}$. Clearly, for every $u_\lambda \in \operatorname{supp}(\nu_\lambda)$, the optimal response against $\bar{e}$ is the same as that of the corresponding $u \in \operatorname{supp}(\nu)$, so ν_λ belongs to $\mathcal{G}_1$ and τ_λ is an equilibrium distribution for ν_λ witnessing this. Moreover, we have

$$D\xi_\lambda(\bar{e}) = \lambda \int_{\mathcal{U}} D_e h(u, \bar{e})\, \mathrm{d}\nu(u) - 1,$$

so $D\xi_\lambda(\bar{e})$ can be zero for at most one value of λ. Hence, it is possible to choose λ as close to 1 as we please, while having $D\xi_\lambda(\bar{e}) \neq 0$. As each ν_λ belongs to $\mathcal{G}_1$, it follows that $\mathcal{G}^*$ is dense in $\mathcal{G}_1$.

Finally, we outline the proof of the existence of pure strategy Nash equilibria for members of the tail of a sequence $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu \in \mathcal{G}^*$. To see this, let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}^*$. Let V, W, h, and ξ be associated with ν and τ as above. For large n, $\operatorname{supp}(\nu_{G_n}) \subset V$, so best replies for any $u \in \operatorname{supp}(\nu_{G_n})$ against any $e \in W$ are given by the function h. The idea now is to set up a fixed-point problem as follows.

Let $\widehat{W}$ be a compact interval such that $e(\tau_A) \in \operatorname{int} \widehat{W}$ and $\widehat{W} \subseteq \operatorname{int} W$. If n is large enough, then for each strategy profile $f \in A^{I_n}$ and each $i \in I_n$,

$$\int a\, \mathrm{d}\tau_{A,f,i}(a) - \int a\, \mathrm{d}\tau_{A,f}(a) + e \in W$$

whenever $e \in \widehat{W}$. Thus, for large n, we can define a map $\Lambda\colon A^{I_n} \times \widehat{W} \to A^{I_n} \times \mathbb{R}$ by setting

$$\Lambda(f, e) = \left(\left\langle h\left(G_n(i), \int a\, \mathrm{d}\tau_{A,f,i}(a) - \int a\, \mathrm{d}\tau_{A,f}(a) + e \right) \right\rangle_{i \in I_n}, \int a\, \mathrm{d}\tau_{A,f}(a) \right)$$

for each $f \in A^{I_n}$ and $e \in \widehat{W}$. If (f, e) is a fixed point of this map, then $e = \int a\, \mathrm{d}\tau_{A,f}(a)$ and, thus, for each $i \in I_n$, $\int a\, \mathrm{d}\tau_{A,f,i}(a)$, that is the externality induced by the actions of the players other than i, equals $\int a\, \mathrm{d}\tau_{A,f,i}(a) - \int a\, \mathrm{d}\tau_{A,f}(a) + e$, implying that $f_i = h(G_n(i), \int a\, \mathrm{d}\tau_{A,f,i}(a))$. Thus, f is a pure strategy Nash equilibrium of (I_n, G_n). It is important to note that Λ is not a self-map. But since $\nu \in \mathcal{G}^*$ and τ witnesses this, we have $D\xi(e(\tau_A)) \neq 0$. In particular, there is neighborhood of $e(\tau_A)$ included in $\widehat{W}$ on which $\xi(e) = 0$ if and only if $e = e(\tau_A)$. Based on these facts, we can show that a fixed-point result due to Mas-Colell (1983) can be applied to the situation (again if n is large enough).

We now mention some of the difficulties in extending these arguments to the general case considered in Theorem 1. First, the externalities set of a player in a finite-player game depends, in general, on the number of players in the game. This is so even when there is a common action set. For example, take $A = [0, 1]$ as above, but take $a \mapsto (a, a^2)$ for the map g; this just adds some notion of dispersion to the average of actions considered above as externality.

Second, in general, the boundary assumption on payoff functions does not guarantee that equilibrium externalities are in the interior of the externality sets. This is so even with a continuum of players and even when players have the same common action sets. As an example, again take $A=[0,1]$ for the common action set of players, and for g, take the map $a\mapsto(a,a^2)\colon A\to\mathbb{R}^2$. To show that, for continuum games, equilibrium externalities are, generically, in the interior of the externality sets, which is needed to be in a convenient position to apply differentiability techniques, makes up for the main difficulty of this paper.

5. Proofs

This section contains the proofs of our results. Here is a short road map.

In Section 3.4, we made a convention on the topology of the space $\mathcal{U}$ of payoff functions. In Lemma 1 below, we prove the existence of this topology, and in Lemma 2, we prove that $\mathcal{U}$ (with this topology) is separable.

We continue with some more preparatory lemmata. We mention Lemma 7, which proves the fact announced in Remark 1 that in our model, every continuum game can be approximated by finite-player games, a fact that is a prerequisite for asymptotic implementability of equilibrium distributions of continuum games.

The set $\mathcal{G}^*$ in the statement of Theorem 1 is defined in part (c) of the proof of this theorem. The proof that this set is dense in $\mathcal{G}$ is delegated to a separate lemma (Lemma 8 after the proof of Theorem 1). This lemma is stated in greater generality than is actually needed in the proof of Theorem 1 so that it can also be used in the proof of Theorem 2.

5.1 *Preliminaries*

Lemma 1. *There is a unique metrizable topology on $\mathcal{U}$ (with which $\mathcal{U}$ is regarded as being endowed) such that a sequence $\langle u_k\rangle$ in $\mathcal{U}$ converges to some $u\in\mathcal{U}$ if and only if the following statements hold:*

(a) We have $\rho_H(\operatorname{dom}u,\operatorname{dom}u_k)\to 0$ and $\rho_H(\partial A_u,\partial A_{u_k})\to 0$;

(b) If $(a,e)\in\operatorname{dom}u$ and $(a_k,e_k)\in\operatorname{dom}u_k$, $k\in\mathbb{N}$, are such that $(a_k,e_k)\to(a,e)$, then $u_k(a_k,e_k)\to u(a,e)$, $Du_k(a_k,e_k)\to Du(e,a)$, and $D^2u_k(a_k,e_k)\to D^2u(e,a)$.

Proof. It suffices to find one metric ρ on $\mathcal{U}$ for which convergence of any sequence in $\mathcal{U}$ is equivalent to the truth of (a) and (b). Write Γ_f for the graph of a function f and define ρ by setting

$$\begin{aligned}\rho\big(u,u'\big)=&\rho_H\big(\operatorname{dom}u,\operatorname{dom}u'\big)+\rho_H(\partial A_u,\partial A_{u'})\\&+\rho_H(\Gamma_u,\Gamma_{u'})+\rho_H(\Gamma_{Du},\Gamma_{Du'})+\rho_H(\Gamma_{D^2u},\Gamma_{D^2u'})\end{aligned}$$

for $u,u'\in\mathcal{U}$. Clearly ρ is a metric on $\mathcal{U}$. Let $u\in\mathcal{U}$ and let $\langle u_k\rangle$ be a sequence in $\mathcal{U}$.

Assume $\rho(u,u_k)\to 0$. Directly from the definition of ρ, we see that (a) must be true. As for (b), let (a,e) and (a_k,e_k), $k\in\mathbb{N}$, be as hypothesized. Note that $\rho_H(\Gamma_u,\Gamma_{u_k})\to 0$ because $\rho(u,u_k)\to 0$. Thus, by the definition of ρ_H, the sequence $\langle(a_k,e_k,u_k(a_k,e_k))\rangle$ must be bounded and any of its cluster points must belong to Γ_u, and, therefore (by the definition of graph), must be of the form $(a,e,u(a,e))$ since $(a_k,e_k)\to(a,e)$. Thus, the assertion of (b) follows (because, in a Euclidean space, a bounded sequence has a cluster point, x say, and is convergent to x if x is the only cluster point). Similarly, the other assertions of (b) follow.

Assume that (a) and (b) are true. Combining the first of the facts in (b) with the fact that $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$, we see that $\Gamma_u \subseteq \operatorname{Li}\Gamma_{u_k}$.[13] Suppose $(a, e, r) \in \operatorname{Ls}\Gamma_{u_k}$. Then for some sequence $\langle n_i \rangle_{i\in\mathbb{N}}$ in $\mathbb{N}$, there are points $(a_{k_i}, e_{k_i}) \in \operatorname{dom} u_{k_i}$, $i \in \mathbb{N}$, such that $(a_{k_i}, e_{k_i}, u_{k_i}(a_{k_i}, e_{k_i})) \to (a, e, r)$. From the fact that $\rho_H(\operatorname{dom} u, \operatorname{dom} u_k) \to 0$, we see that $(a, e) \in \operatorname{dom} u$. Again by this fact, there is a sequence $\langle (a_k, e_k)\rangle_{k\in\mathbb{N}}$ such that $(a_k, e_k) \to (a, e)$ and $(a_k, e_k) \in \operatorname{dom} u_k$ for each k. Define a sequence $\langle (a'_k, e'_k)\rangle_{k\in\mathbb{N}}$ by setting $(a'_k, e'_k) = (a_{k_i}, e_{k_i})$ if $k = k_i$ for some i, and $(a'_k, e'_k) = (a_k, e_k)$ otherwise. Then $(a'_k, e'_k) \in \operatorname{dom} u_k$ for each k and $(a'_k, e'_k) \to (a, e)$. Consequently (b) implies that $u_k(a'_k, e'_k) \to u(a, e)$. In particular, we have $u_{k_i}(u_{k_i}, e_{k_i}) \to u(a, e)$ and, therefore, $r = u(a, e)$. Thus, $\operatorname{Ls}\Gamma_{u_k} \subseteq \Gamma_u$ and it follows that $\Gamma_u = \operatorname{Ls}\Gamma_{u_k} = \operatorname{Li}\Gamma_{u_k}$. Now because $\operatorname{dom} u$ and $\operatorname{dom} u_k$, $k \in \mathbb{N}$, are all included in the compact set $A \times E$, and because the maps u and u_k are continuous, (a) and (b) imply, in particular, that the sets Γ_u and Γ_{u_k}, $k \in \mathbb{N}$, are commonly included in a compact subset of the ambient Euclidean space, so the fact that $\Gamma_u = \operatorname{Ls}\Gamma_{u_k} = \operatorname{Li}\Gamma_{u_k}$ implies that $\rho_H(\Gamma_u, \Gamma_{u_k}) \to 0$.

Similarly, we see that both $\rho_H(\Gamma_{Du}, \Gamma_{Du_k})$ and $\rho_H(\Gamma_{D^2u}, \Gamma_{D^2u_k})$ converge to 0 as $k \to \infty$. By the definition of ρ, we conclude that $\rho(u, u_k) \to 0$. $\square$

Lemma 2. *The space $\mathcal{U}$ is separable.*

Proof. Let $\mathcal{F}_0$ be the set of all nonempty compact subsets of $\mathbb{R}^{n+m}$, let $\mathcal{F}_1$ be the set of all nonempty compact subsets of $\mathbb{R}^{n+m} \times \mathbb{R}$, let $\mathcal{F}_2$ be the set of all nonempty compact subsets of $\mathbb{R}^{n+m} \times \mathbb{R}^{n+m}$, and let $\mathcal{F}_3$ be the set of all nonempty compact subsets of $\mathbb{R}^{n+m} \times \mathbb{R}^{(n+m)^2}$. For each $i = 0, 1, 2, 3$, give $\mathcal{F}_i$ the Hausdorff metric topology, so that each $\mathcal{F}_i$ becomes a separable metric space. Write $\mathcal{F} = \mathcal{F}_0 \times \mathcal{F}_0 \times \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3$ and give $\mathcal{F}$ the product topology. Then $\mathcal{F}$ is a separable metrizable topological space. Consider the map $\phi: \mathcal{U} \to \mathcal{F}$ defined by setting

$$\phi(u) = (\operatorname{dom} u, \partial A_u, \Gamma_u, \Gamma_{Du}, \Gamma_{D^2u})$$

for each $u \in \mathcal{U}$. By definition of the topology of $\mathcal{U}$, ϕ is a homeomorphism from $\mathcal{U}$ onto $\phi(\mathcal{U})$. As $\mathcal{F}$ is separable and metrizable, any subset of $\mathcal{F}$ is separable (in the subspace topology). In particular, $\phi(\mathcal{U})$ is separable, and it follows that $\mathcal{U}$ is separable. $\square$

Lemma 3. *Let C and C_k, $k \in \mathbb{N}$, be compact subsets of $\mathbb{R}^\ell$, all with nonempty interior, such that both $\rho_H(C, C_k) \to 0$ and $\rho_H(\partial C, \partial C_k) \to 0$ as $k \to \infty$. Let $x \in \operatorname{int} C$ and suppose $\langle x_k \rangle$ is a sequence in $\mathbb{R}^\ell$ such that $x_k \to x$. Then $x_k \in \operatorname{int} C_k$ for all sufficiently large k.*

Proof. Otherwise, passing to a subsequence if necessary, we can assume $x_k \notin \operatorname{int} C_k$ for each k. As $\rho_H(C, C_k) \to 0$, we can find a $y_k \in C_k$ for each k so that $y_k \to x$. Now using the

[13]Here and below, $\operatorname{Li}\Gamma_{u_k}$ is the set of limits of sequences $\langle (a_k, e_k, r_k)\rangle$ such that $(a_k, e_k, r_k) \in \Gamma_{u_k}$ for all k, and $\operatorname{Ls}\Gamma_{u_k}$ is the set of cluster points of such sequences.

fact that the C_ks are closed, we can select a $\lambda_k \in [0, 1]$ for each k so that $z_k = (1-\lambda_k)y_k + \lambda_k x_k \in \partial C_k$. As $x_k \to x$ as well as $y_k \to x$, we also have $z_k \to x$. As $\rho_H(\partial C, \partial C_k) \to 0$, it follows that $x \in \partial C$, contradicting the hypothesis about x. □

Lemma 4. *Let C and C_k, $k \in \mathbb{N}$, be compact subsets of $\mathbb{R}^\ell$, all with nonempty interior, such that both $\rho_H(C, C_k) \to 0$ and $\rho_H(\partial C, \partial C_k) \to 0$ as $k \to \infty$. Let K be a compact subset of $\operatorname{int} C$. Then $K \subseteq \operatorname{int} C_k$ for all sufficiently large k.*

Proof. Otherwise, passing to a subsequence if necessary, for each k we can find an $x_k \in K$ such that $x_k \notin \operatorname{int} C_k$. As K is compact, we can assume that $x_k \to x$, again passing to a subsequence, if necessary. Now $x \in K \subseteq \operatorname{int} C$, so by Lemma 3 we must have $x_k \in \operatorname{int} C_k$ for large k, thus getting a contradiction. □

Lemma 5. *(a) For every $\nu \in \mathcal{M}$, $E(\nu)$ is a compact convex subset of E with nonempty interior in $\mathbb{R}^m$. (b) If $\nu_n \to \nu$ in $\mathcal{M}$, then $\rho_H(E(\nu_n), E(\nu)) \to 0$.*

Proof. Clearly $E(\nu)$ is convex for each $\nu \in \mathcal{M}$. As for compactness, fix $\nu \in \mathcal{M}$ and set $Y = \operatorname{supp}(\nu)$. By the definition of $\mathcal{M}$, Y is compact, and by hypothesis, so is the actions universe A. Thus, the set Z of all Borel probability measures on $Y \times A$ is narrowly compact. Evidently the set of those Borel probability measures that matter in the definition of $E(\nu)$ can be regarded as a narrowly closed subset of Z, and, thus, $E(\nu)$ must be compact, because g is continuous.

