Existence of Nash Equilibrium in Games with a Measure Space of Players and Discontinuous Payoff Functions*

Guilherme Carmona[†] Konrad Podczeck[‡]
University of Surrey Universität Wien

April 1, 2014

Abstract

Balder's (2002) model of games with a measure space of players is integrated with the line of research on finite-player games with discontinuous payoff functions which follows Reny (1999). Specifically, we extend the notion of continuous security, introduced by McLennan, Monteiro & Tourky (2011) and Barelli & Meneghel (2013) for finite-players games, to games with a measure space of players and establish the existence of pure strategy Nash equilibrium for such games. A specification of our main existence result is provided which is ready to fit the needs of applications. As an illustration, we consider several optimal income tax problems in the spirit of Mirrlees (1971) and use our game-theoretic result to show the existence of an optimal income tax in each of these problems.

*Thanks to Paulo Barelli, Alex Gershkov, Michael Greinecker, Mário Páscoa, Phil Reny, Hamid Sabourian, Pedro Teles and seminar participants at the University of Cambridge, Università di Napoli Federico II, University of Surrey, Universität Wien and the 20th European Workshop on General Equilibrium (Kracow 2010) for helpful suggestions. Financial support from Fundação para a Ciência e a Tecnologia is gratefully acknowledged.

[†]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

1 Introduction

The line of research initiated by Dasgupta & Maskin (1986) and continued, amongst others, by Reny (1999) has been successful in obtaining equilibrium existence results for finite-player games with discontinuous payoff functions. In this paper we extend this approach to the context of generalized games with a measure space of players, a class of games first considered by Schmeidler (1973), the state of the art now set by Balder (2002). In particular, concerning existence of Nash equilibrium, we bring the branch of game theory dealing with games with a measure space of players on par with that dealing in a systematic way with games with discontinuous payoff functions.[1,2]

Apart from obtaining a unification of important recent game-theoretic results on a general level, a motivation for our analysis is that several economic problems which are addressed in the literature can be modeled as games with a continuum of players, but where payoff functions need not be continuous, and need not even satisfy the assumptions in Balder (2002). As an example, we will consider a version of Mirrlees's (1971) model of optimal taxation (see Section 9).

Our approach to deal with discontinuous payoff functions in the setting of games with a measure space of players is based on the notion of multiply security, which was developed in the context of finite-player games by McLennan, Monteiro & Tourky (2011). More precisely, we take a version of this notion, called continuous security, which was introduced by Barelli & Meneghel (2013), and adapt it to the particular measurability needs arising when there may be a continuum of players. We remark that the notion of multiply security generalizes that of better-reply security, which was introduced in the pioneering paper of Reny (1999).

Based on the notion of continuous security, our result covers, in particular, games where, as in Balder (2002), payoff functions are assumed to be upper semi-continuous and the value functions of the players are assumed to be lower semi-continuous.[3] In fact, when value functions are assumed to be lower semi-continuous, it covers games with payoff functions that are merely weakly upper semi-continuous (as defined in Carmona (2009)).

In addition to the pure strategy existence result of Balder (2002), our result extends that of Khan & Sun (1999). The approaches of Balder (2002) and Khan & Sun (1999) differ in the way how a convexifying effect of aggregation is ensured to deal

[1] Further papers on existence of equilibrium with a measure space of players include Barelli & Duggan (2012), Carmona & Podczeck (2009), Khan (1986), Khan, Rath & Sun (1997), Khan & Sun (1999), Martins da Rocha & Topuzu (2008), Mas-Colell (1984), Páscoa (1993), and Rath (1996).

[2] Recent papers on finite-player games with discontinuous payoff functions include Bagh & Jofre (2006), Balder (2011), Bich (2009), Bich & Laraki (2012), Carmona (2011), Carmona (2012), de Castro (2011), Nessah (2011), Prokopovych (2011), Prokopovych (2013), Reny (2009), Reny (2011), and Reny (2013).

[3] See Section 4.1 for the formal definition of the value function of a player.

with a continuum of negligible players with non-convex action sets. In Balder (2002) it is assumed in this respect that the payoff of each player may depend on the actions of the other players only through a finite-dimensional vector of summary statistics. In Khan & Sun (1999), on the other hand, the entire distribution of the actions of the players may matter for payoffs, but the non-atomicity hypothesis on the measure on the space of players is strengthened. Our result will show that the same strengthening of non-atomicity still allows to obtain pure-strategy Nash equilibria, with non-convex action sets and payoff dependence modeled as in Khan & Sun (1999), when, differently to Khan & Sun (1999), payoff functions need not be continuous.

The paper is organized as follows. We present a motivating example in Section 2. Some notation and terminology is introduced in Section 3. In Section 4 we present the general model and our notion of continuous security. Section 5 contains the statements of our main existence results. The relationship between these results and the pure strategy existence result of Balder (2002) is detailed in Section 6. In Section 7, a special case of the general model is presented. In this special case, each player's payoff is explicitly modeled to depend on his choice, on the choices of the atomic players, and on the vector of the joint distributions of the actions and players' attributes appearing in each one of countably many sub-populations of the atomless players. Sufficient conditions for continuous security are presented in Section 8. Section 9 applies our results to several optimal income taxation problems. In Section 10 we make some concluding remarks. The proofs of our results are in Section 11.

2 A motivating example

Consider a large population of individuals who live at different points in a geographical space and face the following problem. Each individual needs to take one of two bridges to reach his office. His goal is to minimize the total travel time between his home and his office; if he does not arrive on time for a meeting, there is an additional cost resulting from, say, extra time needed to schedule another meeting, or from the loss of a profitable trade. The optimal choice of any individual depends on the choice of all others through their influence on the congestion of the two bridges.

To make the example specific, assume that the population is described by a probability space (T, Σ, ν) where T is the unit interval $[0, 1]$ and ν is Lebesgue measure. Each individual is identified by his address, i.e., individual t lives at geographical point $t \in [0, 1]$, and has to choose one of two bridges, 1 or 2; thus, there is a common action set $A = \{a_1, a_2\}$, where a_i means choosing bridge i, $i = 1, 2$. Bridge 1 is located geographically at $t_1 = 1/3$, and bridge 2 at $t_2 = 2/3$.

The relative frequencies with which the bridges are used are described by the vector $y = (y^1, y^2)$ in the unit simplex Δ of $\mathbb{R}^2$. Assume that the time spent by individual t on the journey from home to work through bridge i is $g_t(a_i, y) = (1+y^i)|t-t_i|$, $i = 1, 2$. The amount of time individual t takes on his journey must be less than $m > 0$ for him to arrive on time for the meeting. His payoff is 1 if he arrives on time, but if he misses the meeting, he suffers a cost $c > 0$ implying that his payoff is $1 - c$. Thus, let individual t's payoff function be defined by setting, for each $x \in A$ and $y \in \Delta$,

$$u_t(x, y) = \begin{cases} 1 & \text{if } g_t(x, y) < m, \\ 1 - c & \text{otherwise.} \end{cases}$$

A strategy profile is a measurable map $f \colon T \to A$. Given such a strategy profile f, let $e(f) = (e^1(f), e^2(f)) \in \Delta$ denote the relative frequencies with which the bridges are used, i.e., $e^i(f) = \nu(\{t \in T \colon f(t) = a_i\})$, $i = 1, 2$. The payoff of individual t is then $u_t(f(t), e(f))$. We have thus a game where for each player (alias "individual") the payoff is determined by the own action and the frequency distribution of the actions chosen by all players. We say that a strategy profile f is a Nash equilibrium if, for almost every $t \in T$, $u_t(f(t), e(f)) = \max\{u_t(x, e(f)) \colon x \in A\}$.

Of course, for small values of m, each player's payoff function has discontinuity points. However, for any $(x, y) \in A \times \Delta$, the set of those players whose payoff function is discontinuous at this point is negligible (in fact, for any $(x, y) \in A \times \Delta$, there are at most two such players). In Section 11.4.3 we shall show that this property implies that our main result applies to yield the existence of a Nash equilibrium.

Some points are worth noting. First, the game is nothing else than a particular case of the situation considered in Theorem 2 of Schmeidler (1973), departing by just relaxing the continuity assumption in this latter result. Second, even though this departure is rather mild, the game is not covered by the results in Balder (2002), which would yield equilibrium existence in case that, for almost all $t \in T$, u_t is upper semi-continuous and the value function w_t, given by $w_t(y) = \max_{x \in A} u_t(x, y)$, is lower semi-continuous. In fact, in the above game, no player has an upper semi-continuous payoff function, so Balder's results cannot be applied.

Of course, the above example can easily be modified so that each player's payoff function is upper semi-continuous, but then the resulting value functions may fail to be lower semi-continuous, and Balder's results still cannot be applied. For instance, let

$$u_t(x, y) = \begin{cases} 1 & \text{if } g_t(x, y) \leq m, \\ 1 - c & \text{otherwise.} \end{cases}$$

Now it is in fact possible that, for a non-negligible subset of players, the value functions are not lower semi-continuous. This can be illustrated as follows: Let $m = 1/2$, which implies that $|t - 2/3| > m$ for all $t \in [0, 1/6)$. Now, for any $y \in \Delta$, a player $t \in [0, 1/6)$ can obtain a payoff of 1 only if he reaches his office in time by using bridge 1, i.e., if $1/3 - t \leq m/(1+y^1)$. Fix $t \in [0, 1/12)$ and let $y \in \Delta$ be such that $1/3 - t = m/(1+y^1)$.

Then $w_t(y) = 1$ but $w_t(y') = 1 - c$ for all $y' \in \Delta$ with $y'^1 > y^1$ and thus w_t is not lower semi-continuous.

To conclude, the above example presents a simple and mild departure from the setting considered in Schmeidler (1973), but which is not covered by the available existence results for large games. A more pronounced departure from the assumption of continuous payoff functions made in Schmeidler (1973) is presented in the optimal taxation framework considered in Section 9.

3 General notation and terminology

In this section we introduce notation and terminology needed for the setup of our model and for the formulation of the results. Additional notation and terminology used in the proofs can be found in the beginning of Section 11.

(a) $\varphi\colon A \twoheadrightarrow B$ denotes a correspondence from the set A to the set B, i.e., a map from A to the power set of B.

(b) We use "usc" as abbreviation for "upper semi-continuous," "lsc" for "lower semi-continuous," and "uhc" for "upper hemi-continuous."

(c) If A and B are topological spaces, a correspondence $\varphi\colon A \twoheadrightarrow B$ is called *well-behaved* if it is uhc and takes non-empty and closed values.

(d) For a topological space X, $\mathcal{B}(X)$ denotes the Borel σ-algebra of X.

(e) A topological space X is called a Souslin space if it is Hausdorff and if there is a continuous surjection from a Polish space onto X (see Schwartz (1973, Chapter II)).

(f) We will often work with the narrow topology on a set Z of Borel measures on a completely regular Souslin space X. Recall that the narrow topology on such a set Z is the smallest topology on Z such that for every continuous bounded $f\colon X \to \mathbb{R}$ the map $\mu \mapsto \int_X f \mathrm{d}\mu$, $\mu \in Z$, is continuous. (Cf. Schwartz (1973, p. 249)).

(g) Products of measurable spaces are always regarded as being endowed with the product σ-algebra, and subsets of measurable spaces with the subspace σ-algebra.

(h) If A and B are as in (c), and (T, Σ, ν) is a measure space, we call a correspondence $\varphi\colon T \times A \twoheadrightarrow B$ a *Caratheodory correspondence* if $\varphi(t, \cdot)$ is well behaved for each $t \in T$, and if for each $a \in A$, the graph of $\varphi(\cdot, a)$ is measurable, i.e., belongs to $\Sigma \otimes \mathcal{B}(B)$.

(i) Given functions $f\colon X \to Y$ and $g\colon X \to Z$, we denote by (f, g) the function $x \mapsto (f(x), g(x))\colon X \to Y \times Z$.

(j) $\mathrm{co}E$ denotes the convex hull of a subset E of a linear space.

(k) A measure space (T, Σ, ν) is called *non-trivial finite* if $0 < \nu(T) < \infty$, and is called *complete* if every ν-null set in T belongs to Σ.

(l) A measure space (T, Σ, ν) is called *separable* if $L_1(\nu)$ (with its usual norm) is separable. We will also say that a measure ν' on a set T' is separable to mean that (T', Σ', ν') is separable if $\Sigma' \subseteq 2^{T'}$ is the domain of ν'.

(m) A measure space (T, Σ, ν) is called *super-atomless* if for every $E \in \Sigma$ with $\nu(E) > 0$, the subspace of $L^1(\nu)$ consisting of the elements of $L^1(\nu)$ vanishing off E is non-separable. (For equivalent definitions, see Podczeck (2009).) We also say that a measure ν' on a set T' is super-atomless to mean that (T', Σ', ν') is super-atomless if $\Sigma' \subseteq 2^{T'}$ is the domain of ν'.

Remark 1. Note that a measure space (T, Σ, ν) is atomless if and only if for every $E \in \Sigma$ with $\nu(E) > 0$, the subspace of $L^1(\nu)$ consisting of the elements of $L^1(\nu)$ vanishing off E is infinite-dimensional. In view of this, it is clear that "super-atomless" is a property which is stronger than "atomless." It is well known that the unit interval with Lebesgue measure is an atomless and separable measure space. Being separable, this measure space is not super-atomless. Examples of super-atomless measure space are atomless Loeb measure spaces or $\{0, 1\}^I$ with the usual coin-flipping product measure when I is uncountable. Furthermore, as follows from Fremlin (2008, Proposition 521P(b)), Lebesgue measure on the unit interval can be extended to a super-atomless probability measure (see Podczeck (2009)).

4 The model

The game-theoretic model we consider is based on (the pure strategy part of) that of Balder (2002, Section 2.2). In Sections 4.1.1–4.1.6 below, we introduce the part of our model that is common with the model of Balder (2002), and we will be rather brief there. Some comments and a simple example which may be helpful for the reader may be found in 4.1.7. Those assumptions in the pure strategy Nash equilibrium existence result in Balder (2002) which we do not included in the setup of our general model may be found in Section 6.

The essential and innovative step in setting up our model is the adaption of the finite-player game notion of continuous security to a context with a measure space of players. The development will be done in Section 4.3.

Following Schmeidler (1973) and Balder (2002), we allow for games with non-convex action sets. The presentation and discussion of the assumptions we make to deal with this aspect of our model are separated in Section 4.2.

4.1 The basic model

4.1.1 Players

There is a measure space (T, Σ, ν) of players. The measure space (T, Σ, ν) may be non-atomic or purely atomic, or may have both an atomic part and a non-atomic part. This allows for non-atomic games as well as for finite-player games as special cases, and in particular covers situations with finitely many large, i.e., atomic players and a continuum of negligible players. Of course, depending on whether a player is atomic or negligible, different assumptions on his characteristics, i.e., action sets and payoff functions, may be appropriate (see Section 4.1.3 below). To accommodate this, the players in Balder's (2002) model are grouped into two measurable sets $\bar{T}$ and $\hat{T}$, with $\bar{T} \cap \hat{T} = \emptyset$ and $\bar{T} \cup \hat{T} = T$, where, to follow Balder's (2002) notation, $\bar{T}$ is taken to contain the atomic players.

The following is supposed to hold.

(A1) (T, Σ, ν) is a complete non-trivial finite measure space.

4.1.2 Actions

Action sets of players are subsets of a universe X where

(A2) X is a Souslin locally convex topological vector space.

Recall from Section 3(e) that a topological space X is called Souslin if it is Hausdorff and if there is a continuous surjection from a Polish space onto X. Thus any Polish space is a Souslin space. Examples of locally convex spaces that matter in several economic models and which are Souslin but not Polish are separable Banach spaces with the weak topology and duals of separable Banach spaces with the weak* topology.

The action set of player $t \in T$ is denoted by X_t, and by Γ_G we denote the graph of the action sets correspondence $t \mapsto X_t$. It is assumed:

(A3) (i) For each $t \in T$, the action set X_t is a non-empty compact subset of X.

(ii) Γ_G is a measurable subset of $T \times X$, i.e., belongs to $\Sigma \otimes \mathcal{B}(X)$.

A strategy profile is a measurable map $f : T \to X$ such that $f(t) \in X_t$ for almost all $t \in T$. By S_G we denote the set of all strategy profiles in the game G. Thus S_G is just the set of all measurable a.e. selections of the action sets correspondence $t \mapsto X_t$.

4.1.3 Externality

The externality of a strategy profile specifies the aggregate via which the strategy profile affects each player. In finite-player games, the strategy profile itself is usually taken for the externality. In non-atomic games, on the other hand, following Schmeidler (1973, Theorem 2), it is common to take some statistics of a strategy profile, like the mean or the distribution, for the externality, so as to get some convexifying effect of large numbers which allows to dispense with the convexity assumptions on action sets and payoff functions which are made in finite-player games. In Balder's (2002) model the two approaches above are integrated in the following way.

Recall first that the set T of players is partitioned into two measurable subsets $\bar{T}$ and $\hat{T}$. For the players in $\bar{T}$, which may be atoms for ν, the convexity assumption on action sets made in finite-player games is retained:

(A4) X_t is convex for every $t \in \bar{T}$.

For players in $\hat{T}$, this assumption is not made and is assumed that

(A5) $\hat{T}$ is included in the non-atomic part of (T, Σ, ν).

Now the externality of a strategy profile is specified as follows. Let $\bar{S}_G = \{f|_{\bar{T}} \colon f \in S_G\}$ be the set of the restrictions of the elements of S_G to $\bar{T}$, or, in other words, the set of all strategy profiles of the players in $\bar{T}$. In addition, let $\hat{\mathcal{C}}$ be a countable set of functions $q \colon \Gamma_G \cap (\hat{T} \times X) \to \mathbb{R}$ such that (i) q is measurable (recall Section 3(g)), (ii) $q(t, \cdot)$ is continuous for each $t \in \hat{T}$, (iii) there is an integrable function $\theta_q \colon \hat{T} \to \mathbb{R}_+$ such that $\sup\{|q(t,x)| \colon x \in X_t\} \leq \theta_q(t)$ for each $t \in \hat{T}$. Let $\bar{e} \colon S_G \to \bar{S}_G$ be given by $\bar{e}(f) = f|_{\bar{T}}$, and $\hat{e} \colon S_G \to \mathbb{R}^{\hat{\mathcal{C}}}$ by $\hat{e}(f) = \langle \int_{\hat{T}} q(t, f(t)) \mathrm{d}\nu(t) \rangle_{q \in \hat{\mathcal{C}}}$. Note that the integrals are indeed defined. Now define $e \colon S_G \to \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ by setting $e(f) = (\bar{e}(f), \hat{e}(f))$ for each $f \in S_G$. The map e is the externality map of the game. Thus, given a strategy profile f, the externality $e(f)$ is the restriction of f to $\bar{T}$ together with the statistics $\hat{e}(f)$ summarizing the actions of the players in $\hat{T}$.

Let $E_G \subseteq \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ denote the image of S_G under e, i.e., $E_G = e(S_G)$. We call the set E_G the externality set or externality space of a game G.

We note that in Balder (2002) it is actually assumed that the set $\hat{\mathcal{C}}$ is finite. We will formally state this as a condition in Section 4.2 below, where also a strengthening of (A5) is presented which gives a convexifying effect of large numbers when $\hat{\mathcal{C}}$ is allowed to be countably infinite.

The reason for us to allow $\hat{\mathcal{C}}$ to be countably infinite is to cover games where, following Mas-Colell (1984), action sets may be uncountable and the externality of a strategy profile is taken to be its distribution, rather than just a finite-dimensional summary statistics. The link between setting up the externality of strategy profiles as in this section and setting up it directly in terms of distributions is explored in a detailed way in Section 7. In that section we show that, actually, with $\hat{\mathcal{C}}$ allowed to be countably infinite, the way the externality of strategy profiles is modeled by Balder (2002) is far more general than just being able to cover their distributions, and thus turns out to be an efficient and very flexible device.

4.1.4 Topology on the externality set

The set E_G is given a topology specified as follows. First, the set $\bar{S}_G$ is given the feeble topology. Recall from Balder (2002) that the feeble topology on $\bar{S}_G$ is the coarsest topology such that the map $h \mapsto \int_{\bar{T}} q(t, h(t)) \mathrm{d}\nu(t) \colon \bar{S}_G \to \mathbb{R}$ is continuous for each $q \in \bar{\mathcal{G}}$, where $\bar{\mathcal{G}}$ is the set of all functions $q \colon \bar{T} \times X \to \mathbb{R}$ such that (i) q is measurable, (ii) $q(t, \cdot)$ is linear and continuous for each $t \in \bar{T}$, (iii) there is an integrable function $\theta_q \colon \bar{T} \to \mathbb{R}_+$ such that $\sup\{|q(t,x)| \colon x \in X_t\} \leq \theta_q(t)$ for each $t \in \bar{T}$. Second, $\mathbb{R}^{\hat{\mathcal{C}}}$ is given the product topology defined from the usual topology of $\mathbb{R}$. Now E_G is given the subspace topology defined from the product topology of $\bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$.

The proofs of the existence results in Balder (2002) require the externality space to be sequential, i.e., that a sequentially closed subset be closed. In Balder (2002) this is guaranteed by assuming that the measure space (T, Σ, ν) of players is separable. Actually, the following assumption suffices (see Lemma 3 in the proofs section below).

(A6) The subspace measure on $\bar{T}$ defined from ν is separable.

We will assume just (A6) because the strengthening of (A5) we will consider in Section 4.2 is incompatible with assuming the entire space (T, Σ, ν) to be separable. Of course, (A6) holds automatically if $\bar{T}$ is the union of atoms, and in particular, if the set $\bar{T}$ is countable.

4.1.5 Payoff functions and constraint correspondences

Each player $t \in T$ has a payoff function $u_t \colon X_t \times E_G \to [-\infty, +\infty]$. Thus, given a strategy profile $f \in S_G$, player t's payoff is determined by his own action $f(t)$ and by the externality $e(f)$.

In addition to the payoff function, for each player $t \in T$ there is a constraint correspondence $A_t \colon E_G \twoheadrightarrow X_t$. The set $A_t(y)$ specifies the actions that are actually available for player t given the externality $y \in E_G$. As elements of E_G represent social outcomes given the choices of all players, the set $A_t(y)$ can be viewed as a socially constrained action set of player t given $y \in E_G$.

The specification of payoff functions and constraint correspondences given here is exactly as in Balder (2002). At a first glance, one may have an interpretational problem with these specifications. Indeed, for a player t in $\bar{T}$, a given value of the externality determines his action, so one may ask for the meaning of $A_t(e(f))$ as a choice set as well as for the meaning of $u_t(x, e(f))$ if $x \neq f(t)$. The problem is easily resolved, though. We simply assume for the players in $\bar{T}$ (as suggested in Balder (2002, Section 2.4), and in conformity with what is standard in finite-player games) that payoff functions and constraint correspondences have factorizations so that, for a player $t \in \bar{T}$, $u_t(x, y)$ can be written as $v_t(x, \pi_t(y))$, and $A_t(y)$ as $B_t(\pi_t(y))$, where the function π_t does not depend on the coordinate of y referring to t, i.e., $\pi_t(y)$ does not specify anything about the action chosen by t; see Remark 4 for details. However, this issue will not play any role in the arguments concerning existence of Nash equilibrium, and therefore, as in Balder (2002), we do not state an explicit assumption in this respect. Note also that for players in $\hat{T}$, this issue does not arise at all under (A5). Indeed, (A5) means that players in $\hat{T}$ are negligible, so by the definition of the externality of a strategy profile the particular choice of an action by an individual player in $\hat{T}$ has no impact on the value of the externality, i.e., for players in $\hat{T}$, externality and action are independent of each other.

4.1.6 Nash equilibrium

We summarize a game as just outlined by a list $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$. Given such a game G, we denote by w_t the value function of player $t \in T$; that is, $w_t\colon E_G \to [-\infty, +\infty]$ is the function defined by setting

$$w_t(y) = \sup\{u_t(x, y)\colon x \in A_t(y)\},\ \ y \in E_G.$$

A strategy profile f is called a Nash equilibrium (for short, an equilibrium) of the game G if $f(t) \in A_t(e(f))$ and $u_t(f(t), e(f)) = w_t(e(f))$ for almost all $t \in T$.

4.1.7 Remarks

Remark 2. The model described in this section is intended as a general framework that can encompass several simpler models. This point has been emphasized in Balder (2002). In this line, we show in Section 7 that the class of games where each player's payoff depends on the own action and the distribution of the actions of all players (considered in Schmeidler (1973) and Mas-Colell (1984)) is included in the framework of this section. The following example shows this for the game in Section 2.

Example 1. Let $G = ((T, \Sigma, \nu), A, \langle u_t\rangle_{t\in T})$ be the game described in Section 2; in particular, the action set A of all players $t \in T$ is the set $A = \{a_1, a_2\}$ and the payoff functions u_t are defined on $A \times \Delta$ where Δ is the unit simplex in $\mathbb{R}^2$. To represent this game in the setting of this section, keep the probability space (T, Σ, ν) as the space of players, let $X = \mathbb{R}$, view A as a subset of $\mathbb{R}$, and for all $t \in T$, let $X_t = A$. Now $\Gamma_G = T \times A$. Define functions $q_i\colon \Gamma_G \to \mathbb{R}$, $i = 1, 2$, by setting

$$q_i(t, x) = \begin{cases} 1 & \text{if } x = a_i \\ 0 & \text{otherwise.} \end{cases}$$

Let $\hat{T} = T$, and let $e\colon S_G \to \mathbb{R}^2$ be the externality map defined from $\hat{\mathcal{C}} = \{q_1, q_2\}$. Write e^i for the ith coordinate function of e, $i = 1, 2$. Note that for any $f \in S_G$,

$e^i(f) = \nu(\{t \in T\colon f(t) = a_i\})$. In particular, because (T, Σ, ν) is a probability space, we have $E_G \equiv e(S_G) = \Delta$, and thus we have the payoff function u_t defined on $X_t \times E_G$ for each $t \in T$. Finally, for each $t \in T$, let $A_t(y) = A$ for all $y \in E_G$. Now we have a game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$ in terms of this section. It is clear that (A1)-(A6) are satisfied (note that $\bar{T} = \emptyset$, and that Lebesgue measure is complete and atomless). It is also clear that a Nash equilibrium of the game G is an equilibrium in the sense demanded in the context of Section 2.

