Pure strategy Nash equilibria in large finite-player games

Guilherme Carmona† Konrad Podczeck‡

March 21, 2019

Abstract

In the context of semi-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 pure strategy Nash equilibria if the number of agents is large. A key feature of our model is that payoff functions have differentiability properties, but no quasi-concavity assumption is imposed on these functions. 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 numbers: C72.

1 Introduction

It is now common in the literature to address economic problems of strategic interaction among many negligible individuals by models of semi-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. A particular and convenient class of models of semi-anonymous games is formed by continuum games (games with a

---

*Thanks to Daniel García, Michael Greinecker, and Klaus Ritzberger for helpful comments and suggestions.
†Address: University of Surrey, School of Economics, Guildford, GU2 7XH, UK; email: g.carmona@surrey.ac.uk.
‡Address: Institut für Volkswirtschaftslehre, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria; email: konrad.podczeck@univie.ac.at
In continuum games, there is no distinction anymore between the “distribution of the actions chosen by the other players” and the “distribution of the actions chosen by all players,” 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). No quasi-concavity assumption on payoff functions is made in these results. After all, of course, continuum games are idealizations of situations with a large but finite number of players, and in this respect some fundamental questions remain unanswered:

(1) 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?

(2) 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 games?

In other words, are continuum games, and equilibria of such games, good idealizations of situations with a large but finite number of players? It is the purpose of this paper to contribute to the solution of this problem.

It is well-known that finite-player games may fail to have pure strategy Nash equilibria unless quasi-concavity assumptions on payoff function are made. In fact, this may 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 2.5). Thus the best one can hope for in regard to the first of the above questions 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 semi-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.

Based on this result, we deal with the second of the above questions. 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 and corresponding pure strategy Nash equilibria which converges (in an appropriate sense) to the given continuum game and its equilibrium.

A byproduct 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 is in contrast with approximate or mixed strategy equilibria where knowledge by a player of the choices made by the other players, or the distribution of these choices, need not determine his/her equilibrium action or equilibrium mixed strategy.

In this regard, our analysis provides a supplement to the known results on asymptotic implementability of equilibria of continuum games by finite-player games, which
are stated in terms of approximate equilibria (e.g., Rashid (1983), Khan and Sun (1999), Kalai (2004), Carmona and Podczeck (2009)). In fact, concerning the problem of asymptotic implementation of equilibria of continuum games, our second result shows that there is no need in the model we set up to consider approximate or mixed strategy equilibria at all.

To note some details of our model, action sets are compact convex subsets of some Euclidean space, with non-empty 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 Balder (2002) or Rauh (2003), by finitely many summary statistics (e.g., the first $k$ non-central moments) of the distribution of these actions. From the viewpoint of applications, this is not a big restriction; in fact, in many applications of semi-anonymous games, e.g., Cournot oligopoly games, it is just the mean action of the other players which determines a player’s payoff in addition to his/her own action. Payoff functions, in our model, are twice continuously differentiability. 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., pairs of action sets and payoff functions) we define a suitable metric; because the actual definition requires some technical constructions, we refer to Section 2.4 and here say only that this metric is defined in terms of graphs of payoff functions to accommodate for the fact that action sets may differ across players. A continuum game is specified as a Borel probability measure with compact support on the space of players’ characteristics. The compact support condition requires that players’ characteristics in a continuum game are not too dissimilar. “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” 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 $\nu$ which have an equilibrium distribution such that the corresponding externality (which is the same for all players in a continuum game) has a neighborhood $W$ on which the correspondence that sends externalities to the externalities determined by the best replies of the players with characteristics in the support of $\nu$ can be identified with a function which is differentiable such that, at each point of $W$, its derivative minus that of the identity
on $W$ has maximal rank.

From a technical viewpoint, a paper related to ours is Mas-Colell (1977), where it is shown, among other things, that, for generic distributions of preferences, finite-agent exchange economies possess Walrasian equilibria if the number of agents is large, without assuming preferences to be convex. However, there are crucial differences to our analysis, so that our proofs cannot be simple translations of those in Mas-Colell (1977). One difference is that in game theory the externality faced by an agent is given directly by the actions of the other players, and thus may differ across agents (in finite-player games), while in Mas-Colell (1977), following standard Walrasian equilibrium theory, the externality is the price system and thus common to all agents. Another difference between the analysis in Mas-Colell (1977) and ours comes from the fact that in the former, consumption sets of agents agree, while in ours, following standard game theory, action sets may differ across agents.

The organization of the paper is as follows. In the next section the model is set up and the results are stated. Proofs can be found in Section 3. An appendix contains some auxiliary material.

2 The model and the results

2.1 General notation and terminology

If $X$ is any metric space, we write $\rho_H$ for the Hausdorff metric on the set of non-empty compact subsets of $X$, and for $A, B \subseteq X$, $\text{int} A$ denotes the interior of $A$, $\partial A$ the boundary of $A$, and $A \setminus B$ set-theoretic subtraction. If $A \subseteq \mathbb{R}^n$, then $\text{co} A$ denotes the convex hull of $A$.

If $\mu$ is a Borel measure on a separable metric space $X$, we write $\text{supp}(\mu)$ for the support of $\mu$, 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 $\mu$ is a Borel measure on a product $X \times Y$ of metric spaces, $\mu_X$ and $\mu_Y$ 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 any number $r > 0$, we write $B(x, r)$ for the open ball of center $x$ and radius $r$, and $\bar{B}(x, r)$ for the closed ball of center $x$ and radius $r$.

Let $X \subseteq \mathbb{R}^k$ be such that $\text{int} X$ is dense in $X$ (which is true, in particular, if $X$ is convex and $\text{int} X \neq \emptyset$). We say that a function $f : X \to \mathbb{R}^\ell$ 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 $\tilde{f} : \tilde{X} \to \mathbb{R}^\ell$ which 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 $\text{int} X$ is

---

1 That is, $\rho_H(A, B) = \max\{\max_{a \in A} d(a, B), \max_{b \in B} d(b, A)\}$ for any two non-empty compact sets $A, B \subseteq X$, writing $d$ for the metric of $X$. Recall that on the set of non-empty 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.
dense in \( X \). In this case, we write \( Df(x) \) for the derivative of \( f \) at \( x \in X \), and \( Df \) for the map \( x \mapsto Df(x) \). If \( f \) 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 \( D^2 f \) 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 \( D_x(x, y) \) 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, \( D^2 f(x, y) \) stands for \( D_x D_x f(x, y) \).

### 2.2 Actions and externality

We will 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 which 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 so as to get suitable sets for the domains of payoff functions.

There is a universe \( A \) of possible actions, which is a non-empty, compact, and convex subset of \( \mathbb{R}^n \) with non-empty interior. Action sets of games will be included in \( A \) and also have non-empty interior.

Summary statistics of a distribution \( \tau_A \) of actions in \( A \) are given by the formulae

\[
\int g_1(a) \, d\tau_A(a), \ldots, \int g_m(a) \, d\tau_A(a)
\]

where \( g_i, i = 1, \ldots, m \), is a continuously differentiable function from \( A \) to \( \mathbb{R} \) (given independently of \( \tau_A \)). We write \( g \) for the vector \( (g_1, \ldots, g_m) \), and \( \int g(a) \, d\tau_A(a) \) for the vector \( (\int g_1(a) \, d\tau_A(a), \ldots, \int g_m(a) \, d\tau_A(a)) \).

A natural example is obtained by setting

\[
g(a) = (a_{(1)}^1, a_{(1)}^2 \ldots a_{(1)}^k, a_{(2)}^2 \ldots a_{(2)}^k, \ldots, a_{(n)}^2 \ldots a_{(n)}^k)
\]

for each \( a \in A \), where the subscript \( (h) \) means the \( h \)th coordinate of \( a \), \( h = 1, \ldots, n \); in this case, \( m = kn \) and \( \int g(a) \, d\tau_A(a) \) is the vector of the first \( k \) non-central moments of the coordinate-distributions determined by \( \tau_A \); 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 \( \int g(a) \, d\tau_A(a) \) is the “mean action” corresponding to the distribution \( \tau_A \) on \( A \).

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

\[
E = \left\{ \int g \, d\tau_A : \tau_A \text{ is a probability measure on } A \right\}.
\]

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 externality for a player in some continuum game. Thus the set \( E \) can be seen as the universe of possible externalities.

To ensure that in games with sufficiently many players the sets of externalities that a player could actually face have dense interior, i.e., are appropriate for differentiability assumptions on payoff functions, we make the following assumption on the map \( g \). There is an \( N \in \mathbb{N} \setminus \{0\} \) such that whenever \( l \in \mathbb{N} \) satisfies \( l \geq N \), then, for
some open dense subset \( \bar{A} \) in the product \( A^l \), the map \( \bar{g}: \bar{A} \to \mathbb{R}^m \), given by setting 
\[
\bar{g}(a_1, \ldots, a_l) = \frac{1}{l} \sum_{i=1}^l g(a_i)
\]
for each \((a_1, \ldots, a_l) \in \bar{A}\), is an open map; see Lemma 10 in the appendix for the desired conclusion. This assumption is satisfied, with \( N = k \), if \( m = kn \) and \( g \) is such that \( \int g \, d\tau_A \) is, as in the example above, the vector of the first \( k \) non-central moments of the coordinate-distributions determined by \( \tau_A \); see Lemma 11 in the appendix. Because the focus of our paper is on large games, there is no problem with an assumption requiring that a game have sufficiently many players.

### 2.3 Payoff functions

A payoff function is a real-valued function \( u \) with \( \text{dom} \, u = A_u \times E_u \) where \( A_u \subseteq A \) and \( E_u \subseteq E \). 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 \( \text{dom} \, u \). 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

\[
\varphi(u,e) = \{ a \in A_u : u(a,e) = \max_{a' \in A_u} u(a', e) \}.
\]

We assume that for a payoff function \( u \), and the associated sets \( A_u \) and \( E_u \), the following are satisfied: (a) \( A_u \) and \( E_u \) are non-empty compact subsets of \( A \) and \( E \), respectively, such that \( A_u \) is convex, with \( \text{int} \, A_u \neq \emptyset \), and \( \text{int} \, E_u \) is dense in \( E_u \); (b) \( u \) is continuous; (c) \( u \) is twice continuously differentiable on \( \text{int} \, A_u \times E_u \); (d) 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 (a) and (b) imply that \( \varphi(u,e) \) is non-empty for each \( e \in E_u \), and that (d) implies that \( \varphi(u,e) \subseteq \text{int} \, A_u \) for each \( e \in E_u \).

**Remark 1.** The convexity assumption on the action sets \( A_u \) is made only for convenience. As with the externalities factors \( E_u \), we could assume the action sets \( A_u \) to just have dense interior and still get domains of payoff functions appropriate for differentiability assumptions. However, we don’t see any point in relaxing the convexity assumption on the action sets, because if these sets are assumed to have non-empty interior, then they necessarily include non-empty convex open sets. On the other hand, in finite-player games the sets \( E_u \) are of the form \( E_u = \frac{1}{l} \sum_{j=1}^l g(A_j) \) (see Section 2.5) and therefore, even when they have dense interior, cannot be convex in general; to see this, take \( A = [0,1] \) and suppose \( g: A \to \mathbb{R}^2 \) is defined by setting \( g(a) = (a, a^2) \) for \( a \in [0,1] \).

**Remark 2.** Payoff functions can be constructed, for instance, in the following way. Choose sets \( \bar{A} \subseteq A, \bar{E} \subseteq E \) such that the requirements in (a) are satisfied. Pick any \( \bar{a} \in \text{int} \, \bar{A} \) and a number \( r > 0 \) such that \( B(\bar{a}, r) \subseteq \bar{A} \). Let \( \bar{u} \) any twice continuously differentiable function from \( \bar{A} \times \bar{E} \) to \( \mathbb{R} \). Set

\[
m_1 = \max\{ \bar{u}(a,e) : a \in \partial \bar{A}, e \in \bar{E} \}.
\]
Choose a number \( k \) with \( k + m_2 > m_1 \) and a twice continuously differentiable function \( \zeta: A \to \mathbb{R}_+ \) such that \( \zeta(a) = k \) for \( a \in B(\bar{a}, r/2) \) and \( \zeta(a) = 0 \) for \( a \in A \setminus B(\bar{a}, r) \).

Now set \( u = \tilde{u} + \zeta, A_u = \tilde{A}, \) and \( E_u = \tilde{E} \). Then \( u \) satisfies conditions (a)–(d).

**Remark 3.** We have required differentiability of a payoff function \( u \) only on \( \text{int} \ A_u \times E_u \) to allow for the possibility that \( \|Du(a_n, e_n)\| \to \infty \) if \( a_n \in \text{int} \ A_u \) but \( a_n \to a \in \partial A_u \). This adds some slight generality without complicating the arguments in the proofs of our results.

### 2.4 Space of payoff functions

The set of payoff functions is denoted by \( U \). We define a metric \( \rho \) on \( U \) as follows. Using Lemma 13 in the appendix, with \( Y = \{A_u: u \in U\} \), equipped with the Hausdorff metric, and taking the map \( u \mapsto A_u: U \to Y \) into consideration, we can choose a family \( \{K_{u,i}\}_{u \in U, i \in \mathbb{N}} \) of compact and convex subsets of \( A \), all with non-empty interior, and with \( K_{u,i} \subseteq \text{int} \ A_u \) for each \( u \in U \) and \( i \in \mathbb{N} \), such that for each fixed \( u \in U \), \( \rho_H(A_u, K_{u,i}) \to 0 \) as \( i \to \infty \), and for each fixed \( i \in \mathbb{N} \), \( \rho_H(K_{u,i}, K_{u_k,i}) \to 0 \) as \( k \to \infty \) whenever \( u \in U \) and \( u_k \in U \), \( k \in \mathbb{N} \), are such that \( \rho_H(A_u, A_{u_k}) \to 0 \). For each \( u \in U \), write \( \Gamma_u \) for the graph of \( u \), and \( \Gamma_{Du}, \Gamma_{D^2u} \) for the graphs of the maps \( Du \) and \( D^2u \) respectively. Now, for any \( u, u' \in U \), let

\[
\rho(u, u') = \rho_H(\text{dom} \, u, \text{dom} \, u') \\
+ \rho_H(\Gamma_u, \Gamma_{u'}) \\
+ \sum_{i \in \mathbb{N}} 2^{-i} \min\{1, \rho_H(\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{n+m}), \Gamma_{Du'} \cap (K_{u',i} \times E_u \times \mathbb{R}^{n+m}))\} \\
+ \sum_{i \in \mathbb{N}} 2^{-i} \min\{1, \rho_H(\Gamma_{D^2u} \cap (K_{u,i} \times E_u \times \mathbb{R}^{(n+m)^2}), \Gamma_{D^2u'} \cap (K_{u',i} \times E_u \times \mathbb{R}^{(n+m)^2}))\}.
\]

In the rest of this paper, \( U \) is regarded as being equipped with the metric \( \rho \). The following is true.

**Lemma 1.** \( U \) is separable.

(See Section 3.2 for the proof.)

The next lemma provides a characterization of \( \rho \) in terms of sequences of actions and externalities. In particular, it shows that the actual choice of the sets \( K_{u,i} \), modulo to the general conditions as stated above, does not matter for the topology of \( U \).

**Lemma 2.** Let \( u \) and \( u_k, k \in \mathbb{N} \), be elements of \( U \). The following are equivalent

(i) \( \rho(u, u_k) \to 0 \).
(ii) (a) $\rho_H(\text{dom } u, \text{dom } u_k) \to 0$; (b) if $(a, e) \in \text{dom } u$ and $(a_k, e_k) \in \text{dom } u_k$, $k \in \mathbb{N}$, are such that $(a_k, e_k) \to (a, e)$ then $u_k(a_k, e_k) \to u(a, e)$; (c) if $(a, e) \in (\text{int } A_u) \times E_u$ and $(a_k, e_k) \in (\text{int } A_{u_k}) \times E_{u_k}$, $k \in \mathbb{N}$, are such that $(a_k, e_k) \to (a, e)$ then both $D u_k(a_k, e_k) \to D u(e, a)$ and $D^2 u_k(a_k, e_k) \to D^2 u(e, a)$.

(See Section 3.2 for the proof.)

**Remark 4.** Lemma 2 shows in particular that on subsets of $U$ consisting of functions with a common domain, the topology induced by the metric $\rho$ is a version of the topology of $C^2$-uniform convergence on compacta. More precisely, if $U_1$ is such a subset of $U$, then for a sequence $\langle u_k \rangle$ in $U_1$ one has $\rho(u, u_k) \to 0$ for some $u \in U_1$ if and only if (a) $u_k \to u$ uniformly and (b), writing $A_1 \times E_1$ for the common domain of the elements of $U_1$, $Du_k \to Du$ as well as $D^2 u_k \to D^2 u$, uniformly on all sets of the form $K \times E_1$ where $K$ is a compact subset of $A_1$.

### 2.5 Finite-player games

A finite-player game is a pair $(I, G)$ where $I$ is a finite set of players, satisfying $\#(I) \geq \max\{2, N\}$, $N$ being the number from the assumption on the map $g$ made in Section 2.2, and $G$ is a map from $I$ to $U$ such that $E_i = \sum_{j \in I(i)} g(A_j)/((\#I) - 1)$, 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 $\sum_{j \in I(i)} \delta_{a_j}/((\#I) - 1)$ where $\delta_{a_j}$ denotes Dirac measure at point $a_j$ in the action set $A_j$ of $j \in I \setminus \{i\}$. Thus the equality $E_i = \sum_{j \in I(i)} g(A_j)/((\#I) - 1)$ 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 $(I, G)$. By Lemma 10(a), $\text{int} \sum_{j \in I(i)} g(A_j)/((\#I) - 1)$ is dense in $\sum_{j \in I(i)} g(A_j)/((\#I) - 1)$, so requiring this equality is consistent with (a) in the assumptions on payoff functions and therefore with the definition of $U$.