For the other claims, consider the correspondence $\theta\colon \mathcal{U} \to 2^{\mathbb{R}^m}$ defined by setting

$$\theta(u) = \operatorname{co} g(A_u)$$

for each $u \in \mathcal{U}$. Then θ has nonempty compact convex values, all with a nonempty interior by Lemma 9(b). The fact that θ has convex values implies that $\int \theta \, d\nu$ is convex for all $\nu \in \mathcal{M}$, and the fact that θ has compact values, all included in the compact set $\operatorname{co} g(A)$, implies that $\int \theta \, d\nu$ is compact for all $\nu \in \mathcal{M}$ (see Hildenbrand 1974, D.II.4, Proposition 7). Note that the correspondence $u \mapsto A_u\colon \mathcal{U} \to 2^A$ is continuous (see Hildenbrand 1974, B.III, Problem 4). Because the map g is continuous, this implies that the correspondence $u \mapsto g(A_u)\colon \mathcal{U} \to 2^{\mathbb{R}^m}$ is continuous. By Hildenbrand (1974, B.III, Propositions 6 and 10), it follows that θ is continuous.

We claim that $E(\nu) = \int \theta(u) \, d\nu(u)$ for each $\nu \in \mathcal{M}$. To see this, fix any $\nu \in \mathcal{M}$ and any $p \in \mathbb{R}^m$. Note that the map $p \circ g$ from A to $\mathbb{R}$ is continuous. Consequently, since the correspondence $u \mapsto A_u$ is continuous, with nonempty compact values, it has a measurable selection h such that $(p \circ g)h(u) = \max(p \circ g)A_u$ for each $u \in \mathcal{U}$ (use the maximum theorem together with Hildenbrand 1974, B.III, Proposition 1 and D.II.2, Lemma 1). We must, therefore, have $\max pE(\nu) = \int_{\mathcal{U}} \max pg(A_u) \, d\nu(u)$, by the definition of $E(\nu)$, and also

$$\int_{\mathcal{U}} \max pg(A_u) \, d\nu(u) = \int_{\mathcal{U}} \max p \operatorname{co} g(A_u) \, d\nu(u) = \max p \int_{\mathcal{U}} \theta(u) \, d\nu(u).$$

As p is an arbitrary element of $\mathbb{R}^m$, and both $E(\nu)$ and $\int_{\mathcal{U}} \theta(u) \, d\nu(u)$ are compact and convex, it follows that $E(\nu) = \int_{\mathcal{U}} \theta(u) \, d\nu(u)$, as claimed.

Now from this equality, we can see that $\operatorname{int} E(\nu) \neq \varnothing$ for each $\nu \in \mathcal{M}$. Indeed, pick any $\nu \in \mathcal{M}$ and any $u' \in \operatorname{supp}(\nu)$. As noted above, $\operatorname{int} \theta(u') \neq \varnothing$, so there is a compact set $K \subseteq \operatorname{int} \theta(u')$ such that $\operatorname{int} K \neq \varnothing$. By Lemma 4, there is an open neighborhood V of u' such that $K \subseteq \theta(u'')$ for each $u'' \in V$. As $u' \in \operatorname{supp}(\nu)$, $\nu(V) > 0$, so the set $\nu(V)K$ has a nonempty interior. Now

$$\nu(V)K + \int_{\mathcal{U}\setminus V} \theta(u)\,\mathrm{d}\nu(u) \subseteq \int_V \theta(u)\,\mathrm{d}\nu(u) + \int_{\mathcal{U}\setminus V} \theta(u)\,\mathrm{d}\nu(u) = \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u),$$

showing that $\operatorname{int} \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u) \neq \varnothing$. Finally, to see that part (b) of the lemma is true, note that since θ is continuous, with nonempty compact values, for each $p \in \mathbb{R}^m$, the map $u \mapsto \max p\theta(u) \colon \mathcal{U} \to \mathbb{R}$ is continuous, by the maximum theorem. Moreover, this map is bounded because the values of θ are included in the compact set $E \subseteq \mathbb{R}^m$. Hence, for each $p \in \mathbb{R}^m$, the map $\nu \mapsto \int_{\mathcal{U}} \max p\theta(u)\,\mathrm{d}\nu(u) \colon \mathcal{M} \to \mathbb{R}$ is continuous. By the facts used above, we see that θ has a measurable selection h such that $ph(u) = \max p\theta(u)$ for each $u \in \mathcal{U}$, implying that $\int_{\mathcal{U}} \max p\theta(u)\,\mathrm{d}\nu(u) = \max p \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u)$, and it follows that for each $p \in \mathbb{R}^m$, the map $\nu \mapsto \max p \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u) \colon \mathcal{M} \to \mathbb{R}$ is continuous. Because $\int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u)$ is nonempty convex and compact for each $\nu \in \mathcal{M}$, it follows that the map $\nu \mapsto \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u)$ is continuous for the Hausdorff metric on the set of all nonempty compact subsets of $\mathbb{R}^m$ (see Castaing and Valadier 1977, II-23). Thus we get (b), again by the equality $E(\nu) = \int_{\mathcal{U}} \theta(u)\,\mathrm{d}\nu(u)$ established above. □

Lemma 6. *Let $\nu \in \mathcal{G}$, and let $\langle (I_n, G_n) \rangle_{n\in\mathbb{N}}$ be a sequence of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$ in $\mathcal{M}$. Let W be a compact subset of $\operatorname{int} E(\nu)$. Then there is an $N \in \mathbb{N}$ such that $W \subseteq E_{G_n(i)}$ for each $i \in I_n$ whenever $n \geq N$.*

Proof. In this proof, we write $E_{n,i}$ for $E_{G_n(i)}$ and write l_n for $\#(I_n) - 1$.

(a) There are numbers $N_1 \in \mathbb{N}$ and $\epsilon > 0$ such that whenever $n \geq N_1$ and $i \in I_n$, then $W \subseteq \operatorname{co} E_{n,i}$ but $\operatorname{dist}(e, \partial \operatorname{co} E_{n,i}) \geq \epsilon$ for each $e \in W$. To see this, for each n and each $i \in I_n$ define $\nu_{n,i} \in \mathcal{M}$ by setting $\nu_{n,i} = (1/l_n)\sum_{j\in I_n\setminus\{i\}} \delta_{G_n(j)}$. Note that $E(\nu_{n,i}) = \operatorname{co} E_{n,i}$ for each n and each $i \in I_n$. Consider any sequence $\langle i_n \rangle$ with $i_n \in I_n$ for each n. The hypothesis that $\nu_{G_n} \to \nu$ in $\mathcal{M}$ implies that $\nu_{n,i_n} \to \nu$ in $\mathcal{M}$. By Lemma 5(b), it follows that $\operatorname{co} E_{n,i_n} \to E(\nu)$, and, therefore, by Lemma 5(a) and Lemma 4, that $W \subseteq \operatorname{int} \operatorname{co} E_{n,i}$ if n is sufficiently large. Thus, were the claim wrong, there would be a subsequence $\langle I_{n_k} \rangle$ of the sequence $\langle I_n \rangle$ such that $\operatorname{dist}(e_k, \partial \operatorname{co} E_{n_k,i_k}) \to 0$ for some points $i_k \in I_{n_k}$ and $e_k \in W$, $k \in \mathbb{N}$. Now because all the sets $E(\nu)$ and $\operatorname{co} E_{n_k,i_k}$, $k \in \mathbb{N}$, are compact, convex, and have a nonempty interior, it follows from the material in Wills (2007) that $\rho_H(\partial E(\nu), \partial \operatorname{co} E_{n_k,i_k})) \to 0$ and, therefore, that $\operatorname{dist}(e_k, \partial E(\nu)) \to 0$. Because W is compact, we can assume that $e_k \to e$ for some $e \in W$. But then, because $\partial E(\nu)$ is closed, we must have $e \in W \cap \partial E(\nu)$, contradicting the hypothesis that $W \subseteq \operatorname{int} E(\nu)$.

(b) There are numbers $\eta > 0$ and N_2, with $N_2 \geq m$ (m being the dimension of the ambient Euclidean space of the externalities universe), such that if $l_n \geq N_2$, then for each $i \in I_n$ and each $J \subseteq I_n \setminus \{i\}$ with $\#(J) = m$, the set $\sum_{j\in J} g(A_{G_n(j)})$ includes a ball of radius η. To see this, consider the correspondence $\theta \colon \operatorname{supp}(\nu) \to 2^A$, given by setting $\theta(u) = A_u$ for each $u \in \operatorname{supp}(\nu)$. Then θ is continuous (see Hildenbrand 1974, B.III, Problem 4), with nonempty compact values, all with nonempty interior. Moreover, $\rho_H(\partial A_u, \partial A_{u_k}) \to 0$ whenever $u_k \to u$ in $\operatorname{supp}(\nu)$. Using Lemma 4 and the fact that $\operatorname{supp}(\nu)$ is compact, we see from these properties that there is a nonempty finite set $\mathcal{C}$ of compact balls in A (all with radius > 0), such that for each $u \in \operatorname{supp}(\nu)$, $\operatorname{int} A_u$ includes some member of $\mathcal{C}$. As $\rho_H(\operatorname{supp}\nu, \operatorname{supp}(\nu_{\mathcal{G}_n})) \to 0$, another invocation of Lemma 4 and the fact that $\operatorname{supp}(\nu)$ is compact now show that there is an $N_2' \in \mathbb{N}$ such that whenever $n \geq N_2'$, then for each $i \in I_n$, $A_{G_n(i)}$ includes some member of $\mathcal{C}$. Let $\mathcal{F}$ be the set of all families $F = (B_{F,1}, \dots, B_{F,m})$, where $B_{F,i} = C$ for some $C \in \mathcal{C}$, $i = 1, \dots, m$. Then $\mathcal{F}$ is finite and, by Lemma 9 in the Appendix, for each $F \in \mathcal{F}$, $\sum_{i=1}^m g(B_{F,i})$ includes a ball of radius $\eta_F > 0$. Set $\eta = \min\{\eta_F : F \in \mathcal{F}\}$. As $\mathcal{F}$ is finite, $\eta > 0$. Thus, setting $N_2 = \max\{N_2', m\}$, the claim follows.

(c) Because the actions universe is compact and g is continuous, there is a number $\delta' > 0$ such that $\operatorname{diam}(g(A_{G_n(i)})) \leq \delta'$ for all $i \in I_n$ and n. Set $\delta_0 = m\delta'$, so that we have $\operatorname{diam}(g(A_{G_n(i)})) \leq \delta_0$ for all $i \in I_n$ and n as well as $\operatorname{diam}(\sum_{j\in J} g(A_{G_n(j)})) \leq \delta_0$ if J is as above. Choose a number $h \in \mathbb{N}$ with $h\eta \geq m\delta_0$. Note that for all n and $i \in I_n$,

$$l_n E_{n,i} = \sum_{j \in I_n \setminus \{i\}} g(A_{G_n(j)}),$$

and that if $l_n \geq hm$, the latter sum can be written in the form

$$\sum_{j\in J_1} g(A_{G_n(j)}) + \cdots + \sum_{j \in J_h} g(A_{G_n(j)}) + \sum_{j \notin \bigcup_{k=1}^h J_k} g(A_{G_n(j)}),$$

where the J_ks, $k = 1, \dots, h$, are pairwise disjoint and $\#(J_k) = m$ for all $k = 1, \dots, h$. By Howe (1979, Proposition 2), it follows that if $n \geq N_2$ is such that $l_n \geq hm$ and $z \in \operatorname{co}(l_n E_{n,i})$ is such that $\operatorname{dist}(z, \partial \operatorname{co}(l_n E_{n,i})) \geq h\delta_0$, then $z \in l_n E_{n,i}$.[14]

Note that (a) implies that whenever $n \geq N_1$ and $i \in I_n$, then $l_n W \subseteq \operatorname{co}(l_n E_{n,i})$ and for each $e \in l_n W$, we have $\operatorname{dist}(e, \partial \operatorname{co}(l_n E_{n,i})) \geq l_n \epsilon$. By the previous paragraph, it follows that if $n \geq \max\{N_1, N_2\}$ is such that both $l_n \geq hm$ and $l_n \epsilon \geq h\delta_0$, then $l_n W \subseteq l_n E_{n,i}$ for each $i \in I_n$, and, thus, $W \subseteq E_{n,i}$ for each $i \in I_n$. As $n \to \infty$ implies $l_n \to \infty$, this establishes the lemma. □

Lemma 7. *Given $\nu \in \mathcal{G}$, there is a sequence $\langle (I_n, G_n)\rangle_{n\in\mathbb{N}}$ of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$ in $\mathcal{M}$.*

Proof. (a) Let $\nu \in \mathcal{G}$ be given. Write $X = \operatorname{supp}(\nu)$. By the law of large numbers (Glivenko–Cantelli version), there is a sequence $\langle u_n \rangle$ in X such that the sequence $\langle \nu_n \rangle$, defined by setting $\nu_n = 1/(n+1)\sum_{i=0}^n \delta_{u_i}$ for each $n \in \mathbb{N}$, converges to ν narrowly. Since $\operatorname{supp}(\nu_n) \subseteq X = \operatorname{supp}(\nu)$, narrow convergence of $\langle \nu_n \rangle$ to ν implies that we also have $\rho_H(\operatorname{supp}(\nu_n), \operatorname{supp}(\nu)) \to 0$. Thus, we have $\nu_n \to \nu$ in the topology of $\mathcal{M}$. For each $n \in \mathbb{N} \setminus \{0\}$ and each $0 \leq i \leq n$, define $\nu_{n,i} \in \mathcal{M}$ by setting $\nu_{n,i} = 1/n \sum_{j \in J_{n,i}} \delta_{u_j}$ where

[14] Note that the h here corresponds to the m in Howe (1979, Proposition 2), while $\dim V$ and ν there are what is called m and η here, respectively.

$J_{n,i} = \{0, \dots, n\} \setminus \{i\}$; set $\nu_{0,0} = \nu_0$. Note that for each $n \in \mathbb{N} \setminus \{0\}$ and each $0 \le i \le n$, we have $\|\nu_n - \nu_{n,i}\|_V \le 2/n$, writing $\|\cdot\|_V$ for the variation norm on $\mathcal{M}$. Consequently, because $\nu_n \to \nu$, we have $\nu_{n,i_n} \to \nu$ as $n \to \infty$ whenever $\langle i_n \rangle$ is a sequence in $\mathbb{N}$ with $0 \le i_n \le n$ for each n. By Lemma 5, it follows that $\rho_H(E(\nu), E(\nu_{n,i_n})) \to 0$ whenever $\langle i_n \rangle$ is as in the previous sentence.