Remark 3. Assuming that all the action sets are included in the same Souslin locally convex space X is not a big restriction. Indeed, suppose for instance that, for two Souslin locally convex spaces X_0 and X_1, we have $X_t \subseteq X_0$ for all $t \in \bar{T}$ and $X_t \subseteq X_1$ for all $t \in \hat{T}$, without imposing any relationship between X_0 and X_1. In this case, we can set $X = X_0 \times X_1$ and identify with X_0 with the subspace $X_0 \times \{0\}$ of X, and X_1 with the subspace $\{0\} \times X_1$, noting that the product of two locally convex Souslin spaces is again a space of this kind (directly from the definition of such spaces).

In fact, it suffices to assume that X_1 is just a completely regular Souslin space (without imposing any linear structure on X_1). The reasons are the following. First, no convexity assumptions are made with respect to the players in $\hat{T}$. Second, if X_1 is a completely regular Souslin space then, writing $M(X_1)$ for the space of all bounded signed Borel measures on X_1 with the narrow topology, $M(X_1)$ is a locally convex Souslin space (Schwartz (1973, p. 387, Corollary)) and the identification of the points in X_1 with the corresponding Dirac measures defines a homeomorphic embedding of X_1 into $M(X_1)$.

Remark 4. Here is a more detailed description of the assumption sketched in the last paragraph of Section 4.1.5. We will look only at payoff functions. Constraint correspondences can be dealt with similarly.

Consider first the special case where $\bar{T}$ is countable and $\{t\} \in \Sigma$ for each $t \in \bar{T}$. In this case, $\bar{S}_G$ is the same as $\prod_{t\in\bar{T}} X_t$. Write π_t for the projection of $\prod_{t'\in\bar{T}} X_{t'} \times \mathbb{R}^{\hat{\mathcal{C}}}$ onto $\prod_{t'\in\bar{T}\setminus\{t\}} X_{t'} \times \mathbb{R}^{\hat{\mathcal{C}}}$. Now the assumption is that for each $t \in \bar{T}$ the payoff function u_t is such that $u_t(x, y) = v_t(x, \pi_t(y))$ for any $x \in X_t$ and $y \in E_G$, where v_t is a function defined on $X_t \times \pi_t(E_G)$. Of course, for any $t \in \bar{T}$, π_t does not depend on the coordinate of y referring to t.

As for the general case, choose a partition of $\bar{T}$ into measurable sets F and E_k, $k \in K$, so that ν is atomless on F and each E_k is an atom for ν. (Note that in general an atom need not be a singleton.) For an element $g \in \bar{S}_G$, write $g^\bullet$ for the ν-equivalence class of g in the space of measurable functions from $\bar{T}$ to X, and let $\bar{S}_G^\bullet$ be the set of all these equivalence classes. Define $\pi\colon \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}} \to \bar{S}_G^\bullet \times \mathbb{R}^{\hat{\mathcal{C}}}$ by setting

$$\pi(y) = (g^\bullet, h),\ y = (g, h) \in \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}.$$

Further, for each $k \in K$, let $\bar{S}_G^k$ be the set of restrictions of the elements of $\bar{S}_G$ to $\bar{T} \backslash E_k$, and define $\pi_k \colon \bar{S}_G \times \mathbb{R}^{\hat{C}} \to \bar{S}_G^k \times \mathbb{R}^{\hat{C}}$ by setting

$$\pi_k(y) = (g|_{\bar{T}\backslash E_k}, h), \ y = (g, h) \in \bar{S}_G \times \mathbb{R}^{\hat{C}}.$$

Now our assumption is that for a player t belonging to $\bar{T}$ the payoff function u_t satisfies $u_t(x, y) = v_t(x, \pi_t(y))$ for every $x \in X_t$ and $y \in E_G$, where $\pi_t = \pi$ if $t \in F$ and $\pi_t = \pi_k$ if $t \in E_k$, $k \in K$, and where v_t is a function defined on $X_t \times \pi_t(E_G)$. Again, for any $t \in \bar{T}$, π_t does not depend on the coordinate of y referring to t. If $t \in E_k$ for some $k \in K$, this is clear. For $t \in F$, just note that $\{t\}$ is a ν-null set as ν is atomless on F, so for any $g \in \bar{S}_G$, $g(t)$ is not determined by $g^\bullet$.

Remark 5. One may think of the players in $\bar{T}$ as large atomic players. But in the literature there are also contexts to which Balder's (2002) model applies with $\bar{T}$ being included in the non-atomic part of the space of players. E.g., in Kim & Yannelis (1997) and Yannelis (2009), non-atomic games are considered where, besides of the own action, the payoff of each player may vary with the "equal almost everywhere"-equivalence class of strategy profiles, rather than just with a summary statistics. It is assumed in these games that action sets of the players are convex. Thus, with $T = \bar{T}$ and ν non-atomic, Balder's (2002) model covers such games (cf. the previous remark).

Remark 6. If $\bar{T}$ is countable and $\{t\} \in \Sigma$ for each $t \in \bar{T}$, then $\bar{S}_G$ is the same as $\prod_{t\in\bar{T}} X_t$. Moreover, if $\nu(\{t\}) > 0$ and the action set X_t is compact for each $t \in \bar{T}$, then the feeble topology on $\bar{S}_G$ is the same as the product topology on $\prod_{t\in\bar{T}} X_t$. This is so for two reasons. First, compactness of X_t means that the weak topology of the locally convex space X coincides on X_t with the given topology of X; thus a net $\langle x_\alpha \rangle$ in X_t converges to some $x \in X_t$ if and only if $p(x_\alpha) \to p(x)$ for each continuous linear function $p \colon X \to \mathbb{R}$. Second, if $\{t\} \in \Sigma$ and X_t is compact for each $t \in \bar{T}$, then, for any such p, the function $q \colon \bar{T} \times X \to \mathbb{R}$, where $q(t, \cdot) = p$ for one $t \in \bar{T}$ and $q(t, \cdot)$ is the zero functional elsewhere in $\bar{T}$, belongs to the set $\bar{\mathcal{G}}$ in the definition of the feeble topology. Thus, given that $\nu(\{t\}) > 0$ for each $t \in \bar{T}$, it follows that if a net $\langle h_\alpha \rangle$ in $\bar{S}_G$ converges to some $h \in \bar{S}_G$ for the feeble topology, then it converges to this h for the product topology of $\prod_{t\in\bar{T}} X_t$. In view of (iii) in the definition of $\bar{\mathcal{G}}$, it is clear that the reverse implication also holds, given that $\bar{T}$ is countable.

Remark 7. The model presented above contains the standard (normal-form) model of finite-player games as a special case (as long as action sets are contained in locally convex Souslin spaces). Indeed, suppose for the measure space (T, Σ, ν) of players that T is finite, $\Sigma = 2^T$, and ν is the counting measure. Set $\bar{T} = T$. Then $S_G = \bar{S}_G$ and thus, by what was noted in the previous remark, S_G is the same as $\prod_{t\in T} X_t$ and the feeble topology on S_G is the same as the product topology on $\prod_{t\in T} X_t$. Concerning the payoff functions, we refer to the second paragraph of Remark 4.

4.2 Convexifying effect of large numbers

As said in Section 4.1.3, a motivation to consider non-atomic games with the externality of strategy profiles being defined by summary statistics is to get some convexifying effect of large numbers so that the usual convexity assumptions on action sets and payoff functions made in finite-player games may be dropped. In the model as described in Section 4.1, the set of players is partitioned into two measurable sets $\bar{T}$ and $\hat{T}$, and while action sets of the players in $\bar{T}$ are assumed to be convex, no such assumption is made for the players in $\hat{T}$. Thus a convexifying effect of large numbers is required for the players in $\hat{T}$. In regard to this, we will consider two scenarios which are described in conditions (S1) and (S2) below.

Recall that the set $\hat{\mathcal{C}}$ in the definition of the externality map e is assumed to be countable. Strengthening this assumption to require $\hat{\mathcal{C}}$ to be finite gives the scenario treated in Balder (2002).

(S1) The set $\hat{\mathcal{C}}$ in the definition of e is finite.

Thus, under (S1), the restriction of the externality map e to $\hat{T}$ takes values in a finite-dimensional vector space. Together with this, the non-atomicity hypothesis in (A5) is sufficient to give a convexifying effect needed to get a (pure strategy) equilibrium without convexity assumptions on action sets or payoff functions of the players in $\hat{T}$. This may be viewed as paralleling Liyapounoff's theorem which says that the range of an atomless vector measure with values in a finite-dimensional vector space is convex.

Now if $\hat{\mathcal{C}}$ may be countably infinite, so that the values of the restriction of the externality map e to $\hat{T}$ may span an infinite-dimensional vector space, then (A5) is not sufficient to guarantee the desired convexifying effect. In fact, with $\hat{\mathcal{C}}$ allowed to be countably infinite, our model covers (as intended, see Section 4.1.3) games where action sets are non-convex and uncountable and where the externality of a strategy profile is its distribution. But for such games it is known that existence of a pure strategy equilibrium may fail even when the space of players is assumed to be non-atomic (see the examples in Khan et al. (1997)). This failure may be viewed as another manifestation of the fact that Liyapounoff's theorem fails with infinite-dimensional codomains (see Diestel & Uhl 1977, IX.1). On the other hand, as shown by the existence results in Carmona & Podczeck (2009, Corollary 4) and Keisler & Sun (2009, Theorem 4.6), the problem can be resolved if the hypothesis that the space of players be atomless is strengthened to require this space to be super-atomless. In the context of the present model, this suggests that in order to handle the case where $\hat{\mathcal{C}}$ is countably infinite, (A5) has to be strengthened as follows.

(S2) The subspace measure on $\hat{T}$ defined from ν is super-atomless.

(See Section 3 for the definition of "super-atomless.") We remark that the existence results in Carmona & Podczeck (2009, Corollary 4) and Keisler & Sun (2009, Theorem

4.6) are established for games where players have a common action set and payoff functions are continuous. In view of this, our existence result with (S2) as stated in Section 5 shows that a convexifying effect of the assumption that the measure space of players be super-atomless becomes manifest in much more general settings.

4.3 Continuous security

The notion of continuous security was introduced in the case of finite-player games by Barelli & Meneghel (2013) (see also Barelli & Soza (2009)), building on the notions of multiply security, which was developed by McLennan et al. (2011), and that of better-reply security, developed by Reny (1999). We first present the definitions of better-reply security and of continuous security for finite-player games and then extend this latter notion to games with a continuum of players.

Consider a game $G = \langle X_i, u_i \rangle_{i \in I}$ with finitely many players, where I is the set of players, X_i is player i's action set, and $u_i : \prod_{j \in I} X_j \to \mathbb{R}$ player i's payoff function. Assume that, for each $i \in I$, X_i is a nonempty, compact, and convex subset of a Hausdorff locally convex topological vector space and u_i is bounded. The game G is said to be *better-reply secure* if for each $y \in \prod_{i \in I} X_i$ which is not a Nash equilibrium of G, and each $\beta \in \mathbb{R}^I$ such that (y, β) is in the closure of the graph of the vector-valued payoff function $u = \prod_{i \in I} u_i$, there exists a player $i \in I$, an action $x_i \in X_i$, a number $\alpha_i > \beta_i$ and a neighborhood U_{-i} of y_{-i} in $\prod_{j \in I \setminus \{i\}} X_j$ such that $u_i(x_i, y'_{-i}) \geq \alpha_i$ for all $y'_{-i} \in U_{-i}$. Here y'_i is the projection of y' onto X_i, and y'_{-i} that onto $\prod_{j \in I \setminus \{i\}} X_j$. It is plain that better-reply security holds whenever payoff functions are continuous. In fact, by Proposition 3.2 in Reny (1999), better-reply security provides a proper generalization of both the usc and the lsc aspect of continuity.

The notion of better-reply security was extended by McLennan et al. (2011) to the notion of multiply security, which in turn was extended by Barelli & Meneghel (2013) to the notion of continuous security. Both extensions are based on a result by McLennan et al. (2011, Lemma 2.5) which shows that better-reply security is equivalent to the following condition: Whenever $y \in \prod_{i \in I} X_i$ is not a Nash equilibrium of G, there is a neighborhood U of y in $\prod_{i \in I} X_i$, a vector $\alpha \in \mathbb{R}^I$, a number $\varepsilon > 0$ and, for every $i \in I$, an action $\bar{x}_i \in X_i$ such that:

(i) For every $i \in I$, $u_i(\bar{x}_i, y'_{-i}) \geq \alpha_i + \varepsilon$ for all $y' \in U$.

(ii) For each $y' \in U$ there is an $i \in I$ such that $u_i(y'_i, y'_{-i}) < \alpha_i - \varepsilon$.

Continuous security generalizes better-reply security as follows. First, rather than requiring single actions (the $\bar{x}_i$'s above) to secure payoffs on U, it is allowed that payoffs can be secured along well-behaved correspondences defined on U; this implies that (i) may be satisfied with a larger α, and hence that (ii) is easier to satisfy. Second, the number ε is dropped, which another time makes both (i) and (ii) easier to satisfy. On the other hand, following McLennan et al. (2011), the notion of continuous security modifies (ii) by adding a convexity assumption on payoff functions, which however is weaker than the assumption that payoff functions be quasi-concave in the own strategy, but strong enough so that no extra quasi-concavity assumption on payoff functions is needed anymore to get existence of Nash equilibrium (differently from Reny's (1999) main existence theorem, which, in addition to better reply security, requires that payoff functions be quasi-concave in the own strategy).

More precisely, the game G is called *continuously secure* if whenever $y \in \prod_{i\in I} X_i$ is not a Nash equilibrium of G, there is a neighborhood U of y in $\prod_{i\in I} X_i$, a vector $\alpha \in \mathbb{R}^I$ and, for every $i \in I$, a well-behaved correspondence $\varphi_i\colon U \twoheadrightarrow X_i$ such that:

(a) For every $i \in I$ and every $y' \in U$, $\varphi_i(y')$ is convex or there is a finite-dimensional subspace of X_i which contains $\varphi_i(y')$.

(b) For every $i \in I$, $u_i(x, y'_{-i}) \geq \alpha_i$ for all $y' \in U$ and $x \in \varphi_i(y')$.

(c) For each $y' \in U$ there is an $i \in I$ such that $y'_i \notin \operatorname{co}\{x \in X_i\colon u_i(x, y'_{-i}) \geq \alpha_i\}$.[4]

We remark that on the level of the primitives of a game, the difference between continuous security and better-reply security (or multiply security for that matter) is not minor. In fact, as it can easily be inferred from Corollary 3 in Carmona (2011) and Proposition 2.4 in Barelli & Meneghel (2013), if the payoff functions u_i in a finite-player game as described above are quasi-concave in the own strategy and usc, and the value functions $y \mapsto \sup\{u_i(x, y_{-i})\colon x \in X_i\}$ from $\prod_{j\in I} X_j$ to $\mathbb{R}$ are lsc, then the game is continuously secure but need not be better-reply secure. Now upper semi-continuity of payoff functions and lower semi-continuity of value functions are the assumptions made in Balder (2002) concerning continuity properties of payoff functions. Thus, as the model of our paper is based on (the pure strategy part of) that of Balder (2002), it is natural to look for an adaption of the notion of continuous security to our model.

One aspect of such an adaptation concerns the sets U, i.e., the sets of strategy profiles at which each player needs to secure some payoff against the actions of the other players. In our setting of games as introduced in Section 4.1, the payoff of a player depends on the actions of the other players through the aggregates $e(f)$ of the

[4]The definition of "continuously secure" in Barelli & Meneghel (2013) is not exactly equal to the one presented here. Actually, Barelli & Meneghel (2013) do not require (a). Unfortunately, the proof of Theorem 2.2 in Barelli & Meneghel (2013) does not go through without (a). The reason is that the correspondence Φ in that proof is not necessarily closed-valued, because the convex hull of a compact set need not be closed in an infinite-dimensional space. To solve this problem, one can, as we did here, require $\varphi_i(y')$ to be convex, as in Barelli & Soza (2009), or to be included in a finite-dimensional subspace of X_i, as in McLennan et al. (2011).

strategy profiles f, and it is therefore natural to take subsets of the space E_G of these aggregates for the sets U.

Another need for adjustment arises because, unlike in the finite-player case, measurability properties are not trivially satisfied when there is a continuum of players, but have to be assumed explicitly. For a notion of continuous security in our setting, this means that the following two specifications are needed. First, the analogs of the correspondences φ_i must, as correspondences taking values in the universal action space X, be linked together over the space of players in a measurable way, which can be done using the notion of Caratheodory correspondence as stated in Section 3. Second, the analog of the vector α must be required to be measurable as a function on the space of players.

A final point concerns condition (c) of the finite-player game definition of continuous security, which invokes just one player at a given strategy profile. Instead we need a formulation in terms of non-negligible sets of players, because an individual player may be negligible in our setting.

Summarizing this discussion leads to the following definition, where E_G is regarded as being endowed with the topology introduced in Section 4.1, and CS abbreviates "continuous security."

Definition 1. A game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ is said to satisfy *CS* if whenever $y \in E_G$ is such that there is no equilibrium strategy profile f with $e(f) = y$, there is a neighborhood U of y in E_G, a Caratheodory correspondence $\varphi\colon T \times U \twoheadrightarrow X$, and a measurable function $\alpha : T \to [-\infty, +\infty]$ such that:

(a) For each $y' \in U$, $\varphi(t, y') \subseteq A_t(y')$ for all $t \in T$.

(b) For all $y' \in U$ and all $t \in \bar{T}$, $\varphi(t, y')$ is convex or there is a finite-dimensional subspace of X which contains $\varphi(t, y')$.

(c) For each $y' \in U$, $u_t(x, y') \geq \alpha(t)$ for almost all $t \in T$ and all $x \in \varphi(t, y')$.

(d) If f is a strategy profile with $e(f) \in U$, $f(t) \in A_t(e(f))$ for almost all $t \in \hat{T}$, and $f(t) \in \mathrm{co}A_t(e(f))$ for almost all $t \in \bar{T}$, then there is a non-negligible set $T' \subseteq T$ such that for every $t \in T' \cap \hat{T}$, $u_t(f(t), e(f)) < \alpha(t)$, and for every $t \in T' \cap \bar{T}$, $f(t) \notin \mathrm{co}\{x \in A_t(e(f))\colon u_t(x, e(f)) \geq \alpha(t)\}$.[5]

The following remark may be useful in applications of CS.

Remark 8. (i) If for each $t \in \bar{T}$, $u_t(\cdot, y)$ is quasi-concave and $A_t(y)$ is convex for all $y \in E_G$, then (d) in this definition is equivalent to the simpler statement:

[5]Note that in (c) the exceptional set of measure zero may vary with y', and that in (d) the set T' may vary with f. Also note that the set T' in (d) is not required to be measurable.

(d') If f is a strategy profile with $e(f) \in U$ and $f(t) \in A_t(e(f))$ for almost all $t \in T$, there is a non-negligible set $T' \subseteq T$ such that $u_t(f(t), e(f)) < \alpha(t)$ for all $t \in T'$.

(ii) In the case where $A_t(y) = X_t$ for all $t \in T$ and $y \in E_G$, (d) in the definition reduces to the statement:

(d") If f is a strategy profile with $e(f) \in U$, then there is a non-negligible set $T' \subseteq T$ such that for all $t \in T' \cap \hat{T}$, $u_t(f(t), e(f)) < \alpha(t)$, and for all $t \in T' \cap \bar{T}$, $f(t) \notin \mathrm{co}\{x \in X_t \colon u_t(x, e(f)) \geq \alpha(t)\}$.

(iii) In particular, if for every $t \in T$ and $y \in E_G$, $A_t(y) = X_t$, and for every $t \in \bar{T}$, X_t is convex and $u_t(\cdot, y)$ is quasi-concave for all $y \in E_G$, then (d) is equivalent to:

(d"') If f is a strategy profile with $e(f) \in U$, then there is a non-negligible set $T' \subseteq T$ such that $u_t(f(t), e(f)) < \alpha(t)$ for all $t \in T'$.

Remark 9. As noted in Remark 7, a special case of the general model of Section 4.1 is that of finite-player games in normal-form (provided that the action spaces of the players are included in a Souslin locally convex space). For this special case, it may be seen that Definition 1 is exactly equivalent to the definition of continuous security presented earlier for finite-player games, given that payoff functions are real-valued and bounded, so that the $\alpha(t)$'s in Definition 1 can be assumed to be real numbers.

Remark 10. We note that the framework of Balder (2002) has been extended in Martins da Rocha & Topuzu (2008) by allowing players to have non-ordered preferences. However, the notion of continuous security requires players to have payoff functions, and this is the reason why we adopted Balder's (2002) model. Recently, two conditions were introduced to deal with discontinuous finite-player games when players may have non-ordered preferences. These are condition B (and B_g) in Barelli & Soza (2009) and the condition of point security in Reny (2013). We leave it as a question for future research whether our results extends to games in the framework of Martins da Rocha & Topuzu (2008) by using some adaptation of these conditions.

5 The main existence results

In this section we state our two main results on existence of Nash equilibrium. They correspond to the two scenarios specified in Section 4.2 by assumptions (S1) and (S2) respectively.

Theorem 1. *Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ be a game satisfying* (A1)-(A6), (S1)*, and* CS*. Then G has a Nash equilibrium.*

Theorem 1 is a special case of Theorem 3 below. We note that Theorem 1 generalizes the pure strategy Nash equilibrium existence result of Balder (2002). Indeed, by Theorem 4 in Section 6, Theorem 1 implies Balder's result. By the example in that section, the converse fails.

Theorem 1 also implies the Nash equilibrium existence result for continuously secure finite-player games by Barelli & Meneghel (2013), provided the action sets of all players are included in a Souslin locally convex space (however, recall footnote 4). Indeed, suppose for the space (T, Σ, ν) of players that T is finite and ν is the counting measure, and set $\bar{T} = T$. Then, by what was pointed out in Remark 7, our setting of games reduces to that of standard normal-form finite-player games. Besides of (A2), the only assumptions of Theorem 1 that are not trivially satisfied in this reduced setting are (A3)(i), (A4), and CS, and these are the assumptions in the result by Barelli & Meneghel (2013) (concerning CS, see Remark 9).

Replacing (S1) in the statement of Theorem 1 by (S2) leads to the statement of our second main result. Recall that (S2) strengthens (A5), so that (A5) is actually not needed when (S2) is assumed.

Theorem 2. *Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ be a game satisfying* (A1)-(A4), (A6), (S2), *and* CS. *Then G has a Nash equilibrium.*

Theorem 2 is again a special case of Theorem 3 below. As may be seen in Section 7, Theorem 2 implies a pure strategy Nash equilibrium result for games where the payoff of each player may depend on the entire distribution of the actions of the players in $\hat{T}$, rather than just on a finite-dimensional summary statistics of these actions as in Theorem 1. In particular, Theorem 2 implies the existence results in Khan & Sun (1999), Carmona & Podczeck (2009), and Keisler & Sun (2009) (see Remark 14 in Section 7 below). In fact, Theorem 2 applies to situations where payoffs may depend on the vector of the distributions of the actions played in each one of countably many sub-populations of $\hat{T}$. This will be explicitly shown in Section 7.

Note that assumptions (S1) and (S2), which are involved in Theorems 1 and 2, respectively, are incomparable; neither of these assumptions implies the other. However, the following condition unifies (S1) and (S2).

(S3) The set $\hat{\mathcal{C}}$ in the definition of the map e is finite or the subspace measure on $\hat{T}$ defined from ν is super-atomless.

Of course, (S3) is not a generalization of (S1) and (S2) at a deeper level. It is an auxiliary assumption, introduced just to state the following theorem, which obviously contains both Theorem 1 and 2 as special cases.

Theorem 3. *Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ be a game satisfying* (A1)-(A6), (S3)*, and* CS. *Then G has a Nash equilibrium.*

(See Section 11.3 for the proof.)

6 Connection to Balder (2002)

In the framework of Section 4.1, consider the following additional assumptions.

(A7) The map $(t,x) \mapsto u_t(x,y)$ from Γ_G to $[-\infty,+\infty]$ is measurable for each $y \in E_G$.

(A8) (i) For each $t \in T$, the correspondence A_t is well-behaved.
(ii) For each $y \in E_G$, the graph of the correspondence $t \mapsto A_t(y)$ is measurable, i.e., belongs $\Sigma \otimes \mathcal{B}(X)$.

(A9) For every $t \in T$, u_t is usc and w_t is lsc.

(A10) For every $t \in \bar{T}$, the set $\{x \in A_t(y) \colon u_t(x,y) = w_t(y)\}$ is convex for all $y \in E_G$.