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, $e_{f,i} = \sum_{j \in I(i)} g(f(j))/((\#I) - 1)$, or, in other words, $e_{f,i} = \int g(a) d \tau_{A,f,i}(a)$, 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 $B \subseteq A$,

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

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

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

8
2.6 Continuum games

It is convenient to start with some definitions. Let $\mathcal{M}$ be the set of all Borel probability measures on $\mathcal{U}$ with compact support. (By Lemma 1, $\mathcal{U}$ is a separable metric space, 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\(^2\) and $\rho_H(\text{supp}(\nu_n), \text{supp}(\nu)) \to 0$, i.e., $\text{supp}(\nu_n) \to \text{supp}(\nu)$ in the Hausdorff metric topology. Note that for any finite-player game, $\nu_G$ as defined in the previous section belongs to $\mathcal{M}$. Given $\nu \in \mathcal{M}$, let

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

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 $u \in \text{supp}(\nu)$. By Lemma 10(b), 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 defined from that 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 $\tau$ on $\mathcal{U} \times A$ such that $\tau_\mathcal{U} = \nu$ and $\text{supp}(\tau) \subseteq \{(u,a) \in \mathcal{U} \times A : a \in \varphi(u,e(\tau_A))\}$. By Mas-Colell (1984), every continuum game $\nu \in \mathcal{G}$ has an equilibrium distribution.

We remark that in the literature there are models of continuum games with an explicitly specified set of players and a map from this set to the set of players’ characteristics. However, in the present paper, we are interested in the connection between continuum games and large finite-player games. Since there is no reasonable topology on sets of players, such a connection cannot be made in terms of graphs of functions from sets of players to a space of players’ characteristics, but only in terms of distributions of players’ characteristics, and for this, the definition of continuum games in terms of such distributions provides all the information needed.

One might be interested in results involving generic properties relative to games in which players’ action sets share some property, e.g., games in which all players

\(^2\)Recall that the narrow topology on the set of Borel measures on a metric space is the topology of pointwise convergence on the bounded continuous functions defined on this space, evaluation being given by integration.
have the same action set. The set of such games could itself be non-generic in the set of all games. In view of this, the following notion will be convenient.

Let $C$ be the set of all closed subsets of the actions universe $A$. For any $C \in \mathcal{C}$, let $S_C = \{ u \in \mathcal{U} : C \subseteq A_u \}$. We call a set of this form a structure on $\mathcal{U}$. (More general structures could be considered in the theorems stated in the next section). Note that $S_C$ is closed in $\mathcal{U}$ for each $C \in \mathcal{C}$. For each $C \in \mathcal{C}$, let $\mathcal{M}_{SC} = \{ \nu \in \mathcal{M} : \text{supp}(\nu) \subseteq S_C \}$ and $\mathcal{G}_{SC} = \mathcal{G} \cap \mathcal{M}_{SC}$.

To give two examples of a structure, if $C = \emptyset$, then $S_C = \mathcal{U}$, and if $C = A$ then $S_C = \{ u \in \mathcal{U} : A_u = A \}$. Note that if $C = A$, then $E(\nu) = E$ whenever $\nu \in \mathcal{M}_{SC}$, and thus $\text{dom } u = A \times E$ for each $u \in \text{supp}(\nu)$ if $\nu \in \mathcal{G}_{SC}$.

**Remark 5.** Clearly $\mathcal{G}_{SC}$ is non-empty for each $C \in \mathcal{C}$; just consider the Dirac measure $\delta_{u_0}$ where $u_0 : A \times E \to \mathbb{R}$ is any twice continuously differentiable function satisfying (d) of the assumptions on payoff functions. To get a slightly more general example, let $C \in \mathcal{C}$, let $A_1, \ldots, A_k$ be closed convex subsets of $A$, all with non-empty interior and including $C$, and let $\alpha_1, \ldots, \alpha_k$ be positive real numbers such that $\sum_{i=1}^k \alpha_i = 1$. Set $\tilde{E} = \sum_{i=1}^k \alpha_i \text{co } g(A_i)$. For each $i = 1, \ldots, k$, choose a $u_i \in \mathcal{U}$ with $A_{u_i} = A_i$ and $E_{u_i} = \tilde{E}$; cf. Remark 2. Set $\nu = \sum_{i=1}^k \delta_{u_i}$. Note that we must have $E(\nu) = \tilde{E}$, by the definition of $E(\nu)$. Consequently $\nu \in \mathcal{G}_{SC}$.

More interesting are continuum games satisfying the following condition. Let us say that a continuum game $\nu \in \mathcal{G}$ has “essentially ordinally non-equivalent payoff functions” if there is a partition of $\text{supp}(\nu)$ into $\nu$-null sets such that two payoff functions belonging to different members of the partition are ordinally non-equivalent (“ordinally non-equivalent” for two payoff functions $u, u' \in \mathcal{U}$ meaning that there is no increasing function $h : u(A_u \times E(\nu)) \to \mathbb{R}$ such that $u' = h \circ u$.) Note that if a $\nu \in \mathcal{G}$ satisfies this condition, then $\nu$ must be zero on singletons, therefore atomless because $\mathcal{U}$ is a separable metric space. In Lemma 8 in Section 3.2 we show that, for any $C \in \mathcal{C}$, the set of continuum games $\nu \in \mathcal{G}_{SC}$ with ordinally non-equivalent payoff functions is dense in $\mathcal{G}_{SC}$.

**Remark 6.** Given $C \in \mathcal{S}_C$ and $\nu \in \mathcal{G}_{SC}$, 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 $\mathcal{M}_{SC}$. In Lemma 9 in Section 3.2 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 non-equivalent) modifications of the countably many payoff functions involved such that the resulting sequences of finite-player games still converge to $\nu$.

### 2.7 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 towards infinity.

**Theorem 1.** There is an open subset $G^*$ of $G$ such that $G^* \cap G_{SC}$ is dense in $G_{SC}$ for every $C \in C$ and whenever $\nu \in G^*$ 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 $\nu_{G_n} \to \nu$, then there is an $N \in \mathbb{N}$ such that $(I_n, G_n)$ has a strict pure strategy Nash equilibrium if $\#(I_n) \geq N$.

(See Section 3.3 for the proof.)

**Remark 7.** A special case of Theorem 1 is given by taking $C = A$. Thus Theorem 1 contains a result about generic existence of pure strategy Nash equilibria for large finite-player games in which all action sets agree.

**Remark 8.** On the level of continuum games, one might be interested, given $C \in C$, in the set $G_{O,SC}$ of elements of $G_{SC}$ with ordinally non-equivalent payoff functions (see Remark 5), so it is worth mentioning explicitly that, in the context of Theorem 1, $G_{O,SC} \cap G_{SC}$ is open and dense in $G_{O,SC}$. This can be seen from Lemma 8 and the general fact that if $X$ is a topological space, $A \subseteq X$ is dense in $X$, and $B \subseteq X$ is open and dense in $X$, then $A \cap B$ is open and dense in $A$.

**Remark 9.** One may ask whether Theorem 1 can be placed into an ordinal setting, i.e., a setting of differentiable preference relations, in order to get a result that is on the level of classical genericity analysis of exchange economies. In the latter context, however, preference relations are topologized under the assumption that they be monotone, an assumption that cannot be made in the game theoretic context of our paper, because action sets and externality sets are assumed to be compact. The main problem of placing Theorem 1 into an ordinal setting would therefore be to find a suitable topology on the set of preference relations. We leave the consideration of this for future research.

**Remark 10.** Without making differentiability assumptions on payoff functions, we show in a companion paper “Nash Equilibria of Large Finite-Player Games and their Relationship to Non-Atomic Games” that, generically, finite-player games have mixed strategy equilibria such that the diameter of the supports of the mixed strategies tends to zero if the number of agents increases towards infinity; in this sense, mixed strategy equilibria become “nearly pure.”

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 $v: A \to \mathbb{R}_+$ be a continuous function such that (a) $v(-1/2) = v(3/2) = 0$, (b) at 0 and 1, $v$ assumes a global maximum, equal to 1, (c) $v$ is twice continuously differentiable on int $A$, with $Dv(1/2) = D^2v(1/2) = v(1/2) = 0$.

Let $g: A \to \mathbb{R}$ be the restriction to $A$ of the identity on $\mathbb{R}$. Then $E = [-1/2, 3/2]$, and
for each \( f: I \to A \) 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 \( u_H: A \times E \to \mathbb{R} \) 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 \( u_J: A \times E \to \mathbb{R} \) 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 \( f: I \to A \) is a strategy profile such that \( f(i) \in \{0, 1\} \) for all \( i \in I \), then \( e_{f,i} = \#\{j \in I \setminus \{i\}: a_j = 1\}/\#(I - 1) \) for each \( i \in I \).

Now suppose \( f: I \to A \) 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 \( e_{f,i'} \geq 1/2 \). On the other hand, 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.

Theorem 1 implies a result saying that, in our model, every equilibrium distribution of every continuum game is the limit of some sequence of finite-player games and corresponding pure strategy Nash equilibria, in the sense of the following definition.

**Definition.** Let \( \nu \in \mathcal{G} \) be a continuum game and \( \tau \) an equilibrium distribution for \( \nu \). 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 \( (\nu, \tau) \) 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 \( \tau \) narrowly. We say that \( (\nu, \tau) \) 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.** Let \( C \in \mathcal{C} \). Every \((\nu, \tau)\), where \( \nu \in \mathcal{G}_{SC} \) is a continuum game and \( \tau \) is an equilibrium distribution for \( \nu \), is asymptotically implementable by a sequence \( \langle (I_n, G_n) \rangle_{n \in \mathbb{N}} \) of finite-player games such that \( \nu_{G_n} \in \mathcal{G}_{SC} \) for each \( n \). (Because \( \emptyset \in \mathcal{C} \), this implies that every \((\nu, \tau)\), where \( \nu \in \mathcal{G} \) is any continuum game and \( \tau \) is an equilibrium distribution for \( \nu \), is asymptotically implementable by a sequence \( \langle (I_n, G_n) \rangle_{n \in \mathbb{N}} \) of finite-player games.)

(See Section 3.4 for the proof.) We emphasize that every continuum game can be taken in Theorem 2, not just one 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. We also emphasize that implementability holds relative to any given structure $C \in \mathcal{C}$.

The notion in the following definition strengthens the notion of asymptotic implementability.

**Definition.** Let $C \in \mathcal{C}$, let $\nu \in \mathcal{G}$ be a continuum game, and $\tau$ an equilibrium distribution for $\nu$. We say that $(\nu, \tau)$ is asymptotically robust in $C$ if each sequence $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ of finite-player games such that $\#(I_n) \to \infty$, $\nu_{G_n} \in \mathcal{M}_{S_C}$ for each $n$, and $\nu_{G_n} \to \nu$ asymptotically implements $(\nu, \tau)$.

**Theorem 3.** There is an open subset $\mathcal{G}^\ast$ of $\mathcal{G}$ such that, for every $C \in \mathcal{C}$, $\mathcal{G}^\ast \cap G_{S_C}$ is dense in $G_{S_C}$ and whenever $\nu \in \mathcal{G}^\ast \cap G_{S_C}$, then there is an equilibrium distribution $\tau$ for $\nu$ such that $(\nu, \tau)$ is asymptotically robust in $C$.

(See Section 3.5 for the proof.)

In Theorem 3 we cannot have $\mathcal{G}^\ast = \mathcal{G}$. This follows from the example after the statement of Theorem 1, because any continuum game has an equilibrium distribution. Indeed, in that example, the distribution of players’ characteristics is independent of the number of players, therefore convergent if this number grows towards infinity.

### 2.8 A special framework

In this section we concentrate on games in which the externality is the mean action, i.e., $g$ is the identity on $A$, and specialize Theorem 1 to a result that easily applies to Cournot oligopoly problems.

**Theorem 4.** Let $C \in \mathcal{C}$ and let $g$ be the identity on $A$. Let $\tilde{S}_C$ be a subset of $S_C$, and $\tilde{G}_{S_C}$ be the subset of $G_{S_C}$ consisting of the elements $\nu$ with $\text{supp}(\nu) \in \tilde{S}_C$. Suppose:

(a) $\mathcal{G}_{S_C}' \cap \tilde{G}_{S_C}$ is dense in $\tilde{G}_{S_C}$, writing $\mathcal{G}_{S_C}'$ for the set of elements of $\mathcal{G}_{S_C}$ with finite support.

(b) For each $u \in \tilde{S}_C$ and each compact set $K \subseteq \text{int} \ A_u$ there is an $\epsilon > 0$ such that $u' \in \tilde{S}_C$ if $u' \in U$ satisfies the following three conditions: (i) $\text{dom} u' = \text{dom} u$, (ii) $\rho(u, u') < \epsilon$, and (iii) $u'(a, e) - u(a, e) = d(a)$ for all $(a, e) \in \text{dom} u$ if $d : A_u \to \mathbb{R}$ vanishes outside $K$.

Then there is an open dense subset $\tilde{G}_{S_C}^\ast$ of $\tilde{G}_{S_C}$ such that if $\nu \in \tilde{G}_{S_C}^\ast$ and $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of finite-player games such that $\nu_{G_n} \to \nu$ and $\#(I_n) \to \infty$, then there is an $N \in \mathbb{N}$ such that $(I_n, G_n)$ has a pure strategy Nash equilibrium if $\#(I_n) \geq N$.

(See Section 3.6 for the proof. Conditions (a) and (b) are used to handle the fact that $\tilde{S}_C$ can be a proper subset of $S_C$; in the case in which $\tilde{S}_C = S_C$, (b) holds automatically by the definition of $S_C$, and (a) is just Lemma 7 stated in Section 3.2.)
We now apply Theorem 4 to get existence of Cournot equilibria for generic distributions of cost functions if the number of firms is large, without making a concavity assumption on revenue functions or a convexity assumptions on cost functions.

We assume that the set of possible outputs that can be produced by an individual firm is the interval $[0, m]$, where $m > 0$; the number $m > 0$ can be interpreted as a capacity constraint. Let $X$ be the set of all continuous functions $v : [0, m] \to \mathbb{R}$ which are twice continuously differentiable on $\text{int}[0, m]$, equipped with the metric $\tilde{\rho}$ given by setting

$$\tilde{\rho}(v, v_1) = \max_{a \in [0, m]} |v(a) - v_1(a)|$$

$$+ \sum_{i=1}^{\infty} \frac{1}{2^i} \min \left\{ 1, \max_{a \in K_i} |Dv(a) - Dv'(a)| \right\}$$

$$+ \sum_{i=1}^{\infty} \frac{1}{2^i} \min \left\{ 1, \max_{a \in K_i} |D^2v(a) - D^2v'(a)| \right\}$$

for any $v, v' \in X$, where $\langle K_i \rangle_{i \in \mathbb{N}[0]}$ is an increasing sequence of intervals in $\mathbb{R}$ such that $K_i \in \text{int}[0, m]$ for each $n \in \mathbb{N}\setminus\{0\}$ and $\bigcup_{i=1}^{\infty} K_i = \text{int}[0, m]$. The inverse demand function is given in terms of output per firm (independently of the actual number of firms in an oligopoly) and specified by an element $p$ of $X$ with $p(e) > 0$ for all $e \in [0, m]$ and $p(0) > p(m) > 0$. Let $Y$ be the subset of $X$ consisting of the elements $v$ of $X$ with $v(0) = 0$ and $Dv(a) > 0$ for all $a \in \text{int}[0, m]$ such that $\lim_{a \to 0} Dv(a)$ exists. Give $Y$ the subspace topology defined from $X$ and let $\mathcal{K}$ be a non-empty open subset of $Y$ such that for each $v \in \mathcal{K}$, $\lim_{a \to 0} Dv(a) < p(e) < v(m)/m$ for all $e \in [0, m]$. The elements of $\mathcal{K}$ are the possible cost functions. A Cournot oligopoly is a pair $(I, G)$ where $I$ is a finite set of firms, with $\#(I) \geq 2$, and $G : I \to \mathcal{K}$ is a map assigning cost functions to firms. A strategy profile $f : I \to [0, m]$ is a Cournot equilibrium of $(I, G)$ if

$$p \left( \frac{1}{\#(I)} f(i) + \frac{1}{\#(I)} \sum_{j \in I \setminus \{i\}} f(j) \right) f(i) - G(i)(f(i))$$

$$\geq p \left( \frac{1}{\#(I)} a + \frac{1}{\#(I)} \sum_{j \in I \setminus \{i\}} f(j) \right) a - G(i)(a)$$

for each $i \in I$ and each $a \in [0, m]$. Let $\mathcal{M}_\mathcal{K}$ be the set of all Borel probability measures on $\mathcal{K}$, with compact support. Give $\mathcal{M}_\mathcal{K}$ the topology analogous to that of $\mathcal{M}$.

**Theorem 5.** There is an open dense subset $\mathcal{M}^*_\mathcal{K}$ of $\mathcal{M}_\mathcal{K}$ such that if $\nu \in \mathcal{M}^*_\mathcal{K}$ and $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of Cournot oligopolies 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 Cournot equilibrium if $\#(I_n) \geq N$.

(See Section 3.7 for the proof.)
3 Proofs

To avoid confusion, the theorems and lemmata from Section 2 are restated when we come to prove them. It will be convenient in the proofs to have fixed some additional terminology.

3.1 Continuous correspondences

Let $X, Y$ be metric spaces. A correspondence $\theta: X \to 2^Y$, i.e., set-valued map from $X$ to $Y$, is called continuous if for each $x \in X$ and each open set $O \subseteq Y$ the following are true: (a) whenever $\theta(x) \subseteq O$, then there is a neighborhood $V$ of $x$ such that $\theta(x') \subseteq O$ for all $x' \in V$, and (b) whenever $\theta(x) \cap O \neq \emptyset$, then there is a neighborhood $V$ of $x$ such that $\theta(x') \cap O \neq \emptyset$ for all $x' \in V$. Note that if $\theta(x)$ is a non-empty and compact subset of $Y$ for each $x \in X$, then this definition is equivalent to saying that whenever $x \in X$ and $x_k \in X, k \in \mathbb{N}$, are such that $x_k \to x$, then $\rho_H(\theta(x), \theta(x_k)) \to 0$ (see Hildenbrand, 1974, B.III, Problem 4).

3.2 Preliminaries

Lemma 1. $\mathcal{U}$ is separable.