Fix $b \in \operatorname{int} E(\nu)$. By Lemma 3, it follows from the conclusion of the previous paragraph that $b \in \operatorname{int} E(\nu_{n,i})$ for each $0 \le i \le n$ if n is large enough; we can assume that this is true for all n. Now, for each n, we can define $r_n > 0$ to be the largest real number $r \le 1$ such that $r(E(\nu_{n,i}) - \{b\}) + \{b\} \subseteq E(\nu)$ for each $0 \le i \le n$. We must have $r_n \to 1$. To see this, fix $0 < r < 1$. Since $E(\nu)$ is convex and $b \in \operatorname{int} E(\nu)$, we have $r(E(\nu) - \{b\}) + \{b\} \subseteq \operatorname{int} E(\nu)$. Using the fact that $\rho_H(E(\nu), E(\nu_{n,i_n})) \to 0$ whenever $\langle i_n \rangle$ is a sequence in $\mathbb{N}$ with $0 \le i_n \le n$ for each n, it follows that if n is large, then $r(E(\nu_{n,i}) - \{b\}) + \{b\} \subseteq \operatorname{int} E(\nu)$ for all $0 \le i \le n$. Thus, $r_n \ge r$ for such n. As $0 < r < 1$ is arbitrary, we conclude that $r_n \to 1$.

(b) For each n and each $i = 0, \dots, n$, define a map $\tilde{u}_{n,i} \colon A_{u_i} \times E(\nu_{n,i_n}) \to \mathbb{R}$ by setting $\tilde{u}_{n,i}(a, e) = u_i(a, r_n(e - b) + b)$ for $(a, e) \in A_{u_i} \times E(\nu_{n,i})$ and note that $\tilde{u}_{n,i} \in \mathcal{U}$. We claim that for any $\epsilon > 0$, there is an n_ϵ such that whenever $n > n_\epsilon$, then $\rho(\tilde{u}_{n,i}, u_i) < \epsilon$ for all $i = 0, \dots, n$ (where ρ is the metric on $\mathcal{U}$ chosen in Section 3.4). Indeed, otherwise there are points $\tilde{u}_{n_k,i_k}$ and u_{i_k}, $k \in \mathbb{N}$, such that $n_k \to \infty$ as $k \to \infty$ and $\rho(\tilde{u}_{n_k,i_k}, u_{i_k}) \ge \epsilon > 0$ for each k. Because $u_{i_k} \in \operatorname{supp}(\nu)$ and $\operatorname{supp}(\nu)$ is compact, we can assume that $u_{i_k} \to \bar{u}$ for some $\bar{u} \in \operatorname{supp}(\nu)$. But then, using Lemma 1 together with the facts that $\rho_H(E(\nu), E(\nu_{n_k,i_{n_k}})) \to 0$ and $r_{n_k} \to 1$, it follows that also $\tilde{u}_{n_k,i_k} \to \bar{u}$, and we get a contradiction.

(c) For each $n \in \mathbb{N} \setminus \{0\}$ and each $0 \le i \le n$, set $E_{n,i} = 1/n \sum_{j \in J_{n,i}} g(A_j)$. Note that $E(\nu_{n,i}) = 1/n \sum_{j \in J_{n,i}} \operatorname{co} g(A_j)$ (cf. the proof of Lemma 5); thus, $E_{n,i} \subseteq E(\nu_{n,i})$. Let $u_{n,i} \colon A_{u_{n,i}} \times E_{n,i} \to \mathbb{R}$ be the restriction of $\tilde{u}_{n,i}$ to $A_{u_{n,i}} \times E_{n,i}$. By Lemma 9(a) and the hypothesis on g made in Section 3.2, there is an $\bar{n} \in \mathbb{N}$ such that if $n \ge \bar{n}$, then for all $0 \le i \le n$, $\operatorname{int} E_{n,i}$ is dense in $E_{n,i}$ and, thus, $u_{n,i} \in \mathcal{U}$. Because all the sets $E_{n,i}$ are included in the compact convex externalities universe E, it follows from the Shapley–Folkman theorem that for each $\epsilon > 0$, there is a $n'_\epsilon \in \mathbb{N}$ such that $\rho_H(E_{n,i}, E(\nu_{n,i})) < \epsilon$ for all $0 \le i \le n$ if $n \ge n'_\epsilon$. Using Lemma 1, it follows from this and (b) that for each $\epsilon > 0$, there is an $n''_\epsilon \in \mathbb{N}$ such that $\rho(u_{n,i}, u_i) < \epsilon$ for all $0 \le i \le n$ whenever $n \ge \max\{\bar{n}, n''_\epsilon\}$, because $u_{n,i}$ is just the restriction of $\tilde{u}_{n,i}$ to $A_{u_{n,i}} \times E_{n,i}$.

Now, for each $n \in \mathbb{N}$ with $n \ge \bar{n}$, set $I_n = \{0, 1, \dots, n\}$ and define $G_n \colon I_n \to \mathcal{U}$ by setting $G(i) = u_{n,i}$ for each $i \in I_n$. For $n < \bar{n}$, let (I_n, G_n) be an arbitrary finite-player game. By (a), $\nu_n = 1/(n+1) \sum_{i=0}^n \delta_{u_i} \to \nu$ narrowly and $\rho_H(\operatorname{supp}(\nu_n), \operatorname{supp}(\nu)) \to 0$, so from the end of the previous paragraph, we see that $\rho_H(\operatorname{supp}(\nu_{G_n}), \operatorname{supp}(\nu)) \to 0$ and that $\int h \, d\nu_{G_n} \to \int h \, d\nu$ whenever $h \colon \mathcal{U} \to \mathbb{R}$ is a bounded uniformly continuous function. By Billingsley (1968, Theorem 2.1), the latter fact implies that $\nu_{G_n} \to \nu$ narrowly. We conclude that $\nu_{G_n} \to \nu$ in the topology of $\mathcal{M}$. □

Remark 3. Inspecting the proof of Lemma 7 shows that, given any $\nu \in \mathcal{G}$, the sequence $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ as guaranteed by Lemma 7 can be chosen so that for each n, $u \in \operatorname{supp}(\nu_{\mathcal{G}_n})$ implies $A_u = A_{u'}$ for some $u' \in \operatorname{supp}(\nu)$. This observation is useful in regard to the proof of Theorem 3.

5.2 *Proofs of the theorems*

Proof of Theorem 1. (a) As noted in Section 3.6, every continuum game $\nu \in \mathcal{G}$ has an equilibrium distribution, i.e., there is a Borel probability measure τ on $\mathcal{U} \times A$ such that $\tau_{\mathcal{U}} = \nu$ and $\operatorname{supp}(\tau) \subseteq \{(u, a) \in \mathcal{U} \times A\colon a \in \varphi(u, e(\tau_A))\}$. By (U3) in the assumptions on payoff functions, we have $\varphi(u, e) \subseteq \operatorname{int} A_u$ for each $u \in \mathcal{U}$ and each $e \in E(\nu)$. Thus if τ is an equilibrium distribution of $\nu \in \mathcal{G}$, then $a \in \operatorname{int} A_u$ for each $(u, a) \in \operatorname{supp}(\tau)$.

(b) Write $\mathcal{G}_1$ for the subset of $\mathcal{G}$ consisting of those ν such that for some equilibrium distribution τ of ν, the following statements hold:

(i) We have $\#(\varphi(u, e(\tau_A))) = 1$ for each $u \in \operatorname{supp}(\nu)$.

(ii) For each $u \in \operatorname{supp}(\nu)$, $D_a^2 u(a_u, e(\tau_A))$ is negative definite, where a_u is the unique element of $\varphi(u, e(\tau_A))$ (note that $a_u \in \operatorname{int} A_u$, so $D_a^2 u(a_u, e(\tau_A))$ is defined).

(iii) We have $e(\tau_A) \in \operatorname{int} E(\nu)$.

Write $\mathcal{U}_c = \{u \in \mathcal{U}\colon E_u \text{ is convex}\}$. We claim that given any $\nu \in \mathcal{G}_1$ and any equilibrium distribution τ of ν such that (i)–(iii) are true, there are open neighborhoods $\tilde{V}$ of $\operatorname{supp}(\nu)$ in $\mathcal{U}$ and W of $e(\tau_A)$ in $\mathbb{R}^m$, with $\operatorname{cl} W \subseteq \operatorname{int} E(\nu)$, such that, setting $\hat{V} = \{u \in \tilde{V}\colon W \subseteq E_u\}$ and $V = \tilde{V} \cap \mathcal{U}_c$, the following relationships hold: $V \subseteq \hat{V}$ and on $\hat{V} \times W$, the best replies of u against e are given by a continuous map $h\colon \hat{V} \times W \to A$ such that (c1) $h(u, \cdot)$ is differentiable for each $u \in \hat{V}$, (c2) the derivative of $h(u, \cdot)$ depends continuously on (u, e), and (c3) $D_a^2 u(h(u, e), e)$ is negative definite for each $(u, e) \in \hat{V} \times W$.

To see that this claim is true, choose a compact neighborhood W_1 of $e(\tau_A)$ such that $W_1 \subseteq \operatorname{int} E(\nu)$, which is possible by (iii). Then by compactness of $\operatorname{supp}(\nu)$, Lemma 1, and Lemma 4, there is a (relatively) open neighborhood V_1 of $\operatorname{supp}(\nu)$ in $\mathcal{U}_c$ such that $W_1 \subseteq \operatorname{int} E_u$ for all $u \in V_1$. Now pick any $u \in \operatorname{supp}(\nu)$. As above, let $a_u \in \operatorname{int} A_u$ be the unique element of $\varphi(u, e(\tau_A))$. Then there is a compact and convex neighborhood U_{a_u} of a_u in $\operatorname{int} A_u$ such that $D_a^2 u(a, e(\tau_A))$ is negative definite for every $a \in U_{a_u}$. Now we can find numbers r_1 and r_2 such that $u(a_u, e(\tau_A)) > r_1 > r_2 > u(a, e(\tau_A))$ for each $a \in A_u \setminus U_{a_u}$. In particular, we must have $r_1 > u(a, e(\tau_A))$ for all $a \in \operatorname{cl}(A_u \setminus U_{a_u})$. Using Lemma 1 and Lemma 4, we see that there is a neighborhood $\tilde{V}_u'$ of u in $\mathcal{U}$ such that $U_{a_u} \subseteq \operatorname{int} A_{u'}$ for each $u' \in \tilde{V}_u'$. Because the actions universe A is compact, Lemma 1 now shows that there are open neighborhoods $\tilde{V}_u$ of u in $\mathcal{U}$ and W_u of $e(\tau_A)$ in $\mathbb{R}^m$, with $\tilde{V}_u \cap \mathcal{U}_c \subseteq \tilde{V}_u' \cap V_1$ and $W_u \subseteq W_1$, such that, setting $\hat{V}_u = \{u' \in \tilde{V}_u\colon W_u \subseteq E_{u'}\}$, $u'(a_u, e) > r_1 > u'(a, e)$ for each $u' \in \hat{V}_u$, $e \in W_u$, and $a \in A_{u'} \setminus U_{a_u}$, and such that $D_a^2 u'(a, e)$ is negative definite for each $u' \in \hat{V}_u$, $e \in W_u$, and $a \in U_{a_u}$. In particular, for each $u' \in \hat{V}_u$ and $e \in W_u$, $u'(\cdot, e)$ is strictly concave on U_{a_u}. Consequently, for each $u' \in \hat{V}_u$ and $e \in W_u$, the best reply of u' against e is unique. Apply this argument to each $u \in \operatorname{supp}(\nu)$. Then by compactness of $\operatorname{supp}(\nu)$, there are $u_1, \ldots, u_k \in \operatorname{supp}(\nu)$ such that $\operatorname{supp}(\nu) \subseteq \tilde{V} = \bigcup_{i=1}^k \tilde{V}_{u_i}$. Set $W = \bigcap_{i=1}^k W_{u_i}$ and $K = \bigcup_{i=1}^k U_{a_i}$. Then for V and $\hat{V}$ (defined relative to the sets $\tilde{V}$ and W according to the previous paragraph), we have $V \subseteq \hat{V}$, and for each $(u, e) \in \hat{V} \times W$, the best reply of u against e is unique and belongs to K. Thus, on $\hat{V} \times W$, the best reply correspondence φ can be identified with a function h taking values in K. Using the fact that K is compact, we see that h is continuous. Note that by construction, $D_a^2 u(h(u, e), e)$ is negative definite for each $(u, e) \in \hat{V} \times W$, i.e., we have (c3). In view of this, the implicit function theorem applied to the maps $(a, e) \mapsto D_a u(a, e)$, $u \in \hat{V}$, shows that (c1) is true. Using Lemma 1 we see that the evaluation maps $(u, a, e) \mapsto D_a^2 u(a, e)$ and $(u, a, e) \mapsto D_e D_a u(a, e)$, which are defined on the set $\{(u, a, e) \in \mathcal{U} \times A \times E\colon a \in \operatorname{int} A_u,\ e \in E_u\}$, are continuous. From this we see that (c2) is true.

Let $\nu \in \mathcal{G}_1$ and let τ be an equilibrium distribution for ν such that (i)–(iii) are satisfied. Let W correspond to τ as above. We can then define a map $\xi_\tau\colon W \to \mathbb{R}^m$ by setting

$$\xi_\tau(e) = \int g\big(h(u, e)\big)\, \mathrm{d}\nu(u) - e$$

for each $e \in W$. Then by the generalized version of Leibniz' rule in Schwartz (1967, Chap IV.11, Theorem 115), ξ_τ is continuously differentiable on W, and we have $D\xi_\tau(e) = \int D_e(g \circ h)(u, e)\, \mathrm{d}\nu(u) - I$, where I is the $(m \times m)$ identity matrix.

(c) Let $\mathcal{G}^*$ be the subset of $\mathcal{G}$ consisting of those $\nu \in \mathcal{G}$ such that for some equilibrium distribution τ of ν, (i)–(iii) of (b) are satisfied and $D\xi_\tau(e(\tau_A))$ has full rank, where ξ_τ is associated with τ as above. (Note that while the choice of the neighborhood W of $e(\tau_A)$, i.e., the domain of ξ_τ, involves some arbitrariness, $D\xi_\tau(e(\tau_A))$ is uniquely determined.) By Lemma 8 below, $\mathcal{G}^*$ is dense in $\mathcal{G}$, and we are now going to show that $\mathcal{G}^*$ is open in $\mathcal{G}$.

(d) Fix $\nu \in \mathcal{G}^*$. We need to show that ν has a neighborhood U in $\mathcal{G}$ such that U is included in $\mathcal{G}^*$. Let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}^*$. Let W, V, h, and ξ_τ be associated with τ as in (b).