In the existence result about pure strategy Nash equilibria in Balder (2002), these assumptions are made in addition to (A1)-(A6) and (S1). The following theorem shows that our notion of continuous security covers this case.

Theorem 4. *Let $G = ((T,\Sigma,\nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$ be a game satisfying* (A1)-(A6) *and* (S1) *or* (S2). *If* (A7)-(A10) *hold in addition, then G satisfies* CS.

(See Section 11.4.1 for the proof.)

Note that (A9) is satisfied whenever all constraint correspondences are lower hemicontinuous and all payoff functions are continuous. Furthermore, if for each $t \in \bar{T}$, A_t takes convex values and $u_t(\cdot, y)$ is quasi-concave for all $y \in E_G$, then (A10) is satisfied. Thus Theorem 4 shows that for games satisfying (A1)-(A8) and (S1) or (S2), continuous security according to our notion holds whenever all constraint correspondences and all payoff functions are continuous, and for players belonging to $\bar{T}$, the constraint correspondences take convex values and the payoff functions are quasi-concave in the own action. In particular, if $T = \hat{T}$, and if there are no constraint correspondences (in other words, if $A_t(y) = X_t$ for all $t \in T$ and $y \in E_G$), then CS holds, as it should be, whenever all payoff functions are continuous.

It is well known that in finite-player games, continuous security does not imply that payoff functions are usc or that value functions are lsc (see, e.g., Carmona (2009)). The same holds in the context of the present paper. In Section 2 we have presented examples of games that satisfy CS and either fail to have usc payoff functions or fail to have lsc value functions. The example below is stronger by presenting a game that satisfies CS and fails in both respects. In particular, the example shows that the converse of the implication in Theorem 4 does not hold in general.

Example 2. *A non-atomic game that satisfies CS where payoff functions are not usc and value functions are not lsc.* Consider the game G constructed in Example 1, just with $A = \{a_1, a_2, a_3\}$ in place of $A = \{a_1, a_2\}$, so that E_G is now the unit simplex in $\mathbb{R}^3$, and with payoff functions $u_t = u$ for all $t \in T$, where $u\colon A \times E_G \to \mathbb{R}$ is given by setting, for all $y \in E_G$ and some $\varepsilon > 0$,

$$u(a_1, y) = \begin{cases} y^1 & \text{if } y^1 \neq 1/2, \\ y^1 + \varepsilon & \text{if } y^1 = 1/2, \end{cases}$$

$$u(a_2, y) = \begin{cases} 1 - y^1 & \text{if } y^1 < 1/2, \\ 1 - y^1 - \varepsilon & \text{if } y^1 \geq 1/2, \end{cases}$$

and

$$u(a_3, y) = \begin{cases} 1/2 & \text{if } y^1 \neq 1/2, \\ 1/2 + \varepsilon & \text{if } y^1 = 1/2. \end{cases}$$

As in Example 1, (A1)-(A6) are satisfied. Clearly, (A7), (A8), and (S1) are also satisfied, and so is (A10) as $\bar{T} = \emptyset$. Evidently u is not usc and the corresponding value function is not lsc. Furthermore, the best-reply correspondence of the game is not uhc. However, the game satisfies CS (see Section 11.4.4 for the proof).

The following definition places a weakening of the condition on payoff functions to be upper semi-continuous, introduced in Carmona (2009), in the context of our model.

Definition 2. A game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ is said to be *weakly upper semi-continuous* (abbreviated "weakly usc" in the sequel) if the following holds for all $t \in T$: Whenever $(x_n, y_n) \to (x, y)$ in $X_t \times E_G$ and $\overline{\lim}\, u_t(x_n, y_n) \neq u_t(x, y)$, there is an $x' \in A_t(y)$ such that $u_t(x', y) > \overline{\lim}\, u_t(x_n, y_n)$.

Remark 11. As $x' \neq x$ is not required in this definition, a game with usc payoff functions and well-behaved constraint correspondences is weakly usc. On the other hand, it is easy to find examples showing that the converse need not hold.

Theorem 5. *Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, e)$ be a game satisfying* (A1)-(A6) *and* (S1) *or* (S2). *If, in addition, G is weakly usc, w_t is lsc for all $t \in T$, and* (A7), (A8), *and* (A10) *hold, then G satisfies* CS.

(See Section 11.4.2 for the proof.)

7 A concretization of the general model

In this section we will present a specification of the model laid out in Section 4.1, illustrating what is covered by the notion of externality map, and in particular aiming to provide a bridge to potential applications.

In typical applications with a measure space of players, large atomic players appear as singletons, and no convexity assumptions are made on the non-atomic part of the space of players. In view of this, we replace (A6) by the following condition.

(A11) The set $\bar{T}$ is countable, and for each $t \in \bar{T}$, $\{t\} \in \Sigma$ with $\nu(\{t\}) > 0$.

Note that by what was stated in Remark 6, (A9) implies that the set $\bar{S}_G$ of restrictions of strategy profiles to $\bar{T}$ is equal to $\prod_{t \in \bar{T}} X_t$ and that if (A3)(i) holds in addition, then the feeble topology on $\bar{S}_G$ is the same as the topology of pointwise convergence, i.e., the product topology of $\prod_{t \in \bar{T}} X_t$.

We are going to present a specification of the externality map so that, in an explicit way, the entire distribution of the actions of the players in $\hat{T}$ may matter for the payoff of each player. However, in some contexts, such a specification is still too narrow. For example, payoffs of players may depend on both the distribution of actions chosen by men and that of the actions chosen by women. We will cover this kind of example by allowing the payoffs of players to depend on the distributions of the actions played in each one of countably many sub-populations of $\hat{T}$.

One may also think of examples of the following kind. Suppose the players in $\hat{T}$ are workers, which may be of different productivity. Now if the total output of workers is relevant for payoffs, then it is not just the distribution of actions, efforts say, of the players in $\hat{T}$ which matters for payoffs, but rather the joint distribution of actions and productivity attributes of these players. To capture this sort of example, we consider a space C of players' attributes (or characteristics) and a map $c\colon \hat{T} \to C$ assigning attributes to the players in $\hat{T}$. The following is supposed to hold.

(A12) (i) C is a completely regular Souslin space.
(ii) The map $c\colon \hat{T} \to C$ is measurable.

In most applications, C will be a Polish space. However, for sake of generality, and for symmetry with Assumption (A2) on the actions universe X, we just assume (A12)(i).

At a first glance it may look odd to have the function c to be defined only on the subset $\hat{T}$ of T. However, this is not a restriction. In fact, the externality of the game will be defined in such a way that the payoff of any player may depend on the entire action profile of the players in $\bar{T}$, and attributes of a player in $\bar{T}$ that are relevant for payoffs of other players may be considered as incorporated already in the identity of this player as point in $\bar{T}$.

Summing up, we want to give a specification of the externality of a game so that, in an explicit way, situations are described where each player's payoff may depend on the strategy profile of the players in $\bar{T}$ and on the vector of the joint distributions of the actions and players' attributes appearing in each one of countably many sub-populations of $\hat{T}$. To this end, write $M^1_+(X \times C)$ for the set of all Borel probability measures on $X \times C$. Let J be a non-empty countable set and suppose that for each $j \in J$ a non-negligible measurable subset T_j of $\hat{T}$ is given. Finally, let $\tilde{e}\colon S_G \to \prod_{t \in \bar{T}} X_t \times \left(M^1_+(X \times C)\right)^J$ be the map given by setting

$$\tilde{e}(f) = \left(f|_{\bar{T}}, \left\langle (1/\nu(T_j))(\nu|_{T_j}) \circ (f|_{T_j}, c|_{T_j})^{-1} \right\rangle_{j \in J}\right)$$

for every $f \in S_G$. Note that if (A12)(ii) holds, then the distributions involved in the above expression are defined. The map $\tilde{e}$ is now taken to be the externality map of a game. Let $\tilde{E}_G$ denote the image of S_G under $\tilde{e}$, i.e., $\tilde{E}_G = \tilde{e}(S_G)$. Now the payoff function of player t is taken to be a function $u_t\colon X_t \times \tilde{E}_G \to [-\infty, +\infty]$, thus being of the form that was intended.

In the context of an externality map $\tilde{e}$ as defined here, we summarize a game by a list $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$, on the understanding that the constraint correspondences A_t are defined on $\tilde{E}_G$, and the payoff functions u_t on the respective sets $X_t \times \tilde{E}_G$. As may be seen from the proof of Theorem 6 below, the form of the map $\tilde{e}$ is just a concrete version of the form in which the externality map was defined in Section 4.1.

As payoff functions are now defined on the respective sets $X_t \times \tilde{E}_G$, and constraint correspondence on $\tilde{E}_G$, we have to adjust the definition of continuous security stated in Section 4.3, putting it into terms of $\tilde{e}$ and $\tilde{E}_G$. In particular, we have to choose a topology on the set $\tilde{E}_G$. With the following choice we will get a statement of continuous security, in terms of $\tilde{e}$ and $\tilde{E}_G$, which will turn out to be topologically equivalent to that in Section 4.3. Assuming that (A2) and (A12)(i) hold, we regard $M^1_+(X \times C)$ as being endowed with the narrow topology, and the action sets X_t of the players in $\bar{T}$ as being endowed with the subspace topology defined from the topology of X. Now we give the set $\tilde{E}_G$ the subspace topology defined from the product topology of $\prod_{t \in \bar{T}} X_t \times \big(M^1_+(X \times C)\big)^J$.

We use the abbreviation CS' to differentiate the following notion of continuous security from the version called CS in Section 4.3.

Definition 3. A game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ is said to satisfy *CS'* if whenever $y \in \tilde{E}_G$ is such that there is no equilibrium strategy profile f with $\tilde{e}(f) = y$, there is a neighborhood U of y in $\tilde{E}_G$, a Caratheodory correspondence $\varphi\colon T \times U \twoheadrightarrow X$, and a measurable function $\alpha : T \to [-\infty, +\infty]$ such that:

(1) For each $y' \in U$, $\varphi(t, y') \subseteq A_t(y')$ for all $t \in T$.

(2) For all $y' \in U$ and all $t \in \bar{T}$, $\varphi(t, y')$ is convex or there is a finite-dimensional subspace of X which contains $\varphi(t, y')$.

(3) For each $y' \in U$, $u_t(x, y') \geq \alpha(t)$ for almost all $t \in T$ and all $x \in \varphi(t, y')$.

(4) If f is a strategy profile with $\tilde{e}(f) \in U$, $f(t) \in A_t(\tilde{e}(f))$ for almost all $t \in \hat{T}$, and $f(t) \in \mathrm{co}A_t(\tilde{e}(f))$ for almost all $t \in \bar{T}$, then there is a non-negligible set $T' \subseteq T$ such that for every $t \in T' \cap \hat{T}$, $u_t(f(t), \tilde{e}(f)) < \alpha(t)$, and for every $t \in T' \cap \bar{T}$, $f(t) \notin \mathrm{co}\{x \in A_t(\tilde{e}(f))\colon u_t(x, \tilde{e}(f)) \geq \alpha(t)\}$.

The next theorem will be proved as a consequence of Theorem 2. The proof will show, in particular, that CS' can be reduced to CS.

Theorem 6. *Let* $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ *be a game satisfying* (A1)-(A4), (A11), (A12), (S2)*, and* CS'*. Then* G *has a Nash equilibrium.*

(See Section 11.5 for the proof.)

Remark 12. The case where there is no attribute function can be regarded as a special case of the framework of this section, by simply letting the attribute space C be any singleton in this case. Thus Theorem 6, as well as Theorem 7 below, continues to be true for a game where there is no attributes function c, and $\tilde{e}$ is defined just in terms of distributions of actions.

Remark 13. As said above, the way the externality map $\tilde{e}$ has been defined is just a concrete version of the form in which the externality map was defined in Section 4.1. In fact, in the proof of Theorem 6 it is shown that, for some choice of an externality map e as defined in Section 4.1, $\tilde{E}_G \equiv \tilde{e}(S_G)$ may be homeomorphically identified with $E_G \equiv e(S_G)$. Thus Theorem 4 continues to hold with CS' in place of CS, with (A1)-(A4), (A11), (A12), (S2), and with (A7)-(A10) formulated in terms of $\tilde{E}_G$ instead of E_G. In particular, by the discussion following the statement of Theorem 4, CS' holds in games as specified in this section if $T = \hat{T}$, if there are no constraint correspondences (i.e., if $A_t(y) = X_t$ for all $t \in T$) and no attribute function, if (A1)-(A3), (A7) and (S2) hold (with (A7) holding for $y \in \tilde{E}_G$), and if all payoff functions are continuous.

The next remark relates our results to those of Khan & Sun (1999), Carmona & Podczeck (2009), and Keisler & Sun (2009).

Remark 14. In these papers, a game is given by a super-atomless complete probability space (T, Σ, ν) of players,[6] a finite partition $\langle T_i \rangle_{i \in I}$ of T into non-negligible measurable sets, a common compact metric action space K for all players, and a payoff function $V(t)$ for each $t \in T$, where $V(t)$ is the value at t of a measurable function $V : T \to C(K \times M^1_+(K)^I)$, denoting by $M^1_+(K)$ the space of Borel probability measures on K, endowed with the narrow topology, and by $C(K \times M^1_+(K)^I)$ the space of real-valued continuous functions on $K \times M^1_+(K)^I$, endowed with the sup-norm. In terms of the present section, this yields a game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t \rangle_{t \in T}, \tilde{e})$ specified as follows, so that Theorem 6 applies. For every $t \in T$, let $X_t = K$. By what was noted in Remark 3, K can be viewed as a subset of a Souslin locally convex space X, so that (A2) holds for G. Concerning the externality, let $\hat{T} = T$ and define $\tilde{e} : S_G \to M^1_+(X)^I$ by setting $\tilde{e}(f) = \big\langle (1/\nu(T_i))(\nu|_{T_i}) \circ (f|_{T_i})^{-1} \big\rangle_{i \in I}$ for each $f \in S_G$. Observe that for each $i \in I$, $(1/\nu(T_i))(\nu|_{T_i})$ is an atomless probability measure on T, so for any $y \in M^1_+(K)$ we have $y = (1/\nu(T_i))(\nu|_{T_i}) \circ (f|_{T_i})^{-1}$ for some $f \in S_G$. Hence,

[6] Recall that an atomless Loeb probability space is super-atomless, and that the notion of saturated probability space in Keisler & Sun (2009) and that of super-atomless probability space are equivalent.

because $\langle T_i \rangle_{i \in I}$ is a partition of T, we have $\tilde{E}_G \equiv \tilde{e}(S_G) = M^1_+(K)^I$. Thus for each $t \in T$, we may set $u_t = V(t)$. Evidently (A1), (A3), and (S2) hold for G. The game G satisfies CS'. Indeed, note first that as the map $t \mapsto u_t \colon T \to C(K \times M^1_+(K)^I)$ is measurable, so is the map $(t, x) \mapsto (u_t, x) \colon T \times K \to C(K \times M^1_+(K)^I) \times K$, and that the map $(u, x) \mapsto u(x, y) \colon C(K \times M^1_+(K)^I) \times K \to \mathbb{R}$ is continuous for each fixed $y \in \tilde{E}_G$. Thus for each $y \in \tilde{E}_G$, the map $(t, x) \mapsto u_t(x, y) \colon T \times K \to \mathbb{R}$ is measurable, being the composition of two measurable maps. Thus G satisfies (A7). Consequently, by what was noted in Remark 13, G satisfies CS'. Finally, as the space (T, Σ, ν) of players is super-atomless, (S2) holds for G. Thus, by Remark 12, Theorem 6 applies, showing that G has an equilibrium. Thus Theorem 6 implies Theorem 1 in Khan & Sun (1999), Corollary 4(4) in Carmona & Podczeck (2009), as well as the necessity part of Theorem 4.6 in Keisler & Sun (2009). In fact, these latter results are implied by Theorem 2, as Theorem 6 is a consequence of Theorem 2 (see Section 11.5).

8 Generalized payoff secure games

In this section we take our model in the version as introduced in the previous section and state a result providing sufficient conditions for CS' to hold. The two conditions in this result, presented in the following definitions, are versions of the notions of "generalized payoff security" and of "better-reply closed" game, introduced for finite-player games by Barelli & Soza (2009) and Carmona (2011), respectively.[7]

Definition 4. A game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ is said to satisfy *GPS* if for all $y \in \tilde{E}_G$ and $\varepsilon > 0$ there is a neighborhood U of y in $\tilde{E}_G$, a Caratheodory correspondence $\varphi : T \times U \twoheadrightarrow X$, and a measurable function $\alpha \colon T \to \mathbb{R}$ such that:

(a) For each $y' \in U$, $\varphi(t, y') \subseteq A_t(y')$ for all $t \in T$.

(b) For all $y' \in U$ and all $t \in \bar{T}$, $\varphi(t, y')$ is convex or there is a finite-dimensional subspace of X which contains $\varphi(t, y')$.

(c) For each $y' \in U$, $u_t(x, y') \geq \alpha(t)$ for almost all $t \in T$ and all $x \in \varphi(t, y')$.

(d) There is a $T_\varepsilon \subseteq T$ with $\nu(T_\varepsilon) < \varepsilon$ such that $\alpha(t) \geq w_t(y) - \varepsilon$ for all $t \in T \backslash T_\varepsilon$.

Definition 5. A game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ is said to satisfy *BRC* if the following holds for any strategy profile f: If there is a sequence $\langle f_n \rangle$ of strategy

[7] For finite-player games, our definition of generalized payoff security is equivalent to the original definition of that notion in Barelli & Soza (2009). Note also that for finite-player games, the property of a game being better reply closed is equivalent to the property of "weak reciprocal upper semi-continuity" at all non-equilibrium strategy profiles; see Carmona (2011, Theorem 5), and see Bagh & Jofre (2006) for the definition of weak reciprocal upper semi-continuity.

profiles with $\tilde{e}(f_n) \to \tilde{e}(f)$ such that, for almost all $t \in T$, (a) $f_n(t) \in A_t(\tilde{e}(f_n))$ for all $n \in \mathbb{N}$, (b) $f(t) \in \operatorname{LS} f_n(t)$, and (c) $\underline{\lim}_n u_t(f_n(t), \tilde{e}(f_n)) \geq w_t(\tilde{e}(f))$, then f is an equilibrium of G.

Here $\operatorname{LS} f_n(t)$ denotes the set of all cluster points of the sequence $\langle f_n(t) \rangle$. Note that under (A11) and (A3)(i), $e(f_n) \to \tilde{e}(f)$ implies $f_n(t) \to f(t)$ for each $t \in \bar{T}$ (see the definition of $\tilde{e}$ and the paragraph after the statement of (A11)), so that (b) in Definition 5 reduces to a condition on the restrictions of strategy profiles to $\hat{T}$.

It is common in applications to assume for an atomic player that his action set is convex, that his constraint correspondence takes convex values, and that his payoff function is quasi-concave in his action. Such convexity properties make the definition of CS' easier as we have noted in Remark 8, and therefore we will assume them in the following theorem. Convexity of the action sets of such players is already part of (A4). Thus we introduce here:

(A13) For every $t \in \bar{T}$, $A_t(y)$ is convex and $u_t(\cdot, y)$ is quasi-concave for all $y \in \tilde{E}_G$.

Theorem 7. *Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ be a game satisfying* (A1)-(A4), (A11)-(A13), (S2), GPS *and* BRC. *Then G also satisfies* CS', *and consequently, by Theorem 6, G has a Nash equilibrium.*

(See Section 11.6 for the proof.)

9 An Application: Optimal Income Taxation

We consider a version of the model of Mirrlees (1971) on optimal income taxation. Specifically, we develop a general framework to include several optimal taxation problems and address the existence of an optimal income tax in this framework. Our existence result for Nash equilibria will be used in the step establishing that the choice set of the government as the agency that chooses the tax is non-empty. Due to implementability constraints, non-emptiness of this set is a non-trivial issue in general; see Example 7 below.

The economy consists of a continuum of individuals, described by a super-atomless complete probability space $(\hat{T}, \hat{\Sigma}, \hat{\nu})$, and a government. There is a single consumption good, which can be produced using labor. Each individual $t \in \hat{T}$ is endowed with one unit of time and is described by his skill level n_t, which is the quantity of labor provided by t per unit of time. We assume that there is an upper bound $\bar{n} > 0$ on the level of skills and an upper bound $\bar{m} \geq \bar{n}$ on consumption. Writing $M = [0, \bar{m}]$ and $L = [0, 1]$, an individual t is further characterized by a continuous utility function $\tilde{u}_t : M \times L \to \mathbb{R}_+$, so that his utility is $\tilde{u}_t(m, l)$ when his individual consumption is m and his effort level is l. Let $N = [0, \bar{n}]$ and let $\hat{n} \colon \hat{T} \to N$ denote the function $t \mapsto n_t$. We make the following assumptions:

(T1) The map $t \mapsto \tilde{u}_t$ from $\hat{T}$ to $C(M \times L)$ is measurable, where $C(M \times L)$ denotes the space of real-valued continuous functions on $M \times L$ endowed with the sup-norm.

(T2) For every $t \in \hat{T}$, $\tilde{u}_t$ is strictly increasing in m, strictly decreasing in l, and $0 \leq \tilde{u}_t(m,l) \leq 1$ for all $(m,l) \in M \times L$.

(T3) The map $\hat{n}$ is measurable.

(T4) The distribution $\hat{\nu} \circ \hat{n}^{-1}$ is atomless.

As in Golosov, Kocherlakota & Tsyvinski (2003), one unit of labor is transformed into one unit of consumption. This assumption is made for simplicity since, normalizing the price of consumption to one, it implies that the equilibrium price of labor is equal to one, too.[8]

The government chooses an income tax, which, as in Mirrlees (1971), is described by a function $\lambda : [0, \bar{n}] \to \mathbb{R}_+$, with the interpretation that someone with income z cannot consumer more that $\lambda(z)$ after tax (note that $[0, \bar{n}]$ is the set of possible incomes). As in Mirrlees (1971), income taxes are non-decreasing and right-continuous (see Proposition 2 in Mirrlees (1971)). In addition, we assume that $\lambda(\bar{n}) \leq \bar{m}$.

The underlying assumption, here as well as in Mirrlees (1971), is that the government can observe the income level of an individual but neither her skill nor her effort level. Thus, we assume that the government observes neither the function $\hat{n}$ assigning skills to individuals nor the effort level chosen by individuals. Specifically, the government observes only the joint distribution of skills, utility functions, consumption and effort levels.

Let Λ be the set of all non-decreasing right-continuous functions $\lambda : [0, \bar{n}] \to \mathbb{R}_+$ satisfying $\lambda(\bar{n}) \leq \bar{m}$. Define a metric ρ on Λ by setting

$$\rho(\lambda, \lambda') = \int_{[0,\bar{n}]} |\lambda(z) - \lambda'(z)| \mathrm{d}z + |\lambda(\bar{n}) - \lambda'(\bar{n})|$$

for $\lambda, \lambda' \in \Lambda$, where the integral is with respect to Lebesgue measure on $[0, \bar{n}]$.[9] (That ρ is indeed a metric, and not just a pseudo-metric, is true because two distinct elements of Λ differ at $\bar{n}$ or, being non-deceasing and right-continuous, differ on some non-empty open subset of $[0, \bar{n})$.) In the sequel, Λ is always regarded as endowed with this metric. By Lemma 5 in Section 11.7.1, Λ is then actually a compact metric space.

We will address the existence of an optimal income tax via a game played by the government and the individuals. Modeling income taxes as above implies that the government's choice set in this game will be compact. This would not be the case would one focus on income taxes that are continuous or on general incentive-feasible

[8] Without this assumption, we would need to add an auctioneer to the game used to show existence of equilibrium.

[9] This form of a metric on Λ was suggested by a referee.

mechanisms (the latter being considered in Golosov et al. (2003)).[10] However, allowing for discontinuous income taxes will imply that the individuals' and the government's payoff functions are discontinuous. Nevertheless, despite of such discontinuities, continuous security will hold and will allow to prove existence of an optimal income tax.

To unify several optimal taxation problems, we consider the possibility that the government in its choice of an income tax is restricted. To this end, let $\hat{C}$ be the closure of the set $\{\tilde{u}_t : t \in \hat{T}\}$ in $C(M \times L)$ and let $C = \hat{C} \times N$. Then C is a complete separable metric space which contains the relevant attributes of the individuals. Let $M_+^1(C \times M \times L)$ be the set of Borel probability measures on $C \times M \times L$. Define $c : \hat{T} \to C$ by $c(t) = (\tilde{u}_t, n_t)$, $t \in \hat{T}$, and let $K = \{\gamma \in M_+^1(C \times M \times L) : \gamma_C = \hat{\nu} \circ c^{-1}\}$ be the set of distributions over attributes and actions. We regard K as being endowed with the subspace topology defined from the narrow topology of $M_+^1(C \times M \times L)$. Let $\Theta : K \twoheadrightarrow \Lambda$ be a correspondence with the interpretation that, given $\gamma \in K$, the government may choose income taxes only from $\Theta(\gamma)$.

Given $\gamma \in K$, the distribution of the map $(u, n, m, l) \mapsto nl$ from $C \times M \times L$ to $[0, \bar{n}]$ is denoted by $\hat{\gamma}$. Thus $\hat{\gamma}$ is the distribution of outputs determined by γ, or, in other words, the pre-tax income distribution given by γ. We write $M_+^1([0, \bar{n}])$ for the set of Borel probability measures on $[0, \bar{n}]$, and $\hat{K} \subseteq M_+^1([0, \bar{n}])$ for the image of K under the map $\gamma \mapsto \hat{\gamma}$. As no confusion can arise, the symbol $\hat{\gamma}$ will also be used to denote generic elements of $\hat{K}$. We give $\hat{K}$ the subspace topology defined from the narrow topology of $M_+^1([0, \bar{n}])$, so that the map $\gamma \mapsto \hat{\gamma}$ becomes continuous.