Proof. Let $\mathcal{F}_0$ be the set of all non-empty compact subsets of $\mathbb{R}^{n+m}$, $\mathcal{F}_1$ the set of all non-empty compact subsets of $\mathbb{R}^{n+m} \times \mathbb{R}$, $\mathcal{F}_2$ the set of all non-empty compact subsets of $\mathbb{R}^{n+m} \times \mathbb{R}^{n+m}$, and $\mathcal{F}_3$ the set of all non-empty 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 topological space. Give $\mathcal{F}_2^\mathbb{N}$ and $\mathcal{F}_3^\mathbb{N}$ the product topology. Write $\mathcal{F} = \mathcal{F}_0 \times \mathcal{F}_1 \times \mathcal{F}_2^\mathbb{N} \times \mathcal{F}_3^\mathbb{N}$ 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

$$\phi(u) = (\text{dom } u, \Gamma_u, (\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{n+m})))_{i \in \mathbb{N}}, (\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{(n+m)^2}))_{i \in \mathbb{N}}$$

for each $u \in \mathcal{U}$. By definition of $\rho$, $\phi$ is a homeomorphism from $\mathcal{U}$ onto $\phi(\mathcal{U})$. Because $\mathcal{F}$ 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.

Lemma 2. Let $u$ and $u_k, k \in \mathbb{N}$, be elements of $\mathcal{U}$. The following are equivalent

(i) $\rho(u, u_k) \to 0$.

(ii) (a) $\rho_H(\text{dom } u, \text{dom } u_k) \to 0$; (b) if $(a, e) \in \text{dom } u$ and $(a_k, e_k) \in \text{dom } u_k, k \in \mathbb{N}$, are such that $(a_k, e_k) \to (a, e)$ then $u_k(a_k, e_k) \to u(a, e)$; (c) if $(a, e) \in (\text{int } A_u) \times E_u$ and $(a_k, e_k) \in (\text{int } A_{u_k}) \times E_{u_k}, k \in \mathbb{N}$, are such that $(a_k, e_k) \to (a, e)$ then both $Du_k(a_k, e_k) \to Du(e, a)$ and $D^2u_k(a_k, e_k) \to D^2u(e, a)$.

Proof. (i)$\Rightarrow$(ii): Part (a) follows directly from the definition of $\rho$. 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
The hypotheses imply that there is an \(i \in \mathbb{N}\) such that \((a,e) \in K_{u,i} \times E_u\) and \((a_k,e_k) \in K_{u_k,i} \times E_{u_k}\) for \(k\) large enough. Indeed, as \(a \in \text{int} A_u\), we can find an \(i^* \in \mathbb{N}\) such that \(a \in \text{int} K_{u,i^*}\), using Lemma 3 together with the fact that \(\rho_H(A_u,K_{u,i}) \to 0\) as \(i \to \infty\) (see the construction of \(\rho\)). By part (a), we have \(\rho_H(A_u,A_{u_k}) \to 0\), and thus \(\rho_H(K_{u,i^*},K_{u_k,i}) \to 0\) as \(k \to \infty\) (again see the construction of \(\rho\)). Another appeal to Lemma 3 therefore shows that \(a_k \in K_{u_k,i^*}\) for large \(k\), because \(a_k \to a\) by hypothesis. Thus the claim above is true.

Now we have both \(\rho_H(\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{n+m}), \Gamma_{Du_k} \cap (K_{u_k,i} \times E_{u_k} \times \mathbb{R}^{n+m})) \to 0\) and \(\rho_H(\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{(n+m)^2}), \Gamma_{Du_k} \cap (K_{u_k,i} \times E_{u_k} \times \mathbb{R}^{(n+m)^2})) \to 0\), because \(\rho(u,u_k) \to 0\), so the assertion of part (c) follows in the same way as that of part (b).

(ii)\(\Rightarrow\)(i): Suppose (a)-(c) of (ii) are true. Combining (a) and (b), it follows directly that \(\Gamma_u \subseteq L_\Gamma u\).\(^3\) Suppose \((a,e,r) \in L_\Gamma u\). Then for some sequence \((n_i)_{i \in \mathbb{N}}\) in \(\mathbb{N}\) there are points \((a_{k_i},e_{k_i}) \in \text{dom} u_{k_i}, i \in \mathbb{N}\), such that \((a_{k_i},e_{k_i},u_{k_i}(a_{k_i},e_{k_i})) \to (a,e,r)\). From (a) we see that \((a,e) \in \text{dom} u\). Again from (a), there is a sequence \(((a_k,e_k))_{k \in \mathbb{N}}\) such that \((a_k,e_k) \to (a,e)\) and \((a_k,e_k) \in \text{dom} u_k\) for each \(k\). Define a sequence \(((a'_{k_i},e'_{k_i}))_{k \in \mathbb{N}}\) by setting \((a'_{k_i},e'_{k_i}) = (a_{k_i},e_{k_i})\) if \(k = k_i\) for some \(i\), and \((a'_{k_i},e'_{k_i}) = (a_k,e_k)\) otherwise. Then \((a'_{k_i},e'_{k_i}) \in \text{dom} u_k\) for each \(k\) and \((a'_{k_i},e'_{k_i}) \to (a,e)\), so (b) implies that \(u_k(a'_{k_i},e'_{k_i}) \to u(a,e)\). In particular, we have \(u_{k_i}(u_{k_i},e_{k_i}) \to u(a,e)\) and therefore \(r = u(a,e)\). Thus \(L_\Gamma u \subseteq \Gamma_u\) and it follows that \(\Gamma_u = L_\Gamma u = L_\Gamma u\). Now because \(\text{dom} u\) and \(u_{k_i}, k \in \mathbb{N}\), are all included in the compact set \(A \times E\), and because the maps \(u\) and \(u_{k_i}\) are continuous, (a) and (b) imply, in particular, that the sets \(\Gamma_u, \Gamma_{u_k}\), \(k \in \mathbb{N}\), are commonly included in a compact subset of the ambient Euclidean space, so the fact that \(\Gamma_u = L_\Gamma u\) implies that \(\rho_H(\Gamma_u,\Gamma_u) \to 0\).

Similarly, we see that \(\rho_H(\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{n+m}), \Gamma_{Du_k} \cap (K_{u_k,i} \times E_{u_k} \times \mathbb{R}^{n+m})) \to 0\) and \(\rho_H(\Gamma_{Du} \cap (K_{u,i} \times E_u \times \mathbb{R}^{(n+m)^2}), \Gamma_{Du_k} \cap (K_{u_k,i} \times E_{u_k} \times \mathbb{R}^{(n+m)^2})) \to 0\) converge to 0 for each \(i \in \mathbb{N}\) as \(k \to \infty\), using the fact that (a) implies that \(\rho_H(K_{u,i},K_{u_k,i}) \to 0\) for any \(i \in \mathbb{N}\) as \(k \to \infty\), just as above, and hence that \(\rho_H(K_{u,i} \times E_u, K_{u_k,i} \times E_{u_k}) \to 0\) for any \(i \in \mathbb{N}\) as \(k \to \infty\). By the definition of \(\rho\), we conclude that (ii)\(\Rightarrow\)(i) is true.

\(^3\)Here an below, \(L_\Gamma u\) is the set of limits of sequences \(\langle(a_k,e_k,r_k)\rangle\) such that \((a_k,e_k,r_k) \in \Gamma u\) for all \(k\), and \(L_\Gamma u\) the set of cluster points of such sequences.

\textbf{Lemma 3.} Let \(C \) and \(C_k, k \in \mathbb{N}\), be non-empty compact convex subsets of \(\mathbb{R}^\ell\), all with non-empty interior, such that \(\rho_H(C,C_k) \to 0\) as \(k \to \infty\). Let \(x \in \text{int} C\), and suppose \(\langle x_k \rangle\) is a sequence in \(\mathbb{R}^\ell\) such that \(x_k \to x\). Then \(x_k \in \text{int} C_k\) for all sufficiently large \(k\).

\textbf{Proof.} Otherwise, passing to a subsequence, if necessary, for each \(k\) we can appeal to the separating hyperplane theorem to find a \(p_k \in \mathbb{R}^\ell\), with \(\|p_k\| = 1\), such that \(p_kx_k \geq p_kC_k\). Again passing to a subsequence, if necessary, we may assume that there is a \(p \in \mathbb{R}^\ell\), with \(\|p\| = 1\), such that \(p_k \to p\). As \(x \in \text{int} C\) there is an \(x' \in C\) such...
that \( px' > px \). As \( \rho_H(C, C_k) \to 0 \), we can find \( x'_k \in C_k, k \in \mathbb{N} \), such that \( x'_k \to x' \). Since \( p_k \to p \) it follows that \( p_k x'_k > p_k x_k \) for large \( k \). But this contradicts the choice of the \( p_k \)'s, because \( x'_k \in C_k \) for all \( k \).

\[ \square \]

**Lemma 4.** Let \( C \) and \( C_k, k \in \mathbb{N} \), be non-empty compact convex subsets of \( \mathbb{R}^\ell \), all with non-empty interior, such that \( \rho_H(C, C_k) \to 0 \) as \( k \to \infty \). Let \( K \) be a compact subset of \( \text{int} \ C \). Then \( K \subseteq \text{int} C_k \) 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 \( x_k \notin \text{int} C_k \). As \( K \) is compact, we can assume that \( x_k \to x \), again passing to a subsequence, if necessary. Now \( x \in K \subseteq \text{int} C \), so by Lemma 3 we must have \( x_k \in \text{int} C_k \) for large \( k \), thus getting a contradiction. \[ \square \]

**Lemma 5.** (a) For every \( \nu \in \mathcal{M} \), \( E(\nu) \) is a compact convex subset of \( E \) with non-empty 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 \( Y = \text{supp}(\nu) \). 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 which 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 \( \theta : \mathcal{U} \to 2^{\mathbb{R}^m} \) defined by setting

\[ \theta(u) = \text{cog}(A_u) \]

for each \( u \in \mathcal{U} \). Then \( \theta \) has non-empty compact convex values, all with non-empty interior by Lemma 10(b). The fact that \( \theta \) has convex values implies that \( \int \theta d\nu \) is convex for all \( \nu \in \mathcal{M} \), and the fact that \( \theta \) has compact values, all included in the compact set \( \text{cog}(A) \), 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 \( u \mapsto A_u : \mathcal{U} \to 2^A \) is continuous. Because the map \( g \) is continuous, this implies that the correspondence \( u \mapsto g(A_u) : \mathcal{U} \to 2^{\mathbb{R}^m} \) is continuous. By Hildenbrand (1974, B.III, Propositions 6 and 10), it follows that \( \theta \) 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 non-empty 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 \( \max pE(\nu) = \int_\mathcal{U} \max p\text{cog}(A_u) d\nu(u) \), by the definition of \( E(\nu) \), and also

\[ \int_\mathcal{U} \max p\text{cog}(A_u) d\nu(u) = \int_\mathcal{U} \max p\theta(u) d\nu(u) = \max p \int_\mathcal{U} \theta(u) d\nu(u) \, . \]

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 \( \text{int} \, E(\nu) \neq \emptyset \) for each \( \nu \in \mathcal{M} \). Indeed, pick any \( \nu \in \mathcal{M} \) and any \( u' \in \text{supp}(\nu) \). By what has been noted above, \( \text{int} \, \theta(u') \neq \emptyset \), so there is a compact set \( K \subseteq \text{int} \, \theta(u') \) such that \( K \neq \emptyset \). By Lemma 4, there is an open neighborhood \( V \) of \( u' \) such that \( K \subseteq \theta(u'') \) for each \( u'' \in V \). As \( u' \in \text{supp}(\nu) \), \( \nu(V) > 0 \), so the set \( \nu(V)K \) has non-empty interior. Now

\[
\nu(V)K + \int_{\nu(V)} \theta(u) \, d\nu(u) \subseteq \int_{\nu} \theta(u) \, d\nu(u) + \int_{\nu(V)} \theta(u) \, d\nu(u) = \int_{\nu} \theta(u) \, d\nu(u),
\]

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

**Lemma 6.** For each \( C \in \mathcal{C} \), \( \mathcal{G}_{SC}^C \) is closed in \( \mathcal{M} \).

**Proof.** The assertion follows from Lemma 5 together with the fact that \( u_n \to u \) in \( \mathcal{U} \) implies \( \rho_H(E(u_n), E(u)) \to 0 \). \( \square \)

**Lemma 7.** For each \( C \in \mathcal{C} \), \( \mathcal{G}_{SC}^C \) is dense in \( \mathcal{G}_{SC} \), writing \( \mathcal{G}_{SC}^C \) for the set of elements of \( \mathcal{G}_{SC} \) with finite support.

**Proof.** (a) Let \( \nu \in \mathcal{G}_{SC} \) be given. Write \( X = \text{supp}(\nu) \). 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 \( \nu \) narrowly. Since \( \text{supp}(\nu_n) \subseteq X = \text{supp}(\nu) \), narrow convergence of \( \langle \nu_n \rangle \) to \( \nu \) implies that we also have \( \rho_H(\text{supp}(\nu_n), \text{supp}(\nu)) \to 0 \). Thus we have \( \nu_n \to \nu \) in the topology of \( \mathcal{M} \). By Lemma 5, \( E(\nu) \) and \( E(\nu_n) \) for each \( n \) are compact, convex, and have non-empty interior; moreover, \( \rho_H(E(\nu), E(\nu_n)) \to 0 \). Fix \( b \in \text{int} \, E(\nu) \). By Lemma 3, \( b \in \text{int} \, E(\nu_n) \) for large \( n \); 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 \( r \leq 1 \) such that \( r(E(\nu_n) - \{ b \}) + \{ b \} \subseteq E(\nu) \). We must have \( r_n \to 1 \). To see this, fix \( 0 < r < 1 \). Since \( E(\nu) \) is convex and \( b \in \text{int} \, E(\nu) \), we have \( r(E(\nu) - \{ b \}) + \{ b \} \subseteq \text{int} \, E(\nu) \). Using the fact that \( \rho_H(E(\nu), E(\nu_n)) \to 0 \), it follows that \( r(E(\nu_n) - \{ b \}) + \{ b \} \subseteq \text{int} \, E(\nu) \) for large \( n \). Thus \( r_n \geq r \) for such \( n \). As \( 0 < r < 1 \) is arbitrary, we conclude that \( r_n \to 1 \).

(b) For each \( n \) and each \( i = 0, \ldots, n \), define \( u_{n,i} : A_{u_i} \times E(\nu_n) \to \mathbb{R} \) by setting \( u_{n,i}(a,e) = u_i(a, r_n(e - b) + b) \) for \( (a,e) \in A_{u_i} \times E(\nu_n) \) and note that \( u_{n,i} \in \mathcal{U} \). We claim that for any \( \epsilon > 0 \) there is an \( n_\epsilon \) such that whenever \( n > n_\epsilon \), then \( \rho(u_{n,i}, u_i) < \epsilon \).
for all $i = 0, \ldots, n$. Indeed, otherwise there are points $u_{n_k, i_k}$ and $u_{i_k}$, $k \in \mathbb{N}$, such that $n_k \to \infty$ as $k \to \infty$ and $\rho(u_{n_k, i_k}, u_{i_k}) \geq \epsilon > 0$ for each $k$. Because $u_{i_k} \in \text{supp}(\nu)$ and $\text{supp}(\nu)$ is compact, we can assume that $u_{i_k} \to \bar{u}$ for some $\bar{u} \in \text{supp}(\nu)$. But then, using Lemma 2, together with the facts that $\rho_H(E(\nu), E(\nu_n)) \to 0$ and $r_n \to 1$, it follows that also $u_{n_k, i_k} \to \bar{u}$, and we get a contradiction.

\begin{enumerate}[(c)]
\item Let $\hat{\nu}_n \in \mathcal{G}$ be defined by setting $\hat{\nu}_n = 1/(n+1) \sum_{i=0}^{n} \delta_{u_{i, n}}$. Because $A_{u_{i, n}} = A_{u_i}$ for all $i = 0, \ldots, n$, we actually have $\hat{\nu}_n \in \mathcal{G}_{S_C}$ for all $n$. From (b) and the fact that $\rho_H(\text{supp}(\nu), \text{supp}(\nu_n)) \to 0$ it follows that $\rho_H(\text{supp}(\hat{\nu}_n), \text{supp}(\nu)) \to 0$. In particular, there is compact $Y \subseteq U$ including $\text{supp}(\nu)$ and all the sets $\text{supp}(\nu_n)$ and $\text{supp}(\hat{\nu}_n)$ as $n$ runs over $\mathbb{N}$. Therefore, because a continuous function on a compact metric space is uniformly continuous, it follows from (b) and the fact that $\nu_n \to \nu$ narrowly, that $\hat{\nu}_n \to \nu$ narrowly, too. We conclude that $\hat{\nu}_n \to \nu$ in the topology of $\mathcal{G}$. \qedhere
\end{enumerate}

**Lemma 8.** For every $C \in \mathcal{C}$ the set of $\nu \in \mathcal{G}_{S_C}$ with ordinally non-equivalent payoff functions is dense in $\mathcal{G}_{S_C}$.

**Proof.** Fix any $C \in \mathcal{C}$. By Lemma 7 it suffices to show that if $\nu \in \mathcal{G}_{S_C}$ is of the form $\nu = \sum_{i=1}^{m} \alpha_i \delta_{u_i}$, then there is a sequence $\nu_n \in \mathcal{G}_{S_C}$ such that $\nu_n \to \nu$ and each $\nu_n$ has ordinally non-equivalent payoff functions. Let such a $\nu$ be given.

Fix any $i \in \{1, \ldots, m\}$. By (d) of the assumptions on payoff functions there is a compact set $K_i \subseteq \text{int} A_{u_i}$ such that $\varphi(u_i, e) \subseteq \text{int} K_i$ for each $e \in E(\nu)$. Choose a compact set $K'_i$ such that $K_i \subseteq \text{int} K'_i$ and $K'_i \subseteq \text{int} A_{u_i}$, and a twice continuously differentiable function $\zeta_i : \mathbb{R}^n \to \mathbb{R}^+$ such that $\zeta_i(x) = 1$ for $x \in K_i$ and $\zeta_i(x) = 0$ for $x \in \mathbb{R}^n \setminus K'_i$. Let $\bar{u}_i$ be an extension of $u_i \upharpoonright (K'_i \times E(\nu))$ to a twice continuously differentiable function from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}$. Let $f$ be the vector $f = (1, 0, \ldots, 0) \in \mathbb{R}^n$. Choose a sequence $\{c_{i, n}\}_{n \in \mathbb{N}}$ of real numbers such that $c_{i, n} \to 0$ and for each $n$, $c_{i, n} > 0$ but small enough so that, for each $t \in [0, 1]$, the function $u_{i, n, t} : A_{u_i} \times E(\nu) \to \mathbb{R}$, defined by setting

$$u_{i, n, t}(a, e) = \xi(a)\bar{u}_i(a + (c_{i, n} t)f, e) + (1 - \xi(a))u_i(a, e),$$

satisfies (d) in the assumptions on payoff functions, and therefore belongs to $U$.