(i) Pick a compact neighborhood W_1 of $e(\tau_A)$ with $W_1 \subseteq W$. Then there is a $k \in \mathbb{N}$ and a neighborhood V_1 of $\operatorname{supp}(\nu)$ in $\mathcal{U}_c$, with $V_1 \subseteq V$, such that $\|D_e(g \circ h)(u, e)\| \leq k$ for each $(u, e) \in V_1 \times W_1$. Indeed, otherwise, for each $k \in \mathbb{N} \setminus \{0\}$, we can find points $e_k \in W_1$ and $u_k \in V$ such that $\|D_e(g \circ h)(u_k, e_k)\| > k$ but $\operatorname{dist}(u_k, \operatorname{supp}(\nu)) < 1/k$. Since W_1 and $\operatorname{supp}(\nu)$ are compact, we may assume that $(u_k, e_k) \to (u, e)$ for some $(u, e) \in \operatorname{supp}(\nu) \times W_1$. Now $D_e(g \circ h)(u_k, e_k) = Dg(h(u_k, e_k)) D_e h(u_k, e_k)$, and because Dg, h, and $D_e h$ are continuous, it follows that $D_e(g \circ h)(u_k, e_k) \to D_e(g \circ h)(u, e)$, and we get a contradiction.

(ii) Write W_2 for the interior of W_1 in $\mathbb{R}^m$. Choose an open neighborhood U_1 of ν in $\mathcal{G}$ such that $\operatorname{supp}(\nu') \subseteq V_1$ for each $\nu' \in U_1$. Note that $W_2 \subseteq E(\nu')$ for each $\nu' \in U_1$. We can, therefore, define a map $\xi_U\colon U_1 \times W_2 \to \mathbb{R}^m$ by setting

$$\xi_U\big(\nu', e\big) = \int g\big(h(u, e)\big)\, \mathrm{d}\nu'(u) - e$$

for each $\nu' \in U_1$ and $e \in W_2$. As above, we see that for each fixed $\nu' \in U_1$, $\xi_U(\nu', \cdot)$ is continuously differentiable on W_2, with $D_e \xi_U(\nu', e) = \int D_e(g \circ h)(u, e)\, \mathrm{d}\nu'(u) - I$, where I is the $(m \times m)$ unit matrix. Now ξ_U is continuous and $D_e \xi_U(\nu', e)$ depends continuously on (ν', e). Indeed, suppose that $e_k \to e$ in W_2 and $u_k \to u$ in V_1. Then $(g \circ h)(u_k, e_k) \to (g \circ h)(u, e)$, because h and g are continuous, and as in (i), we see that $D_e(g \circ h)(u_k, e_k) \to D_e(g \circ h)(u, e)$. Thus, uniformly on compact subsets of V_1, we have both $(g \circ h)(\cdot, e_k) \to (g \circ h)(\cdot, e)$ and $D_e(g \circ h)(\cdot, e_k) \to D_e(g \circ h)(\cdot, e)$. Using Billingsley (1968, Theorem 5.5), it follows that if $\nu_k \to \nu'$ in U_1, then the corresponding sequences of distributions of the maps $(g \circ h)(\cdot, e_k)$ and $D_e(g \circ h)(\cdot, e_k)$ converge narrowly to the distributions of $(g \circ h)(\cdot, e)$ and $D_e(g \circ h)(\cdot, e)$, respectively. As $g \circ h$ takes values in the compact set E, we can now use change of variables to see that $\xi_U(\nu_k, e_k) \to \xi_U(\nu', e)$. Similarly, by (c1), we see that $D_e\xi_U(\nu_k, e_k) \to D_e\xi_U(\nu', e)$. Thus, on $U_1 \times W_2$, ξ_U is continuous and $D_e\xi_U(\nu', e)$ depends continuously on (ν', e), as claimed.

Now as τ is an equilibrium distribution for ν, we have $\xi_U(\nu, e(\tau_A)) = 0$, and since $\nu \in \mathcal{G}^*$, $D_e\xi_U(\nu, e(\tau_A)) \equiv D\xi_\tau(e(\tau_A))$ has full rank. Hence, by a version of the implicit function theorem (see Schwartz, 1967, Chap. III.8, Theorem 25, or Mas-Colell, 1985, Chapter 1, C.3.3), there is an open neighborhood U of ν in $\mathcal{G}$, with $U \subseteq U_1$, and a continuous map $\nu' \mapsto e(\nu')\colon U \to W_2$ such that for each $\nu' \in U$, $\xi_U(\nu', e(\nu')) = 0$. Also, since $D_e\xi_U(\nu', e)$ depends continuously on (ν', e), $D_e\xi_U(\nu', e(\nu'))$ has full rank for each $\nu' \in U$, shrinking U if need be.

Fix any $\nu' \in U$ and set $\tau' = \nu' \circ (\mathrm{id} \times h(\cdot, e(\nu')))^{-1}$. Then

$$\operatorname{supp}(\tau') \subseteq \{(u, a) \in \mathcal{U} \times A\colon a \in \varphi(u, e(\nu'))\}$$

by the choice of h and

$$e(\tau'_A) = \int g(h(u, e(\nu')))\,\mathrm{d}\nu'(u) = \xi_U(\nu', e(\nu')) + e(\nu') = e(\nu').$$

Thus, τ' is an equilibrium distribution for ν'. By the choices of V and W, and since $e(\nu'') \in W_2 \subseteq W$ for each $\nu'' \in U$, (i)–(iii) of (b) are true for τ'. Let V', W', h', and $\xi_{\tau'}$ be associated with τ' as in (b). Then $V' \cap V_1$ is a neighborhood of $\operatorname{supp}(\nu')$ in $\mathcal{U}_c$ and $W' \cap W_2$ is a neighborhood of $e(\tau'_A) = e(\nu')$. Moreover, h and h' agree on $(V' \cap V_1) \times (W' \cap W_2)$, and, hence, so do $\xi_{\tau'}$ and $\xi_U(\nu', \cdot)$. Thus, $D\xi_{\tau'}(e(\tau'_A))$ has maximal rank. It follows that every $\nu' \in U$ belongs to $\mathcal{G}^*$. As $\nu \in \mathcal{G}^*$ is arbitrary, $\mathcal{G}^*$ is open.

(e) Let $\nu \in \mathcal{G}^*$ and let $\langle(I_n, G_n)\rangle_{n\in\mathbb{N}}$ be a sequence of finite-player games such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$ in $\mathcal{M}$. For each n, we can write $I_n = \{1, \dots, k_n\}$, where $k_n = \#(I_n)$. Let $\hat{A}$ be the convex hull of the actions universe A and identify g with a continuous extension to $\hat{A}$. For any map $f\colon I_n \to \hat{A}$ and any $i \in I_n$, we write $\tau_{\hat{A},f}$ for the probability measure on $\hat{A}$ given by setting $\tau_{\hat{A},f}(B) = \#\{j \in I_n\colon f(j) \in B\}/\#(I_n)$ for each Borel set $B \subseteq \hat{A}$, and write $\tau_{\hat{A},f,i}$ for the probability measure on $\hat{A}$ that is given by setting $\tau_{\hat{A},f,i}(B) = \#\{j \in I_n \setminus \{i\}\colon f(j) \in B\}/(\#(I_n) - 1)$ for each Borel set $B \subseteq \hat{A}$.

Write $\|\cdot\|_V$ for the variation norm on the space $M(\hat{A})$ of all signed Borel measures on $\hat{A}$. Note that for any $n \in \mathbb{N}$ and any $f\colon I_n \to \hat{A}$, $\|\tau_{\hat{A},f,i} - \tau_{\hat{A},f}\|_V \le 2/\#(I_n)$ for each $i \in I_n$. Because g is bounded on $\hat{A}$, it follows that for any $\delta > 0$, there is an $N_\delta \in \mathbb{N}$ such that if $n \ge N_\delta$, then $\|\int g(a)\,\mathrm{d}\tau_{\hat{A},f,i}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a)\| < \delta$ for each $f\colon I_n \to \hat{A}$ and each $i \in I_n$.

Let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}^*$. Let $\tilde{V}$, $\hat{V}$, W, and h be as in the paragraph after the statement of (i)–(iii) in (b).

As $\nu \in \mathcal{G}^*$, the derivative of ξ_τ at $e(\tau_A)$ has full rank, which implies that on some convex compact neighborhood W_1 of $e(\tau_A)$ in $\mathbb{R}^m$, with $W_1 \subseteq W$, $\xi_\tau(e) = 0$ if and only if $e = e(\tau_A)$. Let W_2 be a convex compact neighborhood of $e(\tau_A)$ in $\mathbb{R}^m$ such that $W_2 \subseteq \operatorname{int} W_1$.

Now because $\nu_{G_n} \to \nu$ and, therefore, $\rho_H(\mathrm{supp}(\nu_{G_n}), \mathrm{supp}(\nu)) \to 0$, Lemma 6 implies that there is an $N \in \mathbb{N}$ such that whenever $n \geq N$, then $\mathrm{supp}(\nu_{G_n}) \subseteq \hat{V}$ (i.e., $\mathrm{supp}(\nu_{G_n}) \subseteq \tilde{V}$ and $W \subseteq E_{G_n(i)}$ for each $i \in I_n$). By the second paragraph of this part of the proof, we can assume that N is so large that if $n \geq N$, then for each $f \in \hat{A}^{I_n}$ and $i = 1, \ldots, k_n$,

$$\int g(a)\,\mathrm{d}\tau_{\hat{A},f,i}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a) + e \in W_1$$

whenever $e \in W_2$. For $n \geq N$, consider the function $\Lambda\colon \hat{A}^{I_n} \times W_2 \to \hat{A}^{I_n} \times \mathbb{R}^m$ defined by setting

$$\begin{aligned}
&\Lambda(f, e)\\
&\quad = \biggl(h_1\biggl(\int g(a)\,\mathrm{d}\tau_{\hat{A},f,1}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a) + e\biggr), \ldots,\\
&\qquad h_{k_n}\biggl(\int g(a)\,\mathrm{d}\tau_{\hat{A},f,k_n}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a) + e\biggr), \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a)\biggr),
\end{aligned}$$

writing $h_i(\cdot)$ in place of $h(G(i), \cdot)$ for each $i \in \{1, \ldots, k_n\}$. Then a fixed point of Λ gives a strict pure strategy Nash equilibrium of (I_n, G_n).

We claim that there is an $N_1 \geq N$ such that for $n \geq N_1$, the fixed-point theorem stated in the Appendix as Theorem 4 applies to Λ. Clearly Λ is continuous, and for $X = \hat{A}^{I_n}$ and $Y = W_2$, the requirements of Theorem 4 on X and Y are satisfied. With the map ξ_τ, it is also clear that we have (a) of Theorem 4. Let $\gamma > 0$ be such that $\|\xi_\tau(e)\| \geq \gamma$ for each $e \in \partial W_2$. We need to show that for some $N_1 \geq N$ also (b) of that theorem is satisfied for Λ and ξ_τ if $n \geq N_1$.

To this end, fix $n \geq N$ and suppose that $f \in \hat{A}^{I_n}$ and $e \in \partial W_2$ are such that

$$\begin{aligned}
f = \biggl(&h_1\biggl(\int g(a)\,\mathrm{d}\tau_{\hat{A},f,1}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a) + e\biggr),\\
&\ldots, h_{k_n}\biggl(\int g(a)\,\mathrm{d}\tau_{\hat{A},f,k_n}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a) + e\biggr)\biggr).
\end{aligned}$$

Note that

$$\begin{aligned}
\biggl\|\frac{1}{k_n}\sum_{i=1}^{k_n} g\bigl(f(i)\bigr) - e - \xi_\tau(e)\biggr\| &= \biggl\|\frac{1}{k_n}\sum_{i=1}^{k_n} g\bigl(f(i)\bigr) - e - \biggl(\int g\bigl(h(u,e)\bigr)\,\mathrm{d}\nu(u) - e\biggr)\biggr\|\\
&= \biggl\|\frac{1}{n_k}\sum_{i=1}^{n_k} g\bigl(f(i)\bigr) - \int g\bigl(h(u,e)\bigr)\,\mathrm{d}\nu(u)\biggr\|\\
&\leq \biggl\|\frac{1}{k_n}\sum_{i=1}^{k_n} g\bigl(f(i)\bigr) - \frac{1}{k_n}\sum_{i=1}^{k_n} g\bigl(h_i(e)\bigr)\biggr\|\\
&\qquad + \biggl\|\frac{1}{k_n}\sum_{i=1}^{k_n} g\bigl(h_i(e)\bigr) - \int g\bigl(h(u,e)\bigr)\,\mathrm{d}\nu(u)\biggr\|
\end{aligned}$$

$$= \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g\big(f(i)\big) - \frac{1}{k_n} \sum_{i=1}^{k_n} g\big(h_i(e)\big) \right\| + \left\| \int g\big(h(u,e)\big)\,\mathrm{d}\nu_{G_n}(u) - \int g\big(h(u,e)\big)\,\mathrm{d}\nu(u) \right\|.$$

Since $\operatorname{supp}(\nu)$ is compact, there is a neighborhood $\tilde{V}_1$ of $\operatorname{supp}(\nu)$ in $\mathcal{U}$, with $\tilde{V}_1 \subseteq \tilde{V}$, and a $\delta > 0$ such that for $z \in \mathbb{R}^m$ with $e + z \in W$, $\|g(h(u, e + z)) - g(h(u, e))\| < \gamma/3$ whenever $u \in \tilde{V}_1 \cap \hat{V}$, $e \in W_1$, and $\|z\| < \delta$, and as noted earlier, if n is large, then $\|\int g(a)\,\mathrm{d}\tau_{\hat{A},f,i}(a) - \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a)\| < \delta$ for each $f \in \hat{A}^{I_n}$ and each $i \in I_n$. Hence, as $\rho_H(\operatorname{supp}(\nu_{G_n}), \operatorname{supp}(\nu)) \to 0$, we have

$$\left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g\big(f(i)\big) - \frac{1}{k_n} \sum_{i=1}^{k_n} g\big(h_i(e)\big) \right\| < \gamma/3$$

for large n whenever $f \in \hat{A}^{I_n}$ and $e \in \partial W_2$ are as above. Also, combining the first part of the penultimate sentence with the fact that $\nu_{G_n} \to \nu$ narrowly, we see that each $e \in \partial W_2$ has a neighborhood U in ∂W_2 such that for large n, we have $\|\int g(h(u, e'))\,\mathrm{d}\nu_{G_n}(u) - \int g(h(u, e'))\,\mathrm{d}\nu(u)\| < (2\gamma)/3$ for each $e' \in U$. Thus, since ∂W_2 is compact, if n is large, then $\|\int g(h(u, e))\,\mathrm{d}\nu_{G_n}(u) - \int g(h(u, e))\,\mathrm{d}\nu(u)\| < (2\gamma)/3$ for every $e \in \partial W_2$. It follows that for some $N_1 \geq N$,

$$\left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g\big(f(i)\big) - e - \xi_\tau(e) \right\| < \gamma$$

for $n \geq N_1$ whenever $f \in \hat{A}^{I_n}$ and $e \in \partial W_2$ are as above. Consequently, because

$$\frac{1}{k_n} \sum_{i=1}^{k_n} g\big(f(i)\big) = \int g(a)\,\mathrm{d}\tau_{\hat{A},f}(a),$$

(b) of Theorem 4 is satisfied if $n \geq N_1$.