When looking for an optimal income tax, the government faces a constraint in addition to that defined by the correspondence Θ. As in Mirrlees's (1971), the government is constrained by an implementability condition: the allocation g that results from the choice of a given income tax λ must be such that (λ, g) is an equilibrium of the economy. A precise definition of equilibrium is given below. Informally, an equilibrium is a feasible allocation g together with a tax function λ such that all individuals optimize and such that the government runs a balanced budget; however, specific optimal tax problems may require additional properties to be met. Writing $S(E)$ for the set of equilibria of an economy E, the *government's optimization problem* is

$$\max_{(\lambda, g) \in S(E)} \int_{\hat{T}} \tilde{u}_t(g(t)) \mathrm{d}\hat{\nu}(t).$$

An income tax λ^* is an *optimal income tax* if there exists a $g^* : \hat{T} \to M \times L$ such that (λ^*, g^*) is a solution of the government's optimization problem.

There is a commitment aspect implicit in the above definition of an optimal income tax. In fact, it could happen that, while social welfare, i.e., $\int_{\hat{T}} \tilde{u}_t(g(t)) \mathrm{d}\hat{\nu}(t)$, is being

[10]Another reason for focusing on taxes instead of general incentive-feasible mechanisms has to do with decentralization; see, for instance, Section 4.3 in Kocherlakota (2010).

maximized from an ex-ante perspective, ex-post, after individuals have made effort decisions, but before they actually carry out consumption, the government is able to change the tax (in other words, to adjust the redistribution of aggregate output) so as to increase social welfare even further. In Mirrlees (1971), such a case is ignored, i.e., it is assumed that the government can commit to an ex-ante chosen tax system.

Below we will also consider an optimal tax problem where the government is able to revise the tax after individuals have made effort decisions but before consumption is carried out (see Example 7 below). In this case, an optimal income tax should have the property that the government has no incentive to make use of this opportunity. We will take care of this through the definition of equilibrium as follows. We introduce the function $v\colon \Lambda \times K \to \mathbb{R}$, defined by setting $v(\lambda, \gamma) = \int u(\lambda(nl), l)\mathrm{d}\gamma(u, n, m, l)$ for all $(\lambda, \gamma) \in \Lambda \times K$, and require that in an equilibrium (λ^*, g^*), where λ^* is the tax function and $g^* = (m^*, l^*)$ the allocation, λ^* is a solution of the problem $\max_\lambda v(\lambda, \hat{\nu} \circ (c, g^*)^{-1})$ subject to the condition $\lambda \in \Theta(\hat{\nu} \circ (c, g^*)^{-1})$ and to the balanced budget condition. As $v(\lambda^*, \hat{\nu} \circ (c, g^*)^{-1}) = \int_{\hat{T}} \tilde{u}_t(g^*(t))\mathrm{d}\hat{\nu}(t)$ by change of variables, this requirement implies that in an equilibrium (λ^*, g^*) the government cannot increase social welfare by an ex-post change of the tax function subject to the balanced budget condition and the constraints set by Θ.

For convenience of notation, we will require this condition also for an equilibrium in the case where the government can commit to an ex-ante chosen tax function, but in this case with the function v given by $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$, so that the condition effectively reduces to the requirement $\lambda \in \Theta(\hat{\nu} \circ (c, g^*)^{-1})$ and the balanced budget condition. For sake of generality, in view of potential applications, we will actually allow for an abstract function $v\colon \Lambda \times K \to \mathbb{R}$ in our model.

We summarize an *economy* by a list $E = \langle (\hat{T}, \hat{\Sigma}, \hat{\nu}), M, L, N, \hat{n}, \Lambda, \Theta, v, \langle \tilde{u}_t, n_t \rangle_{t \in \hat{T}} \rangle$. An *equilibrium* for an economy E is an income tax λ^* together with a pair $g^* = (m^*, l^*)$, where $m^* : \hat{T} \to M$ and $l^* : \hat{T} \to L$ are measurable functions, such that:

(a) λ^* solves $\max_\lambda v(\lambda, \hat{\nu} \circ (c, g^*)^{-1})$ subject to the conditions $\lambda \in \Theta(\hat{\nu} \circ (c, g^*)^{-1})$ and $\int_{\hat{T}} \lambda(n_t l^*(t))\mathrm{d}\hat{\nu}(t) = \int_{\hat{T}} n_t l^*(t)\mathrm{d}\hat{\nu}(t)$.[11]

(b) For almost all $t \in \hat{T}$, $g^*(t)$ solves $\max_{(m,l) \in M \times L} \tilde{u}_t(m, l)$ subject to $m \leq \lambda^*(n_t l)$.

Conditions (a) and (b) together imply that g^* is a competitive equilibrium allocation. Indeed, by the monotonicity assumption in (T2), we must have $m^*(t) = \lambda^*(n_t l^*(t))$ for almost all $t \in \hat{T}$. Hence $\int_{\hat{T}} m^*(t)\mathrm{d}\hat{\nu}(t) = \int_{\hat{T}} \lambda^*(n_t l^*(t))\mathrm{d}\hat{\nu}(t) = \int_{\hat{T}} n_t l^*(t)\mathrm{d}\hat{\nu}(t)$ and thus market clearing holds.

[11] Abusing notation, we sometimes write m to denote a function from $\hat{T}$ to M, instead a generic element of M; similarly for l and L. The meaning should be clear from the context.

Note that for each individual $t \in \hat{T}$, $n_t l^*(t) - \lambda^*(n_t l^*(t))$ is just the difference between pre-tax income and after-tax income. Thus (a) requires, in particular, that in an equilibrium the government runs a balanced budget.

The following five examples clarify our framework and illustrate the flexibility that our framework allows. These examples also illustrate the assumptions we make on Θ and v, which we now state.

(T5) Θ is well-behaved and takes convex values.

(T6) v is usc, $v(\cdot, \gamma)$ is quasi-concave for all $\gamma \in K$, $v(\lambda, \cdot)$ is continuous for each continuous $\lambda \in \Lambda$, and $v(\lambda, \gamma) \geq 0$ for all $(\lambda, \gamma) \in \Lambda \times K$.

(T7) For all $\gamma \in K$ and $\varepsilon > 0$, there exists an open neighborhood O of γ and a continuous map $\psi : O \to \Lambda$ such that for all $\gamma' \in O$,

  (i) $\psi(\gamma') \in \Theta(\gamma')$,

  (ii) $\int \psi(\gamma')(z) \mathrm{d}\hat{\gamma}'(z) = \int z \mathrm{d}\hat{\gamma}'(z)$,

  (iii) $v(\psi(\gamma'), \gamma') > v(\lambda, \gamma) - \varepsilon$ for all $\lambda \in \Theta(\gamma)$ such that $\int \lambda(z) \mathrm{d}\hat{\gamma}(z) = \int z \mathrm{d}\hat{\gamma}(z)$.

Example 3. Suppose the government is unconstrained in its choice of an income tax and can commit to an income tax it proposes, which is the case considered in Mirrlees (1971). In the notation of our framework this means $\Theta(\gamma) = \Lambda$ and $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. It is clear that (T5) and (T6) are satisfied in this example. As for (T7), simply let $O = K$ and $\psi(\gamma') = \lambda_0$ for all $\gamma' \in O$, where λ_0 is the identity, i.e. $\lambda_0(z) = z$ for all $z \in N$.

Example 4. Let $\Theta(\gamma) = \{\lambda : \lambda(\bar{n} - \xi) \geq \zeta\}$ for all $\gamma \in K$, where $\xi > 0$ is a small number, and $0 < \zeta < \bar{n} - \xi$ a high number. This case can be interpreted as one where the government commits to income taxes that give high work incentives for highly skilled individuals. Alternatively, this case can be regarded as arising because, if taxed at a high tax rate, high skill individuals will choose to evade taxation. Clearly the subset $\{\lambda : \lambda(\bar{n} - \xi) \geq \zeta\}$ of Λ is convex, and by Lemma 6 in Section 11.7, it is closed. Thus (T5) holds. Let $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. Then (T6) holds. As in the previous example, letting $O = K$ and $\psi(\gamma') = \lambda_0$ for all $\gamma' \in O$ shows that (T7) holds, too.

Example 5. In this example we consider the case where the government ceases to function as total output approaches zero. This is modeled by specifying Θ as follows. First, only the 0% income tax λ_0 is allowed if total output is zero, with the interpretation that in this case the government no longer exists and thus, in particular, cannot redistribute income. Second, for total output larger than zero, the income taxes the government can implement are those with a distance to the 0%

income tax not exceeding a number which depends continuously on $\hat{\gamma}$, i.e., on the distribution of outputs.

Specifically, let $\hat{\gamma}_0 \in \hat{K}$ be Dirac measure at $0 \in [0, \bar{n}]$. Set $\Theta(\gamma) = \{\lambda_0\}$ if $\hat{\gamma} = \hat{\gamma}_0$, and for some continuous function $\hat{\gamma} \mapsto \varepsilon(\hat{\gamma})\colon \hat{K} \to \mathbb{R}$, with $\varepsilon(\hat{\gamma}_0) = 0$ and $\varepsilon(\hat{\gamma}) > 0$ for $\hat{\gamma} \neq \hat{\gamma}_0$, set $\Theta(\gamma) = C_{\varepsilon(\hat{\gamma})}(\lambda_0)$ for $\gamma \in K$ with $\hat{\gamma} \neq \hat{\gamma}_0$, writing $C_{\varepsilon(\hat{\gamma})}(\lambda_0)$ for the closed ball of radius $\varepsilon(\hat{\gamma})$ around λ_0 for the metric ρ on Λ. Further, let $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. As in the previous two examples, (T6) and (T7) hold. Clearly Θ is well-behaved, as both the maps $\gamma \mapsto \hat{\gamma}$ and $\hat{\gamma} \mapsto \varepsilon(\hat{\gamma})$ are continuous. Also, $C_{\varepsilon(\hat{\gamma})}(\lambda_0)$ is convex for all $\gamma \in K$, by the definition of ρ. Thus (T5) holds.

We can now consider a fourth example, obtained by specifying Θ so as to capture the basic idea of the credible income taxation problem in Farhi, Sleet, Werning & Yeltekin (2011).

Example 6. As in Golosov et al. (2003), assume, in addition to (T1) and (T2), that $\tilde{u}_t(m, l) = \pi(m) + \eta(l)$ for all $t \in \hat{T}$ and $(m, l) \in M \times L$, where $\pi : M \to \mathbb{R}$ and $\eta : L \to \mathbb{R}$ are functions with π strictly concave and $\pi(0) = \eta(0) = 0$. Let

$$\Theta(\gamma) = \begin{cases} \tilde{\Theta}(\gamma) & \text{if } \pi\left(\int z \mathrm{d}\hat{\gamma}(z)\right) + \int \eta(l) \mathrm{d}\gamma(u, n, m, l) > 0, \\ \Lambda & \text{otherwise,} \end{cases}$$

$\gamma \in K$, where

$$\tilde{\Theta}(\gamma) = \left\{\lambda' \in \Lambda : \int \pi(\lambda'(z)) \mathrm{d}\hat{\gamma}(z) + \delta \int \eta(l) d\gamma(u, n, m, l) \geq (1 - \delta)\pi\left(\int z \mathrm{d}\hat{\gamma}(z)\right)\right\}.$$

Further, let $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. Again, (T6) holds, and in Lemma 13 in Section 11.7.3 it is shown that the same is true for (T5) and (T7), and that in an equilibrium (λ^*, g^*) of E the following must hold, writing $g^*(t) = (m^*(t), l^*(t))$:

$$\text{(1)} \qquad \int_{\hat{T}} \left(\pi(\lambda^*(n_t l^*(t))) + \eta(l^*(t))\right) \mathrm{d}\hat{\nu}(t) \geq (1 - \delta)\left(\pi\left(\int_{\hat{T}} \lambda^*(n_t l^*(t)) \mathrm{d}\hat{\nu}(t)\right) + \int_{\hat{T}} \eta(l^*(t)) \mathrm{d}\hat{\nu}(t)\right).$$

This condition is analogous to the credibility condition of Farhi et al. (2011) and implies that the government credibly commits to λ^* in the following sense. Suppose that in each one of infinitely many periods $k \in \mathbb{N}$, the government and the individuals simultaneously choose an income tax λ_k and consumption/effort pairs $(m_k(t), l_k(t))$, respectively, knowing the entire history of previous taxes and joint distributions over attributes and actions.[12] In addition, assume that for each individual and the government the utility in the repeated interaction is the discounted sum of the period-wise

[12]Note that we are restricting the government to choose income taxes that do not depend on the tax paid previously by individuals, which is something that the government observes. This is an

utilities, with discount factor $\delta \in (0,1)$, where the period-wise utility function of the government is given by $\int_{\hat{T}} \big(\pi(\lambda_k(n_t l_k(t))) + \eta(l_k(t))\big) \mathrm{d}\hat{\nu}(t)$ for all $k \in \mathbb{N}$, λ_k being the income tax and l_k the effort allocation in period k.[13] In this setting, the above credibility condition states that the stationary outcome with (λ^*, g^*) in every period is a subgame perfect equilibrium of the repeated game just described.[14]

At the core of the above example is the inability of the government to commit to an income tax. In fact, the right-hand side of condition (1) describes the best short-run deviation of the government, which, by the fact that individuals' choices are made simultaneously, consists in choosing an income tax that gives all individuals the same after-tax income.

In a final example, we consider another non-commitment problem. Specifically, it is supposed now that, after individuals have made effort decisions, but before they carry out consumption, the government may revise a tax announced earlier.

Example 7. Let Θ be as in Example 5, and set $v(\lambda, \gamma) = \int u(\lambda(nl), l) \mathrm{d}\gamma(u, n, m, l)$ for all $(\lambda, \gamma) \in \Lambda \times K$. Then, as in Example 5, (T5) holds. If, in addition to what is supposed in (T1) and (T2), $\tilde{u}_t(\cdot, l)$ is concave for all $t \in \hat{T}$ and $l \in L$, then by Lemma 14 in Section 11.7.3, (T6) and (T7) hold, too. Now under the assumption stated in the previous paragraph, an equilibrium implementing an optimal income tax must have the property that the income tax maximizes aggregate utility subject to the feasibility constraints set by Θ and the given total output. In view of (a) of the equilibrium definition, this is guaranteed by specifying the government's auxiliary utility function v as above. (Recall in this regard that the specification of Θ in Example 5 means, in particular, that the feasibility sets $\Theta(\gamma)$ of the government do not depend on individuals' consumption, but only on the distribution of outputs.) While in the previous examples an equilibrium can be easily constructed (there is an equilibrium with a 0% income tax in Examples 3-5, and an equilibrium with a 100% income tax in Examples 3 and 6), this is not the case in this example. In particular,

important restriction. In fact, if the outcome in the first period is fully revealing (i.e. individuals with different skills have different income levels) then the first-best could be achieved from period 2 onwards. Alternative assumptions to rule out this case include: (1) the government is legally obliged to tax only individuals' incomes, or (2) each individual lives only for one period and so pays taxes only once. The latter assumption means that individuals are short-lived and the government is long-lived; this poses no difficulties within a repeated-game framework and is, in fact, standard (see, for instance, Mailath & Samuelson (2006, Section 2.7) or Sabourian (1990)). We note that (2) is similar to the overlapping generations assumption in Farhi et al. (2011).

[13] Note that $\int_{\hat{T}} (\pi(\lambda_k(n_t l_k(t))) + \eta(l_k(t))) \mathrm{d}\hat{\nu}(t) = \int_{N \times L} (\pi(\lambda_k(nl)) + \eta(l)) \mathrm{d}\hat{\nu} \circ (\hat{n}, l_k)^{-1}(n, l)$ for all $k \in \mathbb{N}$, so all the government needs to know is the joint distribution of skills and efforts.

[14] Condition (1) is also necessary for a stationary outcome to be subgame perfect (see Chari & Kehoe (1990)) since the payoff of the government in the worst subgame perfect equilibrium from its point of view is $\pi(0) + \eta(0) = 0$.

there cannot be an equilibrium with a 0% income tax given the specification of v, and there cannot be an equilibrium with a 100% income tax given the specification of Θ.

Here is our theorem on the existence of an optimal income tax.

Theorem 8. *If the economy* $E = \langle (\hat{T}, \hat{\Sigma}, \hat{\nu}), M, L, N, \hat{n}, \Lambda, \Theta, v, \langle \tilde{u}_t, n_t \rangle_{t \in \hat{T}} \rangle$ *satisfies* (T1)-(T7)*, then there exists an optimal income tax.*

(See Section 11.7.2 for the proof.)

Remark 15. The problem of existence of an optimal income tax can be decomposed into two parts: Existence of an equilibrium and existence of a solution to the government's optimization problem. While we establish the second part using standard techniques, the first part requires the use of our main existence result.

In many applications considered in the literature, as well as in some of those considered above, an equilibrium can be easily constructed. This is the case, for instance, in the model of Mirrlees (1971), which was considered in Example 3. However, our framework considerably generalizes that in Mirrlees (1971) to include, in particular, the optimal tax model of Farhi et al. (2011) as well as new examples of economically relevant optimal tax problems. Because of this generality, it is not a simple matter to get an equilibrium in our framework. This is precisely illustrated in Example 7 above.

Remark 16. It is important to note that an optimal income tax in our general framework cannot be interpreted as being in any sense superior to the income tax obtained in Mirrlees (1971). This is so because the optimal taxation problem in Mirrlees (1971) is the one which imposes the smallest set of restrictions on the government's choice. Thus, the relevance of our framework and our result does not consist in obtaining "better" optimal income taxes, but rather in obtaining the existence of optimal income taxes in interesting optimal taxation problems, especially in problems where equilibria of the economy cannot be found easily by construction (cf. Example 7).

To establish non-emptiness of the choice set of the government, i.e., of the equilibrium set of the economy, we will use Theorem 7 to establish existence of a Nash equilibrium in a game constructed as follows.

Let $E = \langle (\hat{T}, \hat{\Sigma}, \hat{\nu}), M, L, N, \hat{n}, \Lambda, \Theta, v, \langle \tilde{u}_t, n_t \rangle_{t \in \hat{T}} \rangle$ be an economy satisfying (T1)-(T7). The government will be denoted by player $\bar{t}$, where $\bar{t} \notin \hat{T}$. Let $\bar{T} = \{\bar{t}\}$ and set $T = \bar{T} \cup \hat{T}$, $\Sigma = \hat{\Sigma} \cup \{B \cup \bar{T} : B \in \hat{\Sigma}\}$ and, for all $B \in \Sigma$, $\nu(B) = \hat{\nu}(B \cap \hat{T}) + \chi_B(\bar{t})$, where χ_B is the characteristic function of B. Clearly Assumption (A11) holds, and since $(\hat{T}, \hat{\Sigma}, \hat{\nu})$ is a super-atomless complete probability space by our specification of an economy, so do (A1) and (S2).

Concerning players' action spaces, let $X_{\bar{t}} = \Lambda$ and $X_t = M \times L$ for all $t \in \hat{T}$. Let $X_0 = L_1([0, \bar{n}+1])$, the space of (equivalence classes of) Lebesgue integrable functions on $[0, \bar{n}+1]$, endowed with the usual $\|\cdot\|_1$-norm. Then X_0 is a separable Banach space; in particular, X_0 is Souslin. Identify each $\lambda \in \Lambda$ with the element $f_\lambda \in X_0$ given by setting $f_\lambda(z) = \lambda(z)$ for $z \in [0, \bar{n}]$ and $f_\lambda(z) = \lambda(\bar{n})$ for $z \in (\bar{n}, \bar{n}+1]$. Then $\rho(\lambda, \lambda') = \|f_\lambda - f_{\lambda'}\|_1$ for each λ, $\lambda' \in \Lambda$, so Λ isometrically embeds into X_0. By Lemma 5 in Section 11.7.1, we may therefore regard Λ as a compact subset of X_0. Let $X_1 = \mathbb{R}^2$, and let X be obtained from X_0 and X_1 in the sense of Remark 3. We may then identify Λ and $M \times L$ with subsets of X, so that (A2)-(A4) hold. Also, we may regard strategy profiles as functions from T to X.

We take the space $C = \hat{C} \times N$, which was defined above, for the attribute space of the game, and the function c, which was defined by $c(t) = (\tilde{u}_t, n_t)$ for all $t \in \hat{T}$, for the attribute function. By (T1) and (T3), (A12) is satisfied for C and c.

Now the externality map $\tilde{e}$ and the externality space $\tilde{E}_G$ of the game have the following form. Given $f \in S_G$ (i.e., a measurable $f\colon T \to X$ with $f(\bar{t}) \in \Lambda$ and $f(t) \in M \times L$ for almost all $t \in \hat{T}$), we have $\tilde{e}(f) = (f(\bar{t}), \hat{\nu} \circ (c, \hat{f})^{-1})$ where $\hat{f} = f|_{\hat{T}}$. Thus, with K as defined above, we have $\tilde{E}_G = \Lambda \times K$. In view of this, given $y \in \tilde{E}_G$, we will often write (λ, γ) for y.

Regarding players' payoff functions, define $u_{\bar{t}}\colon \Lambda \times \tilde{E}_G \to \mathbb{R}$ by setting

$$u_{\bar{t}}(\lambda', \lambda, \gamma) = \begin{cases} v(\lambda', \gamma) & \text{if } \int (z - \lambda'(z))\, \mathrm{d}\hat{\gamma}(z) = 0, \\ -1 & \text{otherwise,} \end{cases}$$

for all $\lambda' \in \Lambda$ and $(\lambda, \gamma) \in \tilde{E}_G$. For $t \in \hat{T}$, define $u_t\colon M \times L \times \tilde{E}_G \to \mathbb{R}$ by setting $u_t(m, l, \lambda, \gamma) = \tilde{u}_t(m, l)$ for all $(m, l) \in M \times L$ and $(\lambda, \gamma) \in \tilde{E}_G$.

Finally, we specify the constraint correspondences. Define $A_{\bar{t}}\colon \tilde{E}_G \twoheadrightarrow \Lambda$ by setting $A_{\bar{t}}(y) = \Theta(\gamma)$ for all $y = (\lambda, \gamma) \in \tilde{E}_G$. For $t \in \hat{T}$, define $A_t\colon \tilde{E}_G \twoheadrightarrow M \times L$ by setting $A_t(y) = \{(m, l) \in M \times L : m \leq \lambda(n_t l)\}$ for all $y = (\lambda, \gamma) \in \tilde{E}_G$. Since the three sets $\Theta(\gamma)$, $\{\lambda' \in \Lambda : \int (z - \lambda'(z))\, \mathrm{d}\hat{\gamma}(z) = 0\}$, and $\{\lambda' \in \Lambda : v(\lambda', \gamma) \geq \alpha\}$, $\alpha \in \mathbb{R}$, are convex, it follows that (A13) is satisfied.

Let $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ be the game just defined. It is clear that every equilibrium of the economy E is a Nash equilibrium of G, and that by (T6) and (T7), every Nash equilibrium of G is an equilibrium of E.

We note that the payoff function $u_{\bar{t}}$ in the game G need not be usc, so that the existence result of Balder (2002, Theorem 2.2.1) cannot be applied to conclude the existence of an equilibrium of G. E.g., suppose that $\hat{\nu} \circ \hat{n}^{-1}$ is uniform on $[0, \bar{n}]$, $\Theta(\gamma) = \Lambda$ for all $\gamma \in K$, and $v(\lambda, \gamma) = 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. Define $\lambda' \in \Lambda$ by setting

$$\lambda'(z) = \begin{cases} z \text{ if } z < \bar{n}/2, \\ \bar{n} \text{ if } z \geq \bar{n}/2. \end{cases}$$

Let $\hat{f} : \hat{T} \to M \times L$ and $\hat{f}_k : \hat{T} \to M \times L$, $k \in \mathbb{N} \backslash \{0\}$, be given by $\hat{f}(t) = \hat{f}_k(t) = (n_t, 1)$ if $n_t < \bar{n}/2$, and by $\hat{f}_k(t) = (\bar{n}/2 - n_t/k, \bar{n}/(2n_t) - 1/k)$ and $\hat{f}(t) = (\bar{n}/2, \bar{n}/(2n_t))$

otherwise. Fix any $\lambda \in \Lambda$. Set $y_k = (\lambda, \hat{\nu} \circ (c, \hat{f}_k)^{-1})$ and $y = (\lambda, \hat{\nu} \circ (c, \hat{f})^{-1})$. Then $(\lambda', y_k) \to (\lambda', y)$. Also, $\int (\lambda'(nl) - nl) \, \mathrm{d}\hat{\nu} \circ (c, \hat{f}_k)^{-1}(u, n, m, l) = 0$ for all k, and $\int (\lambda'(nl) - nl) \, \mathrm{d}\hat{\nu} \circ (c, \hat{f})^{-1}(u, n, m, l) = \bar{n}/4 > 0$. Thus $u_{\bar{t}}(\lambda', y_k) = 0$ for all k, but $u_{\bar{t}}(\lambda', y) = -1$, showing that $u_{\bar{t}}$ is not usc. (In fact, $u_{\bar{t}}$ is not even weakly usc and player $\bar{t}$'s best-reply correspondence is not uhc.)