Using Lemma 2(ii)$\Rightarrow$(i) we see that the map $t \mapsto u_{i, n, t}$ from $[0, 1]$ to $U$ is continuous for each $n$. Thus we can speak of the distribution of Lebesgue measure on $[0, 1]$ under this map; moreover, writing $\nu_{i, n}$ for this distribution, $\text{supp}(\nu_{i, n})$ is compact. Again using Lemma 2(ii)$\Rightarrow$(i), we see that $\rho(u_i, u_{i, n, t}) \to 0$ for each $t \in [0, 1]$ as $n \to \infty$, so $\nu_{i, n} \to \nu_i$ narrowly for each $t \in [0, 1]$. In fact, $\rho(u_i, u_{i, n, t}) \to 0$ whenever $\{t_n\}$ is a sequence in $[0, 1]$, from which we see that $\rho_H(\text{supp}(\nu_i), \text{supp}(\nu_{i, n})) \to 0$. Thus $\nu_{i, n} \to \nu_i$ in the topology of $\mathcal{M}$. The fact that $\rho(u_i, u_{i, n, t}) \to 0$ whenever $\{t_n\}$ is a sequence in $[0, 1]$ also implies that for $n$ large enough we have $\varphi(u_{i, n, t, e}) \subseteq \text{int} K_i$ for each $e \in E(\nu)$ and each $t \in [0, 1]$, because $\varphi(u_i, e) \subseteq \text{int} K_i$ by the choice of $K_i$. We may assume that the inclusions $\varphi(u_{i, n, t, e}) \subseteq \text{int} K_i$, $e \in E(\nu)$, $t \in [0, 1]$, are true actually for each $n$.

Do this construction for each $i = 1, \ldots, m$. Set $\nu_n = \sum_{i=1}^{m} \alpha_i \nu_{n, i}$ for each $n$. Then $E(\nu_n) = E(\nu)$ for each $n$ and thus $\nu_n \in \mathcal{G}_{S_C}$ for each $n$. Moreover, $\nu_n \to \nu$. \hspace{1cm} \boxed{19}
Finally, note that for each \( e \in E(\nu) \), each \( n \in \mathbb{N} \), each \( i = 1, \ldots, m \), and each \( t \in [0, 1] \), we have \( \varphi(u_{i,n,t}, e) = \varphi(u_i, e) - \{(c_{i,t})f\} \), by the choice of the function \( u_{i,n,t} \), since both \( \varphi(u_i, e) \) and \( \varphi(u_{i,n,t}, e) \) are included in \( \text{int} K_i \). Because the map \( t \mapsto c_{i,n,t} \) is an injection for each \( i \) and each \( n \), and because \( \varphi(u, e) = \varphi(u', e) \) for two ordinally equivalent payoff functions \( u \) and \( u' \) in any of the continuum games \( \nu_n \), it follows that the sequence \( \langle \nu_n \rangle \) is as desired. \[ \square \]

**Lemma 9.** Given \( C \in S_C \) and \( \nu \in S_{SC} \), there is a sequence \( \langle (I_n, G_n) \rangle_{n \in \mathbb{N}} \) of finite-player games such that \( I_n \rightarrow \infty \) and \( \nu_{G_n} \rightarrow \nu \) in \( M_{SC} \).

**Proof.** Construct a sequence \( \langle \nu_n \rangle_{n \in \mathbb{N}} \) of elements \( \nu_n = 1/(n+1) \sum_{i=0}^{n-1} \delta_{u_i} \) of \( M \) as in (a) of the proof of Lemma 7; in particular, \( \nu_n \rightarrow \nu \) in the topology of \( M \). For each \( n \in \mathbb{N} \setminus \{0\} \) each \( 0 \leq i \leq n \) define \( \nu_{n,i} \in M \) by setting \( \nu_{n,i} = 1/n \sum_{j \in J_n,i} \delta_{u_j} \) where \( J_n,i = \{0, \ldots, n\} \setminus \{i\} \); set \( \nu_0,0 = \nu_0 \). Note that for each \( n \in \mathbb{N} \setminus \{0\} \) and each \( 0 \leq i \leq n \) we have \( \|\nu_n - \nu_{n,i}\|_V \leq 2/n \), writing \( \|q\|_V \) for the variation norm on \( M \). Consequently, because \( \nu_n \rightarrow \nu \), we have \( \nu_{n,i} \rightarrow \nu \) as \( n \rightarrow \infty \) whenever \( \langle i_n \rangle \) is a sequence in \( \mathbb{N} \) with \( 0 \leq i_n \leq n \) for each \( n \). By Lemma 5, it follows that \( \rho_H(E(\nu), E(\nu_{i_n})) \rightarrow 0 \) whenever \( \langle i_n \rangle \) is as in the previous sentence. Based on this fact, we can argue as in the proof of Lemma 7 to obtain elements \( \bar{u}_{n,i} : A_u \times E(\nu_{n,i}) \rightarrow \mathbb{R} \) of \( U \) for each \( n \in \mathbb{N} \setminus \{0\} \) and each \( 0 \leq i \leq n \), so that, setting \( \bar{n}_n = 1/(n+1) \sum_{i=0}^{n-1} \delta_{\bar{u}_{n,i}} \), we have \( \bar{n}_n \rightarrow \nu \) in the topology of \( M \). In particular, by compactness of the supports of \( \nu \) and \( \bar{n}_n \), \( n \in \mathbb{N} \), if \( \bar{u}_{n,i} \in \text{supp}(\bar{n}_n) \) for each \( n \), then there is a subsequence \( \langle \bar{u}_{n_k,i_k} \rangle \) of the sequence \( \langle \bar{u}_{n,i} \rangle \) such that \( \bar{u}_{n_k,i_k} \rightarrow u \) for some \( u \in \text{supp}(\nu) \).

For each \( n \in \mathbb{N} \setminus \{0\} \) and each \( 0 \leq i \leq n \), set \( E_{n,i} = 1/n \sum_{j \in J_n,i} g(A_j) \). Note that \( E(\nu_{n,i}) = 1/n \sum_{j \in J_n,i} \text{cog}(A_j) \) (cf. the proof of Lemma 5); thus \( E_{n,i} \subseteq E(\nu_{n,i}) \). Let \( u_{n,i} : A_{u_{n,i}} \times E_{n,i} \rightarrow \mathbb{R} \) be the restriction of \( \bar{u}_{n,i} \) to \( A_{u_{n,i}} \times E_{n,i} \). 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 \( 0 \leq i \leq n \) if \( n \geq n_\epsilon \). Using Lemma 2, it follows from this and the last sentence of the previous paragraph that for each \( \epsilon > 0 \) there is an \( n'_\epsilon \in \mathbb{N} \) such that \( \rho(u_{n,i}, u_{n'_\epsilon}) < \epsilon \) for all \( 0 \leq i \leq n \) if \( n \geq n'_\epsilon \), because \( u_{n,i} \) is just the restriction of \( u_{n,i} \) to \( A_{u_{n,i}} \times E_{n,i} \), \( n \) being the number from the hypothesis on \( g \) made in Section 2.2.

Now, for each \( n \in \mathbb{N} \) with \( n \geq N \), set \( I_n = \{0, 1, \ldots, n\} \) and define \( G_n : I_n \rightarrow U \) by setting \( G(i) = u_{n,i} \) for each \( i \in I_n \). For \( n < N \), let \( (I_n, G_n) \) be an arbitrary finite-player game. From the fact that \( \rho_H(\text{supp}(\bar{n}_n), \text{supp}(\nu)) \rightarrow 0 \) and the conclusion of the previous paragraph we see that \( \rho_H(\text{supp}(\nu_{G_n}), \text{supp}(\nu)) \rightarrow 0 \). Thus, as in (c) of the proof of Lemma 7, we can conclude that \( \nu_{G_n} \rightarrow \nu \) in the topology of \( M \). \[ \square \]

### 3.3 Proof of Theorem 1

**Theorem 1.** There is an open subset \( \mathcal{G}^* \) of \( \mathcal{G} \) such that \( \mathcal{G}^* \cap \mathcal{G}_{SC} \) is dense in \( \mathcal{G}_{SC} \) for every \( C \in C \) and 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 \( \#(I_n) \rightarrow \infty \) and \( \nu_{G_n} \rightarrow \nu \), then there is an \( N \in \mathbb{N} \) such that \( (I_n, G_n) \) has a strict pure strategy Nash equilibrium if \( \#(I_n) \geq N \).
Proof. (a) As noted in Section 2.6, every continuum game $\nu \in \mathcal{G}$ has an equilibrium distribution, i.e., there is a Borel probability measure $\tau$ on $\mathcal{U} \times A$ such that $\tau_{u} = \nu$ and $\text{supp}(\tau) \subseteq \{(u, a) \in \mathcal{U} \times A : a \in \varphi(u, e(\tau))\}$. By (d) in the assumptions on payoff functions, we have $\varphi(u, e) \subseteq \text{int } A_u$ for each $u \in \mathcal{U}$ and each $e \in E(\nu)$. Thus if $\tau$ is an equilibrium distribution of $\nu$ in $\mathcal{G}$, then $a \in \text{int } A_u$ for each $(u, a) \in \text{supp}(\tau)$.

(b) Write $\mathcal{G}_1$ for the subset of $\mathcal{G}$ consisting of those $\nu$ such that for some equilibrium distribution $\tau$ of $\nu$,

(i) $\#(\varphi(u, e(\tau))) = 1$ for each $u \in \text{supp}(\nu)$;

(ii) for each $u \in \text{supp}(\nu)$, $D_u^2 a(u, e(\tau))$ is negative definite, where $a_u$ is the unique element of $\varphi(u, e(\tau))$ (note that $a_u \in \text{int } A_u$, so $D_u^2 a_u u, e(\tau))$ is defined);

(iii) $e(\tau)$ is defined).

We claim that given any $\nu \in \mathcal{G}_1$ and any equilibrium distribution $\tau$ of $\nu$ such that (i)–(iii) are satisfied, there are open neighborhoods $V$ of $\text{supp}(\nu)$ in $\mathcal{U}$ and $W$ of $e(\tau)$ in $\mathbb{R}^m$, with $W \subseteq E_u$ for all $u \in V$, such that, on $V \times W$, the best replies of $u$ against $e$ are given by a continuous map $h: V \times W \to A$ such that (1) $h(u, \cdot)$ is differentiable for each $u \in V$, (2) the derivative of $h(u, \cdot)$ depends continuously on $(u, e)$, and (3) $D_u^2 a(u, e(\tau))$ is negative definite for each $(u, e) \in V \times W$.

To see that this claim is true, choose a compact neighborhood $W_1$ of $e(\tau)$ such that $W_1 \subseteq \text{int } E(\nu)$, which is possible by (iii). Then by compactness of $\text{supp}(\nu)$, Lemma 2, and Lemma 4, there is an open neighborhood $V_1$ of $\text{supp}(\nu)$ in $\mathcal{U}$ such that $W_1 \subseteq \text{int } E_u$ for all $u \in V_1$. Now pick any $u \in \text{supp}(\nu)$. As above, let $a_u \in \text{int } A_u$ be the unique element of $\varphi(u, e(\tau))$. Then there is a compact and convex neighborhood $U_{a_u}$ of $a_u$ in $\text{int } A_u$ such that $D_u^2 a_u u(a, e(\tau))$ is negative definite for every $a \in U_{a_u}$. Now we can find numbers $r_1$, $r_2$ such that $u(a_u, e(\tau)) > r_1 > r_2 > u(a, e(\tau))$ for each $a \in A_u \setminus U_{a_u}$.

In particular, we must have $r_1 > u(a, e(\tau))$ for all $a \in \text{cl}(A_u \setminus U_{a_u})$.

Using Lemma 2 and Lemma 4 we see that there is a neighborhood $V'_u$ of $u$ such that $U_{a_u} \subseteq \text{int } A_u$ for each $u' \in V'_u$. Because the actions universe $A$ is compact, Lemma 2 now shows that there are open neighborhoods $V_u$ of $u$ in $\mathcal{U}$ and $W_u$ of $e(\tau)$ in $\mathbb{R}^m$, with $V_u \subseteq V'_u \cap V_1$ and $W_u \subseteq W_1$, $u'(a_u, e) > r_1 > u'(a, e)$ for each $u' \in V_u$, $e \in W_u$, and $a \in A_u \setminus U_{a_u}$, and such that $D_u^2 a'(e_1, a)$ is negative definite for each $u' \in V_u$, $e \in W_u$, and $a \in U_{a_u}$. In particular, for each $u' \in V_u$ and $e \in W_u$, $u'(a, e)$ is strictly concave on $U_{a_u}$.

Consequently for each $u' \in V_u$ and $e \in W_u$, the best reply of $u'$ against $e$ is unique. Apply this argument to each $u \in \text{supp}(\nu)$. Then by compactness of $\text{supp}(\nu)$ there are $u_1, \ldots, u_k \in \text{supp}(\nu)$ such that $\text{supp}(\nu) \subseteq V = \bigcup_{i=1}^k V_{u_i}$. Set $W = \bigcup_{i=1}^k W_{u_i}$ and $K = \bigcup_{i=1}^k U_{u_i}$. Then, for each $(u, e) \in V \times W$, the best reply of $u$ against $e$ is unique and belongs to $K$. Thus, on $V \times W$, the best reply correspondence $\varphi$ 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, $D_u^2 a(h(u, e), e)$ is negative definite for each $(u, e) \in V \times W$, i.e., we have (3). In view of this, the implicit function theorem applied to the maps $(a, e) \mapsto D_u a(u, e)$, $u \in V$, shows that (1) is true. Using Lemma 2 we see that the evaluation maps $(u, a, e) \mapsto D_u^2 a(u, e)$ and $(u, a, e) \mapsto D_u D_a a(u, e)$, which are
defined on the set \( \{(u, a, e) \in U \times A \times E : a \in \text{int } A_u, e \in E_u \} \), are continuous. From this we see that (2) is true.

Let \( \nu \in G_1 \) and \( \tau \) an equilibrium distribution for \( \nu \) such that (i)–(iii) are satisfied. Let \( W \) correspond to \( \tau \) as above. We can then define a map \( \xi_\tau : W \to \mathbb{R}^m \) by setting

\[
\xi_\tau(e) = \int g(h(u, e))
\]

for each \( e \in W \); then by the generalized version of Leibniz’ rule in Schwartz (1967, Chap IV.11, Theorem 115), \( \xi_\tau \) is continuously differentiable on \( W \), and we have

\[
D\xi_\tau(e) = \int D_e(g \circ h)(u, e)\, d\nu(u) - I
\]

where \( I \) is the \((m \times m)\)-identity matrix.

(c) Let \( G^* \) be the subset of \( G \) consisting of those \( \nu \in G \) such that for some equilibrium distribution \( \tau \) of \( \nu \), (i)–(iii) of (b) are satisfied and \( D\xi_\tau(e(\tau_\lambda)) \) has full rank, where \( \xi_\tau \) is associated with \( \tau \) as above. (Note that while the choice of the neighborhood \( W \) of \( e(\tau_\lambda) \), i.e., the domain of \( \xi_\tau \), involves some arbitrariness, \( D\xi_\tau(e(\tau_\lambda)) \) is uniquely determined.) In (d) below we will show that \( G^* \) is open, and in (e) prove the denseness part.

(d) Fix \( \nu \in G^* \). We need to show that \( \nu \) has a neighborhood that is included in \( G^* \). Let \( \tau \) be an equilibrium distribution for \( \nu \), witnessing that \( \nu \in G^* \). Let \( W, V, h, \) and \( \xi_\tau \) be associated with \( \tau \) as in (b).

(i) Choose a compact neighborhood \( W_1 \) of \( e(\tau_\lambda) \) with \( W_1 \subseteq W \). Then there is a \( k \in \mathbb{N} \) and neighborhood \( V_1 \subseteq V \) of \( \text{supp(} \nu \text{)} \) 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 \( k \in \mathbb{N} \backslash \{0\} \), 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 \( \text{dist}(u_k, \text{supp}(\nu)) < 1/k \). Since \( W_1 \) and \( \text{supp}(\nu) \) are compact we may assume that \( (u_k, e_k) \to (u, e) \) for some \( (u, e) \in \text{supp}(\nu) \times W_1 \). 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 \( U \) of \( \nu \) in \( G \) such that \( \text{supp}(\nu') \subseteq V_1 \) for each \( \nu' \in U \). Note that \( W_2 \subseteq E(\nu') \) for each \( \nu' \in U \). We can therefore define a map \( \xi_U : U \times W_2 \to \mathbb{R}^m \) by setting

\[
\xi_U(\nu', e) = \int g(h(u, e))
\]

for each \( \nu' \in U \) and \( e \in W_2 \). As above we see that for each fixed \( \nu' \in U \), \( \xi_U(\nu', \cdot) \) is continuously differentiable on \( W_2 \), with \( D_e \xi_U(\nu', \cdot) = \int D_e(g \circ h)(u, e)\, d\nu'(u) - I \) where \( I \) is the \((m \times m)\)-unit matrix. Now \( \xi_U \) is continuous and \( D_e \xi_U(\nu', \cdot) \) depends continuously on \( (\nu', e) \). 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 \( U \), 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 $\xi_U(v_k, e_k) \to \xi_U(v', e)$. Similarly, by (1), we see that $D_v \xi_U(v_k, e_k) \to D_v \xi_U(v', e)$.

Thus, on $U \times W_2$, $\xi_U$ is continuous and $D_v \xi_U(v', e)$ depends continuously on $(v', e)$, as claimed.