We can conclude that for $n \geq N_1$, Λ has a fixed point and, thus, (I_n, G_n) has a strict pure strategy Nash equilibrium. □

Lemma 8. *Let $\mathcal{G}^*$ be defined as in the proof of Theorem 1. Let $\nu \in \mathcal{G}$ and let τ be an equilibrium distribution for ν. Then there is a sequence $\langle \nu_k \rangle$ of elements of $\mathcal{G}^*$ and a sequence $\langle \tau_k \rangle$ of corresponding equilibrium distributions such that $\nu_k \to \nu$ in the topology of $\mathcal{G}$, $\tau_k \to \tau$ narrowly, and for each k, τ_k witnesses that $\nu_k \in \mathcal{G}^*$.*

Proof. In the sequel, for elements ν and ν_k, $k \in \mathbb{N}$, in $\mathcal{G}$, we write $\nu_k \to \nu$ to mean convergence of the sequence $\langle \nu_k \rangle$ to ν in the topology of $\mathcal{G}$; for equilibrium distributions τ and τ_k, $k \in \mathbb{N}$, we write $\tau_k \to \tau$ to mean convergence of the sequence $\langle \tau_k \rangle$ to τ in the narrow topology. Given $\nu \in \mathcal{G}$, we often write Y for $\operatorname{supp}(\nu)$. The set $\mathcal{G}_1 \subseteq \mathcal{G}$ is defined as in the proof of Theorem 1. We write $\mathcal{G}_2$ for the set of elements of $\mathcal{G}$ that have an equilibrium distribution satisfying (i) and (ii) in the definition of $\mathcal{G}_1$, and write $\mathcal{G}_3$ for the subset of $\mathcal{G}_2$ consisting of the elements ν of $\mathcal{G}_2$ such that for some $\bar{u} \in \operatorname{supp}(\nu)$, there is a decreasing sequence $\langle W_l \rangle$ of compact subsets of Y, with $A_u = A_{\bar{u}}$ for each $u \in W_0$, such that $\bigcap_{l=0}^{\infty} W_l = \{\bar{u}\}$, and $\nu(W_l) > 0$ for each l. The proof of the lemma is organized in the following steps.

(a) If $\nu \in \mathcal{G}$ and τ is an equilibrium distribution for ν, there is a sequence $\langle \nu_k \rangle$ in $\mathcal{G}_2$ and a sequence $\langle \tau_k \rangle$ of equilibrium distributions for the ν_ks such that $\nu_k \to \nu$, $\tau_k \to \tau$, and for each k, τ_k witnesses that $\nu_k \in \mathcal{G}_2$.

(b) If $\nu \in \mathcal{G}_2$ and τ is an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_2$, then there is a sequence $\langle \nu_k \rangle$ in $\mathcal{G}_3$ and a sequence $\langle \tau_k \rangle$ of equilibrium distributions for the ν_ks such that $\nu_k \to \nu$, $\tau_k \to \tau$, and for each k, τ_k witnesses that $\nu_k \in \mathcal{G}_3$.

(c) If $\nu \in \mathcal{G}_3$ and τ is an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_2$, then there is a sequence $\langle \nu_k \rangle$ in $\mathcal{G}_1$ and a sequence $\langle \tau_k \rangle$ of equilibrium distributions for the ν_ks such that $\nu_k \to \nu$, $\tau_k \to \tau$, and for each k, τ_k witnesses that $\nu_k \in \mathcal{G}_1$.

(d) If $\nu \in \mathcal{G}_1$ and τ is an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_1$, then there is a sequence $\langle \nu_k \rangle$ in $\mathcal{G}^*$ and a sequence $\langle \tau_k \rangle$ of equilibrium distributions for the ν_ks such that $\nu_k \to \nu$, $\tau_k \to \tau$, and for each k, τ_k witnesses that $\nu_k \in \mathcal{G}^*$.

Putting (a)–(d) together, proves the lemma.

Step (a). Let $\nu \in \mathcal{G}$ and τ any equilibrium distribution for ν. Choose a twice continuously differentiable map $\rho\colon \mathbb{R}^n \to \mathbb{R}$ such that $\rho(0) = 0 > \rho(x)$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and $D^2\rho(0)$ is negative definite. For each $k \in \mathbb{N}$, $a \in A$, and $u \in Y$, define a function $u_{k,a}\colon A_u \times E(\nu) \to \mathbb{R}$ by setting

$$u_{k,a}(a', e) = u(a', e) + \tfrac{1}{k+1}\bigl(\rho(a - a')\bigr)$$

for each $(a', e) \in A_u \times E(\nu)$. Evidently (U1) and (U2) in the definition of payoff functions are satisfied for each $u_{k,a}$. Using Lemmata 1 and 3, we see from compactness of Y, A, and $E(\nu)$ that there is a $k_0 \in \mathbb{N}$ such that also (U3) in the definition of payoff functions is satisfied whenever $k \geq k_0$. We can assume that $k_0 = 0$.

Now, for each k, define a map $\lambda_k\colon Y \times A \to \mathcal{U}$ by setting $\lambda_k(u, a) = u_{k,a}$ for each $u \in Y$ and $a \in A$. Write proj_Y for the projection of $Y \times A$ onto Y. Using Lemma 1, we see that λ_k is continuous for each k and that the sequence $\langle \lambda_k \rangle$ converges uniformly to proj_Y as $k \to \infty$. Because A and Y are compact, and λ_k is continuous for each k, we can define an element ν_k in $\mathcal{M}$ for each k by setting $\nu_k = \tau \circ \lambda_k^{-1}$. Evidently $E(\nu_k) = E(\nu)$ for each k, so each ν_k actually belongs to $\mathcal{G}$. The fact that $\langle \lambda_k \rangle$ converges uniformly to proj_Y implies that $\nu_k \to \nu$, because $\tau_{\mathcal{U}} = \nu$ by the fact that τ is an equilibrium distribution for ν.

Write proj_A for the projection of $\mathcal{U} \times A$ onto A. For each k, define a continuous map $\kappa_k\colon Y \times A \to \mathcal{U} \times A$ by setting $\kappa_k = \lambda_k \times \operatorname{proj}_A$ and set $\tau_k = \tau \circ \kappa_k^{-1}$. Then $\tau_{k,\mathcal{U}} = \nu_k$ and $\tau_{k,A} = \tau_A$ for each k. As $\langle \lambda_k \rangle$ converges uniformly to proj_Y, $\langle \kappa_k \rangle$ converges uniformly to the identity on $Y \times A$, and, thus, we have $\tau_k \to \tau$. Fix any k. Note that because $Y \times A$ is compact and κ_k is continuous, $\operatorname{supp}(\tau_k) = \{(\kappa_k(u, a)\colon (u, a) \in Y \times A\}$; that is, $\operatorname{supp}(\tau_k) = \{(u_{k,a}, a)\colon (u, a) \in Y \times A\}$. Thus if $(u', a') \in \operatorname{supp}(\tau_k)$, then for some $(u, a) \in \operatorname{supp}(\tau)$, $u' = u_{k,a}$ and $a' = a$. But if $(u, a) \in \operatorname{supp}(\tau)$, then $a \in \varphi(u, e(\tau_A))$ since τ is an equilibrium distribution; in particular, $D_a^2 u(a, e(\tau_A))$ is negative semidefinite. Consequently, by the choice of the functions $u_{k,a}$, if $(u', a') \in \operatorname{supp}(\tau_k)$, then $\varphi(u', e(\tau_A)) = \{a'\}$

and $D_a^2 u'(a', e(\tau_A))$ is negative definite. As $\tau_{k,\mathcal{U}} = \nu_k$ and $\tau_A = \tau_{k,A}$, it follows that τ_k is an equilibrium distribution for ν_k and that $\nu_k \in \mathcal{G}_2$.

Step (b). Let $\nu \in \mathcal{G}_2$ and let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_2$. If there is a $u \in Y$ with $\nu(\{u\}) > 0$, we can simply set $\nu_k = \nu$ and $\tau_k = \tau$ for each κ. Suppose that $\nu(\{u\}) = 0$ for each $u \in Y$. Pick an arbitrary point $\bar{u}$ in Y. By Lemma 11 in the Appendix, there is an increasing sequence $\langle A_{\bar{u}_k} \rangle$ of nonempty compact subsets of $A_{\bar{u}}$, all with dense interior, such that $A_{\bar{u}_k} \subseteq \operatorname{int} A_{\bar{u}}$ for each k, and both $\rho_H(A_{\bar{u}}, A_{\bar{u}_k}) \to 0$ and $\rho_H(\partial A_{\bar{u}}, \partial A_{\bar{u}_k}) \to 0$. Now by Lemma 4, for each k, there is a closed neighborhood V'_k of $\bar{u}$ in $\mathcal{U}$ such that $A_{\bar{u}_k} \subseteq \operatorname{int} A_u$ for each $u \in V'_k$. For each k, set $V_k = V'_k \cap Y$. Then $\nu(V_k) > 0$ for each k, because $\bar{u} \in Y$ $(\equiv \operatorname{supp}(\nu))$. We can assume that the sequence $\langle V_k \rangle$ is decreasing and that $\bigcap_{k=0}^{\infty} V_k = \{\bar{u}\}$.

For each k and each $u \in Y$, define a map u'_k by setting $u'_k = u \restriction (A_{\bar{u}_k} \times E(\nu))$ if $u \in V_k$ and $u'_k = u$ otherwise. Then all these maps are twice continuously differentiable on their domains. As $\rho_H(A_{\bar{u}}, A_{\bar{u}_k}) \to 0$ and $\rho_H(\partial A_{\bar{u}}, \partial A_{\bar{u}_k}) \to 0$, we can see, as in Step (a), that if k is sufficiently large, then (U3) in the assumptions on payoff functions is satisfied by all the maps u'_k. We can assume that this is true for each k, so that we can define a map $\lambda'_k \colon Y \to \mathcal{U}$ for each k by setting $\lambda'_k(u) = u'_k$ for each $u \in Y$. Note that the restrictions of λ'_k to V_k and $Y \setminus V_k$ are both continuous. Thus, for each k, λ'_k is measurable, so $\nu'_k = \nu \circ \lambda'_k$ is defined. Evidently, for each k, $\operatorname{supp}(\nu'_k)$ is compact, so $\nu'_k \in \mathcal{M}$. Using Lemma 1 and the choice of the sequence $\langle A_{\bar{u}_k} \rangle$, it follows that the sequence $\langle \lambda'_k \rangle$ converges uniformly to the identity on Y, which implies that $\nu'_k \to \nu$ in the topology of $\mathcal{M}$.

Now by Lemma 5, $\rho_H(E(\nu), E(\nu'_k)) \to 0$. Moreover, $E(\nu'_k) \subseteq E(\nu)$ for each k. Indeed, fix any k. Then

$$\int_{V_k} \operatorname{co} g(A_{\bar{u}_k}) \, \mathrm{d}\nu = \nu(V_k) \operatorname{co} g(A_{\bar{u}_k}) = \nu'_k\big(\{\bar{u}'_k\}\big) \operatorname{co} g(A_{\bar{u}_k}),$$

so by the proof of Lemma 5 and the choice of ν'_k,

$$\begin{aligned} E(\nu) = \int_Y \operatorname{co} g(A_u) \, \mathrm{d}\nu(u) &\supseteq \int_{Y \setminus V_k} \operatorname{co} g(A_u) \, \mathrm{d}\nu(u) + \int_{V_k} \operatorname{co} g(A_{\bar{u}_k}) \, \mathrm{d}\nu \\ &= \int_{Y \setminus V_k} \operatorname{co} g(A_u) \, \mathrm{d}\nu'_k(u) + \nu'_k\big(\{\bar{u}'_k\}\big) \operatorname{co} g(A_{\bar{u}_k}) = E\big(\nu'_k\big). \end{aligned}$$

For each $u \in Y$, we can, therefore, define a map $u_k \colon A_u \times E(\nu'_k) \to \mathbb{R}$ by setting $u_k(a, e) = u(a, e)$ for $(a, e) \in A_u \times E(\nu'_k)$, and for each $u \in V_k$, we can define a map $\tilde{u}_k \colon A_{\bar{u}_k} \times E(\nu'_k) \to \mathbb{R}$ by setting $\tilde{u}_k(a, e) = u(a, e)$ for $(a, e) \in A_{\bar{u}_k} \times E(\nu'_k)$. For each k, define $\lambda_k \colon Y \to \mathcal{U}$ by setting $\lambda_k(u) = u_k$ for $u \in Y \setminus V_k$ and setting $\lambda_k(u) = \tilde{u}_k$ for $u \in V_k$. As with λ'_k, we see that λ_k is measurable for each k, so $\nu_k = \nu \circ \lambda_k^{-1}$ is defined. Clearly $\operatorname{supp}(\nu_k)$ is compact and $E(\nu_k) = E(\nu'_k)$ for each k, so ν_k belongs to $\mathcal{G}$. Since $\rho_H(E(\nu), E(\nu'_k)) \to 0$, we see that $\langle \lambda_k \rangle$ converges uniformly to the identity on Y, using Lemma 1 and the choice of the sequence $\langle A_{\bar{u}_k} \rangle$. Thus, $\nu_k \to \nu$.

Since $\nu \in \mathcal{G}_2$ and τ witnesses this, we see as in (i) in the proof of Theorem 1 that there is continuous map $f \colon Y \to A$ such that $\varphi(u, e(\tau_A)) = \{f(u)\}$ for all $u \in Y$; in particular, $\tau = \nu \circ (\mathrm{id}_Y \times f)^{-1}$. By (U3) in the assumptions on payoff functions, we have

$f(\bar{u}) \in \operatorname{int} A_{\bar{u}}$. Consequently, by Lemma 3 and the choice of the sequence $\langle A_{\bar{u}_k} \rangle$, we must have $f(\bar{u}) \in \operatorname{int} A_{\bar{u}_k}$ for large k, and, therefore, by the choice of the sequence $\langle V_k \rangle$, continuity of f implies that $f(u) \in A_{\bar{u}_k}$ for each $u \in V_k$ if k is large enough. We can assume that this is true for all k. Now for each k, set $\tau_k = \nu \circ (\lambda_k \times f)^{-1}$. Then $\tau_{k,\mathcal{U}} = \nu_k$ for each k. Because Y and $V_k \subseteq Y$ are compact, and the maps f, $u \mapsto u_k \colon Y \to \mathcal{U}$, and $u \mapsto \tilde{u}_k \colon V_k \to \mathcal{U}$ are continuous, the set $\{(u_k, f(u)) \colon u \in Y\} \cup \{(\tilde{u}_k, f(u)) \colon u \in V_k\}$ is closed and, therefore, includes $\operatorname{supp}(\tau_k)$. Note that $e(\tau_{k,A}) = \int g(f(u))\, d\nu(u) = e(\tau_A)$. By the hypothesis that τ is an equilibrium distribution for ν satisfying (i) and (ii) in the definition of $\mathcal{G}_1$, together with the choice of the u_ks and $\tilde{u}_k$s, it follows that τ_k is an equilibrium distribution for ν_k, also satisfying (i) and (ii) in the definition of $\mathcal{G}_1$. Clearly the sequence $\langle \lambda_k, f \rangle$ converges uniformly to $\mathrm{id}_Y \times f$, so $\tau_k \to \tau$, because $\tau = \nu \circ (\mathrm{id}_Y \times f)^{-1}$.