As announced, we will establish existence of a Nash equilibrium in the game G by appealing to Theorem 7. To this end, we need to show that G satisfies GPS and BRC. This step of the proof is isolated in the following lemma.

Lemma 1. *The game G satisfies* GPS *and* BRC.

(The proof may be found in Section 11.7.1.) Now in view of this lemma and of what was noted above, the game G satisfies all the assumptions of Theorem 7. Thus:

Theorem 9. *The game G has an equilibrium.*

10 Conclusion

We have established the existence of a pure-strategy Nash equilibrium for (the pure-strategy part) of Balder's (2002) model of large games, which has been extended as follows. First, we have relaxed the continuity assumptions on players' payoff functions made in Balder (2002), adapting Barelli & Meneghel's (2013) notion of continuous security to games with a measure space of players. Thus, in particular, we have extended the existence result of Barelli & Meneghel (2013) by dispensing the requirement that the set of players be finite. Second, under a strengthening of the notion of non-atomicity, we have generalized the way the payoff of each player depends on the actions chosen by all players, covering, in particular, the case where players' payoff functions depend on the distribution of actions chosen by all players. Thus we have generalized results in the line of Khan & Sun (1999) by allowing for discontinuous payoff functions. In short, our main existence results simultaneously generalize the existence results of Barelli & Meneghel (2013) (by dispensing with finite set of players), Balder (2002) (by allowing discontinuous payoff functions and by generalizing the way players' payoffs depend on the actions chosen by all players), and Khan & Sun (1999) (by allowing both for discontinuous payoff functions and for action sets that need not be common to all players and are just subsets of a Souslin space, rather than of a Polish space).

Our main existence result has been applied to two classes of economic problems. First, we have considered a game with a non-atomic space of players, a finite action space, and dispersed discontinuities. Non-atomic games with finite action spaces are often found in applications with the additional assumption of continuous payoff functions, e.g. Karni & Levin (1994) and Roughgarden & Tardos (2004). Concerning this class of games, our results can be used to cover a broader class of economic applications, a point which we have illustrated with our simple congestion example in Section 2. Second, we have presented a general framework for optimal taxation problems, which includes the original model of Mirrlees (1971) (Example 3) and the more recent model of Farhi et al. (2011) (Example 6). Several other examples illustrate the flexibility of our approach to address economically relevant optimal tax problems. Of these examples, Example 7 is noteworthy because it illustrates that other known approaches to establish existence of equilibrium are inapplicable.

It is conceivable that our existence results have interesting applications in related models. Especially relevant are the settings of Rath (1996), Kim & Yannelis (1997), Milgrom & Weber (1985), and Balder (1988), as these are covered by the pure strategy existence result of Balder (2002) which we have extended in this paper.

11 Appendix: Proofs

11.1 A general roadmap

In this appendix we present the proofs of our results. They are grouped according to the topic of the results into subsections. Section 11.3 is devoted to our main existence results (Theorems 1 and 2, unified by Theorem 3), Section 11.4 to the results concerning the relationship between our existence results and the pure strategy existence result of Balder (2002) (Theorems 4 and 5 and Example 2), Sections 11.5 and 11.6 to the results for the concretization of our framework as described in Section 7 (Theorems 6 and 7 respectively), and Section 11.7 contains the proofs concerning our optimal taxation application. Section 11.2 introduces some additional notation.

Let us note here also that completeness of a measure space is used below without specific reference.

11.2 Additional notation

Let Z be a topological space. For a sequence $\langle z_n \rangle$ in Z, $\mathrm{LS}\, z_n$ denotes the set of its cluster points. For a sequence $\langle A_n \rangle$ of subsets of Z, $\mathrm{KLS}\, A_n$ denotes the Kuratowski limes superior, i.e., the set of those points z in Z such that for every neighborhood U of z there are infinitely many n with $A_n \cap U \neq \emptyset$.

11.3 Proof of the main results

11.3.1 Basic idea and road map

The basic idea to prove our main existence results, Theorems 1 and 2, which are unified in Theorem 3, is to combine arguments as used in the existence proof of McLennan et al. (2011) with Young measure techniques as used by Balder (2002), the link being made by a purification argument. As in McLennan et al. (2011), we argue by contradiction. We assume there is no Nash equilibrium and, via a fixed-point argument, obtain a strategy profile with properties contradicting CS. The difficulty is to fit this argument to our context with a measure space of players. Several measurability issues have to be taken care of. In particular, we cannot follow McLennan et al. (2011) in applying the fix point argument to a map operating on the product $\prod_{t \in T} X_t$ of the action sets of the players. To get in the position for a fixed point argument, we follow Balder (2002) and place the analysis into a context of Young measures defined on T. Roughly, Young measures on T are profiles of mixed strategies. To link them to our setting of pure strategies we need in particular a purification argument. For players in $\bar{T}$, this does not make a problem as actions set of these players are convex; as in Balder (2002) we can purify by simply switching to the barycenter of a mixed strategy. For players in $\hat{T}$, the situation is more involved. In fact, in Theorem 2 the component of the externality space referring to $\hat{T}$ may be infinite-dimensional, so, for the matter of this theorem, we cannot follow Balder (2002) by appealing to Liyapounoff's theorem for Young measures (cf. the discussion in Section 4.2). A purification result that applies to infinite-dimensional externality spaces may be found in Podczeck (2009). However, this latter result needs a common action space for all players, so does not directly apply to our setting, and therefore we need some extension.

Now the roadmap of the proof is as follows. In 11.3.2 we collect some basic facts about Young measures, taken from Balder (1989) and Balder (2000). In 11.3.3 we state and prove the purification result we need. In 11.3.4 we start by establishing the link between Young measures on the one side and strategy profiles and externalities on the other. The link is specified by a map h whose properties are summarized in Lemma 2. It is in the proof of that lemma where the purification result of 11.3.3 is applied. After that we present that part of our proof which is the analog in our context of the existence proof in McLennan et al. (2011).

11.3.2 Young measures

We start by stating some definitions. Let us fix a complete measure space (T, Σ, ν) with $0 < \nu(T) < \infty$ and a completely regular Souslin space X. Write $M^1_+(X)$ for the set of all Borel probability measures on X.

Recall that a Young measure from T to X is just a function $g\colon T \to M^1_+(X)$ which is measurable for the narrow topology of $M^1_+(X)$. Recall that this property is equivalent to the property that the map $t \mapsto g(t)(B)$ is measurable for each $B \in \mathcal{B}(X)$, and also to the property that, for each bounded continuous $p\colon X \to \mathbb{R}$, the map $t \mapsto \int p \mathrm{d}g(t)$ is measurable (cf. Balder (2002, p. 441)). The first equivalence shows in particular that if $f\colon T \to X$ is a measurable map, then the map $t \mapsto \delta_{f(t)}$ is a Young measure, where $\delta_{f(t)}$ denotes Dirac measure at $f(t)$.

Let $\mathcal{R}$ denote the set of all Young measures from T to X, endowed with the narrow topology for Young measures. Recall that this topology is defined to be the coarsest topology on $\mathcal{R}$ such that for each $q \in \mathcal{G}$ the functional

$$g \mapsto \int_T \int_X q(t,x)\mathrm{d}g(t)\mathrm{d}\nu(t)\colon \mathcal{R} \to \mathbb{R}$$

is continuous, where $\mathcal{G}$ is the set of all measurable functions $q\colon T \times X \to \mathbb{R}$ such that $q(t,\cdot)$ is continuous for each $t \in T$ and such that, for some integrable $\theta_q\colon T \to \mathbb{R}_+$, $\sup\{|q(x)|\colon x \in X\} \leq \theta_q(t)$ for each $t \in T$. It should be noted that, in general, the narrow topology for Young measures is not a Hausdorff topology.

In the sequel, if $\kappa\colon T \twoheadrightarrow X$ is a correspondence, then $\mathcal{R}_\kappa$ denotes the subset of $\mathcal{R}$ defined by setting

$$\mathcal{R}_\kappa = \{g \in \mathcal{R}\colon \operatorname{supp} g(t) \subseteq \kappa(t) \text{ for almost all } t \in T\}.$$

The fact presented in the following theorem is fundamental for the fixed point argument in the proof of our equilibrium existence result. The compactness part in this theorem is a deep result due to Balder (1989).

Theorem 10. *Let $\kappa\colon T \twoheadrightarrow X$ be a correspondence with a measurable graph such that $\kappa(t)$ is non-empty and compact for all $t \in T$. Then the subset $\mathcal{R}_\kappa$ of $\mathcal{R}$ is non-empty, closed, compact, and sequentially compact.*

Proof. As noted above, if $f\colon T \to X$ is measurable, then the map $t \mapsto$ "Dirac measure at $f(t)$" belongs to $\mathcal{R}$. Thus non-emptiness of $\mathcal{R}_\kappa$ is implied by the von Neumann-Aumann-Sainte Beuve measurable selection theorem (see Castaing & Valadier (1977, Theorem III.22)). From Balder (1989, Theorem 2.2 and Remark 2.4) it follows that $\mathcal{R}_\kappa$ is both relatively compact and relatively sequentially compact. Hence, given that $\mathcal{R}_\kappa$ is closed, it is both compact and sequentially compact. That $\mathcal{R}_\kappa$ is indeed closed may be seen as follows. Recall that a Souslin space is separable and that on any such space, if it is completely regular, there is metric that gives a topology weaker than the original one but such that the Borel sets for both topologies are the same.[15] Choose

[15]Recall that a Souslin space X is Hausdorff by definition, and therefore, if it is completely regular, the set of (bounded) continuous functions on X separates the points of X. But on a Souslin space X,

such a metric, say ρ, on the space X under consideration. In particular, X is separable for ρ. Define $q_\rho\colon T \times X \to \mathbb{R}$ by setting

$$q_\rho(t,x) = \min\{1, \rho(x, \kappa(t))\},\ (t,x) \in T \times X.$$

Using Castaing & Valadier (1977, Theorems III.22, III.9, and Lemma III.14) it may be seen that q belongs to the set $\mathcal{G}$ defined above. Now as κ takes closed values, an element $g \in \mathcal{R}$ belongs to $\mathcal{R}_\kappa$ if and only if $\int_T \int_X q_\rho(t,x)\mathrm{d}g(t)\mathrm{d}\nu(t) = 0$. Thus, by the definition of the narrow topology for Young measures, $\mathcal{R}_\kappa$ is closed in $\mathcal{R}$. $\square$

The next statement is just a translation of Theorem 4.15 in Balder (2000) into our notation.

Theorem 11. *If $\kappa\colon T \twoheadrightarrow X$ is as in Theorem 10 and $g_n \to g$ in $\mathcal{R}_\kappa$, then* $\operatorname{supp} g(t) \subseteq \operatorname{KLS} \operatorname{supp} g_n(t)$ *for almost all $t \in T$.*

11.3.3 The purification result

The purification result we need will be proved as a consequence of the purification result in Podczeck (2009). As in the previous section, (T, Σ, ν) is a complete measure space with $0 < \nu(T) < \infty$, and X a completely regular Souslin space.

Theorem 12. *Let $\kappa\colon T \twoheadrightarrow X$ be as in Theorem* 10. *Writing Γ_κ for the graph of κ, let $\mathcal{C}$ be a countable set of functions $q\colon \Gamma_\kappa \to \mathbb{R}$ such that* (i) *q is measurable for the subspace σ-algebra of Γ_κ defined from $\Sigma \otimes \mathcal{B}(X)$,* (ii) *$q(t, \cdot)$ is continuous on $\kappa(t)$ for each $t \in T$,* (iii) *there is an integrable $\theta_q\colon T \to \mathbb{R}_+$ such that* $\sup\{|q(t,x)|\colon x \in \kappa(t)\} \le \theta_q(t)$ *for almost all $t \in T$. Suppose (T, Σ, ν) is super-atomless. Then given any $g \in \mathcal{R}_\kappa$, there is a measurable $f\colon T \to X$ such that*

(1) *$f(t) \in \operatorname{supp} g(t)$ for almost all $t \in T$;*

(2) *$\int_T \int_X q(t,x)\mathrm{d}g(t)(x)\mathrm{d}\nu(t) = \int_T q(t, f(t))\mathrm{d}\nu(t)$ for all $q \in \mathcal{C}$;*

(3) *$\int_T g(t)(B)\mathrm{d}\nu(t) = \nu(f^{-1}(B))$ for all $B \in \mathcal{B}(X)$.*

Proof. By Podczeck (2009), the theorem is true in the special case where X is a compact metric space and the maps q are defined on all of $T \times X$, and where (i) assumed to hold with $T \times X$ in place of Γ_κ, and (ii) and (iii) with X in place of $\kappa(t)$. We will show that the situation of the present theorem can be reduced to this case.

any set of continuous functions that separates the points of X contains a countable subset with the same property (Castaing & Valadier (1977, III.31)). This yields the assertion concerning the metric. Now the assertion concerning the Borel sets follows from Schwartz (1973, p. 101 Corollary 2).

As in the proof of Theorem 10, recall that a Souslin space is separable and that on any such space there is metric that gives a topology weaker than the original one but such that the Borel sets for both topologies are the same.

This fact implies that we may view X as a subset of a compact metric space K such that the inclusion map $i\colon X \to K$ is continuous and such that $\mathcal{B}(X)$ coincides with the subspace σ-algebra defined from $\mathcal{B}(K)$ (use Engelking (1989, Theorem 3.5.2), which is a compactification result, and for the assertion on the Borel σ-algebras, use Castaing & Valadier (1977, Lemma III.20) in addition). In particular, for each $t \in T$, since $\kappa(t)$ is a compact subset of X, $\kappa(t)$ is also compact as a subset of K, and the topologies of X and K give the same subspace topology on $\kappa(t)$. Furthermore, a map $f\colon T \to X$ is measurable for $\mathcal{B}(X)$ if and only if it is measurable for $\mathcal{B}(K)$ when viewed as a map into K.

Fix any $g \in \mathcal{R}_\kappa$. As the inclusion $i\colon X \to K$ is continuous, we may define a map $g_1\colon T \to M^1_+(K)$ by setting $g_1(t) = g(t) \circ i^{-1}$ for each $t \in T$ (i.e., $g_1(t)$ is the image measure of $g(t)$ under i). In particular, for any continuous map $p\colon K \to \mathbb{R}$, we have $\int_K p \mathrm{d} g_1(t) = \int_X p \circ i \mathrm{d} g(t)$ for every $t \in T$, and therefore the map $t \mapsto \int_K p \mathrm{d} g_1(t)$ is measurable, by the fact that g is a Young measure. Moreover, because $\operatorname{supp} g(t) \subseteq \kappa(t)$ for almost $t \in T$ by definition of $\mathcal{R}_\kappa$, we must have $\operatorname{supp} g_1(t) = \operatorname{supp} g(t)$ for almost all $t \in T$, by what was noted in the previous paragraph about the sets $\kappa(t)$.

We assert the following.

Claim: For each $q \in \mathcal{C}$ there is a $q'\colon T \times K \to \mathbb{R}$ such that (a) $q'|_{\Gamma_\kappa} = q$, (b) q' is measurable, (c) $q'(t, \cdot)$ is continuous for each $t \in T$, (d) $\sup\{|q'(t,x)|\colon x \in K\} \leq \theta_q(t)$ for almost all $t \in T$.

Assuming the claim has been established, the theorem can be proved as follows. By Podczeck (2009, Corollary and Lemma 2), the fact that $t \mapsto \int_K p \mathrm{d} g_1(t)$ is measurable for each continuous $p\colon K \to \mathbb{R}$ implies that there is a measurable $f\colon T \to K$ such that (1)-(3) of the theorem hold with g_1 substituted for g, $\mathcal{B}(K)$ for $\mathcal{B}(X)$, and with each $q \in \mathcal{C}$ replaced by an element q' associated with q according to the claim. Now since $\operatorname{supp} g_1(t) = \operatorname{supp} g(t)$ for almost all $t \in T$, (1) of the theorem must also hold with f and g, i.e., $f(t) \in \operatorname{supp} g(t)$ for almost all $t \in T$, and therefore, in view of (a) of the claim, (2) must actually hold with f, g, and the given $\mathcal{C}$, because $\operatorname{supp} g(t) \subseteq \kappa(t)$ for almost all $t \in T$, and because $g_1(t) = g(t) \circ i^{-1}$ for each $t \in T$. Note that since $f(t) \in \kappa(t)$ for almost all $t \in T$, we have $f(t) \in X$ for almost all $t \in T$. Changing f on a null set of T, if necessary, we may assume that f takes all of its values in X. Now, because for every $B \in \mathcal{B}(X)$ there is a $B' \in \mathcal{B}(K)$ with $B = B' \cap X = i^{-1}(B')$, f must be measurable for $\mathcal{B}(X)$, and the fact that (3) of the theorem holds with f, g_1, and $\mathcal{B}(K)$ implies that (3) of the theorem holds with f, g, and $\mathcal{B}(X)$ as well.

Thus it remains to establish the above claim. Take any $q \in \mathcal{C}$ and note first that continuity of $q(t,\cdot)$ on $\kappa(t)$ as subspace of X, which holds for each $t \in T$ by hypothesis, implies continuity on $\kappa(t)$ as a subspace of K, by what was noted above. Define $q_1 : \Gamma_\kappa \to \mathbb{R}$ by setting

$$q_1(t,x) = 3/2 + (1/2)\arctan q(t,x), \ (t,x) \in \Gamma_\kappa.$$

Then q_1 is measurable for the subspace σ-algebra on Γ_κ defined from $\Sigma \otimes \mathcal{B}(X)$, because q is. Also, $q_1(t,\cdot)$ is continuous on $\kappa(t)$ for each $t \in T$, and q_1 takes all of its values in $(1,2)$.

Let ρ denote the metric of K. As $\kappa(t)$ is compact and therefore closed in K for each $t \in T$, and as q_1 takes values in $(1,2)$, we can define a function $q_2 : T \times K \to \mathbb{R}$, also taking values in $(1,2)$, by setting $q_2(t,x) = q_1(t,x)$ if $x \in \kappa(t)$ and

$$q_2(t,x) = \inf\left\{ \frac{q_1(t,u)\rho(x,u)}{\rho(x,\kappa(t))} : u \in \kappa(t) \right\}$$

otherwise. Note that for each $t \in T$, $q_2(t,\cdot)$ is a continuous extension of $q_1(t,\cdot)$ to K (see, e.g., Mandelkern (1990)). We claim that q_2 is measurable. As $q_2(t,\cdot)$ is continuous for each $t \in T$, to establish this claim it suffices by Castaing & Valadier (1977, Lemma III.14) to show that $q(\cdot,x)$ is measurable for each $x \in K$.

To this end, we appeal to Castaing & Valadier (1977, Theorem III.22) to choose a countable set $\{h_i : i \in I\}$ of measurable functions $h_i : T \to X$ such that $\{h_i(t) : i \in I\}$ is a dense subset of $\kappa(t)$ for each $t \in T$. By the measurability property of q_1 mentioned before, measurability of the h_i's implies in particular that the maps $t \mapsto q_1(t, h_i(t))$ are measurable. By what was said in the third paragraph of this proof, each h_i is measurable also when viewed as map into K, and for each $t \in T$, $\{h_i(t) : i \in I\}$ is dense in $\kappa(t)$ also for the topology of K. Taking some $x \in K$ now as given, it follows that for each $i \in I$ the map $t \mapsto \rho(x, h_i(t))$ is measurable, and therefore the map $t \mapsto \rho(x,\kappa(t))$ must be measurable as well. Thus if we set $T_1 = \{t \in T : \rho(x,\kappa(t)) > 0\}$ and define $e_i : T \to \mathbb{R}$, $i \in I$, by setting

$$e_i(t) = \begin{cases} \dfrac{q_1(t,h_i(t))\rho(x,h_i(t))}{\rho(x,\kappa(t))} & \text{if } t \in T_1, \\ q_1(t,x) & \text{if } t \in T \setminus T_1, \end{cases}$$

then e_i is measurable. Now because $\{h_i(t) : i \in I\}$ is dense in $\kappa(t)$, and because the function $u \mapsto q_1(t,u)\rho(x,u)/\rho(x,\kappa(t))$ is continuous on $\kappa(t)$ for each $t \in T_1$, we must have $q_2(t,x) = \inf\{e_i(t) : i \in I\}$ for each $t \in T$. As I is countable, it follows that $q_2(\cdot,x)$ is measurable.

Recalling that q_2 takes values in $(1,2)$, define $q_3 : T \times K \to \mathbb{R}$ by setting

$$q_3(t,x) = \tan(2q_2(t,x) - 3), \ (t,x) \in T \times K.$$

Then, because q_2 is measurable, so is q_3, and as $q_2(t,\cdot)$ is continuous for each $t \in T$, so is $q_3(t,\cdot)$ for each $t \in T$. Thus (b) and (c) of the claim above hold for q_3. By construction, for each $t \in T$ we have $q_3(t,x) = q(t,x)$ if $x \in \kappa(t)$, i.e., (c) of the claim holds for q_3. As for (d), consider any $t \in T$, and note that the choice of $q_2(t,\cdot)$ implies that for any $x \in K \backslash \kappa(t)$,

$$\inf\{q_1(t,y)\colon y \in \kappa(t)\} \leq q_2(t,x) \leq \sup\{q_1(t,y)\colon y \in \kappa(t)\},$$

and therefore, by choice of $q_3(t,\cdot)$, since $q_3(t,x) = q(t,x)$ for all $y \in \kappa(t)$,

$$\inf\{q(t,y)\colon y \in \kappa(t)\} \leq q_3(t,x) \leq \sup\{q(t,y)\colon y \in \kappa(t)\},$$

from which it follows that $\sup\{|q_3(t,x)|\colon x \in K\} = \sup\{|q(t,x)|\colon x \in \kappa(t)\} \leq \theta_q(t)$. Thus (d) of the claim holds for q_3. This completes the proof. $\square$

11.3.4 Proof of Theorem 3

Let us fix a game $G = ((T,\Sigma,\nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$ such that (A1)-(A6) and (S3) hold. Thus X is now a locally convex space in addition to being Souslin.

As before, $\mathcal{R}$ denotes the set of all Young measures from T to X, endowed with the narrow topology for Young measures. By $\mathcal{R}_G$ we denote the subset of $\mathcal{R}$ defined as

$$\mathcal{R}_G = \{g \in \mathcal{R}\colon \operatorname{supp} g(t) \subseteq X_t \text{ for almost all } t \in T\}.$$

Now by Balder (2002, Corollary 4.2.1), (A3)(i) and (A4) together imply that there is a map $h_1\colon \mathcal{R}_G \to \bar{S}_G$ such that for each $g \in \mathcal{R}_G$, $h_1(g)(t)$ is the barycenter of $g(t)$ for almost all $t \in \bar{T}$, i.e., the unique element $x \in X_t$ for which $p(x) = \int_{X_t} p\mathrm{d}g(t)$ for every continuous linear functional p on X. Define $h_2\colon \mathcal{R}_G \to \mathbb{R}^{\hat{\mathcal{C}}}$ by setting $h_2(g) = \left\langle \int_{\hat{T}} \int_{X_t} q(t,x)\mathrm{d}g(t)(x)\mathrm{d}\nu(t)\right\rangle_{q\in\hat{\mathcal{C}}}$ for each $g \in \mathcal{R}_G$, where $\hat{\mathcal{C}}$ is the countable set involved in the externality map e. Finally define $h\colon \mathcal{R}_G \to \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ by setting $h(g) = (h_1(g), h_2(g))$ for $g \in \mathcal{R}_G$.

Recall from Section 4.1 that $\mathbb{R}^{\hat{\mathcal{C}}}$ is endowed with the product topology, $\bar{S}_G$ with the feeble topology, $\bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ with the corresponding product topology, and the set $E_G \equiv e(S_G) \subseteq \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ with the subspace topology.

Lemma 2. *The following hold for the map h.*

(a) *h is a surjection onto a dense subset of E_G.*

(b) *h is continuous as map from $\mathcal{R}_G$ to E_G*

(c) *Given any $g \in \mathcal{R}_G$, there is an $f \in S_G$, with $e(f) = h(g)$, such that for almost all $t \in \hat{T}$, $f(t) \in \operatorname{supp} g(t)$, and for almost all $t \in \bar{T}$, $f(t) \in \overline{\mathrm{co}}\operatorname{supp} g(t)$.*

Proof. We first show that (c) holds. Consider any $g \in \mathcal{R}_G$. If the set $\hat{\mathcal{C}}$ is finite, then by Lyapunov's theorem for Young measures (see Balder (2002, Theorem, 4.2.3)), Assumption (A5) implies that there is a measurable function $\hat{f}: \hat{T} \to X$ such that $h_2(g) = \big\langle \int_{\hat{T}} q(t, \hat{f}(t)) \mathrm{d}\nu(t) \big\rangle_{q \in \hat{\mathcal{C}}}$ and such that $\hat{f}(t) \in \operatorname{supp} g(t)$ for almost all $t \in \hat{T}$. If $\hat{\mathcal{C}}$ is countably infinite, then by (S3), the subspace measure on $\hat{T}$ defined from ν is super-atomless, and we get an $\hat{f}: \hat{T} \to X$ with the same properties by Theorem 12 (with T there replaced by $\hat{T}$, and $\mathcal{C}$ by $\hat{\mathcal{C}}$), noting that by (A3), X_t is compact for each $t \in \hat{T}$ and the graph of the correspondence $t \mapsto X_t: \hat{T} \to X$ is measurable.