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

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

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

by the choice of $h$, and

$$e(\tau'_A) = \int g(h(u, e(\nu')) \, du(\nu') = \xi_U(\nu', e(\nu')) + e(\nu') = e(\nu').$$

Thus $\tau'$ is an equilibrium distribution for $v'$. By the choices of $V$ and $W$, and since $e(\nu') \in W_2 \subseteq W$ for each $\nu' \in U_1$, (i)–(iii) of (b) are true for $\tau'$. Let $V', W', h'$, and $\xi_{\tau'}$ be associated with $\tau'$ as in (b). Then $V' \cap V_1$ is a neighborhood of $\supp(\nu')$, and $W' \cap W_2$ a neighborhood of $e(\xi'_A) = e(\nu')$. 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 $D_h \xi_U(e(\tau'_A))$ has maximal rank. It follows that every $\nu'$ in the neighborhood $U_1$ of $\nu$ belongs to $\mathcal{G}^*$. As $\nu \in \mathcal{G}^*$ is arbitrary, $\mathcal{G}^*$ is open.

(e) Fix $C \in \mathcal{C}$. We next show that $\mathcal{G}^* \cap \mathcal{G}_{SC}$ is dense in $\mathcal{G}_{SC}$. Note first that by Lemma 7 the set $\mathcal{G}'$ of $\nu \in \mathcal{G}_{SC}$ such that $\supp(\nu)$ is finite is dense in $\mathcal{G}_{SC}$. Let $\nu \in \mathcal{G}'$; thus $\nu = \sum_{i=1}^k \alpha_i \delta_{u_i}$ with $\alpha_i > 0$, $i = 1, \ldots, k$, and $\sum_{i=1}^k \alpha_i = 1$. Let $\pi$ be an equilibrium distribution for $\nu$. Then $\tau_A = \sum_{i=1}^k \alpha_i \tau_{A,i}$ for probability measures $\tau_{A,i}$ on $A$ such that for each $i$ and each $a \in \supp(\tau_{A,i})$ we have $a \in \varphi(u_i, e(\tau_A))$. Note that

$$e(\tau_A) \equiv \int g(a) \, d\tau_A(a) = \sum_{i=1}^k \alpha_i \int g(a) \, d\tau_{A,i}(a).$$

Now, for each $i = 1, \ldots, k$,

$$\int g(a) \, d\tau_{A,i}(a) = \int_{\supp(\tau_{A,i})} g \, d\tau_{A,i} \in \text{co}(g(\supp(\tau_{A,i}))),$$

so by Carathéodory’s theorem there are points $a_{i,h} \in \supp(\tau_{A,i})$ and numbers $\alpha_{i,h} > 0$, $h = 1, \ldots, h_i$, with $\sum_{h=1}^{h_i} \alpha_{i,h} = 1$, such that $\int g(a) \, d\tau_{A,i}(a) = \sum_{h=1}^{h_i} \alpha_{i,h} g(a_{i,h})$.

Choose numbers $0 < \delta_1 < \delta_2 \leq 1$ so that for the closed balls $B(a_{i,h}, \delta_1)$ and $\bar{B}(a_{i,h}, \delta_2)$ we have $\bar{B}(a_{i,h}, \delta_1) \subseteq B(a_{i,h}, \delta_2) \subseteq \text{int} A_{u_i}$ for each $i = 1, \ldots, k$ and each $h = 1, \ldots, h_i$. Choose twice continuously differentiable functions $\rho_{i,h}: A_{u_i} \to \mathbb{R}$ such
that $0 \leq \rho_{i,h}(a) \leq 1$ for each $a \in A_{u_i}$, \( \rho_{i,h}(a) = 1 \) for each $a \in B(a_{i,h}, \delta_1)$, and \( \rho_{i,h}(a) = 0 \) for each $a \in A_{u_i} \setminus B(a_{i,h}, \delta_2)$.

For each $0 < \lambda < 1$ and each \((i,h)\), define a payoff function $u_{i,h,\lambda} \in U$, with $\text{dom } u_{i,h,\lambda} = \text{dom } u_i$, so that $u_{i,h,\lambda} \in S_C$, by setting

$$u_{i,h,\lambda}(a,e) = u_i(a,e) + \lambda \rho_{i,h}(a) \left( 1 - \frac{1}{\delta_2} \|a_{i,h} - a\|^2 \right)$$

(Euclidean norm)

for each \((a,e) \in \text{dom } u_i = A_{u_i} \times E(\nu)\). (Clearly \(a\), \(b\), and \(c\) in the assumption on payoff functions are satisfied by $u_{i,h,\lambda}$, and since $\lambda \rho_{i,h}(a) \left( 1 - \frac{1}{\delta_2} \|a_{i,h} - a\|^2 \right) \geq 0$ for each $a \in A_{u_i}$, and $u_{i,h,\lambda}(a,e) = u_i(a,e)$ if $a \in \partial A_{u_i}$, so is \(d\).) Note that for each \((i,h,\lambda)\), the fact that $a_{i,h} \in \varphi(u_i,e(\tau_A))$, which in particular means that $D^2 u_i(a_{i,h},e(\tau_A))$ is negative semi-definite, implies that $\varphi(u_{i,h,\lambda},e(\tau_A)) = \{a_{i,h}\}$ and that

$$D^2 u_{i,h,\lambda}(a_{i,h},e(\tau_A)) = D^2 u_i(a_{i,h},e(\tau_A)) - \lambda \frac{2}{\delta_2} I$$

is negative definite, as before writing $I$ for the $(m \times m)$-identity matrix. Now for each $0 < \lambda < 1$, define $\nu_\lambda \in \mathcal{G}'$ by setting $\nu_\lambda = \sum_{i=1}^k \alpha_i \sum_{h=1}^{h_i} \alpha_{i,h} \delta_{i,h,\lambda}$ and set $\tau_\lambda = \sum_{i=1}^k \alpha_i \sum_{h=1}^{h_i} \alpha_{i,h} \delta_{i,h,\lambda}$. Note that $\int g(a) \, d\tau_{\lambda,A}(a) = \int g(a) \, d\tau_A(a)$, or, in other words, that $e(\tau_{\lambda,A}) = e(\tau_A)$. It follows that for each $0 < \lambda < 1$, $\tau_\lambda$ is an equilibrium distribution for $\nu_\lambda$ such that (i) and (ii) from (b) are satisfied. Making $\lambda$ small, we get $\nu_\lambda$ as close to $\nu$ as we please (use Lemma 2(ii)$\Rightarrow$(i)). Thus, setting

$$\mathcal{G}_2 = \{ \nu \in \mathcal{G}': \nu \text{ has an equilibrium distribution satisfying (i) and (ii) of (b)} \},$$

$\mathcal{G}_2$ is dense in $\mathcal{G}'$, therefore dense in $\mathcal{G}_{S_C}$.

Let $\nu \in \mathcal{G}_2$ and let $\tau$ be an equilibrium distribution for $\nu$ such that (i) and (ii) of (b) are satisfied. We can write $\nu = \sum_{i=1}^k \alpha_i \delta_{u_i}$ and $\tau = \sum_{i=1}^k \alpha_i \delta_{u_i,a_i}$, where $\alpha_i > 0$ for each $i = 1, \ldots, k$ and $\sum_{i=1}^k \alpha_i = 1$. Consider $i = 1$ and note that $a_1 \in \text{int } A_{u_1}$. Let $(W_n)$ be a non-increasing sequence of compact convex neighborhoods of $a_1$ in $A_{u_1}$ such that $\bigcap_{n=1}^\infty W_n = \{a_1\}$. By Lemma 10(b), $\text{int } \text{co}(W_n)$ is non-empty for each $n$. For each $n$ fix a point $e_{1,n} \in \text{int } \text{co}(W_n)$. Using Caratheodory’s theorem, for each $n$ we can find points $a_{1,n,h} \in W_n$ such that $e_{1,n} = \sum_{h=1}^{m+1} \beta_{n,h} g(a_{1,n,h})$ for some numbers $\beta_{n,h}$ with $\beta_{n,h} \geq 0$ and $\sum_{h=1}^{m+1} \beta_{n,h} = 1$. Note that $E(\nu) = \sum_{i=1}^k \alpha_i \text{co}(A_{u_i})$. Thus, setting $e_n = \alpha_1 e_{1,n} + \sum_{i=2}^k \alpha_i g(a_i)$ we have $e_n \in \text{int } E(\nu)$ for each $n$. Also, $e_n \to e(\tau_A)$, by continuity of $g$, because $a_{1,n,h} \to a_1$ for each $h$ if $n \to \infty$, by choice of the points $a_{1,n,h}$.

Now for each $i = 1, \ldots, k$, we have $a_i \in \text{int } A_{u_i}$. We can therefore find numbers $0 < r_1 < r_2$ such that $B(a_i, r_1) \subseteq B(a_i, r_2) \subseteq \text{int } A_{u_i}$ for each $i = 1, \ldots, k$. For each $i$, let $\rho_i : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable map such that $\rho_i(a) = 1$ if $a \in B(a_i, r_1)$, $0 \leq \rho_i(a) \leq 1$ for all $a \in \mathbb{R}^n$, and $\rho_i(a) = 0$ if $a \notin B(a_i, r_2)$. As $a_{1,n,h} \to a_1$ for each $h = 1, \ldots, m + 1$ if $n \to \infty$, we can assume for each $n$ and each $h$ that $a_{1,n,h} \in B(a_1, r_1)$ and that $a + \rho_1(a)(a_1 - a_{1,n,h}) \in \text{int } A_{u_1}$ whenever $a \in B(a_1, r_2)$. For each $i = 1, \ldots, k$, let $\tilde{E}_i$ be an open set in $\mathbb{R}^m$, including $E(\nu)$, and
\( \tilde{u}_i \) is a twice continuously differentiable map which coincides with \( u_i \) on \( \text{int} \, A_{u_i} \times \tilde{E}_i \rightarrow \mathbb{R} \). Since \( e_n \rightarrow e(\tau_A) \) and \( E(\nu) \) is compact, we can assume that for all \( n, e - \rho_i(a)(e_n - e(\tau_A)) \in \tilde{E}_i \) for each \( i = 1, \ldots, k \) whenever \( a \in A_{u_i} \) and \( e \in E(\nu) \).

We can now define a function \( u_{1,n,h} : A_{u_1} \times E(\nu) \rightarrow \mathbb{R} \) for each \( h = 1, \ldots, m + 1 \) and each \( n \) by setting

\[
 u_{1,n,h}(a,e) = \begin{cases} 
 \tilde{u}_1(a + \rho_1(a)(a_1 - a_{1,n,h}), e - \rho_1(a)(e_n - e(\tau_A))) & \text{if } a \in \text{int} \, A_{u_1} \\
 u_1(a,e) & \text{if } a \in \partial A_{u_1}, 
\end{cases}
\]

and for each \( i = 2, \ldots, k \) and each \( n \), we can define a function \( u_{i,n,h} : A_{u_i} \times E(\nu) \rightarrow \mathbb{R} \) by setting

\[
 u_{i,n,h}(a,e) = \begin{cases} 
 \tilde{u}_i(a,e - \rho_i(a)(e_n - e(\tau_A))) & \text{if } a \in \text{int} \, A_{u_i} \\
 u_i(a,e) & \text{if } a \in \partial A_{u_i}. 
\end{cases}
\]

Then all these functions are continuous on their domains. Moreover, for each \( n \), \( u_{1,n,h} \) is twice continuously differentiable on \( \text{int} \, A_{u_1} \times E(\nu) \) for each \( h = 1, \ldots, m + 1 \), and so is \( u_{i,n} \) on \( \text{int} \, A_{u_i} \times E(\nu) \) for each \( i = 2, \ldots, k \). Further, for each \( i = 1, \ldots, k \) and each \( h = 1, \ldots, m + 1 \), since \( \partial A_{u_i} \) and \( E(\nu) \) are compact, we can assume for each \( n \) that whenever \( a \in \partial A_{u_i} \) and \( e \in E(\nu) \), then for some \( a' \in A_{u_i} \), \( u_{i,n,h}(a,e) < u_{i,n,h}(a',e) \) if \( i = 1 \) and \( u_{i,n}(a,e) < u_{i,n}(a',e) \) if \( i \neq 1 \) respectively; otherwise, the \( u_i \)'s could not satisfy (d) of the assumptions on payoff functions, as can be seen arguing by contradiction. Thus \( u_{1,n,h} \in \mathcal{U} \) for each \( h = 1, \ldots, m + 1 \) and each \( n \), as well as \( u_{i,n} \in \mathcal{U} \) for each \( n \) and each \( i = 2, \ldots, k \).

By Lemma 2(ii) \( \Rightarrow \) (i), if \( n \rightarrow \infty \), then \( u_{1,n,h} \rightarrow u_1 \) for each \( h = 1, \ldots, m + 1 \), and \( u_{i,n} \rightarrow u_i \) for each \( i = 2, \ldots, k \). Thus, setting \( \nu_n = \alpha_1 \sum_{h=1}^{m+1} \beta_{n,h} \delta_{u_{1,n,h}} + \sum_{i=2}^{k} \alpha_i \delta_{u_{i,n}} \), we have \( \nu_n \rightarrow \nu \). Set \( \tau_n = \alpha_1 \sum_{h=1}^{m+1} \beta_{n,h} \delta_{u_{1,n,h}}(a_{1,n,h}) + \sum_{i=2}^{k} \alpha_i \delta_{u_{i,n}}(a_{i,n}) \). Then, for large \( n \), \( \tau_n \) is an equilibrium distribution for \( \nu_n \), satisfying (i)–(iii) of (b). To see this, note first that \( e(\tau_n,\nu_n) = e_n \) for each \( n \). Next, recall that \( \tau \) satisfies (ii) of (b). Therefore, for any \( h = 1, \ldots, m + 1 \), as \( u_{1,n,h} \rightarrow u_1 \) and \( e(\tau_n,\nu_n) \rightarrow e(\tau_n) \), there is a compact neighborhood \( U_h \) of \( a_1 \) in \( \text{int} \, A_{u_1} \) such that \( D^2 u_{1,n,h}(a, e(\tau_n,\nu_n)) \) is negative definite for each \( a \in U_h \) if \( n \) is large enough. Similarly, for each \( i = 2, \ldots, k \), there is a compact neighborhood \( U_i \) of \( a_i \) in \( \text{int} \, A_{u_i} \) such that \( D^2 u_{i,n}(a, e(\tau_n,\nu_n)) \) is negative definite for each \( a \in U_i \) if \( n \) is large enough. As \( \tau \) also satisfies (i) of (b), and \( a_{1,n,h} \rightarrow a_{1} \in \text{int} \, A_{u_1} \) for each \( h = 1, \ldots, m + 1 \), we can now see, using the facts noted in the previous two sentences, that if \( n \) is large enough, then \( \tau_n \) is an equilibrium distribution for \( \nu_n \) satisfying (i) and (ii) of (b). As \( e(\tau_n,\nu_n) = e_n \in \text{int} \, E(\nu) \) for each \( n \) by construction, we have (iii) of (b) in addition, and it follows that \( \mathcal{G}_1 \cap \mathcal{G}_2 \) is dense in \( \mathcal{G}_2 \), therefore dense in \( \mathcal{G}_{SC} \).

Pick any \( \nu \in \mathcal{G}_1 \cap \mathcal{G}_2 \) and let \( \tau \) be an equilibrium distribution of \( \nu \) such that (i)–(iii) from (b) are satisfied. Write \( \nu = \sum_{i=1}^{h} \alpha_i \delta_{\tilde{a}_i} \), let \( \tilde{a}_1, \ldots, \tilde{a}_h \) be the optimal actions, and
write \( \bar{e} = e(\tau_A) \). Let \( \xi_\tau \) correspond to \( \tau \) as in (b). Observe that \( D\xi_\tau(\bar{e}) \) is of the form

\[
\sum_{i=1}^{h} A_i B_i^{-1} C_i - I
\]

where \( I \) is the \( m \times m \)-unit matrix, and for each \( i \), \( A_i \) is an \( m \times n \)-matrix, \( B_i \) an \( n \times n \)-matrix, and \( C_i \) an \( n \times m \)-matrix. If \( \det(\sum_{i=1}^{h} A_i B_i^{-1} C_i - I) \neq 0 \), then \( \nu \in \mathcal{G}^* \). Otherwise, pick any \( 0 < \lambda < 1 \). For each \( i \) replace \( u_i \) by \( u_{i,\lambda} \), defined by setting \( u_{i,\lambda}(a,e) = u_i(a, (1-\lambda)\bar{e} + \lambda e) \) for each \( (a,e) \in A_{u_i} \times E(\nu) \), so that, in particular, \( \text{dom} u_{i,\lambda} = \text{dom} u_i \) and thus \( u_{i,\lambda} \in \mathcal{S}_C \). Then \( \nu_\lambda = \sum_{i=1}^{h} \alpha_i \delta_{u_{i,\lambda}} \) belongs to \( \mathcal{G}' \) and has an equilibrium distribution \( \tau_\lambda \) with the same optimal actions \( \bar{a}_1, \ldots, \bar{a}_h \) and with \( e(\tau_{\lambda,A}) = \bar{e} \). In particular, (i)–(iii) of (b) are satisfied for \( \tau_\lambda \). Thus \( \nu_\lambda \in \mathcal{G}_1 \cap \mathcal{G}_2 \). Now \( D\xi_\tau(\bar{e}) \) equals

\[
\lambda \sum_{i=1}^{h} A_i B_i^{-1} C_i - I,
\]

with the same \( A_i, B_i, C_i \) as above. Because the characteristic polynomial of the matrix \( \sum_{i=1}^{h} A_i B_i^{-1} C_i \) can have only finitely many zeros, we have \( \det(\sum_{i=1}^{h} A_i B_i^{-1} C_i - I) \neq 0 \) for all sufficiently large \( 0 < \lambda < 1 \) and hence \( \det(\lambda \sum_{i=1}^{h} A_i B_i^{-1} C_i - I) \neq 0 \) for such \( \lambda \).