Finally to see that each ν_k actually belongs to $\mathcal{G}_3$, fix any k. Note first that $\lambda_k(\bar{u}) \in \operatorname{supp}(\nu_k)$. Indeed, note that λ_k is continuous on V_k and that $\bar{u}$ belongs to the relative interior of V_k in Y. Hence whenever O is an open neighborhood of $\lambda_k(\bar{u})$, there is a relatively open neighborhood U of $\bar{u}$ in Y, with $U \subseteq V_k$, such that $\lambda_k(U) \subseteq O$. Consequently, for any such O and U,

$$\nu_k(O) = \nu(\{u \in Y \colon \lambda_k(u) \in O\} \geq \nu(U) > 0,$$

since $\bar{u} \in Y$ $(\equiv \operatorname{supp}(\nu))$.

Now set $W_l = \lambda_k(V_{k+l}) \cap \operatorname{supp}(\nu_k)$ for each $l \in \mathbb{N}$, so that each W_l is a closed subset of $\operatorname{supp}(\nu_k)$, because V_{k+l} is compact and λ_k is continuous on V_{k+l}. By the choice of the sequence $\langle V_k \rangle$, we have $\bigcap_{l=0}^{\infty} V_{k+l} = \{\bar{u}\}$, therefore $\bigcap_{l=0}^{\infty} \lambda_k(V_{k+l}) = \{\lambda_k(\bar{u})\}$ since λ_k is continuous on V_k, and hence $\bigcap_{l=0}^{\infty} W_l = \{\lambda_k(\bar{u})\}$. Also, for each l, the relative interior of V_{k+l} in Y is nonempty, therefore $\nu(V_{k+l}) > 0$ since $V_{k+l} \subseteq \operatorname{supp}(\nu)$, and hence $\nu_k(\lambda_k(V_{k+l})) > 0$. Thus $\nu_k(W_l) > 0$ for each l. By construction, $A_u = A_{\lambda_k(\bar{u})}$ for each $u \in W_0$, and it follows that ν_k satisfies the requirements to be a member of $\mathcal{G}_3$.

Step (c). Let $\nu \in \mathcal{G}_3$ and let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_2$. As in (b), there is a continuous map $f \colon Y \to A$ such that $\varphi(u, e(\tau_A)) = \{f(u)\}$ for all $u \in Y$, and we have $e(\tau_A) = \int g(f(u))\, d\nu(u)$ and $\tau = \nu \circ (\mathrm{id}_Y \times f)^{-1}$.

(i) Suppose first that there is a $\bar{u} \in Y$ with $\nu(\{\bar{u}\}) > 0$. Write α for $\nu(\{\bar{u}\})$. Note that $f(\bar{u}) \in \operatorname{int} A_{\bar{u}}$. Let $\langle W_k \rangle$ be a nonincreasing sequence of compact convex neighborhoods of $f(\bar{u})$ in $A_{\bar{u}}$ such that $\bigcap_{k=1} W_k = \{f(\bar{u})\}$. By Lemma 9(b), $\operatorname{int} \operatorname{co} g(W_k)$ is nonempty for each k. For each k, fix a point $e_{\bar{u},k} \in \operatorname{int} \operatorname{co} g(W_k)$. Using Caratheodory's theorem, for each k, we can find points $a_{k,h}$, $h = 1, \ldots, m+1$, in W_k such that $e_{\bar{u},k} = \sum_{h=1}^{m+1} \beta_{k,h} g(a_{k,h})$ for some numbers $\beta_{k,h}$ with $\beta_{k,h} \geq 0$ and $\sum_{h=1}^{m+1} \beta_{k,h} = 1$. Note that $E(\nu) = \alpha \operatorname{co}(g(A_{\bar{u}})) + \int_{Y \setminus \{\bar{u}\}} \operatorname{co} g(A_u)\, d\nu(u)$. Thus, setting $e_k = \alpha e_{\bar{u},k} + \int_{Y \setminus \{\bar{u}\}} g(f(u))\, d\nu(u)$, we have $e_k \in \operatorname{int} E(\nu)$ for each k. Also, $e_k \to e(\tau_A)$ by continuity of g, since $a_{k,h} \to f(\bar{u})$ for each h if $k \to \infty$ by choice of the points $a_{k,h}$.

As $f(\bar{u}) \in \operatorname{int} A_{\bar{u}}$, we can find numbers $0 < r_1 < r_2$ such that $\bar{B}(f(\bar{u}), r_1) \subseteq \bar{B}(f(\bar{u}), r_2) \subseteq \operatorname{int} A_{u_i}$. Let $\rho \colon \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable map such that $\rho(a) = 1$ if $a \in B(f(\bar{u}), r_1)$, $0 \leq \rho(a) \leq 1$ for all $a \in \mathbb{R}^n$, and $\rho(a) = 0$ if $a \notin B(f(\bar{u}), r_2)$. As $a_{k,h} \to f(\bar{u})$ for each $h = 1, \ldots, m+1$ if $k \to \infty$, we can assume for each k and each h that $a_{k,h} \in$

$B(f(\bar{u}), r_1)$ and that $a + \rho(a)(f(\bar{u}) - a_{k,h}) \in \operatorname{int} A_{\bar{u}}$ whenever $a \in B(f(\bar{u}), r_2)$. Define a map $\bar{u}_{k,h} \colon A_{\bar{u}} \times E(\nu) \to \mathbb{R}$ for each $h = 1, \ldots, m+1$ and each k by setting

$$\bar{u}_{k,h}(a, e) = \bar{u}\big(a + \rho(a)\big(f(\bar{u}) - a_{k,h}\big), e\big) \\ - \bar{u}\big(a + \rho(a)\big(f(\bar{u}) - a_{k,h}\big), e_k\big) + \bar{u}\big(a + \rho(a)\big(f(\bar{u}) - a_{k,h}\big), e(\tau_A)\big)$$

for each $(a, e) \in A_{\bar{u}} \times E(\nu)$; for each $u \in Y$ and each k, define a map $u_k \colon A_u \times E(\nu) \to \mathbb{R}$ by setting

$$u_k(a, e) = u(a, e) - u(a, e_k) + u\big(a, e(\tau_A)\big)$$

for each $(a, e) \in A_u \times E(\nu)$. Then, for each k, all the maps u_k and $\bar{u}_{k,h}$, $h = 1, \ldots, m+1$, are twice continuously differentiable on their domains. Using the facts that $e_k \to e(\tau_A)$ and that $a_{k,h} \to f(\bar{u})$ for each $h = 1, \ldots, m+1$, we can assume, as in (a), that they all satisfy (U3) in the assumptions on payoff functions.

Now for each k, define $\lambda_k \colon Y \to \mathcal{U}$ by setting $\lambda_k(u) = u_k$ for each $u \in Y$ and note that λ_k is continuous. For each k, set $\nu' = \nu - \alpha\delta_{\bar{u}}$ and $\nu_k = \alpha \sum_{h=1}^{m+1} \beta_{k,h}\delta_{\bar{u}_{k,h}} + \nu' \circ \lambda_k^{-1}$. Because Y is compact and λ_k is continuous, $\operatorname{supp}(\nu' \circ \lambda_k^{-1})$ is compact for each k and, hence, so is $\operatorname{supp}(\nu_k)$ for each k. Thus, $\nu_k \in \mathcal{M}$ for each k. Because the distribution of action sets induced by ν is the same as that induced by ν_k, we have $E(\nu_k) = E(\nu)$, and, thus, ν_k actually belongs to $\mathcal{G}$ for each k. Using Lemma 1, we see that λ_k converges uniformly to the identity on Y, and from this, we see that $\nu' \circ \lambda_k^{-1} \to \nu'$ narrowly and that $\rho_H(\operatorname{supp}(\nu'), \operatorname{supp}(\nu' \circ \lambda_k^{-1})) \to 0$. Also $\alpha \sum_{h=1}^{m+1} \beta_{k,h}\delta_{\bar{u}_{k,h}} \to \alpha\delta_{\bar{u}}$ narrowly and $\rho_H(\operatorname{supp}(\alpha\delta_{\bar{u}}), \operatorname{supp}(\alpha \sum_{h=1}^{m+1} \beta_{k,h}\delta_{\bar{u}_{k,h}})) \to 0$ by the facts that $a_{k,h} \to f(\bar{u})$ for each $h = 1, \ldots, m+1$ and that $e_k \to e(\tau_A)$. Consequently we have $\nu_k \to \nu$.

Set $\tau_k = \alpha \sum_{h=1}^{m+1} \beta_{k,h}\delta_{\bar{u}_{k,h}, a_{k,h}} + \nu' \circ (\lambda_k \times f)^{-1}$ for each k. We claim that if k is large enough, then τ_k is an equilibrium distribution for ν_k such that such that (i)–(iii) in the definition of $\mathcal{G}_1$ are true. Indeed, it is clear that $\tau_{k,\mathcal{U}} = \nu_k$. Note next that by the facts that f and λ_k are continuous and Y is compact, the set

$$\big\{(\bar{u}_{k,h}, a_{k,h}) \colon h = 1, \ldots, m+1\big\} \cup \big\{\big(\lambda_k(u), f(u)\big) \colon u \in Y\big\}$$

is closed and must, therefore, include $\operatorname{supp}(\tau_k)$. Note also that

$$e(\tau_{k,A}) = \alpha \sum_{h=1}^{m+1} \beta_{k,h} g(a_{k,h}) + \int g\big(f(u)\big)\, \mathrm{d}\nu'(u) = e_k.$$

Now for any $u \in Y$, we have $\lambda_k(u)(a, e_k) = u(a, e(\tau_A))$ for each $a \in A_u$ by the choice of $\lambda_k(u)$; hence, $\varphi(\lambda_k(u), e_k) = \{f(u)\}$ and $D^2\lambda_k(u)(f(u), e_k)$ is negative definite, because τ is an equilibrium distribution for ν satisfying (i) and (ii) in the definition of $\mathcal{G}_1$. From the second property of τ, we also see that there is a compact neighborhood $\bar{U}$ of $f(\bar{u})$ in $\operatorname{int} A_{\bar{u}}$ such that $D_a^2 \bar{u}_{k,h}(a, e_k)$ is negative definite for each $a \in \bar{U}$ and each $h = 1, \ldots, m+1$ if k is large enough, because $\bar{u}_{k,h} \to \bar{u}$ for each $h = 1, \ldots, m+1$. Since $a_{k,h} \to f(\bar{u})$ and $D_a \bar{u}_{k,h}(a_{k,h}, e_k) = D_a \bar{u}(f(\bar{u})), e(\tau_A)) = 0$ for each $h = 1, \ldots, m+1$, it now follows from first of the properties noted for the equilibrium distribution τ that $\varphi(\bar{u}_{k,h}, e_k) = \{a_{kh}\}$

for each $h = 1, \dots, m+1$ if k is large enough. By construction, $e(\tau_{k,A}) = e_k \in \operatorname{int} E(\nu) = \operatorname{int} E(\nu_k)$ for each k and, thus, the above claim is true.

Clearly, the sequence $\langle \alpha \sum_{h=1}^{m+1} \beta_{k,h} \delta_{\bar{u}_{k,h}, a_{k,h}} \rangle$ of measures on $\mathcal{U}$ converges narrowly to $\alpha\delta_{\bar{u}, f(\bar{u})}$, and since $\langle \lambda_k \times f \rangle$ converges uniformly to $\mathrm{id}_Y \times f$, the sequence $\langle \nu' \circ (\lambda_k \times f)^{-1} \rangle$ of measures on $\mathcal{U}$ converges narrowly to $\nu' \circ (\mathrm{id}_Y \times f)^{-1}$. Since $\alpha\delta_{\bar{u}, f(\bar{u})} = \alpha\delta_{\bar{u}} \circ (\mathrm{id}_Y \times f)^{-1}$ and $\nu' \circ (\mathrm{id}_Y \times f)^{-1} = (\nu - \alpha\delta_{\bar{u}})(\mathrm{id}_Y \times f)^{-1}$, it follows that $\tau_k \to \nu \circ (\mathrm{id}_Y \times f)^{-1}$ narrowly; that is, $\tau_k \to \tau$, because $\tau = \nu \circ (\mathrm{id}_Y \times f)^{-1}$. Thus the assertion of (c) is true in case there is a $\bar{u} \in Y$ with $\nu(\bar{u}) > 0$.

(ii) Now suppose $\nu(\{u\}) = 0$ for each $u \in Y$. As $\nu \in \mathcal{G}_3$, we can choose a $\bar{u} \in Y$ and a decreasing sequence $\langle W_l \rangle$ of compact subsets of Y such that $\bigcap_{l=0}^{\infty} W_l = \{\bar{u}\}$, $\nu(W_l) > 0$ for each l, and $A_u = A_{\bar{u}}$ for each $u \in W_0$.

For each l, define a map $f_l \colon Y \to A$ by setting $f_l = 1_{Y \setminus W_l} f + 1_{W_l} f(\bar{u})$ and set $e_l = \int_{Y \setminus W_l} g(f(u)) \,\mathrm{d}\nu(u) + \nu(W_l) g(f(\bar{u}))$. Note that $e_l \to e(\tau_A)$ as $l \to \infty$ and that $e_l \in E(\nu)$ for each l. Now for each l and each $u \in Y$, define a map $u_l \colon A_u \times E(\nu \to \mathbb{R}$ by setting

$$u_l(a, e) = u(a, e) - u(a, e_l) + u\big(a, e(\tau_A)\big)$$

for $(a, e) \in A_u \times E(\nu)$, and for each $u \in W_l$ and each l, define $\bar{u}_l \colon A_{\bar{u}_l} \times E(\nu) \to \mathbb{R}$ by setting

$$\bar{u}_l(a, e) = \bar{u}(a, e) - \bar{u}(a, e_l) + \bar{u}\big(a, e(\tau_A)\big)$$

for $(a, e) \in A_{\bar{u}_l} \times E(\nu)$. As in (i), all these maps are twice continuously differentiable on their domains and we may assume that they all satisfy (U3) in the assumptions on payoff functions. For each l, define $\lambda_l \colon Y \to \mathcal{U}$ by setting $\lambda_l(u) = u_l$ for $u \in Y \setminus W_l$ and $\lambda_l(u) = \bar{u}_l$ for $u \in W_l$. As in (b), λ_l is measurable for each i, so $\nu_l = \nu \circ \lambda_l^{-1}$ is defined. As in (i), it follows that $\nu_l \in \mathcal{G}$ for each l. Using Lemma 1 and the fact that $e_l \to e(\tau_A)$, we see that $\langle \lambda_l \rangle$ converges uniformly to the identity on Y. This implies that $\nu_l \to \nu$. Moreover, we have $\nu_l(\{\bar{u}\}) \geq \nu(W_l) > 0$ for each l.