Define $f: T \to X$ by setting $f(t) = h_1(g)(t)$ for $t \in \bar{T}$, and $f(t) = \hat{f}(t)$ for $t \in \hat{T}$. Thus $f|_{\bar{T}} = h_1(g) \in \bar{S}_G$. In particular, $f|_{\bar{T}}$ is measurable and we have $(f|_{\bar{T}})(t) \in X_t$ for almost all $t \in \bar{T}$. As $g \in \mathcal{R}_G$ means $\operatorname{supp} g(t) \subseteq X_t$ for almost all $t \in T$, the fact that $\hat{f}(t) \in \operatorname{supp} g(t)$ for almost all $t \in \hat{T}$ implies that we have $f(t) \in X_t$ also for almost all $t \in \hat{T}$. Thus $f(t) \in X_t$ for almost all $t \in T$, and because $\bar{T}$ and $\hat{T}$ are measurable subsets of T, it follows that f is measurable. That is, $f \in S_G$. In particular, $e(f)$ is defined. Recalling that the definition of e says $e(f) = (f|_{\bar{T}}, \big\langle \int_{\hat{T}} q(t, \hat{f}(t)) \mathrm{d}\nu(t) \big\rangle_{q \in \hat{\mathcal{C}}})$, we may conclude that $e(f) = (h_1(g), h_2(g)) = h(g)$. Now by construction, we have $f(t) \in \operatorname{supp} g(t)$ for almost all $t \in \hat{T}$. Also, because $f(t) = h_1(g)(t)$ for $t \in \bar{T}$, and because by the definition of h_1, $h_1(g)(t)$ is the barycenter of $g(t)$ for almost all $t \in \bar{T}$, we must have $f(t) \in \overline{\mathrm{co}} \operatorname{supp} g(t)$ for almost all $t \in \bar{T}$, by the definition of barycenter and by the separation theorem. Thus (c) holds.

Now as $E_G = e(S_G)$ by definition, (c) implies that h takes values in E_G. Evidently the definition of h implies that if f is any element of S_G, and $g \in \mathcal{R}_G$ is defined by $g(t) = \delta_{f(t)}$, where $\delta_{f(t)}$ is Dirac measure at $f(t)$, then $h(g) = e(f)$, except perhaps on a null set in $\bar{T}$. This shows that (a) holds. As for (b), by Balder (2002, Theorem 4.2.2) it follows that the map h_1 continuous, and by the argument in Step 1 in the proof of Theorem 2.2.1 in Balder (2002), it follows that for each $q \in \hat{\mathcal{C}}$ the map $g \mapsto \int_{\hat{T}} \int_{X_t} q(t, x) \mathrm{d}g(t) \mathrm{d}\nu(t): \mathcal{R}_G \to \mathbb{R}$ is continuous. Thus, by choice of the topologies of $\bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ and E_G, (b) follows. □

The following fact is also used outside the proof of Theorem 3 and for this reason singled out as a lemma.

Lemma 3. *The set E_G is compact and pseudo-metrizable.*

Proof. Assumptions (A2), (A3), and (A6) imply that $\bar{S}_G$ is pseudo-metrizable (see Balder (2002, Remark 4.3.1)), and because $\hat{\mathcal{C}}$ is countable, $\mathbb{R}^{\hat{\mathcal{C}}}$ is metrizable. Thus $E_G \subseteq \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ is pseudo-metrizable. Consequently, Theorem 10 and Lemma 2(a), (b) show that E_G is compact. □

We are now ready for the proof of Theorem 3. Suppose by way of contradiction that the game G has no equilibrium. Then by CS and compactness of E_G, there is a finite family $\langle U_j, \varphi_j, \alpha_j \rangle_{j\in J}$ where $\langle U_j \rangle_{j\in J}$ is a covering of E_G by open subsets, and φ_j, α_j correspond to U_j according to CS for each $j \in J$.

Let ρ denote a pseudo-metric for E_G and recall that Lebesgue's covering theorem holds in compact pseudo-metric spaces. We can therefore find an $\varepsilon > 0$ so that each closed ρ-ball in E_G of radius 2ε is included in some member of the open covering $\langle U_j \rangle_{j\in J}$ of E_G. Let $\langle B_i \rangle_{i\in I}$ be a finite covering of E_G by closed ρ-balls of radius ε. Then the choice of ε implies that whenever $\langle B_i \rangle_{i\in I'}$ is subfamily of $\langle B_i \rangle_{i\in I}$ with $\bigcap_{i\in I'} B_i \neq \emptyset$, there is a $j \in J$ such that $\bigcup_{i\in I'} B_i \subseteq U_j$.

Fix any $i \in I$. Define a Caratheodory correspondence $\varphi^i \colon T \times B_i \twoheadrightarrow X$ as follows. Let $H = \{j \in J \colon B_i \subseteq U_j\}$. By the previous paragraph, H is non-empty. Using the fact that the functions α_j are measurable, we can find a finite partition $\langle T_k \rangle_{k\in K}$ of T into measurable sets and a corresponding family $\langle j_k \rangle_{k\in K}$ of elements of H such that if $t \in T_k$ then $\alpha_{j_k}(t) \geq \alpha_j(t)$ for each $j \in H$. Now define $\varphi^i \colon T \times B_i \twoheadrightarrow X$ by setting $\varphi^i(t, y) = \varphi_{j_k}(t, y)$ for $(t, y) \in T_k \times B_i$, $k \in K$. It is clear that $\varphi^i(t, \cdot)$ is well-behaved for each $t \in T$, with compact values. Furthermore, for every $y \in B_i$, we have

$$\operatorname{graph}(\varphi^i(\cdot, y)) = \bigcup_{k\in K} \left(\operatorname{graph}(\varphi_{j_k}(\cdot, y)) \cap (T_k \times X)\right),$$

showing that the graph of $\varphi^i(\cdot, y)$ is measurable, because K is finite. Thus φ^i is a Caratheodory correspondence. By (a) in the definition of CS, $\varphi^i(t, y) \subseteq A_t(y)$ for each $(t, y) \in T \times B_i$. Also, given any $y \in B_i$, (c) in the definition of CS implies that, for almost all $t \in T$, if $x \in \varphi^i(t, y)$ then $u_t(x, y) \geq \alpha_j(t)$ for each j with $B_i \subseteq U_j$.

Do this construction for each $i \in I$. For each $y \in E_G$, set $I^y = \{i \in I \colon y \in B_i\}$. Let $\varphi \colon T \times E_G \twoheadrightarrow X$ be the correspondence defined by setting

$$\varphi(t, y) = \bigcup_{i\in I^y} \varphi^i(t, y), \ (t, y) \in T \times E_G.$$

Then, because each φ^i is a Caratheodory correspondence with compact values, so is φ. (Indeed, first it is clear that φ takes non-empty values, and as I is finite, φ takes compact values. To see that $\varphi(t, \cdot)$ is uhc for each $t \in T$, fix $t \in T$ and $y \in E_G$, and consider an open $O \subseteq X$ such that $\varphi(t, y) \equiv \bigcup_{i\in I^y} \varphi^i(t, y) \subseteq O$. As each $\varphi^i(t, \cdot)$ is uhc and I^y is finite, there is a neighborhood V of y in E_G such that $\bigcup_{i\in I^y} \varphi^i(t, y') \subseteq O$ for each $y' \in V$. As $I \backslash I^y$ is finite and all the B_i's are closed, setting $V' = V \backslash (\bigcup_{i\in I\backslash I^y} B_i)$, V' is still a neighborhood of y, but such that $I^{y'} \subseteq I^y$ for all $y' \in V'$, implying that $\varphi(t, y') \subseteq O$ for all $y' \in V'$. Finally, for any $y \in E_G$, it is clear that since each $\varphi^i(\cdot, y)$ has a measurable graph, so does $\varphi(\cdot, y)$, again since I is finite, and since the graph of $\varphi(\cdot, y)$ is the union over I^y of the graphs of the correspondences $\varphi^i(\cdot, y)$.)

We claim that there are a $y^* \in E_G$ and a measurable $f^* \colon T \to X$ such that $y^* = e(f^*)$ and for almost all $t \in T$, $f^*(t) \in \varphi(t, y^*)$ if $t \in \hat{T}$, and $f^*(t) \in \overline{\operatorname{co}}\, \varphi(t, y^*)$

if $t \in \bar{T}$. Assuming for the time being that this has been established, let us see how to finish the proof.

Note first that by definition of φ, we have $f^*(t) \in \bigcup_{i \in I^{y^*}} \varphi^i(t, y^*)$ for almost all $t \in \hat{T}$, and $f^*(t) \in \overline{\mathrm{co}} \bigcup_{i \in I^{y^*}} \varphi^i(t, y^*)$ for almost all $t \in \bar{T}$. By (b) in CS, if $t \in \bar{T}$ then $\varphi^i(t, y^*)$ is convex or included in a finite-dimensional subspace of X, in addition to being compact, so we actually have $f^*(t) \in \mathrm{co} \bigcup_{i \in I^{y^*}} \varphi^i(t, y^*)$ for almost all $t \in \bar{T}$ (by the general fact that the convex hull of the union of finitely many such subsets of a Hausdorff topological vector space is closed; e.g., combine Lemma 5.29 and Corollary 5.33 in Aliprantis & Border (2006)). Now by definition of the sets I^y, applied to I^{y^*}, we have $y^* \in \bigcap_{i \in I^{y^*}} B_i$. According to what was noted in the last sentence of the second paragraph of this proof, this means there is a $j^* \in J$ such that $B_i \subseteq U_{j^*}$ for all $i \in I^{y^*}$. But from the third paragraph, if $y^* \in B_i \subseteq U_{j^*}$ then, for almost all $t \in T$, $x \in \varphi^i(t, y^*)$ implies $u_t(x, y^*) \geq \alpha_{j^*}(t)$. As $x \in \varphi^i(t, y^*)$ also implies $x \in A_t(y^*)$, we may conclude that $f^*(t) \in \{x \in A_t(y^*) \colon u_t(x, y^*) \geq \alpha_{j^*}(t)\}$ for almost all $t \in \hat{T}$, and that $f^*(t) \in \mathrm{co}\{x \in A_t(y^*) \colon u_t(x, y^*) \geq \alpha_{j^*}(t)\}$ for almost all $t \in \bar{T}$. However, by (d) in the definition of CS, this is impossible because $e(f^*) = y^* \in U_{j^*}$, and this contradiction establishes the theorem.

Thus it remains to be shown that the above claim is correct. To this end, consider the correspondence $\varphi_1 \colon E_G \twoheadrightarrow \mathcal{R}_G$ defined by setting

$$\varphi_1(y) = \{g \in \mathcal{R}_G \colon \operatorname{supp} g(t) \subseteq \varphi(t, y) \text{ for almost all } t \in T\},\ y \in E_G.$$

Then by Theorem 10, the fact that φ is a Caratheodory correspondence with compact values implies that φ_1 takes non-empty closed values. The fact that φ is a Caratheodory correspondence means, in particular, that $\varphi(t, \cdot)$ is well-behaved for each $t \in T$, which implies that if $y \in E_G$ is a limit of a sequence $\langle y_n \rangle$ in E_G, then $\mathrm{KLS}\, \varphi(t, y_n) \subseteq \varphi(t, y)$ for each $t \in T$ (because the codomain X_t of $\varphi(t, \cdot)$ is compact Hausdorff, therefore regular). Consequently, by Theorem 11, φ_1 has a sequentially closed graph in $E_G \times \mathcal{R}_G$. As E_G is pseudo-metrizable by Lemma 3, and $\mathcal{R}_G$ is sequentially compact by Theorem 10, it follows that φ_1 is uhc as may readily be seen.

Let $\varphi_2 \colon \mathcal{R}_G \twoheadrightarrow \mathcal{R}_G$ be the composition $\varphi_2 = \varphi_1 \circ h$, where h is the map from Lemma 2. Then φ_2 is uhc, because φ_1 is and because by Lemma 2, h is continuous. Also, φ_2 takes non-empty closed values. As $\mathcal{R}_G$ is compact by Theorem 10, it follows that φ_2 has a fixed point, g^* say (see the discussion in steps 2 and 6 in the proof of Theorem 4.1.2 in Balder (2002)). Choose an element $f^* \in S_G$ which corresponds to g^* according to Lemma 2(c), and set $y^* = e(f^*)$. Then $g^* \in \varphi_1(y^*)$, and by the definition of φ_1 it follows that f^* and y^* are as required in the claim above. This completes the proof.

11.4 Proofs related to Sections 2 and 6

11.4.1 Proof of Theorem 4

Fix a game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$ satisfying assumptions (A1)-(A10) and (S1) or (S2).

The idea of the proof is to adapt the following argument. Consider a finite-player game $(X_i, u_i)_{i\in I}$ as described in Section 4.3, and assume that value functions are lsc, and payoff functions usc and quasi-concave in the own action, so that, in particular, best reply correspondences are well-behaved and take convex values. Now if $y \in \prod_{j\in I} X_j$ is not an equilibrium strategy profile in this game, then for some $\varepsilon > 0$, some player i (a) can ensure a payoff at least $u_i(y) + \varepsilon$ along his best reply correspondence on some neighborhood U_1 of y_{-i} in $\prod_{j\neq i} X_j$, by lower semi-continuity of his value function, but (b) gets a playoff smaller than $u_i(y) + \varepsilon$ on some neighborhood U_2 of y in $\prod_{j\in I} X_j$, by upper semi-continuity of his payoff function. It follows at once from this that the game is continuously secure, with the φ_i's taken to be certain restrictions of the best-reply correspondences.

The main difficulty with carrying out this kind of argument in the context of the proof to be given here concerns the analog of (b). The problem is that CS is defined in terms of neighborhoods of points in the externality space E_G and these neighborhoods do not put any restrictions on the actions of any player belonging to the non-atomic part of T. The crucial step to deal with the problem is separated in the following lemma, because the fact established is also needed in other contexts later on.

Lemma 4. *Let $\langle f_n\rangle$ be a sequence in S_G with $e(f_n) \to y$ for some $y \in E_G$. Then there is an $f \in S_G$ such that $e(f) = y$ and such that for almost all $t \in \hat{T}$, $f(t) \in \mathrm{LS}\, f_n(t)$, and for almost all $t \in \bar{T}$, $f(t) \in \overline{\mathrm{co}}\,\mathrm{LS}\, f_n(t)$.*

Proof. Let $\mathcal{R}_G$ and h be as in Section 11.3.4, and for each n define $g_n \in \mathcal{R}_G$ by setting $g_n(t) = \delta_{f_n(t)}$ for $t \in T$, where $\delta_{f_n(t)}$ denotes Dirac measure at $f_n(t)$. By Theorem 10 there are a $g \in \mathcal{R}_G$ and a subsequence $\langle g_k\rangle$ of the sequence $\langle g_n\rangle$ such that $g_k \to g$ in $\mathcal{R}$. By Theorem 11, we have $\mathrm{supp}\, g(t) \subseteq \mathrm{KLS}\,\mathrm{supp}\, g_k(t) \subseteq \mathrm{KLS}\,\mathrm{supp}\, g_n(t)$ for almost all $t \in T$, and by Lemma 2(b), $h(g_k) \to h(g)$. Choose $f \in S_G$ such that f corresponds to g according to Lemma 2(c). Observe that $h(g_n) = e(f_n)$ for each n, except perhaps on a null set in $\bar{T}$, and that $\mathrm{KLS}\,\mathrm{supp}\, g_n(t) = \mathrm{LS}\, f_n(t)$ for each $t \in T$ by definition of the g_n's. Thus, changing f on a null set in $\bar{T}$, if necessary, we have an f as required. □

Let $\varphi\colon T \times E_G \twoheadrightarrow X$ denote the best reply correspondence of the game G. Thus $\varphi(t, y) = \{x \in A_t(y)\colon u_t(x, y) = w_t(y)\}$ for each $(t, y) \in T \times E_G$. We will show that CS holds with certain restrictions of φ.

Recall first that a compact subset of a Souslin space is metrizable (for the subspace topology); see Schwartz (1973, p. 96, Theorem 3, and p. 106, Corollary 2). Thus (A2) and (A3)(i) imply that the sets X_t are metrizable. Using this fact, it is easily seen that upper semi-continuity of the payoff functions and (A8)(i) imply the following:

$$(*)\quad \begin{array}{l} \text{If } y_n \to y \text{ in } E_G,\ x_n \in A_t(y_n) \text{ for each } n, \text{ and } \underline{\lim}\, u_t(x_n, y_n) \geq w_t(y), \\ \qquad \text{then } x \in \mathrm{LS}\, x_n \text{ implies } x \in A_t(y) \text{ and } u_t(x, y) = \overline{\lim}\, u_t(x_n, y_n) = w_t(y). \end{array}$$

From this fact together with (A3)(i) and the hypothesis that the functions w_t are lsc it follows that for any $t \in T$, w_t is actually continuous and $\varphi(t, \cdot)$ is well-behaved, because E_G is pseudo-metrizable. In particular, $\varphi(t, \cdot)$ is closed for all $t \in T$ and $y \in E_G$. Also, by (A7) and (A8)(ii), for any $y \in E_G$, the map $t \mapsto w_t(y)$ is measurable and the correspondence $\varphi(\cdot, y)$ has a measurable graph (see Castaing & Valadier (1977, Lemma III.39 and remarks in the sequel)[16]). Thus φ is a Caratheodory correspondence.

Now fix any $y \in E_G$ and assume there is no equilibrium strategy profile f such that $e(f) = y$. Recall from Lemma 3 that E_G is compact and pseudo-metrizable. Let ρ be a corresponding pseudo-metric. For $n \in \mathbb{N}\backslash\{0\}$, write $B_{1/n}(y)$ for the open ρ-ball around y of radius $1/n$. Also, for each $n \in \mathbb{N}\backslash\{0\}$ and each $y' \in B_{1/n}(y)$, let

$$\varphi^n(t, y') = \{x \in A_t(y') \colon u_t(x, y') \geq w_t(y) - 1/n\}.$$

We claim that there is an integer $n_1 > 0$ such that (d) in the definition of CS holds with $U = B_{1/n_1}(y)$, $\alpha \colon T \to [-\infty, +\infty]$ given by $\alpha(t) = w_t(y) - 1/n_1$, and with $T' \subseteq T$ having outer measure at least 2^{-n_1}. Indeed, otherwise there would be a sequence $\langle f_n \rangle$ of strategy profiles, with $e(f_n) \to y$, such that for each n there is a $T_n \subseteq T$ with $\nu(T_n) \leq 2^{-n}$ such that $f_n(t) \in \varphi^n(t, e(f_n))$ for almost all $t \in \hat{T}\backslash T_n$, and $f_n(t) \in \mathrm{co}\varphi^n(t, e(f_n))$ for almost all $t \in \bar{T}\backslash T_n$. Now by Lemma 4 there is a strategy profile f, with $e(f) = y$, such that $f(t) \in \mathrm{LS}\, f_n(t)$ for almost all $t \in \hat{T}$, and $f(t) \in \overline{\mathrm{co}}\, \mathrm{LS}\, f_n(t)$ for almost all $t \in \bar{T}$. Noting that the sequence $\left\langle \bigcup_{n \geq m} T_n \right\rangle_{m \in \mathbb{N}}$ of sets is decreasing with $\nu\left(\bigcup_{n \geq m} T_n\right) \to 0$, it follows that for almost all $t \in \hat{T}$, $f(t) \in \mathrm{KLS}\, \varphi^n(t, e(f_n))$, and for almost all $t \in \bar{T}$, $f(t) \in \overline{\mathrm{co}}\, \mathrm{KLS}\, \mathrm{co}\varphi^n(t, e(f_n))$.

By $(*)$ above, we have $\mathrm{KLS}\, \varphi^n(t, e(f_n)) \subseteq \varphi(t, y)$ for all $t \in T$. It follows that $f(t) \in \varphi(t, y)$ for almost all $t \in \hat{T}$. But the same must hold for almost all $t \in \bar{T}$. Indeed, let $t \in \bar{T}$. As noted above, $\varphi(t, y)$ is closed and by (A10), $\varphi(t, y)$ is convex. Thus if $f(t) \notin \varphi(t, y)$, then the separation theorem gives an open half-space $H \subseteq X$ (i.e., a set of form $H = \{z \in X : p(z) < r\}$ where p is a continuous linear functional on X, and r a real number) such that $\varphi(t, y) \subseteq H$ and $f(t) \notin \bar{H}$ where $\bar{H}$ is the closure of H. Now X_t and hence $X_t \backslash H$ are compact, so $\mathrm{KLS}\, \varphi^n(t, e(f_n)) \subseteq \varphi(t, y)$ implies that for large n, $\varphi^n(t, e(f_n) \subseteq H$ and hence also $\mathrm{co}\varphi^n(t, e(f_n) \subseteq H$. Consequently $\mathrm{KLS}\, \mathrm{co}\varphi^n(t, e(f_n)) \subseteq \bar{H}$, therefore also $\overline{\mathrm{co}}\, \mathrm{KLS}\, \mathrm{co}\varphi^n(t, e(f_n)) \subseteq \bar{H}$, so $f(t) \notin \varphi(t, y)$ implies $f(t) \notin \overline{\mathrm{co}}\, \mathrm{KLS}\, \mathrm{co}\varphi^n(t, e(f_n))$.

[16] We mention that this reference involves a measurable selection theorem.

Thus $f(t) \in \varphi(t,y)$ for almost all $t \in T$. As $e(f) = y$, this means f is an equilibrium strategy profile, and we get a contradiction to the assumption made about y, thus establishing the claim above.

Choose and fix an integer n_1 according to this claim. We next claim that there is an integer $n_2 > 0$ and a $T^{n_2} \subseteq T$ such that $\nu(T^{n_2}) > \nu(T) - 2^{-n_1}$ and such that for each $t \in T^{n_2}$, if $y' \in B_{1/n_2}(y)$ and $x \in \varphi(t,y')$ then $u_t(x,y') > w_t(y) - 1/n_1$. To see this, for each $n \in \mathbb{N}\backslash\{0\}$ let

$$T^n = \{t \in T \colon \inf_{y' \in B_{1/n}(y)} w_t(y') \geq w_t(y) - 1/n_1\}.$$

As was noted above, the map w_t is continuous for each $t \in T$, and the map $t \mapsto w_t(y')$ is measurable for each $y' \in E_G$. Also, as E_G, being compact and pseudo-metrizable, is separable, $B_{1/n}(y)$ contains a countable dense subset. Combining these facts, it follows that the map $t \mapsto \inf_{y' \in B_{1/n}(y)} w_t(y')$ is measurable, and hence that the set T^n is measurable. Now as each w_t is continuous, we have $T^n \uparrow T$ as $n \to \infty$, and it follows that $\nu(T^n) > \nu(T) - 2^{-n_1}$ for n large enough. Thus, since $x \in \varphi(t,y')$ means $u_t(x,y') = w_t(y')$, n_2 and T^{n_2} with the desired properties do exist.

Choose such n_2 and T^{n_2}, and set $n_3 = \max\{n_1, n_2\}$. Let φ^y be the restriction of φ to $T \times B_{1/n_3}(y)$. Then, since φ is a Caratheodory correspondence, so is φ^y. Modify the function α established above, if necessary, so as to get

$$\alpha(t) = \begin{cases} w_t(y) - 1/n_1 & \text{if } t \in T^{n_2} \\ -\infty & \text{otherwise.} \end{cases}$$

Then (d) of CS still holds with $U = B_{1/n_3}(y)$, as $1/n_3 \leq 1/n_1$ and $\nu(T\backslash T^{n_2}) < 2^{-n_1}$. Also, by construction, (a) and (c) of CS hold for $U = B_{1/n_3}(y)$, φ^y, and α. Finally, by (A10), (b) of CS holds for φ^y, too.

11.4.2 Proof of Theorem 5

With the hypothesis that a game be weakly usc, it still follows that $(*)$ in the proof of Theorem 4 holds, and therefore that proof still does the job (again with φ being the best reply correspondence of the game).

11.4.3 Proof of that the game is Section 2 has an equilibrium

Let $G = ((T,\Sigma,\nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, e)$ be the game constructed in Example 1. As noted in Example 1, this game satisfies (A1)-(A6), and an equilibrium of this game gives an equilibrium in the sense of Section 2. The construction of G shows that (S1) is also satisfied. If we have shown that G satisfies CS, then the existence of an equilibrium for the game in Section 2 follows from Theorem 1.

To verify CS, let us first recapitulate from the construction of G as representation of the game given in Section 2 that $X_t = A = \{a_1, a_2\}$ for all $t \in T$, that E_G is the unit simplex in $\mathbb{R}^2$, that $T = \hat{T}$, and that $A_t(y) = A$ for all $t \in T$ and all $y \in E_G$. Note also that $w_t(y) = \max_{i=1,2} u_t(a_i, y)$ for all $t \in T$ and all $y \in E_G$.

Now fix any $y \in E_G$ and assume there is no equilibrium strategy profile f such that $e(f) = y$.

Let $T^1 = \{t \in T \colon g_t(a_1, y) > m > g_t(a_2, y)\}$, $T^2 = \{t \in T \colon g_t(a_2, y) > m > g_t(a_1, y)\}$, and for each $k \in \mathbb{N} \setminus \{0\}$, let $T_k^1 = \{t \in T \colon g_t(a_1, y) > m + 1/k,\ g_t(a_2, y) < m - 1/k\}$ and $T_k^2 = \{t \in T \colon g_t(a_2, y) > m + 1/k,\ g_t(a_1, y) < m - 1/k\}$.