Thus we have \( \nu_\lambda \in \mathcal{G}^* \) for all sufficiently large \( 0 < \lambda < 1 \). Moreover, making \( 0 < \lambda < 1 \) large, we get \( \nu_\lambda \) as close to \( \nu \) as we please. Thus \( \mathcal{G}' \cap \mathcal{G}_1 \cap \mathcal{G}_2 \) is dense in \( \mathcal{G}_1 \cap \mathcal{G}_2 \), therefore dense in \( \mathcal{G}_{SC} \) by the previous paragraph. As \( \mathcal{G}_2 \subseteq \mathcal{G}_{SC} \) we conclude that \( \mathcal{G}' \cap \mathcal{G}_{SC} \) is dense in \( \mathcal{G}_{SC} \).

(f) 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 \( I_n = \{1, \ldots, k_n\} \), where \( k_n = \#(I_n) \). Also, for any map \( f : I_n \to A \), and any \( i \in I_n \), we write \( \tau_{A,f,i} \) for the probability measure on \( A \) given by setting \( \tau_{A,f,i}(B) = \#\{j \in I_n : f(j) \in B\}/\#(I_n) \) for each Borel set \( B \subseteq A \), and \( \tau_{A,f,i} \) for the probability measure on \( A \) which is given by setting \( \tau_{A,f,i}(B) = \#\{j \in I_n \setminus \{i\} : f(j) \in B\}/(\#(I_n) - 1) \) for each Borel set \( B \subseteq A \).

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

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

As \( \nu \in \mathcal{G}^* \), the derivative of \( \xi_\tau \) 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 \( W_2 \subseteq \text{int} W_1 \).

Now because \( \nu_{G_n} \to \nu \), and therefore \( \rho_H(\text{supp}(\nu_{G_n}), \text{supp}(\nu)) \to 0 \), and by what was noted in the second paragraph of this part of the proof, there is an \( N \in \mathbb{N} \) such
that for \( n \geq N \), \( \text{supp}(\nu_{G_n}) \subseteq V \) and for each \( f \in A^{I_n} \) and \( i = 1 \ldots, k_n \),
\[
\int g(a) \, d\tau_{A,f,i}(a) - \int g(a) \, d\tau_{A,f}(a) + e \in W_1
\]
whenever \( e \in W_2 \). For \( n \geq N \), consider the function \( \Lambda : A^{I_n} \times W_2 \to A^{I_n} \times \mathbb{R}^m \) defined by setting
\[
\begin{align*}
\Lambda(f, e) &= \left(h_1\left(\int g(a) \, d\tau_{A,f,1}(a) - \int g(a) \, d\tau_{A,f}(a) + e\right), \ldots, \\
&\quad h_{k_n}\left(\int g(a) \, d\tau_{A,f,k_n}(a) - \int g(a) \, d\tau_{A,f}(a) + e\right), \int g(a) \, d\tau_{A,f}(a)\right),
\end{align*}
\]
writing \( h_i(\cdot) \) in place of \( h(G(i), \cdot) \) for each \( i \in \{1, \ldots, k_n\} \). Then a fixed point of \( \Lambda \) is 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 6 applies to \( \Lambda \). Clearly \( \Lambda \) is continuous, and for \( X = A^{I_n} \) and \( Y = W_2 \) the requirements of Theorem 6 on \( X \) and \( Y \) are satisfied. With the map \( \xi_\tau \) it is also clear that we have (a) of Theorem 6. 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 \( \Lambda \) and \( \xi_\tau \) if \( n \geq N_1 \).

To this end, fix \( n \geq N \) and suppose that \( f \in A^{I_n} \) and \( e \in \partial W_2 \) are such that
\[
f = \left(h_1\left(\int g(a) \, d\tau_{A,f,1}(a) - \int g(a) \, d\tau_{A,f}(a) + e\right), \ldots, \\
&\quad h_{k_n}\left(\int g(a) \, d\tau_{A,f,k_n}(a) - \int g(a) \, d\tau_{A,f}(a) + e\right)\right).
\]
Note that
\[
\left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - e - \xi_\tau(e) \right\| = \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - e - \left(\int g(h(u,e)) \, d\nu(u) - e\right) \right\|
\]
\[
= \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - \int g(h(u,e)) \, d\nu(u) \right\|
\]
\[
\leq \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - \frac{1}{k_n} \sum_{i=1}^{k_n} g(h_i(e)) \right\| + \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(h_i(e)) - \int g(h(u,e)) \, d\nu(u) \right\|
\]
\[
= \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - \frac{1}{k_n} \sum_{i=1}^{k_n} g(h_i(e)) \right\| + \left\| \int g(h(u,e)) \, d\nu_{G_n}(u) - \int g(h(u,e)) \, d\nu(u) \right\|
\]
Since \( \text{supp}(\nu) \) is compact, there is a neighborhood \( V_1 \subseteq V \) of \( \text{supp}(\nu) \) and a \( \delta > 0 \) such that for \( z \in \mathbb{R}^m \) with \( e + z \in W \) we have \( \|g(h(u,e+z)) - g(h(u,e))\| < \gamma/3 \) whenever \( u \in V_1, e \in W_1 \), and \( \|z\| < \delta \), and by what was noted earlier, if \( n \) is large we must have
\[ \left\| \int g(a) \, d\tau_{A,f,i}(a) - \int g(a) \, d\tau_{A,f}(a) \right\| < \delta \text{ for each } f \in A^I \text{ and each } i \in I_n. \]

Hence, since \( \rho_H(\text{supp}(\nu_{G_n}), \text{supp}(\nu)) \to 0 \), we have \( \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - \frac{1}{k_n} \sum_{i=1}^{k_n} g(h_i(e)) \right\| < \gamma/3 \) for large \( n \) whenever \( f \in A^I \) and \( e \in \partial W_2 \) are as above. On the other hand, 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 \( \partial W_2 \) such that for large \( n \) we have \( \left\| \int g(h(u,e')) \, d\nu_{G_n}(u) - \int g(h(u,e)) \, d\nu(u) \right\| < (2\gamma)/3 \) for each \( e' \in U \). Thus, since \( \partial W_2 \) is compact, if \( n \) is large then \( \left\| \int g(h(u,e)) \, d\nu_{G_n}(u) - \int g(h(u,e)) \, d\nu(u) \right\| < (2\gamma)/3 \) for every \( e \in \partial W_2 \). It follows that for some \( N_1 \geq N, \left\| \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) - e - \xi_{\tau}(e) \right\| < \gamma \) for \( n \geq N_1 \) whenever \( f \in A^I \) and \( e \in \partial W_2 \) are as above. Consequently, because \( \frac{1}{k_n} \sum_{i=1}^{k_n} g(f(i)) = \int g(a) \, d\tau_{A,f}(a) \), (b) of Theorem 6 is satisfied if \( n \geq N_1 \).

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

### 3.4 Proof of Theorem 2

**Theorem 2.** Let \( C \in C \). Every \((\nu, \tau)\), where \( \nu \in \mathcal{G}_{SC} \) is a continuum game and \( \tau \) is an equilibrium distribution for \( \nu \), is asymptotically implementable by a sequence \( \langle (I_n, G_n) \rangle_{n \in \mathbb{N}} \) of finite-player games such that \( \nu_{G_n} \in \mathcal{G}_{SC} \) for each \( n \). (Because \( \emptyset \in C \), this implies that every \((\nu, \tau)\), where \( \nu \in \mathcal{G} \) is any continuum game and \( \tau \) is an equilibrium distribution for \( \nu \), is asymptotically implementable by a sequence \( \langle (I_n, G_n) \rangle_{n \in \mathbb{N}} \) of finite-player games.)

**Proof.** (a) We first show that if \( \nu \in \mathcal{G}_{SC} \) and \( \tau \) is an equilibrium distribution for \( \nu \), then there exists a sequence \( \langle \hat{\nu}_n \rangle \) in \( \mathcal{G}_{SC} \), along with a sequence \( \langle \hat{\tau}_n \rangle \) of corresponding equilibrium distributions, such that (i) \( \text{supp} \langle \hat{\tau}_n \rangle \), and hence also \( \text{supp} \langle \hat{\nu}_n \rangle \), is finite for each \( n \), (ii) \( \hat{\tau}_n \to \tau \) narrowly, and (iii) \( \hat{\nu}_n \to \nu \).

To see this, fix \( \nu \in \mathcal{G}_{SC} \) and let \( \tau \) be an equilibrium distribution for \( \nu \). Write \( Y \) for \( \text{supp} \langle \nu \rangle \) and note that \( Y \) is a compact metric space in the subspace topology defined from \( \mathcal{U} \). Write \( \hat{\theta} \) for the correspondence \( u \mapsto A_u : Y \to 2^{\mathbb{R}^n} \) and note that \( \hat{\theta} \) is continuous and that \( C \subseteq \theta(u) \) for all \( u \in Y \).

(1) Since \( Y \) is compact and the correspondence \( \theta \) is continuous, with compact and convex values, all with non-empty interior, we can find a continuous correspondence \( \theta'_n : Y \to 2^{\mathbb{R}^n} \) for each \( n \in \mathbb{N} \) with compact and convex values, so that for each \( u \in Y \), \( \text{int} \theta'_n(u) \neq \emptyset \), \( \theta'_n(u) \subseteq \text{int} \theta(u) \), and \( \rho_H(\theta(u), \theta'_n(u)) \leq 1/(n+1) \) (see Lemma 13 in the appendix). Using Lemmata 2 and 3, together with the fact that \( Y \) is compact, we see from (d) of the definition of payoff functions that there must be an \( n_1 \in \mathbb{N} \) such that if \( n \geq n_1 \) then for all \( u \in Y \) and all \( e \in E(\nu) \), whenever \( a \in \theta(u) \setminus \text{int} \theta'_n(u) \) we have \( u(a', e) > u(a, e) \) for some \( a' \in \theta'_n(u) \). Thus if \( n \geq n_1 \), then for each \( u \in Y \) the restriction of \( u \) to \( \theta'_n(u) \times E(\nu) \) is in \( \mathcal{U} \), and \( (u, a) \in \text{supp} \langle \tau \rangle \) implies \( a \in \theta'_n(u) \) because \( \tau \) is an equilibrium distribution for \( \nu \). We may assume for the rest of this proof that \( n_1 = 0 \).

Now for each \( n \) and each \( u \in Y \), we can find a compact set \( K'_{n,u} \) in \( \mathbb{R}^n \) such that \( \theta'_n(u) \subseteq \text{int} \ K'_{n,u} \) and \( K'_{n,u} \subseteq \text{int} \theta(u) \), because \( \theta'_n(u) \) is compact and \( \theta'_n(u) \subseteq \text{int} \theta(u) \). As \( \theta \) and \( \theta'_n \) are continuous, it follows that for each \( u \in Y \) and each \( n \) there is an open
neighborhood \(O_{n,u}\) of \(u\) in \(Y\) such that \(\theta_n'(u') \subseteq K_{n,u}' \subseteq \theta(u')\) for each \(u' \in O_{n,u}\) (for the second inclusion, use Lemma 4). We can assume that \(\text{diam}(O_{n,u}) \leq 1/(n+1)\) for each \(O_{n,u}\). As \(Y\) is compact, it follows that for each \(n\) there is a finite partition \(\mathcal{P}_n\) of \(Y\) into Borel sets \(B_{n,i}, i = 1, \ldots, m_n\), such that for each \(i\), \(\text{diam}(B_{n,i}) \leq 1/(n+1)\) and for some compact set \(K_{n,i}'\) in \(\mathbb{R}^n\) we have \(\theta_n'(u') \subseteq K_{n,i}' \subseteq \theta(u')\) for each \(u \in B_{n,i}\).

For each \((n,i), n \in \mathbb{N}, i = 1, \ldots, m_n\), set \(K_{n,i} = \text{co}(K_{n,i}' \cup C)\), and for each \(n\) define a correspondence \(\theta_n : Y \rightarrow 2^{\mathbb{R}^n}\) by setting \(\theta_n(u) = \text{co}(\theta_n'(u) \cup C)\) for each \(u \in Y\). Then, for each \(n, C \subseteq \theta_n(u) \subseteq K_{n,i} \subseteq \theta(u)\) for each \(u \in B_{n,i}, i = 1, \ldots, m_n\). Evidently, for each \(n\), the correspondence \(u \mapsto \theta_n(u) \cup C : Y \rightarrow 2^{\mathbb{R}^n}\) is continuous, so by Hildenbrand (1974, B.III, Propositions 6 and 10) the same is true of \(\theta_n\). Observe also that \(\rho_H(\theta(u), \theta_n(u)) \leq 1/(n+1)\) each \(u \in Y\) and each \(n\).

Note that whenever \(\langle E_n \rangle\) is a sequence of non-empty compact subsets of \(E(\nu)\) such that \(\rho_H(E(\nu), E_n) \rightarrow 0\), and \(\langle u_n \rangle\) is a sequence in \(\mathcal{U}\) such that \(\rho(u, u_n) \rightarrow 0\), then, writing \(u_n'\) for the restriction of \(u_n\) to \(\theta_n(u) \times E_n\), we have \(\rho(u, u_n') \rightarrow 0\) (see Lemma 2).

(2) For each \((n,i), n \in \mathbb{N}, i = 1, \ldots, m_n\), let \(\tau_{n,i}\) be the Borel measure on \(Y \times A\) given by setting \(\tau_{n,i}(B) = \tau(B \cap (B_{n,i} \times A))\) for each Borel set \(B\) in \(Y \times A\). Note that \(\text{supp}(\tau_{n,i}) \subseteq \text{supp}(\tau)\) for each \((n,i)\). By Lemma 12 in the appendix, with \(g \circ \text{proj}_A\) for \(h\), for each \((n,i)\) there is a sequence \(\langle \tau_{n,i,k} \rangle_{k \in \mathbb{N}}\) of Borel measures on \(Y \times A\) such that \(\tau_{n,i,k} \rightarrow \tau_{n,i}\) narrowly, and for each \((n,i,k), \tau_{n,i,k}(Y \times A) = \tau_{n,i}(Y \times A)\), \(\text{supp}(\tau_{n,i,k})\) is a finite subset of \(\text{supp}(\tau_{n,i})\), and

\[
\int g \circ \text{proj}_A d\tau_{n,i,k} = \int g \circ \text{proj}_A d\tau_{n,i}.
\]

For each \((n,k)\) set \(\tau_{n,k} = \sum_{i=1}^{m_n} \tau_{n,i,k}\). Then for each \((n,k)\), \(\tau_{n,k}\) is a probability measure on \(Y \times A\), and for each fixed \(n\) we have both \(\tau_{n,k} \rightarrow \tau\) narrowly as \(k \rightarrow \infty\) and \(\tau_{n,k}(B_{n,i} \times A) = \tau(B_{n,i} \times A)\) for each \(k\) and each element \(B_{n,i}\) of the partition \(\mathcal{P}_n\). Moreover, for each \((n,k)\), \(\text{supp}(\tau_{n,k})\) is finite, and we have

\[
\text{supp}(\tau_{n,k}) \subseteq \bigcup_{i=1}^{m_n} \text{supp}(\tau_{n,i,k}) \subseteq \bigcup_{i=1}^{m_n} \text{supp}(\tau_{n,i}) \subseteq \text{supp}(\tau)
\]

as well as

\[
\int g \circ \text{proj}_A d\tau_{n,k} = \sum_{i=1}^{m_n} \int g \circ \text{proj}_A d\tau_{n,i,k} = \sum_{i=1}^{m_n} \int g \circ \text{proj}_A d\tau_{n,i} = \sum_{i=1}^{m_n} \int_{B_{n,i} \times A} g \circ \text{proj}_A d\tau = \int g \circ \text{proj}_A d\tau = e(\tau_A).
\]

Now we have \(\text{diam}(B_{n,i}) \leq 1/(n+1)\) for each \(n\) and each \(B_{n,i} \in \mathcal{P}_n\), so for each \(n\) we can pick an element \(\tau_n\) from the sequence \(\langle \tau_{n,k} \rangle_{k \in \mathbb{N}}\) so as to get a sequence \(\langle \tau_n \rangle_{n \in \mathbb{N}}\) such that both \(\tau_n \rightarrow \tau\) narrowly and \(\text{diam}(B_{n,i}) \leq 1/(n+1)\) for each \(n\) and \(B_{n,i} \in \mathcal{P}_n\). For each \(n\) write \(\nu_n\) for the marginal measure of \(\tau_n\) on \(Y\) and note that \(\nu_n \rightarrow \nu\) narrowly, because the marginal measure of \(\tau\) is \(\nu\). By construction, we have \(\text{supp}(\tau_n) \subseteq \text{supp}(\tau)\).
for each $n$, therefore $\text{supp}(\nu_n) \subseteq \text{supp}(\nu)$ for each $n$. From this and the fact that $\nu_n \to \nu$ narrowly, it follows that $\rho_H(\text{supp}(\nu), \text{supp}(\nu_n)) \to 0$. Note also that we have $\nu_n(B_{n,i}) = \nu(B_{n,i})$ for each $B_{n,i} \in \mathcal{P}_n$ and each $n$.

(3) For each $n$ write $R_n : Y \to \mathcal{U}$ for the restriction operator $u \mapsto u \upharpoonright (\theta_n(u) \times E(\nu))$. Then $R_n$ is continuous for each $n$, so $\nu'_n = \nu_n \circ R_n^{-1}$ is defined as an element of $\mathcal{M}$ for each $n$. By the last paragraph of (1), the sequence $\langle R_n \rangle$ converges uniformly to the identity on $Y$, from which we see both that $\nu'_n \to \nu$ narrowly and that $\rho_H(\text{supp}(\nu_n), \text{supp}(\nu'_n)) \to 0$. The latter and the fact that $\rho_H(\text{supp}(\nu), \text{supp}(\nu_n)) \to 0$ imply that $\rho_H(\text{supp}(\nu), \text{supp}(\nu'_n)) \to 0$. Thus we have $\nu'_n \to \nu$ in the topology of $\mathcal{M}$. In particular, by Lemma 5, $\rho_H(\nu', \nu) \to 0$.