For each l, set $\tau_l = \nu \circ (\lambda_l \times f_l)^{-1}$. Then $\tau_{l,\mathcal{U}} = \nu_l$ for each l. For reasons as in (b), we have $\operatorname{supp}(\tau_l) \subseteq \{(u_l, f(u)) \colon u \in Y\} \cup \{(\bar{u}_l, f(\bar{u})\}$. Noting that $e(\tau_{l,A}) = \int g \circ f_l \,\mathrm{d}\nu = e_l$, it follows from the hypothesis that τ is an equilibrium distribution for ν satisfying (i) and (ii) in the definition of $\mathcal{G}_1$, together with the choice of the u_ls and $\bar{u}_l$s, that τ_l is an equilibrium distribution for ν_l, also satisfying (i) and (ii) in the definition of $\mathcal{G}_1$. Finally, note that $\tau = \nu \circ (\mathrm{id}_Y \times f)^{-1}$ and that the sequence $\langle \lambda_l \times f_l \rangle$ converges uniformly to $\mathrm{id}_Y \times f$. Consequently, $\tau_l \to \tau$.

(iii) Combining (i) and (ii), proves the assertion of (c).

Step (d). Let $\nu \in \mathcal{G}_1$ and let τ be an equilibrium distribution for ν, witnessing that $\nu \in \mathcal{G}_1$. Let V, W, h, and ξ_τ be associated with τ as in (b) of the proof of Theorem 1. Write $\bar{e} = e(\tau_A)$. If $\det D\xi_\tau(\bar{e}) \neq 0$, then $\nu \in \mathcal{G}^*$. Otherwise, pick any $0 < \lambda < 1$. For each $u \in Y$, define $u_\lambda \in \mathcal{U}$ by setting $u_\lambda(a, e) = u(a, (1-\lambda)\bar{e} + \lambda e)$ for each $(a, e) \in A_u \times E(\nu)$, so that, in particular, $\operatorname{dom} u_\lambda = \operatorname{dom} u$. Note that for each $u \in Y$ and $a \in A_u$, we have $u_\lambda(a, \bar{e}) = u(a, \bar{e})$. Define a map $\kappa_\lambda \colon Y \to \mathcal{U}$ by setting $\kappa_\lambda(u) = u_\lambda$. By Lemma 1, κ_λ is continuous. Thus, $\nu_\lambda = \nu \circ \kappa_\lambda^{-1}$ has compact support (since ν has) and, therefore, ν_λ belongs to $\mathcal{M}$. Because the distribution of action sets induced by ν_λ is the same as that induced by ν, we have $E(\nu_\lambda) = E(\nu)$, and, thus, ν_λ actually belongs to $\mathcal{G}$. Set $\tau_\lambda = \nu \circ (\kappa_\lambda \times h(\cdot, \bar{e}))^{-1}$. Clearly $\tau_{\lambda,\mathcal{U}} = \nu_\lambda$. Moreover,

$$e(\tau_{\lambda,A}) = \int_{\mathcal{U}} g\big(h(u, \bar{e})\big)\, \mathrm{d}\nu(u) = \bar{e}.$$

Further, if $(u', a') \in \operatorname{supp}(\tau_\lambda)$, then for some $u \in Y$, $u' = \kappa_\lambda(u) = u_\lambda$ and $a' = h(u, \bar{e})$. Since $u_\lambda(a, \bar{e}) = u(a, \bar{e})$ for each $u \in Y$ and $a \in A_u$, it follows that τ_λ is an equilibrium distribution for ν_λ satisfying (i) and (ii) in the definition of $\mathcal{G}_1$. Because $\bar{e} \in \operatorname{int} E(\nu)$ (by hypothesis), and since $E(\nu_\lambda) = E(\nu)$ and $e(\tau_{\lambda,A}) = \bar{e}$, also (iii) in the definition of $\mathcal{G}_1$ is true for τ_λ. Thus, τ_λ is an equilibrium distribution for ν_λ, witnessing that $\nu_\lambda \in \mathcal{G}_1$.

Using Lemma 1 and the fact that Y is compact, we see that whenever $\langle \lambda_k \rangle$ is a sequence in $(0, 1)$ such that $\lambda_k \to 1$, then the sequence $\langle \kappa_{\lambda_k} \rangle$ converges uniformly to id_Y and, thus, $\rho_H(\operatorname{supp}(\nu), \operatorname{supp}(\nu_{\lambda_k})) \to 0$. Hence, for large $\lambda \in (0, 1)$, $\operatorname{supp}(\nu_\lambda) \subseteq V$ and, therefore, $h(u_\lambda, \cdot)\colon W \to A$ is defined for each $u \in Y$; in particular, $D_e u_\lambda(h(u_\lambda, \bar{e}), \bar{e})$ must be defined. Clearly, for such λ, we have $h(u_\lambda, \bar{e}) = h(u, \bar{e})$ for each $u \in Y$, and, in addition, we must have $D_a^2 u_\lambda(h(u_\lambda, \bar{e}), \bar{e}) = D_a^2 u(h(u, \bar{e}), \bar{e})$ and $D_e u_\lambda(h(u_\lambda, \bar{e}), \bar{e}) = \lambda D_e u(h(u, \bar{e}), \bar{e}))$; therefore,

$$D_e(g \circ h)(u_\lambda, \bar{e}) = \lambda D_e(g \circ h)(u, \bar{e}).$$

Consequently,

$$D\xi_{\tau_\lambda}(\bar{e}) = \int_{\mathcal{U}} D_e(g \circ h)(u_\lambda, \bar{e})\, \mathrm{d}\nu(u) - I = \lambda \int_{\mathcal{U}} D_e(g \circ h)(u, \bar{e})\, \mathrm{d}\nu(u) - I.$$

Because the characteristic polynomial of the matrix $\int_{\mathcal{U}} D_e(g \circ h)(u, \bar{e})\, \mathrm{d}\nu(u)$ can have only finitely may zeros, we have

$$\det\left(\int_{\mathcal{U}} D_e(g \circ h)(u, \bar{e})\, \mathrm{d}\nu(u) - \frac{1}{\lambda} I\right) \neq 0$$

for all sufficiently large $0 < \lambda < 1$, and, hence, $\det(\lambda \int_{\mathcal{U}} D_e(g \circ h)(u, \bar{e})\, \mathrm{d}\nu(u) - I) \neq 0$ for such numbers λ.

As earlier, write id_Y for the identity on Y. Let $\langle \lambda_k \rangle$ be a sequence in $(0, 1)$ such that $\lambda_k \to 1$. Then, as noted above, the sequence $\langle \kappa_{\lambda_k} \rangle$ converges uniformly to id_Y, so $\nu_{\lambda_k} \to \nu$. The fact that $\langle \kappa_{\lambda_k} \rangle$ converges uniformly to id_Y implies also that the sequence $\langle \kappa_{\lambda_k} \times h(\cdot, \bar{e}) \rangle$ converges uniformly to $\mathrm{id}_Y \times h(\cdot, \bar{e})$, and, thus, we have $\tau_{\lambda_k} \to \tau$. By what was also noted above, each τ_{λ_k} belongs to $\mathcal{G}_1$. Combining these facts with the conclusion of the previous paragraph, we see that there is a sequence $\langle \nu_k \rangle$ in $\mathcal{G}^*$ and a sequence $\langle \tau_k \rangle$ of equilibrium distributions for the ν_ks such that $\nu_k \to \nu$, $\tau_k \to \tau$, and, for each k, τ_k witnesses that $\nu_k \in \mathcal{G}^*$. □

Proof of Theorem 2. (a) Let $\mathcal{G}^*$ be defined as in the proof of Theorem 1. Fix any $\nu \in \mathcal{G}^*$ and let τ be an equilibrium distribution for ν such that the requirements in (c) of the proof of Theorem 1 are satisfied. Suppose $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of finite-player games such that $\#(I_n) \to \infty$, $\nu_{G_n} \in \mathcal{M}$ for each n, and $\nu_{G_n} \to \nu$. Let $\hat{V}$,

W, W_1, and $h\colon \hat{V} \times W \to A$ be as in (e) of the proof of Theorem 1. Observe that $\tau = \nu \circ (\mathrm{id}_{\hat{V}} \times h(\cdot, e(\tau_A)))^{-1}$. Let $\langle W_{2,k} \rangle$ be a nonincreasing sequence of compact convex neighborhoods of $e(\tau_A)$, with $W_{2,k} \subseteq \operatorname{int} W_1$ for each k, such that $\bigcap_{k=0}^{\infty} W_{2,k} = \{e(\tau_A)\}$. Instead with a fixed W_2, the argument in (f) in the proof of Theorem 1 can be applied with each member of the sequence $\langle W_{2,k} \rangle$ to yield an increasing sequence $\langle n_k \rangle$ in $\mathbb{N}$, and for each $k \in \mathbb{N}$, an equilibrium f_k of the game (I_{n_k}, G_{n_k}) such that f_k can be written in the form $f_k(i) = h(G_{n_k}(i), e_k + z_k(i))$, $i \in I_{n_k}$, where $e_k \in W_{2,k}$ and $\|z_k(i)\| \leq \epsilon_k$ for each $i \in I_{n_k}$ and $\epsilon_k \to 0$ as $k \to \infty$.

For each k, let $f'_k\colon I_{n_k} \to A$ be the map defined by setting $f'_k(i) = h(G_{n_k}(i), e_k)$ for each $i \in I_{n_k}$. Let τ_k be the distribution of $G_{n_k} \times f_k$, and let τ'_k be that of $G_{n_k} \times f'_k$. As $e_k \in W_{2,k}$ for all k and $\bigcap_{k=0}^{\infty} W_{2,k} = \{e(\tau_A)\}$, we have $e_k \to e(\tau_A)$. From this we see that $\mathrm{id}_V \times h(\cdot, e_k) \to \mathrm{id}_V \times h(\cdot, e(\tau_A))$ uniformly on compact subsets of V, because h is continuous. Consequently,

$$\tau'_k = \nu_{G_{n_k}} \circ \big(\mathrm{id}_V \times h(\cdot, e_k)\big)^{-1} \to \nu \circ \big(\mathrm{id}_V \times h\big(\cdot, e(\tau_A)\big)\big)^{-1} = \tau,$$

i.e., the sequence $\langle \tau'_k \rangle$ of distributions of the maps $G_{n_k} \times f'_k$ converges to τ narrowly. Now note that if $\langle u'_k \rangle$ is any sequence in $\mathrm{supp}(\nu)$, and $\langle z_k \rangle$ is a sequence in $\mathbb{R}^m$ such that $h(u'_k, e_k + z_k)$ is defined and $\|z_k\| \to 0$, then

$$\big\| h\big(u'_k, e_k + z_k\big) - h\big(u'_k, e_k\big) \big\| \to 0,$$

because $\mathrm{supp}(\nu)$ is compact, h continuous, and $e_k \to e(\tau_A)$. Since $\nu_{G_n} \to \nu$ and, thus, $\rho_H(\mathrm{supp}(\nu_{G_{n_k}}), \mathrm{supp}(\nu)) \to 0$, it follows that for every $\epsilon' > 0$, there is a $k_{\epsilon'} \in \mathbb{N}$ such that whenever $k \geq k_{\epsilon'}$, then

$$\big\| h\big(G_{n_k}(i), e_k + z_k(i)\big) - h\big(G_{n_k}(i), e_k\big) \big\| \leq \epsilon'$$

for all $i \in I_{n_k}$, i.e., $\|f_k(i) - f'_k(i)\| \leq \epsilon'$ for all $i \in I_{n_k}$, and, thus, for some product metric $\tilde{\rho}$ on $\mathcal{U} \times A$ (recall that $\mathcal{U}$ can be regarded as a metric space), we have

$$\tilde{\rho}\big(\big(G_{n_k}(i), f_k(i)\big), \big(G_{n_k}(i), f'_k(i)\big)\big) \leq \epsilon'$$

for all $i \in I_{n_k}$ whenever $k \geq k_{\epsilon'}$. In view of this, we can conclude, using Billingsley (1968, Theorem 4.1), that the fact that the sequence $\langle \tau'_k \rangle$ of distributions of the maps $G_{n_k} \times f'_k$ converges narrowly to τ implies that the sequence $\langle \tau_k \rangle$ of distributions of the maps $G_{n_k} \times f_k$ converges narrowly to τ, too.

(b) By Lemma 7, given $\nu \in \mathcal{G}$, a sequence $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ of finite-player games such that $\nu_{G_n} \to \nu$, $\nu_{G_n} \in \mathcal{M}$ for each n, and $\#(I_n) \to \infty$ does exist. Putting this fact together with (a) and Lemma 8 proves the theorem. □

Proof of Theorem 3. Fix $C \in \mathcal{C}$. Let $\mathcal{G}^*$ be defined as in the proof of Theorem 1. Set $\mathcal{G}^*_C = \mathcal{G}^* \cap \mathcal{G}_C$. Then $\mathcal{G}^*_C$ is relatively open in $\mathcal{G}_C$. As for density, note that the only step in the proof of Lemma 8 that requires perturbation of action set is step (b) there, and this step is not needed if $\nu \in \mathcal{G}_C$. Thus, the assertion of Lemma 8 is true with $\mathcal{G}_C$ substituted for $\mathcal{G}$ and $\mathcal{G}^*_C$ substituted for $\mathcal{G}^*$; in particular, $\mathcal{G}^*_C$ is dense in $\mathcal{G}_C$. Now, with the appropriate substitutions, (e) of the proof of Theorem 1 yields the first assertion of Theorem 3, and the proofs of Theorems 2 , together with the fact noted in Remark 3, apply to establish the second assertion. □

Appendix

Lemma 9. *Let A be a nonempty subset of $\mathbb{R}^n$, with dense interior, and let $g\colon A \to \mathbb{R}^m$ be a continuously differentiable function. Suppose that $g(O)$ affinely spans $\mathbb{R}^m$ whenever O is a nonempty open set in $\mathbb{R}^n$ such that $O \subseteq A$. Then the following statements hold:*

(a) If $A_1, \dots, A_l$ are nonempty closed subsets of A such that $\operatorname{int} A_i$ is dense in A_i for each $i = 1, \dots, l$, then $\operatorname{int} \sum_{i=1}^{l} g(A_i)/l$ is dense in $\sum_{i=1}^{l} g(A_i)/l$ if $l \geq m$.