We claim that there is a $k \in \mathbb{N}$ and an open neighborhood U_1 of y such that for any $y' \in U_1$ there is no strategy profile f with $\nu(T_k^1 \cap f^{-1}(\{a_1\})) = \nu(T_k^2 \cap f^{-1}(\{a_2\})) = 0$ and $e(f) = y'$. Indeed, otherwise there is a sequence $\langle f_k \rangle$ of strategy profiles such that $e(f_k) \to y$ and such that for each $k \in \mathbb{N}$, $f_k(t) = a_2$ for almost all $t \in T_k^1$ and $f_k(t) = a_1$ for almost all $t \in T_k^2$. Now by Lemma 4, there is a strategy profile f with $e(f) = y$ such that $f(t) \in \operatorname{LS} f_k(t)$ for almost all $t \in T$. As the sequence $\langle T_k^1 \rangle$ is increasing with $\bigcup_{k \geq 1} T_k^1 = T^1$ we must have $f(t) = a_2$ for almost all $t \in T^1$. Similarly, we have $f(t) = a_1$ for almost all $t \in T^2$. By the definition of the sets T^1 and T^2, and the definition of the payoff functions, it follows that $u_t(f(t), e(f)) = w_t(e(f))$ for almost all $t \in T^1 \cup T^2$. Also by these definitions, and since the set of those $t \in T$ for which $g_t(a_1, y) = m$ or $g_t(a_2, y) = m$ is a null-set by definition of the functions g_t, we must have $u_t(f(t), e(f)) = w_t(e(f))$ for almost all $t \in T \setminus (T^1 \cup T^2)$. It follows that f is an equilibrium. But this contradicts the assumption made about y, thus establishing the claim above.

Choose k and U_1 according to the claim. Note that the maps $(t, a_i, y') \mapsto g_t(a_i, y')$ are uniformly continuous. There is therefore an open neighborhood U_2 of y such that for all $y' \in U_2$, $g_t(a_1, y') > m > g_t(a_2, y')$ for all $t \in T_k^1$, and $g_t(a_2, y') > m > g_t(a_1, y')$ for all $t \in T_k^2$. Set $U = U_1 \cap U_2$ and define $\varphi : T \times U \twoheadrightarrow X$ by setting, for all $(t, y) \in T \times U$,

$$\varphi(t, y) = \begin{cases} \{a_2\} & \text{if } t \in T_k^1, \\ \{a_1\} & \text{if } t \in T_k^2, \\ A & \text{otherwise.} \end{cases}$$

Furthermore, define $\alpha \colon T \times U \to X$ by setting, for all $(t, y) \in T \times U$,

$$\alpha(t) = \begin{cases} 1 & \text{if } t \in T_k^1 \cup T_k^2, \\ 1 - c & \text{otherwise.} \end{cases}$$

It is easily seen that the pair (φ, α) satisfies the requirements in condition CS. For (d) of CS, just note that if f is any strategy profile with $e(f) \in U$, then, setting $T' = (T_k^1 \cap f^{-1}(\{a_1\})) \cup (T_k^2 \cap f^{-1}(\{a_2\}))$, the set T' is non-negligible by the choice of k and U_1, and $u_t(f(t)) = 1 - c$ for all $t \in T'$ by definition of the payoff functions and of the sets T_k^i, $i = 1, 2$.

11.4.4 Proof that the game in Example 2 satisfies CS

Note first that if $f \in S_G$ satisfies $e^1(f) = 1/2$, then f is a Nash equilibrium of this game if and only if $e^3(f) = 1/2$.

Let $y \in E_G$ be such that there is no equilibrium strategy profile f with $e(f) = y$. The case $y^1 \neq 1/2$ is easy because, in this case, there is a neighborhood V of y in E_G such that u is continuous on $A \times V$. Hence, assume $y^1 = 1/2$. Since $y \neq e(f)$ for any equilibrium strategy profile f, then $y^2 > 0$.

Let $U = \{y' \in E_G \colon y'^1 > 0 \text{ and } y'^2 > 0\}$, so that U is a neighborhood of y in E_G, and let $\alpha\colon T \to \mathbb{R}$ be given by $\alpha(t) = 1/2$ for all $t \in T$, and $\varphi\colon T \times U \twoheadrightarrow X$ by $\varphi(t, y') = \{a_3\}$ for all $(t, y') \in T \times U$, so that α is measurable and φ is a Caratheodory correspondence. Clearly, with this choice of φ, (a) of CS holds, and since $\bar{T} = \emptyset$, so does (b) of CS. Also, $u_t(x, y') \geq 1/2$ for all $t \in T$, $y' \in U$ and $x \in \varphi(t, y')$, so (c) of CS is satisfied for φ and α. But (d) of CS is satisfied as well: If f is a strategy profile with $e(f) \in U$, then $e^1(f) > 0$ and $e^2(f) > 0$; hence, as $e^i(f) = \nu(f^{-1}\{a_i\})$, setting

$$T' = \begin{cases} f^{-1}(\{a_1\}) \text{ if } e^1(f) < 1/2, \\ f^{-1}(\{a_2\}) \text{ if } e^1(f) \geq 1/2, \end{cases}$$

it follows that $\nu(T') > 0$ and that $u_t(f(t), e(f)) < \alpha(t)$ for all $t \in T'$.

11.5 Proof of Theorem 6

Fix a game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t \rangle_{t \in T}, \tilde{e})$ as described in Section 7 such that the assumptions of Theorem 6 hold. We will show that for a suitably chosen set $\hat{\mathcal{C}}$ as involved in the definition of externality as stated in Section 4.1.3, E_G is such that there is a homeomorphism $h\colon \tilde{E}_G \to E_G$ with $e = h \circ \tilde{e}$. Given such an h, we may identify $\tilde{E}_G$ with E_G via h. In particular, we may view the constraint correspondences A_t as being defined on E_G, and the payoff functions u_t as being defined on the respective sets $X_t \times E_G$. Moreover, under this identification, CS' is equivalent to CS. The theorem under proof is therefore implied by Theorem 2, noting that (A11) implies (A6).

The idea to get such a homeomorphism $h\colon \tilde{E}_G \to E_G$ is simple. Indeed, suppose that $T = \hat{T}$ and ν is an atomless probability measure, that there is common action set for all players, say $X_t = K$ for all $t \in T$ where K is a compact metric space, and that $\tilde{e}(f) = \nu \circ f^{-1}$ for each $f \in S_G$, ignoring the attribute space C. This is just the standard scenario of large games. Now $\tilde{E}_G \equiv \tilde{e}(S_G)$ is equal to the space $M_+^1(K)$ of Borel probability measures on K with the narrow topology, and in particular is compact (note that as ν is atomless, every $y \in M_+^1(K)$ is the distribution of some $f \in S_G$).

Choose a countable family $\langle p_i \rangle_{i \in i}$ of continuous functions defined on K which separate the points of $M^1_+(K)$. For each $i \in I$, define $q_i \colon T \times K \to \mathbb{R}$ by $q_i(t,x) = p_i(x)$. Let $\hat{\mathcal{C}} = \{q_i \colon i \in I\}$, and let e and E_G be defined relative to $\hat{\mathcal{C}}$ according to Section 4.1.3. Define $h \colon \tilde{E}_G \to \mathbb{R}^I$ by setting $h(y) = \langle \int p_i \mathrm{d}y \rangle_{i \in I}$. Evidently $h \circ \tilde{e} = e$ and it follows that h is a bijection from $\tilde{E}_G$ onto E_G. By definition of the narrow topology, h is continuous, and since $\tilde{E}_G$ is compact, h must be a homeomorphism.

But for the proof of Theorem 6, this argument needs to be expanded. In particular, it is not assumed in Theorem 6 that the action sets of the players are included in a common compact subset of X. Further, we have to take care of the attribute space C which is involved in the definition of the map $\tilde{e}$. We proceed as follows.

Recall first that any Hausdorff locally convex space is completely regular. Hence by (A2) and (A12)(i), $X \times C$ is a completely regular Souslin space. By what was noted in footnote 15, there is a countable family $\langle p_i \rangle_{i \in I}$ of continuous bounded functions on $X \times C$ which separates the points of $X \times C$. We may assume that the family $\langle p_i \rangle_{i \in I}$ is stable with respect to multiplication. Then by Schwartz (1973, p. 388, Corollary 1), the family of maps $\gamma \mapsto \int p_i \mathrm{d}\gamma \colon M^1_+(X \times C) \to \mathbb{R}$, $i \in I$, separates the points of $M^1_+(X \times C)$. For each $j \in J$ and $i \in I$, define a map $q_{ij} \colon \Gamma_G \cap (\hat{T} \times X) \to \mathbb{R}$ by setting

$$q_{ij}(t,x) = \begin{cases} \frac{1}{\nu(T_j)} p_i(x, c(t)) & \text{if } t \in T_j \\ 0 & \text{if } t \in \hat{T} \setminus T_j \end{cases}$$

(where the sets T_j are the subsets of $\hat{T}$ from the definition of the externality map $\tilde{e}$).

Let $\hat{\mathcal{C}} = \{q_{ij} \colon i \in I, j \in J\}$, and with this choice of $\hat{\mathcal{C}}$ let $e \colon S_G \to \bar{S}_G \times \mathbb{R}^{\hat{\mathcal{C}}}$ be the externality mapping of the general model as developed in Section 4.1. Note that $\hat{\mathcal{C}}$ satisfies the requirements in that model; in particular, $\hat{\mathcal{C}}$ is countable. As in Section 4.1, let $E_G = e(S_G)$, endowed with the same topology as there.

Define $h' \colon \prod_{t \in \bar{T}} X_t \times \left(M^1_+(X \times C)\right)^J \to \prod_{t \in \bar{T}} X_t \times \mathbb{R}^{\hat{\mathcal{C}}}$ by setting

$$h'\big(z, \langle \gamma_j \rangle_{j \in J}\big) = \left(z, \left\langle \left\langle \int_{X \times C} p_i \mathrm{d}\gamma_j \right\rangle_{i \in I} \right\rangle_{j \in J}\right), \; z \in \prod_{t \in \bar{T}} X_t \,, \; \langle \gamma_j \rangle_{j \in J} \in \big(M^1_+(X \times C)\big)^J.$$

Recall from Section 7 that we are viewing each X_t as being endowed with the subspace topology defined from X, $M^1_+(X \times C)$ as being endowed with the narrow topology, and all products involving these spaces as being endowed with the product topology. Thus h' is continuous. Note also that by choice of the family $\langle p_i \rangle_{i \in I}$, h' is an injection. Let h be the restriction of h' to $\tilde{E}_G$.

For each $f \in S_G$ and each $j \in J$, set $\gamma_j(f) = (1/\nu(T_j))(\nu|_{T_j}) \circ (f|_{T_j}, c|_{T_j})^{-1}$. Note that for any $f \in S_G$, and each $i \in I$ and $j \in J$, we have

$$\int_{X \times C} p_i \mathrm{d}\gamma_j(f) = \int_{T_j} \frac{1}{\nu(T_j)} p_i(f(t), c(t)) \mathrm{d}\nu(t) = \int_{\hat{T}} q_{ij}(t, f(t)) \mathrm{d}\nu(t).$$

Using this fact, it follows that $e = h \circ \tilde{e}$, by the definitions of the three maps involved. In particular, as $E_G = e(S_G)$ by definition of E_G, h is a surjection from $\tilde{E}_G$ onto E_G.

We next show that for every $j \in J$ there is a compact set $K_j \subseteq M^1_+(X \times C)$ such that $\{(1/\nu(T_j))(\nu|_{T_j}) \circ (f|_{T_j}, c|_{T_j})^{-1} \colon f \in S_G\} \subseteq K_j$. To this end, let $\kappa \colon \hat{T} \twoheadrightarrow X \times C$ be defined by setting $\kappa(t) = X_t \times \{c(t)\}$ for each $t \in \hat{T}$ and note that κ takes non-empty compact values. As C is a Souslin space according to (A12)(i), measurability of the map c (Assumption (A12)(ii)) implies that the correspondence $t \mapsto \{c(t)\}$ has a measurable graph (see Castaing & Valadier (1977, p. 74)), and according to (A3)(ii), so does the correspondence $t \mapsto X_t$. By the fact that $\mathcal{B}(X \times C) \supseteq \mathcal{B}(X) \otimes \mathcal{B}(C)$, it is now elementary to check that κ has a measurable graph. Fix any $j \in J$ and let κ_j be the restriction of κ to T_j. Note that κ_j has again a measurable graph. Let $\mathcal{R}_{\kappa_j}$ denote the set of all Young measures g from T_j to $X \times C$ with $\operatorname{supp} g(t) \subseteq \kappa(t)$ for almost all $t \in T_j$, endowed with the narrow topology for Young measures (cf. 11.3.2). For each $g \in \mathcal{R}_{\kappa_j}$, define $\mu_g \in M^1_+(X)$ by setting $\mu_g(B) = \frac{1}{\nu(T_j)} \int_{T_j} g(t)(B) \mathrm{d}(\nu|_{T_j})(t)$ for every $B \in \mathcal{B}(X)$. (That μ_g is indeed countably additive may be easily seen with the help of the monotone convergence theorem). Note that, for any $f \in S_G$, setting $g(t) = \delta_{(f(t),c(t))}$ for $t \in T_j$, where $\delta_{(f(t),c(t))}$ is Dirac measure at $(f(t), c(t))$, defines an element g of $\mathcal{R}_{\kappa_j}$ for which $\mu_g = (1/\nu(T_j))(\nu|_{T_j}) \circ (f|_{T_j}, c|_{T_j})^{-1}$. Now the map $g \mapsto \mu_g$ from $\mathcal{R}_{\kappa_j}$ to $M^1_+(X)$ is continuous for the narrow topology of $M^1_+(X)$ (see Balder (2002, proof of Theorem 3.1.1)), and thus the assertion follows from Theorem 10 in Section 11.3.2.

Now by (A3), $\prod_{t \in \bar{T}} X_t$ is compact. From this and the previous paragraph it follows that there is a compact $K \subseteq \prod_{t \in \bar{T}} X_t \times \left(M^1_+(X \times C)\right)^J$ such that $\tilde{E}_G \subseteq K$. The facts that K is compact and h' is a continuous injection defined on $\prod_{t \in \bar{T}} X_t \times \left(M^1_+(X \times C)\right)^J$ mean that the restriction of h' to K is a homeomorphism from K onto $h'(K)$. As $\tilde{E}_G \subseteq K$ and h is the restriction of h' to $\tilde{E}_G$, it follows that h is a homeomorphism from $\tilde{E}_G$ onto $h(\tilde{E}_G)$ (recall that the topology of $\tilde{E}_G$ is the subspace topology defined from $\prod_{t \in \bar{T}} X_t \times \left(M^1_+(X \times C)\right)^J$). By what was stated in Remark 6, (A11) means that the feeble topology on $\bar{S}_G \equiv \prod_{t \in \bar{T}} X_t$, which is involved in the definition of E_G, is the same as the product topology of $\prod_{t \in \bar{T}} X_t$. By the fact that $h(\tilde{E}_G) = E_G$, we may now conclude that h is a homeomorphism from $\tilde{E}_G$ to E_G. This completes the proof.

11.6 Proof of Theorem 7

Recall first from the proof of Theorem 6 that, for some choice of an externality map e as defined in Section 4.1, $\tilde{E}_G \equiv \tilde{e}(S_G)$ may be homeomorphically identified with $E_G \equiv e(S_G)$. Consequently, Lemmas 3 and 4 continue to hold with $\tilde{e}$ in place of e, and $\tilde{E}_G$ in place of E_G. In particular, $\tilde{E}_G$ is pseudo-metrizable, therefore first-countable. (In fact, by (A11), $\tilde{E}_G$ is metrizable.)

Let $y \in \tilde{E}_G$ be such that there is no equilibrium strategy profile f of G with $\tilde{e}(f) = y$. We claim that there is an $\varepsilon > 0$ and a neighborhood V of y such that if $f \in S_G$ satisfies $\tilde{e}(f) \in V$ and $f(t) \in A_t(\tilde{e}(f))$ for almost all $t \in T$, there is $T' \subseteq T$ with outer measure larger than ε such that $u_t(f(t), \tilde{e}(f)) < w_t(y) - \varepsilon$ for all $t \in T'$.

Indeed, otherwise there is a sequence $\langle f_k \rangle$ in S_G with $\tilde{e}(f_k) \to y$ such that for each k, $f_k(t) \in A_t(\tilde{e}(f_k))$ for almost all $t \in T$, and for some $T_k \subseteq T$ with $\nu(T_k) < 2^{-k}$, $u_t(f_k(t), \tilde{e}(f_k)) \geq w_t(y) - 2^{-k}$ for all $t \in T \backslash T_k$. Now the sequence $\left\langle \bigcup_{n \geq k} T_n \right\rangle_{k \in \mathbb{N}}$ of sets is decreasing with $\nu\left(\bigcup_{n \geq k} T_n\right) \to 0$, and thus we must have $\underline{\lim}_k u_t(f_k(t), \tilde{e}(f_k)) \geq w_t(y)$ for almost every $t \in T$. Lemma 4 gives an $f \in S_G$ such that $f(t) \in \mathrm{LS}\, f_k(t)$ for almost all $t \in \hat{T}$ and such that $\tilde{e}(f) = y$. Note that $\tilde{e}(f_k) \to \tilde{e}(f)$ implies $f(t) = \lim_k f_k(t)$ for $t \in \bar{T}$; see the paragraph following the statement of Assumption (A11). Thus we must have $f(t) \in \mathrm{LS}\, f_k(t)$ for almost all $t \in T$. By BRC, it follows that f is an equilibrium of G, and as $\tilde{e}(f) = y$ we thus get a contradiction to the assumption made about y.

Fix a neighborhood V of y and a number $\varepsilon > 0$ as just established. Relative to this ε, let U, φ, and α be chosen according to GPS. Then (U, φ, α) satisfies (1)-(3) in CS'. By (A13), (4) in CS' is equivalent to the following statement: Whenever f is a strategy profile with $\tilde{e}(f) \in U$ and $f(t) \in A_t(\tilde{e}(f))$ for almost all $t \in T$, then there is a non-negligible set $T' \subseteq T$ such that $u_t(f(t), \tilde{e}(f)) < \alpha(t)$ for all $t \in T'$. Thus, shrinking the set U, if necessary, so that $U \subseteq V$, (4) in CS' must hold because of (d) in GPS.

11.7 Proofs related to Section 9

11.7.1 Proof of Lemma 1

Here is a short roadmap. Lemmas 5 to 8 provide auxiliary results. Verification of BRC will be done in Lemma 9, and verification of GPS in Lemmas 10 and 11.

To give an overview of the argument establishing BRC, let f and f_n, $n \in \mathbb{N}$, be strategy profiles. BRC requires that if (a)-(c) in its statement hold for these strategy profiles, then in the limit f almost every player realizes a best reply in $A_t(\tilde{e}(f))$ against $\tilde{e}(f)$. That this requirement is satisfied for players in $\hat{T}$ follows because these players have continuous payoff functions and well-behaved budget constraint correspondences (the latter fact will be established in Lemma 8). For player $\bar{t}$, i.e., the government, a difficulty arises because the payoff function $u_{\bar{t}}$ need not be usc (see the example prior to the statement of Theorem 1). Resolving this difficulty needs some work which occupies the main part of the proof of Lemma 9. The argument, which uses Lemma 7, consists in showing that if almost all players in $\hat{T}$ realize a best reply in the limit f, then this must hold for player $\bar{t}$, too.

To establish GPS, we apply (T7) to deal with player $\bar{t}$. The main step concerning the players in $\hat{T}$ is isolated in Lemma 10. The idea is to show that given a player

$t \in \hat{T}$ and an externality $y = (\lambda, \gamma)$, the value $w_t(y)$ can be approximated by the payoff of an action that is available in the budget sets for all tax functions λ' near to λ. To do so, we will use fact that a tax function, being non-decreasing, can have at most countably many discontinuity points, and the fact, established in Lemma 6, that the map $(\lambda, z) \mapsto \lambda(z)$, which sends tax functions and pre-tax incomes to after tax incomes, is continuous at a point (λ, z) whenever z is a continuity point of λ.

Lemma 5. Λ *is compact.*

Proof. Let $\langle \lambda_k \rangle$ be a sequence in Λ. As Λ is uniformly bounded, we may assume $\lambda_k(\bar{n}) \to a$ for some number a. Now using the version of Helly's selection theorem in Billingsley (1968, p. 227)), we may find a $\lambda \in \Lambda$ and a subsequence $\langle \lambda_i \rangle$ of $\langle \lambda_k \rangle$ such that $\lambda(\bar{n}) \leq a$ and $\lambda_i(z) \to \lambda(z)$ if $z \in (0, \bar{n})$ is a continuity point of λ. Define $\lambda_1 \colon [0, \bar{n}] \to \mathbb{R}_+$ by setting $\lambda_1(\bar{n}) = a$ and $\lambda_1(z) = \lambda(z)$ for $z \in [0, \bar{n})$. Then $\lambda_i(\bar{n}) \to \lambda_1(\bar{n})$, and as $a \geq \lambda(\bar{n})$, λ_1 is still in Λ. Being non-decreasing, λ_1 has only countably many discontinuity points, and it follows that $\langle \lambda_i \rangle$ converges to λ pointwise almost everywhere with respect to Lebesgue measure. Using Lebesgue's dominated convergence theorem, we may conclude that $\rho(\lambda_i, \lambda_1) \to 0$. □

Lemma 6. *Suppose* $\lambda_k \to \lambda$ *in* Λ *(i.e.,* $\rho(\lambda, \lambda_k) \to 0$*) and* $z_k \to z$ *in* $[0, \bar{n}]$*. Then* $\lambda(z) \geq \overline{\lim}_k \lambda_k(z_k)$*. If* $z > 0$ *and* z *is a continuity point of* λ*, then* $\lambda(z) = \lim_k \lambda_k(z_k)$*.*

Proof. If $z = \bar{n}$, then, as $\lambda_k \to \lambda$ implies $\lambda_k(\bar{n}) \to \lambda(\bar{n})$ by definition of ρ, we have

$$\overline{\lim}_k \lambda_k(z_k) \leq \overline{\lim}_k \lambda_k(\bar{n}) = \lambda(\bar{n}),$$

just because the members of Λ are non-decreasing.

Now by the fact that the members of Λ are right-continuous and non-decreasing, it is elementary to check that $\lambda_k(z^*) \to \lambda(z^*)$ if $z^* \in (0, \bar{n})$ is a continuity point of λ, and as λ can have only countably many discontinuity points, the set of continuity points of λ is dense in $[0, \bar{n}]$. Using these facts, we can continue as follows.

Suppose $z < \bar{n}$. Then, for any continuity point z^* of λ with $z^* \in (z, \bar{n}]$,

$$\overline{\lim}_k \lambda_k(z_k) \leq \overline{\lim}_k \lambda_k(z^*) = \lambda(z^*),$$

and thus $\overline{\lim}_k \lambda_k(z_k) \leq \lambda(z)$ by right-continuity of λ.

Suppose $z > 0$ is a continuity point of λ. Then given $\varepsilon > 0$ there is a continuity point z^* of λ with $z^* < z$ such that $\lambda(z^*) > \lambda(z) - \varepsilon$. Now

$$\underline{\lim}_k \lambda_k(z_k) \geq \underline{\lim}_k \lambda_k(z^*) = \lambda(z^*) > \lambda(z) - \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, $\underline{\lim}_k \lambda_k(z_k) \geq \lambda(z)$. □

For the next lemma, recall that if $\gamma \in K$ then $\hat{\gamma} \in \hat{K}$ denotes the distribution of the map $(u, n, m, l) \mapsto nl \colon C \times M \times N \to [0, \bar{n}]$, and that the map $\gamma \mapsto \hat{\gamma}$ is continuous.

Lemma 7. *Let $\lambda_k \to \lambda$ in Λ and $\gamma_k \to \gamma$ in K. Then (a) $\int u(\lambda(nl), l) \mathrm{d}\gamma(u, n, m, l) \geq \overline{\lim}_k \int u(\lambda_k(nl), l) \mathrm{d}\gamma_k(u, n, m, l)$ and (b) $\int \pi(\lambda(z)) \mathrm{d}\hat{\gamma}(z) \geq \overline{\lim}_k \int \pi(\lambda_k(z)) \mathrm{d}\hat{\gamma}_k(z)$ if $\pi \colon [0, \bar{m}] \to \mathbb{R}$ is continuous and non-decreasing.*

Proof. (b) Note that $\gamma_k \to \gamma$ implies $\hat{\gamma}_k \to \hat{\gamma}_k$. Let μ be Lebesgue measure on $[0, 1]$. By Skorokhod's Theorem we can select measurable maps h, $h_k \colon [0, 1] \to [0, \bar{n}]$, $k \in \mathbb{N}$, such that $\hat{\gamma} = \mu \circ h^{-1}$ and $\hat{\gamma}_k = \mu \circ h_k^{-1}$ for each k and such that $h_k(a) \to h(a)$ for almost all $a \in [0, 1]$. Using Fatou's lemma, Lemma 6, and the properties of π, we get

$$\begin{aligned} \overline{\lim}_k \int \pi(\lambda_k(z)) \mathrm{d}\hat{\gamma}_k(z) = \overline{\lim}_k \int \pi(\lambda_k(h_k(a))) \mathrm{d}\mu(a) &\leq \int \overline{\lim}_k \pi(\lambda_k(h_k(a))) \mathrm{d}\mu(a) \\ &\leq \int \pi(\lambda(h(a))) \mathrm{d}\mu(a) = \int \pi(\lambda(z)) \mathrm{d}\hat{\gamma}(z). \end{aligned}$$

This proves (b). Part (a) follows similarly, but this time with the maps h and h_k going to $\hat{C} \times N \times M \times N$, observing that if $(u_k, n_k, m_k, l_k) \to (u, n, m, l)$ in $\hat{C} \times N \times M \times N$ and $\lambda_k \to \lambda$ in Λ, then $\overline{\lim}_k u_k(\lambda_k(n_k l_k), l_k) \leq u(\lambda(nl), l)$, by Lemma 6 and because $(u, n, m, l) \in \hat{C} \times N \times M \times N$ implies that u is non-decreasing in m. □

Lemma 8. *For each $t \in T$, the correspondence A_t is well-behaved. Furthermore, for each $y \in \tilde{E}_G$, the correspondence $t \mapsto A_t(y)$ has a measurable graph.*

Proof. It follows from (T5) that $A_{\bar{t}}$ is well-behaved. Consider any $t \in \hat{T}$. As $\lambda(z) \in \mathbb{R}_+$ for all $z \in [0, \bar{n}]$ and $\lambda \in \Lambda$, we have $(0, 0) \in A_t(y)$ for all $y \in \tilde{E}_G$, i.e., A_t takes non-empty values. Let $y = (\lambda, \gamma) \in \tilde{E}_G$, $(m, l) \in M \times L$, and suppose $\langle(\lambda_k, \gamma_k)\rangle$ and $\langle(m_k, l_k)\rangle$ are sequences in $\tilde{E}_G$ and $M \times L$, respectively, with $(\lambda_k, \gamma_k) \to (\lambda, \gamma)$, $(m_k, l_k) \to (m, l)$ and $(m_k, l_k) \in A_t(\lambda_k, \gamma_k)$ for all k. Then by Lemma 6, we have $m = \lim_k m_k \leq \overline{\lim}_k \lambda(n_t l_k) \leq \lambda(n_t l)$, so $(m, l) \in A_t(\lambda, \gamma)$. Thus A_t is closed and therefore well-behaved as $M \times L$ is compact.