We claim that $E(\nu'_n) \subseteq E(\nu)$ for each $n$. To see this, note that the definition of $\nu'_n$ and the fact that $\nu_n$, and therefore also $\nu'_n$, has finite support imply that

$$
\sum_{u \in \text{supp}(\nu'_n)} \nu'_n(\{u\}) \cos g(\theta_n(u)) = \sum_{u \in \text{supp}(\nu_n)} \nu_n(\{u\}) \cos g(\theta_n(u)).
$$

Thus by the construction in (2), and the argument in the proof of Lemma 5,

$$
E(\nu'_n) = \int_{\text{supp}(\nu'_n)} \cos g(\theta_n(u)) \, d\nu'_n(u)
= \sum_{u \in \text{supp}(\nu'_n)} \nu'_n(\{u\}) \cos g(\theta_n(u))
= \sum_{u \in \text{supp}(\nu_n)} \nu_n(\{u\}) \cos g(\theta_n(u))
\subseteq \sum_{i=1}^{m_n} \left( \sum_{u \in B_{n,i} \cap \text{supp}(\nu_n)} \nu_n(\{u\}) \right) \cos g(K_{n,i})
= \sum_{i=1}^{m_n} \nu_n(B_{n,i}) \cos g(K_{n,i})
= \sum_{i=1}^{m_n} \nu(B_{n,i}) \cos g(K_{n,i}) \subseteq \int_Y \cos g(\theta(u)) \, d\nu(u) = E(\nu),
$$

establishing the claim above.

We can now define an operator $S_n : Y \to \mathcal{U}$ by setting $S_n(u) = u \upharpoonright (\theta_n(u) \times E(\nu'_n))$. Clearly $S_n$ is continuous for each $n$, so we can set $\hat{\nu}_n = \nu_n \circ S_n^{-1}$. Then, for each $n$, $E(\hat{\nu}_n) = E(\nu'_n)$, and thus $\hat{\nu}_n \in \mathcal{G}_{\mathcal{C}}$.

(4) (a) For each $n$ set $T_n = S_n \times id_A$, where $id_A$ is the identity on $A$. Define a probability measure $\hat{\tau}_n$ on $\mathcal{U} \times A$ for each $n$ by setting $\hat{\tau}_n = \tau_n \circ T_n^{-1}$. Evidently $\hat{\tau}_{Y,n} = \hat{\nu}_n$ for each $n$. As $Y$ and $A$ are compact, we see from the last paragraph of (1) that the sequence $\langle T_n \rangle$ converges uniformly to the identity on $Y \times A$; in particular,
narrowly, and that \( \hat{\nu}_n \to \nu \) in the topology of \( \mathcal{M} \). Because \( \text{supp}(\tau_n) \) is finite for each \( n \), so is \( \text{supp}(\hat{\nu}_n) \) for each \( n \). Evidently \( \hat{\tau}_{A,n} = \tau_{A,n} \) for each \( n \), so \( e(\hat{\tau}_{A,n}) = e(\tau_A) \) for each \( n \), by the construction of \( \tau_n \) in (2). Finally, note that if \( (\hat{u}, a) \in \text{supp}(\hat{\tau}_n) \), then there is a \( u \in Y \) such that \((u, a) \in \text{supp}(\tau_n) \) and \( \hat{u} \) is the restriction of \( u \) to \( \theta_n(u) \times \hat{\nu}_n \).

Because \( \text{supp}(\tau_n) \subseteq \text{supp}(\tau) \) and \( e(\hat{\tau}_{A,n}) = e(\tau_A) \) it follows that \( \hat{\tau}_n \) is an equilibrium distribution for \( \hat{\nu}_n \).

(b) Let \( G^* \) be defined as in the proof of Theorem 1. Inspecting (e) in that proof we see that whenever \( \nu \in G_{SC}^* \) has finite support and \( \tau \) is an equilibrium distribution for \( \nu \) such that \( \text{supp}(\tau) \) is also finite, then there is a sequence \( \langle \nu_n \rangle \) in \( G^* \cap G_{SC} \) such that \( \nu_n \to \nu \), and a sequence \( \langle \tau_n \rangle \) of corresponding equilibrium distributions such that \( \tau_n \to \tau \) narrowly and each \( \tau_n \) satisfies the requirements listed in (c) of the proof of Theorem 1 (note that if \( \text{supp}(\tau) \) is finite, then, in (e) of the proof of Theorem 1, the step involving Caratheodory’s theorem can be omitted). Putting this fact together with (a) shows that for any \( \nu \in G_{SC}^* \) and any equilibrium distribution \( \tau \) for \( \nu \) there is a sequence \( \langle \nu_n \rangle \) in \( G^* \cap G_{SC} \) such that \( \nu_n \to \nu \), and a sequence \( \langle \tau_n \rangle \) of corresponding equilibrium distributions such that \( \tau_n \to \tau \) narrowly and each \( \tau_n \) satisfies the requirements in (c) of the proof of Theorem 1.

(c) Fix any \( \nu \in G^* \cap G_{SC} \) and let \( \tau \) be an equilibrium distribution for \( \nu \) 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 \#(\( I_n \)) \( \to \infty \), \( \nu_{G_n} \in \mathcal{M}_{SC} \) for each \( n \), and \( \nu_{G_n} \to \nu \). Let \( V, W, W_1 \), and \( h : V \times W \to A \) be as in (f) of the proof of Theorem 1. Observe that \( \tau = \nu \circ (id_V \times h(\cdot, e(\tau_A)))^{-1} \). Let \( \langle W_{2,k} \rangle \) be a non-increasing sequence of compact neighborhoods of \( e(\tau_A) \) in \( in E \) such that \( \bigcap_{k=0}^\infty W_{2,k} = \{ e(\tau_A) \} \). Instead with a fixed \( W_2 \), the argument in (f) of 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 \( N \), 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 \( f'_k : I_{n_k} \to A \) be the map defined by setting \( f'_k(i) = h(G_{n_k}(i), e_k) \) for each \( i \in I_{n_k} \). Let \( \tau_k \) be the distribution of \( G_{n_k} \times f_k \), and \( \tau'_k \) that of \( G_{n_k} \times f'_k \). As \( \nu_{G_{n_k}} \to \nu \), we can assume, considering only large \( k \), if necessary, that \( \text{supp}(\nu_{G_{n_k}}) \subseteq V \) for all \( k \), and therefore write \( \tau'_k = \nu_{G_{n_k}} \circ (id_V \times h(\cdot, e_k))^{-1} \) for all \( k \). Now because \( e_k \in W_{2,k} \) 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 \( id_V \times h(\cdot, e_k) \to id_V \times h(\cdot, e(\tau_A)) \), uniformly on compact subsets of \( V \), because \( h \) is continuous. Consequently

\[
\tau'_k = \nu_{G_{n_k}} \circ (id_V \times h(\cdot, e_k))^{-1} \to \nu \circ (id_V \times h(\cdot, e(\tau_A)))^{-1} = \tau,
\]

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

\[
\| h(u'_k, e_k + z_k) - h(u'_k, e_k) \| \to 0,
\]
In view of Theorem 1, it needs only be shown that ˜ν is an open dense subset of ˜G. By Lemma 9, given ν ∈ ˜G, a sequence ((In, Gn))n∈N of finite player games such that νGn → ν, νGn ∈ MSc for each n, and #(In) → ∞ does exist. Putting this fact together with (b) and (c) proves the theorem.

\[ \|h(G_{nk}(i), e_k + z_k(i)) - h(G_{nk}(i), e_k)\| \leq \epsilon' \]

for all i ∈ In, i.e., \( \|f_k(i) - f'_k(i)\| \leq \epsilon' \) for all i ∈ In, and thus, for some standard product metric ˜ρ on U × A,

\[ ˜\rho(((G_{nk}(i), f_k(i)), (G_{nk}(i), f'_k(i))) \leq \epsilon' \]

for all i ∈ In. In view of this, we can conclude, using Billingsley (1968, Theorem 4.1), that the fact that the sequence 〈τk〉 of distributions of the maps Gnk × f_k converges narrowly to τ implies that the sequence 〈τk〉 of distributions of the maps Gnk × f_k converges narrowly to τ, too.

(b) For each u ∈ S_C and each compact set K ⊆ A_u there is an ϵ > 0 such that u′ ∈ S_C if u′ ∈ U satisfies the following three conditions: (i) dom u′ = dom u, (ii) ˜ρ(u, u′) < ϵ, and (iii) u′(a, e) − u(a, e) = d(a) for all (a, e) ∈ dom u if d: A_u → R vanishes outside K.

Then there is an open dense subset GSc* of GSc such that if ν ∈ GSc* and 〈(In, Gn)〉n∈N is a sequence of finite-player games such that νGn → ν and #(In) → ∞, then there is an N ∈ N such that (In, Gn) has a pure strategy Nash equilibrium if #(In) ≥ N.

Proof. Let G* be defined as in the proof of Theorem 1 and apply (c) in the proof of Theorem 2.

3.5 Proof of Theorem 3

Theorem 3. There is an open subset G* of G such that, for every C ∈ C, G* ∩ GSc is dense in GSc and whenever ν ∈ G* ∩ GSc, then there is an equilibrium distribution τ for ν such that (ν, τ) is asymptotically robust in C.

Proof. Let G* be defined as in the proof of Theorem 1 and apply (c) in the proof of Theorem 2.

3.6 Proof of Theorem 4

Theorem 4. Let C ∈ C and let g be the identity on A. Let S_C be a subset of Sc, and GSc be the subset of GSc consisting of the elements ν with supp(ν) ∈ S_C. Suppose:

(a) GSc* ∩ GSc is dense in GSc, writing GSc* for the set of elements of GSc with finite support.

(b) For each u ∈ S_C and each compact set K ⊆ A_u there is an ϵ > 0 such that u′ ∈ S_C if u′ ∈ U satisfies the following three conditions: (i) dom u′ = dom u, (ii) ˜ρ(u, u′) < ϵ, and (iii) u′(a, e) − u(a, e) = d(a) for all (a, e) ∈ dom u if d: A_u → R vanishes outside K.

Then there is an open dense subset GSc* of GSc such that if ν ∈ GSc* and 〈(In, Gn)〉n∈N is a sequence of finite-player games such that νGn → ν and #(In) → ∞, then there is an N ∈ N such that (In, Gn) has a pure strategy Nash equilibrium if #(In) ≥ N.

Proof. Let G* ⊆ G be defined as in the proof of Theorem 1 and set GSc* = G* ∩ GSc. In view of Theorem 1, it needs only be shown that GSc* is dense in GSc.
Let \( G_1 \subseteq G \) be defined as in the proof of Theorem 1. Note for the following that because \( g \) is the identity on \( A \), (iii) of the definition of \( G_1 \) holds automatically for any \( \nu \) and any equilibrium distribution \( \tau \) of \( \nu \). (To see this, note that convexity of the \( A_u \)'s and \( g \) being the identity on \( A \) imply that \( E(\nu) = \int_A A_u \, d\nu(u) \) and that, by condition (d) in the definition of payoff functions, if \( \tau \) is an equilibrium distribution for \( \nu \), then \( (u, a) \in \text{supp}(\tau) \) implies that \( a \in \text{int} \, A_u \). Now use the facts that, for any \( p \in \mathbb{R}^n \), \( \max pE(\nu) = \int_A \max pA_u \, d\nu(u) \) and that for any compact convex set \( B \subseteq \mathbb{R}^n \), \( \max pB \) is attained at a boundary point of \( B \). For the former fact, see the proof of Lemma 5.)

Let \( \nu \in \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \) and write \( \nu \) in the form \( \nu = \sum_{i=1}^{k} \alpha_i \delta_{\nu_i} \) where \( \alpha_i > 0 \) for each \( i = 1, \ldots, k \). From the first part of (e) in the proof of Theorem 1 we see that there is a sequence \( \{\nu_n\} \) of elements of \( \mathcal{G}_1 \cap \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \) such that \( \nu_n \to \nu \) and for each \( n \), \( \nu_n \) is of the form \( \nu_n = \sum_{i=1}^{k} \alpha_i \sum_{h=1}^{h_n} \alpha_i \delta_{\nu_i + \lambda_n g_{i,h}} \) where \( \lambda_n \to 0 \) as \( n \to \infty \) and each \( g_{i,h} \) is a twice continuously differentiable non-negative real-valued function on \( A \) vanishing outside of some compact subset \( K \) of \( \text{int} \, A \) where \( K \) is independent of \( n \). Because the set of functions \( u_i \) and \( g_{i,h} \) which are involved is finite, the fact that \( \lambda_n \to 0 \) implies by condition (b) of the theorem that if \( n \) is large, then \( \nu_n \) belongs to \( \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \). It follows that \( \mathcal{G}_1 \cap \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \) is dense in \( \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \), therefore dense in \( \tilde{G}_{SC} \) by condition (a) of the theorem.

The rest of this proof is very similar to the argumentation in the last paragraph of (e) of the proof of Theorem 1. Let \( \nu = \sum_{i=1}^{k} \alpha_i \delta_{\nu_i} \in \mathcal{G}_1 \cap \mathcal{G}'_{SC} \cap \tilde{G}_{SC} \). Let \( \tau \) be an equilibrium distribution for \( \nu \), witnessing that \( \nu \in \mathcal{G}_1 \). Let \( a_1, \ldots, a_k \) be the corresponding optimal actions, and write \( \tilde{e} = e(\tau_\lambda) \). Then \( \tau \) can be written in the form \( \tau = \sum_{i=1}^{k} \alpha_i \delta_{\nu_i(a_i)} \), and we have \( \varphi(u_i, \tilde{e}) = \{\tilde{a}_i\} \) for each \( i \); moreover, \( \tilde{a}_i \in \text{int} \, A \) for each \( i \). Observe that in the present context, \( D\xi(\tilde{e}) \) is equal to

\[
\sum_{i=1}^{k} B_i^{-1} C_i - I = \sum_{i=1}^{k} (D^2 u_i(\tilde{a}_i, \tilde{e}))^{-1} (-\alpha_i D_e D_u u_i(\tilde{a}_i, \tilde{e})) - I,
\]

where \( I \) is the \((m \times m)\)-identity matrix. If \( \det(\sum_{i=1}^{k} B_i^{-1} C_i - I) \neq 0 \), then \( \nu \in \mathcal{G}^* \).

Otherwise, pick any \( 0 < \lambda < 1 \). Now for each \( i = 1, \ldots, k \), we have \( \tilde{a}_i \in \text{int} \, A_{u_i} \) and \( \text{dom} \, u_i = \text{dom} \, u_i \Lambda \) with \( \text{dom} \, u_i \Lambda = \text{dom} \, u_i \Lambda \). Then \( \nu_\lambda = \sum_{i=1}^{k} \alpha_i \delta_{u_i \Lambda(\tilde{a}_i, \tilde{e})} \) belongs to \( \mathcal{G}_{SC} \) and has an equilibrium distribution \( \tau_\lambda \) with the same optimal actions \( \tilde{a}_1, \ldots, \tilde{a}_h \) and with \( e(\tau_\lambda, A) = \tilde{e} \). In particular, \( \nu_\lambda \in \mathcal{G}_1 \cap \mathcal{G}_{SC} \). Now \( D\xi(\tilde{e}) \) equals

\[
\frac{1}{1 + \lambda} \sum_{i=1}^{k} B_i^{-1} C_i - I,
\]

with the same \( B_i, C_i \) as above. Because the characteristic polynomial of the matrix \( \sum_{i=1}^{k} B_i^{-1} C_i \) can have only finitely may zeros, we have \( \det(\sum_{i=1}^{k} B_i^{-1} C_i - I) \neq 0 \).
for all $\lambda > 0$ sufficiently close to 0. Thus we have $\nu_\lambda \in \mathcal{G}^*$ for all $\lambda > 0$ sufficiently close to 0. Making $\lambda > 0$ small, we get $\nu_\lambda$ as close to $\nu$ as we please. Thus by condition (b) of the theorem, $\nu_\lambda \in \mathcal{G}^* \cap \mathcal{G}_1 \cap \mathcal{G}_{SC}$ if $\lambda > 0$ is small enough. Consequently, $\mathcal{G}^* \cap \mathcal{G}_1 \cap \mathcal{G}'_{SC} \cap \mathcal{G}_{SC}$ is dense in $\mathcal{G}_1 \cap \mathcal{G}'_{SC} \cap \mathcal{G}_{SC}$, therefore dense in $\mathcal{G}_{SC}$ by the penultimate paragraph. We conclude that $\mathcal{G}^* \cap \mathcal{G}_{SC}$ is dense in $\mathcal{G}_{SC}$. \qed

3.7 Proof of Theorem 5

Theorem 5. There is an open dense subset $\mathcal{M}^*_K$ of $\mathcal{M}_K$ such that if $\nu \in \mathcal{M}^*_K$ and $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ is a sequence of Cournot oligopolies 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 Cournot equilibrium if $\#(I_n) \geq N$.

Proof. With $A = [0, m]$ and $g$ being the restriction to $A$ of the identity on $\mathbb{R}$, let $\mathcal{U}$ and $\mathcal{M}$ be defined as in the general model. Note that $E = g(A) = [0, m]$. Let $C = A$, so that $E(\nu) = E$ for all $\nu \in \mathcal{M}_{SC}$ because $E = g(A)$.

To each $v \in \mathcal{K}$ associate $u_v \in \mathcal{U}$ by setting $u_v(a, e) = p(e)a - v(a)$ for $a \in A$ and $e \in E$. Define a map $\kappa: \mathcal{K} \to \mathcal{U}$ by setting $\kappa(v) = u_v$ for each $v \in \mathcal{X}$. Note that $\kappa$ is a homeomorphism from $\mathcal{K}$ onto $\kappa(\mathcal{K})$. Let $\tilde{\kappa}: \mathcal{M}_K \to \mathcal{G}$ be defined by setting $\tilde{\kappa}(\nu) = \nu \circ \kappa^{-1}$ for each $\nu \in \mathcal{M}_K$. As $\kappa$ is a homeomorphism from $\mathcal{K}$ onto $\kappa(\mathcal{K})$, $\tilde{\kappa}$ is a homeomorphism from $\mathcal{M}_K$ onto $\tilde{\kappa}(\mathcal{M}_K)$.