(b) If A' is a nonempty closed subset of A such that $\operatorname{int} A'$ is dense in A', then $\operatorname{int} \operatorname{co} g(A') \neq \varnothing$.

Proof. (a) Fix an integer $l \geq m$. Let $\tilde{g}\colon A^l \to \mathbb{R}^m$ be the function defined by setting $\tilde{g}(a_1, \dots, \alpha_l) = \sum_{i=1}^{l} g(a_i)$ for each $(a_1, \dots, a_l) \in A^l$. Then $\tilde{g}$ is continuously differentiable, with derivative $D\tilde{g}(a_1, \dots, a_l) = (Dg(a_1), \dots, Dg(a_l))$ at each $(a_1, \dots, a_l) \in A^l$. Let $\tilde{A}$ be the set of elements of $\operatorname{int} A^l$ at which the derivative of $\tilde{g}$ has rank m. Then $\tilde{A}$ is open. Moreover, the restriction of $\tilde{g}$ to $\tilde{A}$ is an open map (see Guillemin and Pollack 1974, p. 25, Exercise 1).

We next show that $\tilde{A}$ is dense in $\operatorname{int} A^l$. We first claim that given any nonempty open and convex set $O \subseteq \mathbb{R}^n$ such that $O \subseteq A$, the set C of all column vectors of the matrices $Dg(a)$ as a runs over O linearly spans $\mathbb{R}^m$. Suppose, if possible, otherwise. Then C is included in a linear subspace L of $\mathbb{R}^m$ with $\dim L < m$. Fix any $a_0 \in O$. Note that as O is convex, we have $ta + (1-t)a_0 \in O$ for all $a \in O$. Consequently, for each $a \in O$, we have

$$g(a) - g(a_0) = \int_{[0,1]} \frac{d}{dt} g\big(ta + (1-t)a_0\big)\, \mathrm{d}t = \int_{[0,1]} Dg\big(ta + (1-t)a_0\big)(a - a_0)\, \mathrm{d}t \in L.$$

But this implies that $g(O) \subseteq L + \{g(a_0)\}$, contradicting the hypotheses about g.

Let $(a_1, \dots, a_l) \in \operatorname{int} A^l$, and let $U \subseteq \operatorname{int} A^l$ be a neighborhood of this point. We can assume that U is of the form $O_1 \times \cdots \times O_l$, where O_i is a convex neighborhood of a_i, $i = 1, \dots, l$. By the previous paragraph, we can choose an $a'_1 \in O_1$ such that $Dg(a'_1)$ has a nonzero column vector v_1. Again by the previous paragraph, we can choose an $a'_2 \in O_2$ such that $Dg(a'_2)$ has a column vector v_2 such that v_1 and v_2 are linearly independent. Continuing in this fashion, we find points $a'_1, \dots, a'_m$ such that each matrix $Dg(a'_i)$ has a column vector v_i such that the matrix $(v_1, \dots, v_m)$ has rank m. If $l > m$, let a'_i be an arbitrary point of O_i for $m < i \leq l$. Then the matrix $(Dg(a'_1), \dots, Dg(a'_l))$ has rank m. Thus, $\tilde{A}$ is dense in $\operatorname{int} A^l$.

Now let $A_1, \dots, A_l$ be nonempty closed subsets of A, all with dense interior. Then $\operatorname{int}(A_1 \times \cdots \times A_l) = \operatorname{int} A_1 \times \cdots \times \operatorname{int} A_l \neq \varnothing$, so $\operatorname{int}(A_1 \times \cdots \times A_l)$ is dense in $A_1 \times \cdots \times A_l$, and, thus, $\tilde{A} \cap \operatorname{int}(A_1 \times \cdots \times A_l)$ is open and dense in $A_1 \times \cdots \times A_l$. Because g is continuous, it follows that the set

$$\left\{\frac{1}{l}\sum_{k=1}^{l} g(a_i)\colon (a_1,\dots,a_l)\in \tilde{A}\cap \operatorname{int}(A_1\times\cdots\times A_l)\right\}$$

is dense in $\sum_{i=1}^{l} g(A_i)/l$, and since $\tilde{g}$ is an open map, it follows that the former set is open. We conclude that $\operatorname{int}\sum_{i=1}^{l} g(A_i)/l$ is dense in $\sum_{i=1}^{l} g(A_i)/l$.

(b) By (a), $\operatorname{int}\sum_{i=1}^{m} g(A')/m\neq\varnothing$ and, of course, $\sum_{i=1}^{m} g(A')/m\subseteq \operatorname{co} g(A')$. □

Lemma 10. *Let A be a nonempty subset of $\mathbb{R}^n$, with dense interior, let $m=kn$, where $k\in\mathbb{N}$, and let $g\colon A\to\mathbb{R}^m$ be given by setting*

$$g(a)=\left(a_{(1)},a_{(1)}^2\dots a_{(1)}^k,a_{(2)},a_{(2)}^2\dots a_{(2)}^k,\dots,a_{(n)},\dots,a_{(n)}^k\right)$$

for each $a\in A$, where the subscript (h) means the hth coordinate of a, $h=1,\dots,n$. Then $g(O)$ affinely spans $\mathbb{R}^m$ whenever O is a nonempty open set in $\mathbb{R}^n$ with $O\subseteq A$.

Proof. Fix a nonempty open set $O\subseteq\mathbb{R}^n$ with $O\subseteq A$ and choose elements $a_1,\dots,a_k$ in O such that for each coordinate $h=1,\dots,n$, the points $a_{1,(h)},a_{2,(h)},\dots,a_{k,(h)}$ are distinct. Reorder the columns of the matrix $(Dg(a_1),\dots,Dg(a_k))$ so as to get a block diagonal matrix

$$\begin{pmatrix} B_1 & 0 & \dots & 0\\ 0 & B_2, & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & B_n\end{pmatrix},$$

where B_h, $h=1,\dots,n$, is a $(k\times k)$ matrix of the form

$$B_h=\begin{pmatrix} 1 & 1 & 1 & \dots & 1\\ 2a_{1,(h)} & 2a_{2,(h)} & 2a_{3,(h)} & \dots & 2a_{k,(h)}\\ 3a_{1,(h)}^2 & 3a_{2,(h)}^2 & 3a_{3,(h)}^2 & \dots & 3a_{k,(h)}^2\\ \dots & \dots & \dots & \dots & \dots\\ ka_{1,(h)}^{k-1} & ka_{2,(h)}^{k-1} & ka_{3,(h)}^{k-1} & \dots & ka_{k,(h)}^{k-1}\end{pmatrix}.$$

Now for each h, the determinant of the matrix B_h is just a positive multiple of the Vandermonde determinant and, thus, is nonzero because the points $a_{1,(h)},a_{2,(h)},\dots,a_{k,(h)}$ are distinct. This shows that the matrix $(Dg(a_1),\dots,Dg(a_k))$ has rank kn and, therefore, since $kn=m$, that $g(O)$ cannot be included in an affine subspace of dimension smaller than m. □

Lemma 11. *Let K be a nonempty compact subset of $\mathbb{R}^n$, with dense interior. Then there is a nondecreasing sequence $\langle K_k\rangle$ of nonempty compact subsets of K, all with dense interior, such that $K_k\subseteq\operatorname{int}K$ for each k and both $\rho_H(K,K_k)\to 0$ and $\rho_H(\partial K,\partial K_k)\to 0$.*

Proof. For each $k \in \mathbb{N}$, let $K'_k = \{x \in K\colon \operatorname{dist}(x, \partial K)\} > 1/(k+1)$ and let $K_k = \operatorname{cl} K'_k$. Then K'_k is open for each k. Also, for sufficiently large k, K'_k is nonempty; we may assume that this is true for each k. Further, $\partial K_k \subseteq \{x \in K\colon \operatorname{dist}(x, \partial K) = 1/(k+1)\}$ for each k, so, as ∂K is compact, we need only show that given any $y \in \partial K$ and any $\epsilon > 0$, we have $\operatorname{dist}(y, \partial K_k) < \epsilon$ if k is large enough. To see this, fix $y \in \partial K$ and $\epsilon > 0$. As $\operatorname{int} K$ is dense in K, there is an $x \in B(y, \epsilon) \cap \operatorname{int} K$. In particular, there is a $\bar{k} \in \mathbb{N}$ such that $x \in K'_k$ if $k > \bar{k}$. Pick any $k > \bar{k}$. Consider the line segment $Z = \{\lambda x + (1 - \lambda)y\colon 0 \le \lambda \le 1\}$. Clearly $Z \subseteq B(y, \epsilon)$. But also, $Z \cap K_k$ is closed, and since $x \in K'_k$ and $y \notin K_k$, Z must contain a boundary point of K_k. □

The following theorem is a special version of a result by Mas-Colell (1983).

Theorem 4. *Let $X \subseteq \mathbb{R}^\ell$ and $Y \subseteq \mathbb{R}^m$ be compact convex sets with nonempty interior. Let $\Lambda\colon X \times Y \to X \times \mathbb{R}^m$ be a continuous function; write Λ_X for* $\operatorname{proj}_X \circ \Lambda$ *and Λ_Y for* $\operatorname{proj}_{\mathbb{R}^m} \circ \Lambda$. *Suppose there is an open set $U \subseteq \mathbb{R}^m$, with $Y \subseteq U$, and a continuously differentiable function $\zeta\colon U \to \mathbb{R}^m$ such that, setting $\gamma = \min\{\|\zeta(y)\|\colon y \in \partial Y\}$, the following conditions hold:*

(a) For some $y^* \in \operatorname{int} Y$, $D\zeta(y^*)$ has full rank and $\zeta(y) = 0$ if and only if $y = y^*$ (so that, in particular, $\gamma > 0$).

(b) If $y \in \partial Y$ and $x = \Lambda_X(x, y)$, then $\|\Lambda_Y(x, y) - y - \zeta(y)\| < \gamma$.

Then Λ has a fixed point, i.e., there is an $(x, y) \in X \times Y$ such that $\Lambda(x, y) = (x, y)$.

References

Araújo, Aloisio, Jaime Orrillo, and Mario Páscoa (2000), "Equilibrium with default and endogenous collateral." *Mathematical Finance*, 10, 1–21. [1057]

Balder, Erik (2002), "A unifying pair of Cournot-Nash equilibrium existence results." *Journal of Economic Theory*, 102, 437–470. [1057, 1060, 1061]

Billingsley, Patrick (1968), *Convergence of Probability Measures*. Wiley, New York. [1076, 1078, 1088]

Carmona, Guilherme and Konrad Podczeck (2009), "On the existence of pure strategy Nash equilibria in large games." *Journal of Economic Theory*, 144, 1300–1319. [1058]

Carmona, Guilherme and Konrad Podczeck (2012), "Ex-post stability of Bayes-Nash equilibria of large games." *Games and Economic Behavior*, 74, 418–430. [1058]

Carmona, Guilherme and Konrad Podczeck (2019), "Strict pure strategy Nash equilibria in large finite-player games." University of Surrey and Universität Wien. [1058]

Carmona, Guilherme and Konrad Podczeck (2020a), "Approximation and characterization of Nash equilibria of large games." (Forthcoming, Economic Theory). [1059]

Carmona, Guilherme and Konrad Podczeck (2020b), "Pure strategy Nash equilibria of large finite-player games and their relationship to non-atomic games." *Journal of Economic Theory*, 187, 105015. [1059]

Castaing, C. and M. Valadier (1977), *Convex Analysis and Measurable Multifunctions*, volume 580 of Lecture Notes Mathematics. Springer-Verlag, New York. [1074]

Dierker, Hildegard (1975), "Smooth preferences and the regularity of equilibria." *Journal of Mathematical Economics*, 2, 43–63. [1059, 1060]

Guillemin, Victor and Alan Pollack (1974), *Differential Topology*. Prentice-Hall, Englewood Cliffs, N.J. [1089]

Hildenbrand, W. (1974), *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton. [1073, 1074]

Housman, David (1988), "Infinite player noncooperative games and the continuity of the Nash equilibrium correspondence." *Mathematics of Operations Research*, 13, 488–496. [1059]

Howe, Roger (1979), "On the tendency toward convexity of the vector sum of sets." Cowles Foundation Discussion Paper No. 538. [1075]

Kalai, Ehud (2004), "Large robust games." *Econometrica*, 72, 1631–1665. [1055, 1058]

Khan, M. Ali and Yeneng Sun (1996), "Nonatomic games on Loeb spaces." *Proceedings of the National Academy of Sciences*, 93, 15518–15521. [1056]

Khan, M. Ali and Yeneng Sun (1999), "Non-cooperative games on hyperfinite Loeb spaces." *Journal of Mathematical Economics*, 31, 455–492. [1055, 1058]

Khan, Mohammed Ali, Kali Rath, Haomiao Yu, and Yongchao Zhang (2017), "On the equivalence of large individualized and distributionalized games." *Theoretical Economics*, 12, 533–554. [1060]

Loeb, Peter, and Manfred Wolff, eds (2015), *Nonstandard Analysis for the Working Mathematician*, Second edition. Springer, Dordrecht. [1056]

Mas-Colell, Andreu (1977), "Regular, nonconvex economies." *Econometrica*, 45, 1387–1407. [1057, 1059, 1060]

Mas-Colell, Andreu (1983), "Walrasian equilibria as limits of noncooperative equilibria. Part I: Mixed strategies." *Journal of Economic Theory*, 30, 153–170. [1070, 1091]

Mas-Colell, Andreu (1984), "On a theorem by Schmeidler." *Journal of Mathematical Economics*, 13, 201–206. [1055, 1056, 1058, 1060, 1065, 1067]

Mas-Colell, Andreu (1985), *The Theory of General Economic Equilibrium: A Differentiable Approach*. Econometric Society Monographs. Cambridge University Press. [1079]

Rashid, Salim (1983), "Equilibrium points of non-atomic games: Asymptotic results." *Economics Letters*, 12, 7–10. [1058]

Rauh, Michael T. (2003), "Non-cooperative games with a continuum of players whose payoffs depend on summary statistics." *Economic Theory*, 21, 901–906. [1057, 1061, 1062]

Schmeidler, David (1973), "Equilibrium points of nonatomic games." *Journal of Statistical Physics*, 7, 295–300. [1055, 1060, 1067]

Schwartz, Laurent (1967), *Cours d'Analyse*, volume I. Paris: Hermann. [1078, 1079]

Trockel, Walter (1984), *Market Demand: An Analysis of Large Economies With Non-Convex Preferences*, volume 223 of Lecture Notes in Economics and Mathematical Systems. Springer Verlag, Berlin. [1059]

Wills, Michael D. (2007), "Hausdorff distance and convex sets." *Journal of Convex Analysis*, 14, 109–117. [1063, 1074]

Yu, Haomiao and Wei Zhu (2005), "Large games with transformed summary statistics." *Economic Theory*, 26, 237–241. [1057]

Co-editor Thomas Mariotti handled this manuscript.

Manuscript received 10 October, 2019; final version accepted 15 September, 2020; available online 21 September, 2020.