Let $\mathcal{G}_{SC} = \tilde{\kappa}(\mathcal{M}_K)$. Let $\mathcal{M}_K'$ be the subset of $\mathcal{M}_K$ consisting of the elements with finite support, and note $\mathcal{M}_K'$ is dense in $\mathcal{M}_K$, because $\mathcal{K}$ (in the subspace topology defined from $\mathcal{X}$) is a separable metric space. Thus, because $\tilde{\kappa}: \mathcal{M}_K \to \mathcal{G}_{SC}$ is a homeomorphism, $\tilde{\kappa}(\mathcal{M}_K')$ is dense in $\mathcal{G}_{SC}$, i.e., (a) of Theorem 4 is satisfied. Using the definition of $\mathcal{K}$, we see that (b) of Theorem 4 is satisfied too.

Let $\mathcal{G}_{SC}^* \subseteq \mathcal{G}_{SC}$ be as guaranteed by Theorem 4. Then $\mathcal{M}_K^* = \tilde{\kappa}^{-1}(\mathcal{G}_{SC}^*)$ is open and dense in $\mathcal{M}_K$, because $\tilde{\kappa}$ is a homeomorphism. Let $\nu \in \mathcal{M}_K^*$ and let $\langle (I_n, G_n) \rangle_{n \in \mathbb{N}}$ be a sequence of Cournot oligopolies such that $\#(I_n) \to \infty$ and $\nu_{G_n} \to \nu$. By continuity of $\tilde{\kappa}$, we have $\tilde{\kappa}(\nu_{G_n}) \to \tilde{\kappa}(\nu)$. For each $n$ and each $i \in I_n$, define $u_{n,i} \in \mathcal{U}$ by setting $u_{n,i}(a, e) = p(\frac{1}{\#(I_n)}a + \#(I_n-1)/\#(I_n))a - G_n(i)(a)$. For each $n$, let $\tilde{\nu}_n = \sum_{i \in I_n} \frac{1}{\#(I_n)}\delta_{u_{n,i}}$, so that $\tilde{\nu}_n \in \mathcal{G}_{SC}$. Note that since $\tilde{\kappa}(\nu_n) \to \tilde{\kappa}(\nu)$ and $\#(I_n) \to \infty$, we also have $\tilde{\nu}_n \to \tilde{\kappa}(\nu)$ (use Billingsley (1968, Theorem 4.1)). Because $\tilde{\kappa}(\nu) \in \mathcal{G}_{SC}^*$, it follows from Theorem 4 (and the definition of Nash equilibrium for finite-player games as stated before the formulation of Theorem 1) that $(I_n, G_n)$ has a Cournot equilibrium if $n$ is large. \qed

4 Appendix

Lemma 10. Let $A$ be a compact convex set in $\mathbb{R}^n$ such that $\text{int } A \neq \emptyset$, and $g: A \to \mathbb{R}^m$ a continuous function. Suppose there is an $N \in \mathbb{N} \setminus \{0\}$ such that whenever $l \in \mathbb{N}$ satisfies $l \leq N$ then there is an open dense subset $\tilde{A}$ in the product $A^l$ such that the map $\tilde{g}: \tilde{A} \to \mathbb{R}^m$, given by $\tilde{g}(a_1, \ldots, a_l) = \frac{1}{l} \sum_{i=1}^l g(a_i)$, is an open map. Then

(a) $\text{int } \frac{1}{l} \sum_{i=1}^l g(A_i)$ is dense in $\frac{1}{l} \sum_{i=1}^l g(A_i)$ whenever $A_1, \ldots, A_l$ are closed convex subsets of $A$ such that $\text{int } A_i \neq \emptyset$ for each $i = 1, \ldots, l$ and $l \geq N$;

34
(b) \( \text{int} \, \text{cog}(A') \neq \emptyset \) whenever \( A' \) is a closed convex subset of \( A \) such that \( \text{int} \, A' \neq \emptyset \).

**Proof.** (a) Because \( A_1, \ldots, A_l \) are closed and convex, so is the product \( A_1 \times \cdots \times A_l \). As \( \text{int}(A_1 \times \cdots \times A_l) = \text{int} A_1 \times \cdots \times \text{int} A_l \neq \emptyset \), it follows that \( \text{int}(A_1 \times \cdots \times A_l) \) is dense in \( A_1 \times \cdots \times A_l \). Suppose \( l \geq N \) and let \( \tilde{A} \) be as hypothesized. Then \( \tilde{A} \cap \text{int}(A_1 \times \cdots \times A_l) \) is open and dense in \( A_1 \times \cdots \times A_l \). Since \( g \) is continuous, it follows that the set
\[
\left\{ \frac{1}{l} \sum_{k=1}^{l} g(a_k) : (a_1, \ldots, a_l) \in \tilde{A} \cap \text{int}(A_1 \times \cdots \times A_l) \right\}
\]
is dense in \( \frac{1}{l} \sum_{i=1}^{l} g(A_i) \), by the hypothesis on \( \tilde{g} \), the former set is open. We conclude that \( \text{int} \left( \frac{1}{l} \sum_{i=1}^{l} g(A_i) \right) \) if \( l \geq N \).

(b) Note that \( \frac{1}{N} \sum_{i=1}^{N} g(A') \subseteq \text{cog}(A') \), and that \( \text{int} \frac{1}{N} \sum_{i=1}^{N} g(A') \neq \emptyset \), by (a). \( \square \)

**Lemma 11.** Let \( A \subseteq \mathbb{R}^n \) be convex with non-empty interior, let \( m = kn \), and let \( g: A \to \mathbb{R}^m \) be given by setting
\[
g(a) = (a_{1,1}, a_{1,2}, \ldots, a_{1,k}, a_{2,1}, a_{2,2}, \ldots, a_{2,k}, \ldots, a_{n,1}, a_{n,2}, \ldots, a_{n,k})
\]
for each \( a \in A \), where the subscript \((h)\) means the \( h \)th coordinate of \( a \), \( h = 1, \ldots, n \). Then for any \( l \in \mathbb{N} \) with \( l \geq k \) there is an open dense subset \( \tilde{A} \) in the product \( A^l \) such that the map \( \tilde{g}: \tilde{A} \to \mathbb{R}^m \), given by \( \tilde{g}(a_1, \ldots, a_l) = \frac{1}{l} \sum_{i=1}^{l} g(a_i) \), is an open map.

**Proof.** Fix \( l \geq k \). Note first that because \( A^l \) is convex and \( \text{int} A^l \) is non-empty, \( \text{int} A^l \) is dense in \( A^l \). Let
\[
\Delta = \{(a_1, \ldots, a_l) \in A^l : a_{i,(h)} \neq a_{j,(h)}, \forall i, j \in \{1, \ldots, l\} \text{ with } i \neq j, \forall h \in \{1, \ldots, n\}\}.
\]
Evidently \( \Delta \) is open and dense in \( A^l \). Setting \( \tilde{A} = \Delta \cap \text{int} A^l \), it follows that \( \tilde{A} \) is open and dense in \( A^l \), too. Note that, at any \( (a_1, \ldots, a_l) \in \tilde{A} \), the derivative of \( \tilde{g} \) has a matrix representation of the form
\[
D \tilde{g}(a_1, \ldots, a_l) = \frac{1}{l} \begin{pmatrix}
B_1 & 0 & \cdots & 0 \\
0 & B_2 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & B_n
\end{pmatrix}
\]
where \( B_h, h = 1, \ldots, n \), is a \((k \times l)\)-matrix of the form
\[
B_h = \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
2a_{1,(h)} & 2a_{2,(h)} & 2a_{3,(h)} & \cdots & 2a_{l,(h)} \\
3a_{1,(h)}^2 & 3a_{2,(h)}^2 & 3a_{3,(h)}^2 & \cdots & 3a_{l,(h)}^2 \\
\vdots & \vdots & \vdots & \cdots & \vdots \\
ka_{1,(h)}^{k-1} & ka_{2,(h)}^{k-1} & ka_{3,(h)}^{k-1} & \cdots & ka_{l,(h)}^{k-1}
\end{pmatrix}.
\]
Now the determinant of the first \( k \) columns of the matrix \( B_h \) is just a positive multiple of the Vandermonde determinant, and thus non-zero because the points \( a_{1,(h)}, a_{2,(h)}, \ldots, a_{k,(h)} \) are distinct, by the choice of \( \tilde{A} \). Thus, at any \((a_1, \ldots, a_l) \in \tilde{A} \), the derivative of \( \tilde{g} \) has maximal rank, which implies that \( \tilde{g} \) is an open map (see Guillemin and Pollack, 1974, p.25, Exercise 1). \( \square \)
therefore find a probability measure $e$ with compact and convex values, all with non-empty interior. Then there exists a

Lemma 12. Let $X$ be a compact metric space, $\mu$ a Borel measure on $X$, and let $h: X \to \mathbb{R}^m$ be a continuous function. Then there is a sequence $\langle \mu_n \rangle$ of Borel measures on $X$ such that $\text{supp}(\mu_n)$ is a finite subset of $\text{supp}(\nu)$ for each $n$, $\mu_n(X) = \mu(X)$ for each $n$, $\mu_n \to \mu$ narrowly, and $\int h d\mu_n = \int h d\mu$ for each $n$.

Proof. Without loss of generality we may assume that $\mu$ is a probability measure with $\text{supp}(\mu) = X$. For each $n \in \mathbb{N}$ let $\mathcal{P}_n$ be a finite partition of $X$ into Borel sets $B_{n,i}$ such that $\text{diam}(B_{n,i}) \leq 1/(1 + n)$ for each $B_{n,i} \in \mathcal{P}_n$. Fix any $n$. For each $B_{n,i} \in \mathcal{P}_n$ let $\tilde{\mu}_{n,i}$ be the Borel measure given by setting $\tilde{\mu}_{n,i}(B) = \mu(B \cap B_{n,i})$ for each Borel set $B$ in $X$. Note that $\tilde{\mu}_{n,i}(B_{n,i}) = \mu(B_{n,i})$ for each $B_{n,i}$, and that $\sum_{i=1}^{k_n} \tilde{\mu}_{n,i} = \mu$, writing $k_n = \#(\mathcal{P}_n)$. Now if $\tilde{\mu}_{n,i}(B_{n,i}) > 0$ set $\tilde{\mu}_{n,i} = \frac{1}{\tilde{\mu}_{n,i}(B_{n,i})} \tilde{\mu}_{n,i}$, so that $\tilde{\mu}_{n,i}$ is a probability measure on $X$: if $\tilde{\mu}_{n,i}(B_{n,i}) = 0$, let $\tilde{\mu}_{n,i}$ be an arbitrary Borel measure on $X$. For each $B_{n,i}$, write $\tilde{\mu}_{n,i}$ for the closure of this set. Fix $B_{n,i}$ such that $\tilde{\mu}_{n,i}(B_{n,i}) > 0$. Write $e_{n,i} = \int h d\tilde{\mu}_{n,i}$. Note that $e_{n,i} \in \text{co}(h(B_{n,i}))$. Using Caratheodory’s theorem, we can therefore find a probability measure $\mu'_{n,i}$ on $X$ such that $\text{supp}(\mu'_{n,i})$ is a finite subset of $\tilde{\mu}_{n,i}$ and $\int h d\mu'_{n,i} = e_{n,i}$. Set $\mu_n = \sum_{i=1}^{k_n} \tilde{\mu}_{n,i}(B_{n,i}) \mu'_{n,i}$, where $\mu'_{n,i}$ is an arbitrary Borel measure on $X$ if $\tilde{\mu}_{n,i}(B_{n,i}) = 0$. Evidently $\mu_n$ is a Borel probability measure on $X$ such that $\text{supp}(\mu_n)$ finite, and we have

$$
\int h d\mu_n = \sum_{i=1}^{k_n} \left( \tilde{\mu}_{n,i}(B_{n,i}) \int h d\mu'_{n,i} \right) = \sum_{i=1}^{k_n} \int h d\tilde{\mu}_{n,i} = \int h d\mu,
$$

because $\tilde{\mu}_{n,i}(B_{n,i}) \tilde{\mu}_{n,i} = \tilde{\mu}_{n,i}$ for each $B_{n,i}$, and $\sum_{i=1}^{k_n} \tilde{\mu}_{n,i} = \mu$.

To see that $\mu_n \to \mu$ narrowly, fix any continuous map $p: X \to \mathbb{R}$. Fix $\epsilon > 0$. Note that since $X$ is compact, $p$ is uniformly continuous. Consequently there is an $n_\epsilon$ such that whenever $n \geq n_\epsilon$, then $|p(x) - p(x')| \leq \epsilon$ for all $x, x' \in B_{n,i}$ and all $B_{n,i} \in \mathcal{P}_n$. As $p$ is continuous, it follows that whenever $n \geq n_\epsilon$, then for each $B_{n,i} \in \mathcal{P}_n$ we actually have $|p(x) - p(x')| \leq \epsilon$ for all $x, x' \in B_{n,i}$. Thus, for any $n \geq n_\epsilon$, using the facts that $\mu = \sum_{i=1}^{k_n} \tilde{\mu}_{n,i} = \sum_{i=1}^{k_n} \tilde{\mu}_{n,i}(B_{n,i}) \tilde{\mu}_{n,i}$ and that $\tilde{\mu}_{n,i}(B_{n,i}) = \mu(B_{n,i})$ for each $B_{n,i}$, together with the definition of $\mu_n$, we see that

$$
\left| \int p d\mu - \int p d\mu_n \right| \leq \sum_{i=1}^{n_k} \left( \tilde{\mu}_{n,i}(B_{n,i}) \left| \int_{B_{n,i}} p d\tilde{\mu}_{n,i} - \int_{B_{n,i}} p d\mu'_{n,i} \right| \right) 
$$

$$
\leq \sum_{i=1}^{n_k} \left( \tilde{\mu}_{n,i}(B_{n,i}) \epsilon \right) = \epsilon \sum_{i=1}^{n_k} \tilde{\mu}(B_{n,i}) = \epsilon \sum_{i=1}^{n_k} \mu(B_{n,i}) = \epsilon.
$$

As $\epsilon > 0$ is arbitrary, it follows that $\int p d\mu_n \to \int p d\mu$. As $p$ is an arbitrary continuous map from $X$ to $\mathbb{R}$, we conclude that $\mu_n \to \mu$ narrowly. \◻

Lemma 13. Let $Y$ be a metric space, and $\theta: Y \to 2^{\mathbb{R}^n}$ a continuous correspondence, with compact and convex values, all with non-empty interior. Then there exists a
sequence \( \langle \theta_n \rangle \) of continuous correspondences from \( Y \) to \( 2^{\mathbb{R}^n} \) such that for each \( n \) and each \( y \in Y \), \( \theta_n(y) \) is compact, convex, has non-empty interior, and is included in \( \text{int} \theta(y) \), and for each \( y \in Y \), \( \rho_H(\theta(y), \theta_n(y)) \to 0 \) as \( n \to \infty \). If \( Y \) is compact, then for any \( \epsilon > 0 \) there is an \( n_\epsilon \) such that \( \rho_H(\theta(y), \theta_n(y)) < \epsilon \) for each \( y \in Y \) if \( n > n_\epsilon \).

**Proof.** Note first that as \( \theta \) is continuous, so is the correspondence \( y \mapsto \partial \theta(y) : Y \to 2^{\mathbb{R}^n} \) (see Wills, 2007). Using Lemma 3 we see that whenever \( O \subseteq \mathbb{R}^n \) is open, then the set \( \{ y \in Y : O \cap \text{int} \theta(y) \neq \emptyset \} \) is open in \( Y \). Consequently, because \( \text{int} \theta(y) \) is convex and non-empty for each \( y \in Y \), by Michael (1956, Theorem 3.1\(^m\)(c) and p. 372) there is a continuous map \( f : Y \to \mathbb{R}^n \) such that \( f(y) \in \text{int} \theta(y) \) for all \( y \in Y \). Now by the maximum theorem, the function \( y \mapsto \text{dist}(f(y), \partial \theta(y)) : Y \to \mathbb{R} \) is continuous. For each \( y \in Y \), set \( r(y) = (1/2) \text{dist}(f(y), \partial \theta(y)) \). Define a correspondence \( F : Y \to 2^{\mathbb{R}^n} \) by setting \( F(y) = \{ f(y) \} + \overline{B}(0, r(y)) \) for each \( y \in Y \). Then \( F \) is continuous, and for each \( y \in Y \), \( F(y) \) is compact, convex, has non-empty interior, and is included in \( \text{int} \theta(y) \). Now, for each \( n \), define the correspondence \( \theta_n : Y \to 2^{\mathbb{R}^n} \) by setting

\[
\theta_n(y) = \frac{1}{n+1} F(y) + (1 - \frac{1}{n+1}) \theta(y)
\]

for \( y \in Y \). Then, for each \( n \), \( \theta_n \) is continuous (see Hildenbrand, 1974, B.III, Propositions 5 and 9), and it follows that each \( \theta_n \) is of the required form. Moreover, if \( \langle n_k \rangle_{k \in \mathbb{N}} \) is any sequence in \( \mathbb{N} \) and \( y, y_{n_k}, k \in \mathbb{N} \), are points in \( Y \) such that \( y_{n_k} \to y \) as \( k \to \infty \), then \( \rho_H(\theta(y), \theta_{n_k}(y_{n_k})) \to 0 \), by the definition of \( \rho_H \). Using these facts, we conclude that the sequence \( \langle \theta_n \rangle \) is as desired. \( \square \)

The following theorem is a special version of a fixed point result due to Mas-Colell (1983).

**Theorem 6.** Let \( X \subseteq \mathbb{R}^f \) and \( Y \subseteq \mathbb{R}^m \) be compact convex sets with non-empty interior. Let \( \Lambda : X \times Y \to X \times \mathbb{R}^m \) be a continuous function; write \( \Lambda_X \) for \( \text{proj}_X \circ \Lambda \) and \( \Lambda_Y \) for \( \text{proj}_m \circ \Lambda \). Suppose there is an open set \( U \subseteq \mathbb{R}^m \), with \( Y \subseteq U \), and a continuously differentiable function \( \zeta : U \to \mathbb{R}^m \) such that, setting \( \gamma = \min\{ \| \zeta(y) \| : y \in \partial Y \} \),

(a) for some \( y^* \in \text{int} Y \), \( D\zeta(y^*) \) 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 \( \Lambda \) has a fixed point, i.e., there is an \( (x, y) \in X \times Y \) such that \( \Lambda(x, y) = (x, y) \).

**References**