Let $y = (\lambda, \gamma) \in \tilde{E}_G$. To show that $t \mapsto A_t(y)$ has a measurable graph, it suffices to show that Γ_A is measurable, where Γ_A is the graph of the restriction of $t \mapsto A_t(y)$ to $\hat{T}$. Define $p : \hat{T} \times M \times L \to \mathbb{R}$ by $p(t, m, l) = \lambda(n_t l) - m$ for all $(t, m, l) \in \hat{T} \times M \times L$. Then p is measurable and $\Gamma_A = p^{-1}(\mathbb{R}_+)$. Thus Γ_A is measurable. □

Lemma 9. *The game satisfies* BRC.

Proof. Let $f \in S_G$ and $\langle f_k \rangle$ a sequence in S_G such that (a) $\tilde{e}(f_k) \to \tilde{e}(f)$ and, for almost all $t \in T$, (b) $f_k(t) \in A_t(\tilde{e}(f_k))$ for all $k \in \mathbb{N}$, (c) $f(t) \in \mathrm{LS}\, f_k(t)$, and (d) $\underline{\lim}_k u_t(f_k(t), \tilde{e}(f_k)) \geq w_t(\tilde{e}(f))$. We have to show that for almost all $t \in T$, $f(t) \in A_t(\tilde{e}(f))$ and $u_t(f(t), \tilde{e}(f(t))) = w_t(\tilde{e}(f))$.

By Lemma 8, (b) and (c) imply that $f(t) \in A_t(\tilde{e}(f))$ for almost all $t \in T$. As $\tilde{u}_t$ is continuous for each $t \in \hat{T}$, (c) implies $\tilde{u}_t(f(t)) \in \mathrm{LS}\, \tilde{u}_t(f_k(t))$ for almost all $t \in \hat{T}$. Hence, by the definitions of u_t and w_t for $t \in \hat{T}$, (d) implies $u_t(f(t), \tilde{e}(f(t))) = w_t(\tilde{e}(f))$ for almost all $t \in \hat{T}$.

It remains to see that $u_{\bar{t}}(\lambda, \tilde{e}(f)) = w_{\bar{t}}(\tilde{e}(f))$. To this end, set $\gamma = \hat{\nu} \circ (\hat{n}, \hat{f})^{-1}$, and $\gamma_k = \hat{\nu} \circ (\hat{n}, \hat{f}_k)^{-1}$ for $k \in \mathbb{N}$. Write $\lambda_k = f_k(\bar{t})$, $k \in \mathbb{N}$, and $\lambda = f(\bar{t})$.

Now (a) implies both $\lambda_k \to \lambda$ and $\gamma_k \to \gamma$, so $\int nl \mathrm{d}\gamma_k(u, n, m, l) \to \int nl \mathrm{d}\gamma(u, n, m, l)$ and $\int m \mathrm{d}\gamma_k(u, n, m, l) \to \int m \mathrm{d}\gamma(u, n, m, l)$. Moreover, from Lemma 7(b), with π there being the identity on $[0, \bar{m}]$, we have $\int \lambda(nl) \mathrm{d}\gamma(u, n, m, l) \geq \overline{\lim}_k \int \lambda_k(nl) \mathrm{d}\gamma_k(u, n, m, l)$.

Next note that for all k, we have $\int m \mathrm{d}\gamma_k(u, n, m, l) \leq \int \lambda_k(nl) d\gamma_k(u, n, m, l)$ as $f_k(t) \in A_t(\tilde{e}(f_k))$ for almost all $t \in \hat{T}$. Also, $\int m \mathrm{d}\gamma(u, n, m, l) = \int \lambda(nl) d\gamma(u, n, m, l)$ because for almost all $t \in \hat{T}$, both $f(t) \in A_t(\tilde{e}(f))$ and $u_t(f(t), \tilde{e}(f(t))) = w_t(\tilde{e}(f))$, and because the functions u_t are strictly increasing in m. Hence, by the last two facts stated in the previous paragraph, $\int \lambda_k(nl) d\gamma_k(u, n, m, l) \to \int \lambda(nl) \mathrm{d}\gamma(u, n, m, l)$.

Finally, note that by (T7) and the definition of $u_{\bar{t}}$, we have $w_{\bar{t}}(\tilde{e}(f)) \geq 0$ since v is non-negative by (T6). Thus $\underline{\lim}_k u_{\bar{t}}(\lambda_k, \tilde{e}(f_k)) \geq w_{\bar{t}}(\tilde{e}(f))$ implies that we must have $u_{\bar{t}}(\lambda_k, \tilde{e}(f_k)) = v(\lambda_k, \gamma_k)$ as well as $\int \lambda_k(nl) \mathrm{d}\gamma_k(u, n, m, l) = \int nl \mathrm{d}\gamma_k(u, n, m, l)$ for all sufficiently large k, again by definition of $u_{\bar{t}}$ and non-negativity of v. By the fact that $\int nl \mathrm{d}\gamma_k(u, n, m, l) \to \int nl \mathrm{d}\gamma(u, n, m, l)$ and the conclusion of the previous paragraph, it follows that $\int \lambda(nl) \mathrm{d}\gamma(u, n, m, l) = \int nl \mathrm{d}\gamma(u, n, m, l)$ and hence that $u_{\bar{t}}(\lambda, \tilde{e}(f)) = v(\lambda, \gamma)$. As v is usc by (T6), we can conclude that

$$u_{\bar{t}}(\lambda, \tilde{e}(f)) = v(\lambda, \gamma) \geq \overline{\lim}_k v(\lambda_k, \gamma_k) = \overline{\lim}_k u_{\bar{t}}(\lambda_k, \tilde{e}(f_k)) \geq w_{\bar{t}}(\tilde{e}(f)),$$

so that $u_{\bar{t}}(\lambda, \tilde{e}(f)) = w_{\bar{t}}(\tilde{e}(f))$. This completes the proof. $\square$

For the following, recall that $u_t(x, y) = \tilde{u}_t(x)$ for all $t \in \hat{T}$, $x \in M \times L$ and $y \in \tilde{E}_G$.

Lemma 10. *Given $y = (\lambda, \gamma) \in \tilde{E}_G$ and $\varepsilon > 0$, there are measurable maps $\beta\colon \hat{T} \to \mathbb{R}$ and $f\colon \hat{T} \to M \times L$ and a neighborhood W of λ in Λ such that:*

1. *$f(t) \in A_t(y')$ for all $t \in \hat{T}$ and all $y' = (\lambda', \gamma) \in \tilde{E}_G$ with $\lambda' \in W$.*
2. *$u_t(f(t), y') \geq \beta(t)$ for all $t \in \hat{T}$ and all $y' = (\lambda', \gamma) \in \tilde{E}_G$ with $\lambda' \in W$.*
3. *For some $T_\varepsilon \subseteq \hat{T}$ with $\nu(T_\varepsilon) < \varepsilon$, $\beta(t) \geq w_t(y) - \varepsilon$ for all $t \in \hat{T} \backslash T_\varepsilon$.*

Proof. Let $y = (\lambda, \gamma) \in \tilde{E}_G$ and $\varepsilon > 0$ be given. Note first that (T1) implies that if $h\colon \hat{T} \to M \times L$ is measurable, then so is the map $t \mapsto \tilde{u}_t(h(t))\colon \hat{T} \to \mathbb{R}$. Moreover, (T1) also implies that the map $(t, x) \mapsto \tilde{u}_t(x)\colon \hat{T} \times M \times L \to \mathbb{R}$ is measurable. Thus, by Lemma 8, since $\tilde{u}_t$ is continuous for each $t \in \hat{T}$, the map $t \mapsto w_t(y)$ is measurable and there is a measurable $g = (m, l)\colon \hat{T} \to M \times L$ such that $m(t) \leq \lambda(n_t l(t))$ and $\tilde{u}_t(g(t)) = w_t(y)$ for all $t \in \hat{T}$ (see Castaing & Valadier (1977, Lemma III.39 and the remarks following its proof)). Let $D \subseteq N$ be the set of discontinuity points of λ, and note that D is countable as λ is non-decreasing. Let $D_1 = \{d_1, d_2, \dots\}$ be a countable dense subset of N such that $D \cap D_1 = \emptyset$. Let $T_1 = \{t \in \hat{T} \colon m(t) > 0, n_t > 0, n_t \notin D\}$ and note that T_1 is a measurable subset of $\hat{T}$. For $i, j \in \mathbb{N}\backslash\{0\}$, set

$$T_{ij} = \big\{t \in T_1 : d_i/n_t \leq 1,\ \tilde{u}_t\big(\tfrac{j-1}{j} m(t), d_i/n_t\big) > w_t(y) - \varepsilon,\ \tfrac{j-1}{j} m(t) < \lambda(d_i)\big\}.$$

Then T_{ij} is measurable, and by the definition of T_1 and the fact that λ is non-decreasing and right-continuous, we have $\bigcup_{i,j\in\mathbb{N}\backslash\{0\}} T_{ij} = T_1$. We can therefore find a measurable map $g_1 = (m_1, l_1) \colon T_1 \to M \times L$ such that for all $t \in T_1$, $\tilde{u}_t(g_1(t)) > w_t(y) - \varepsilon$, $m_1(t) < \lambda(n_t l_1(t))$, and $n_t l_1(t) \in D_1$. Now as each $d \in D_1$ is a continuity point of λ, and as D_1 is countable, Lemma 6 implies that there are a neighborhood W of λ and a measurable $T_\varepsilon \subseteq T_1$, with $\nu(T_\varepsilon) < \varepsilon$, such that $m_1(t) < \lambda'(n_t l_1(t))$ for each $t \in T_1\backslash T_\varepsilon$ and $\lambda' \in W$. Now define $f \colon \hat{T} \to M \times L$ by setting $f(t) = g_1(t)$ if $t \in T_1\backslash T_\varepsilon$ and $f(t) = (0,0)$ otherwise, and $\beta \colon \hat{T} \to \mathbb{R}$ by setting $\beta(t) = \tilde{u}_t(f(t))$ for all $t \in \hat{T}$. Concerning $t \in \hat{T}\backslash T_1$, note that if $m(t) = 0$ then $w_t(y) = \tilde{u}_t(0,0)$ by (T2), and that $\{t \in \hat{T} \colon n_t \in D \text{ or } n_t = 0\}$ is a null set by (T4), as D is countable. □

Lemma 11. *The game satisfied* GPS.

Proof. Let $y = (\lambda, \gamma) \in \tilde{E}_G$ and $\varepsilon > 0$. Let W, f, and β be chosen according to Lemma 10, and let O and ψ be chosen according to Assumption (T7). Set $U = W \times O$. Define a correspondence $\varphi \colon T \times U \twoheadrightarrow (M \times L) \cup \Lambda$ as follows. Given $y' = (\lambda', \gamma') \in U$, set $\varphi(\bar{t}, y') = \{\psi(\gamma')\}$, and set $\varphi(t, y') = \{f(t)\}$ for each $t \in \hat{T}$. Define a function $\alpha \colon T \to \mathbb{R}$ by setting $\alpha(\bar{t}) = w_{\bar{t}}(y) - \varepsilon$, and by setting $\alpha(t) = \beta(t)$ for each $t \in \hat{T}$. It is readily seen that (U, φ, α) satisfies the requirements in GPS. □

11.7.2 Proof of Theorem 8

We will establish Theorem 8 by appeal to Weierstrass' theorem. The main task is to find a representation of the government's optimization problem so that we have both continuity of the objective function and compactness of the choice set. We will obtain such a representation in terms of the game $G = ((T, \Sigma, \nu), X, \langle X_t, u_t, A_t\rangle_{t\in T}, \tilde{e})$ specified in Section 9.

We start by recording the following fact.

Lemma 12. *For all $y \in \tilde{E}_G$, w_t is lsc at y for almost all $t \in \hat{T}$.*

Proof. Fix $y = (\lambda, \gamma) \in \tilde{E}_G$. Clearly, if $w_t(y) = \tilde{u}_t(0,0)$ then w_t is lsc at y because $(0,0) \in A_t(y')$ for every $y' \in \tilde{E}_G$. Now by the proof of Lemma 10, for almost all $t \in T$, if $w_t(y) > \tilde{u}_t(0,0)$ then, given $\varepsilon > 0$, there is a point $(m, l) \in M \times L$, with $m < n_t l$ and $\tilde{u}_t(m, l) > w_t(y) - \varepsilon$, such that $n_t l$ is a continuity point of λ. In view of Lemma 6, this shows that the assertion holds. □

Now let

$$\tilde{S}(E) = \{\tilde{e}(f) : f \text{ is an equilibrium of } G\}$$

and let $F\colon \tilde{S}(E) \to \mathbb{R}$ be defined by setting $F(\lambda, \gamma) = \int u(m,l) d\gamma(u,n,m,l)$ for all $y = (\lambda, \gamma) \in \tilde{S}(E)$. Consider the problem $\max_{y \in \tilde{S}(E)} F(y)$. By change of variables, this problem has a solution if and only if the problem $\max_{f \in S(E)} \int_{\hat{T}} \tilde{u}_t(f(t)) \mathrm{d}\nu(t)$ has a solution, where $S(E)$ is the set of all equilibria of G. As pointed out in Section 9, every equilibrium of the game G defines an equilibrium of the economy E, with the same utilities for the individuals in $\hat{T}$, and vice versa. Thus it suffices to show that the problem $\max_{y \in \tilde{S}(E)} F(y)$ has a solution.

To this end, note by Theorem 9, $\tilde{S}(E)$ is nonempty. Moreover, as a subset of the compact and metrizable space $\tilde{E}_G$, $\tilde{S}(E)$ is closed and therefore compact. To see this, let $y \in \tilde{E}_G$ and $\langle y_k \rangle$ a sequence in $\tilde{S}(E)$ such that $y_k \to y$. For each $k \in \mathbb{N}$, let f_k be an equilibrium of G such that $\tilde{e}(f_k) = y_k$. By Lemma 4 in 11.4.1, there is an $f \in S_G$ such that $\tilde{e}(f) = y$ and $f(t) \in \mathrm{LS}\, f_k(t)$ for almost all $t \in T$. Since w_t is lsc for almost all $t \in T$ (by Lemma 12 and (T7)), we must have

$$\underline{\lim}_k\, u_t(f_k(t), \tilde{e}(f_k)) = \underline{\lim}_k\, w_t(\tilde{e}(f_k)) \geq w_t(\tilde{e}(f))$$

for almost all $t \in T$. It now follows from Lemma 9 that f is an equilibrium of G, and thus $y \in \tilde{S}(E)$.

Finally, note that by (T1) the map $(u,n,m,l) \mapsto u(m,l)\colon \tilde{S}(E) \to \mathbb{R}$ is continuous. As $\tilde{S}(E)$ is compact, it follows first that the map F is continuous, and then that the problem $\max_{y \in \tilde{S}(E)} F(y)$ has a solution. This completes the proof.

11.7.3 Lemmas for Examples 6 and 7

Recall for the proofs below that if $\gamma \in K$ then $\hat{\gamma} \in \hat{K}$ denotes the distribution of the map $(u,n,m,l) \mapsto nl\colon C \times M \times N \to [0, \bar{n}]$, and that the map $\gamma \mapsto \hat{\gamma}$ is continuous.

Lemma 13. *Let Θ and v be as in Example 6. Then (T5)-(T7) hold. Moreover, (1) holds for any equilibrium (λ^*, g^*) of $E = \langle(\hat{T}, \hat{\Sigma}, \hat{\nu}), M, L, N, \hat{n}, \Lambda, \Theta, \langle \tilde{u}_t, n_t \rangle_{t \in \hat{T}}\rangle$.*

Proof. For (T5), note first that as π is concave, Θ has convex values. To show that Θ is well-behaved, it suffices to show that Θ is closed. To see this, note first that the two functions π and η must be continuous. Because, in addition, π is increasing, Lemma 7 in Section 11.7.1 implies that the map $(\lambda, \gamma) \mapsto \int \pi(\lambda(z)) \mathrm{d}\hat{\gamma}(z)\colon \Lambda \times K \to \mathbb{R}$ is usc. Using these facts, it is easily seen that Θ has a closed graph.

Concerning (T7), fix $\gamma \in K$ and $\varepsilon > 0$. Let $O = K$ and define $\psi\colon O \to \Lambda$ by setting $\psi(\gamma') = \left(\int z \mathrm{d}\hat{\gamma}'(z)\right)\chi_N$ for all $\gamma' \in O$, where χ_N is the characteristic function of N. Then ψ is continuous and (T7)(ii) holds. Also, $\int \pi(\psi(\gamma')(z)) \mathrm{d}\hat{\gamma}'(z) = \pi\left(\int z \mathrm{d}\hat{\gamma}'(z)\right)$. Using this fact, it follows that (T7)(i) holds. Finally, it is clear that (T7)(iii) holds, as v constantly takes value 0.

As for the last part of the lemma, suppose (λ^*, g^*) is an equilibrium of E and set $\gamma = \hat{\nu} \circ (c, g^*)^{-1}$. Assume first that $\pi\left(\int z d\hat{\gamma}(z)\right) + \int \eta(l) d\gamma(u, n, m, l) > 0$. By (a) of the equilibrium definition, (1) is equivalent to $\int \pi(\lambda^*(z)) d\hat{\gamma}(z) + \delta \int \eta(l) d\gamma(u, n, m, n) \geq (1-\delta)\pi\left(\int z d\hat{\gamma}(z)\right)$. Thus (1) holds, since $\lambda^* \in \Theta(\gamma)$ by definition of equilibrium.

Assume $\pi\left(\int z d\hat{\gamma}(z)\right) + \int \eta(l) d\gamma(u, n, m, l) \leq 0$. Note that (b) of the equilibrium definition implies $\pi(\lambda^*(n_t l^*(t))) + \eta(l^*(t)) \geq 0$ for almost all $t \in \hat{T}$, because $0 \leq \lambda^*(0)$ and $\pi(0) + \eta(0) = 0$. Thus, by concavity of π and by (a) of the equilibrium definition,

$$
\begin{aligned}
0 &\leq \int \pi(\lambda^*(z)) \mathrm{d}\hat{\gamma}(z) + \int \eta(l) \mathrm{d}\gamma(u, n, m, l) \\
&\leq \pi\left(\int \lambda^*(z) \mathrm{d}\hat{\gamma}(z)\right) + \int \eta(l) \mathrm{d}\gamma(u, n, m, l) \\
&= \pi\left(\int z \mathrm{d}\hat{\gamma}(z)\right) + \int \eta(l) \mathrm{d}\gamma(u, n, m, l) \leq 0.
\end{aligned}
$$

Now this shows that the first two sums must be zero, from which (1) follows. □

Lemma 14. *Let Θ and v be as in Example 7 and assume that $\tilde{u}_t(\cdot, l)$ is concave for all $t \in \hat{T}$ and $l \in L$. Then (T6) and (T7) hold.*

Proof. To see that (T6) holds, note first that by Lemma 7 in 11.7.1, v is usc. Now if $\lambda \in \Lambda$ is continuous, then the bounded map $(u, n, m, l) \mapsto u(\lambda(nl), l) \colon C \times M \times L \to \mathbb{R}$ is continuous, and hence the map $\gamma \mapsto \int u(\lambda(nl), l) \mathrm{d}\gamma(u, n, m, l) \equiv v(\lambda, \gamma) \colon K \to \mathbb{R}$ is continuous (by definition of the narrow topology). By the assumption that $\tilde{u}_t(\cdot, l)$ is concave for all $t \in \hat{T}$ and $l \in L$, it follows that $v(\cdot, \gamma)$ is quasi-concave for all $\gamma \in K$. Finally, by (T2), $v(\lambda, \gamma) \geq 0$ for all $(\lambda, \gamma) \in \Lambda \times K$. Thus (T6) holds.

As for (T7), fix $\gamma \in K$ and $\varepsilon > 0$. Recall that λ_0 is the element of Λ satisfying $\lambda_0(z) = z$ for all $z \in [0, \bar{n}]$. Let $B = \{\lambda \in \Lambda \colon \int \lambda(z) \mathrm{d}\hat{\gamma}(z) = \int z \mathrm{d}\hat{\gamma}(z)\}$ and note that $\lambda_0 \in B$. Let $S = \{\lambda \in B \colon \rho(\lambda, \lambda_0) \leq \varepsilon(\hat{\gamma})\}$ (recall that ρ is the metric on Λ defined in Section 9). Set $s = \sup\{v(\lambda, \gamma) \colon \lambda \in S\}$.

Suppose $\int z \mathrm{d}\hat{\gamma}(z) = 0$. Then $v(\lambda_0, \gamma) = s$. Since λ_0 is continuous, by (T6) there is a neighborhood O of γ such that $v(\lambda_0, \gamma') > v(\lambda_0, \gamma) - \varepsilon$ for all $\gamma' \in O$. Set $\psi(\gamma') = \lambda_0$ for all $\gamma' \in O$. Evidently ψ is as required in (T7).

Suppose $\int z \mathrm{d}\hat{\gamma}(z) > 0$ which implies $\varepsilon(\hat{\gamma}) > 0$. It suffices to find a $\lambda \in B$ with the following properties: (a) $v(\lambda, \gamma) > s - \varepsilon$, (b) $\rho(\lambda, \lambda_0) < \varepsilon(\hat{\gamma})$, and (c) λ is continuous.

Indeed, given such a λ, by (c) we have $\int \lambda(z) \mathrm{d}\hat{\gamma}'(z) > 0$ for all γ' in some neighborhood O of γ, and thus $\int \alpha(\gamma')\lambda(z) \mathrm{d}\hat{\gamma}'(z) = \int z \mathrm{d}\hat{\gamma}'(z)$ for some number $\alpha(\gamma') > 0$. Note that the map $\gamma' \mapsto \alpha(\gamma')$ is continuous with $\alpha(\gamma) = 1$. Consider a sequence $\langle \gamma_k \rangle$ in O with $\gamma_k \to \gamma$; in particular, $\alpha(\gamma_k) \to 1$. Now (c) implies that we have $u_k(\alpha(\gamma_k)\lambda(n_k l_k), l_k) \to u(\lambda(nl), l)$ whenever $(u_k, n_k, m_k, l_k) \to (u, n, m, l)$. Using Billingsley (1968, Theorem 5.5, p. 34) and the definition of v, it follows that

$v(\alpha(\gamma_k)\lambda, \gamma_k) \to v(\lambda, \gamma)$. In view of this and (a), shrinking the set O, if necessary, we have $v(\alpha(\gamma')\lambda, \gamma') > s - \varepsilon$ for all $\gamma' \in O$. Shrinking the set O another time, if necessary, (b) implies that we can arrange to also have $\rho(\alpha(\gamma')\lambda, \lambda_0) < \varepsilon(\hat{\gamma}')$ for each $\gamma' \in O$, using the fact that the maps $\gamma' \mapsto \hat{\gamma}'$ and $\hat{\gamma}' \mapsto \varepsilon(\hat{\gamma}')$ are continuous, together with continuity of a metric. Now define $\psi\colon O \to \Lambda$ by setting $\psi(\gamma') = \alpha(\gamma')\lambda$ for each $\gamma' \in O$. Then ψ is as required in (T7).

Now to find a λ as desired, choose some $\lambda_1 \in S$ such that (a) holds for λ_1. Note that the function $\alpha \mapsto v(\alpha\lambda_1 + (1-\alpha)\lambda_0, \gamma)\colon [0,1] \to \mathbb{R}$ is continuous, and that $\rho(\alpha\lambda_1 + (1-\alpha)\lambda_0, \lambda_0) < \varepsilon(\hat{\gamma})$ for each $0 < \alpha < 1$, by definition of the metric ρ. Thus, there is a $\lambda_2 \in S$ such that (b) holds in addition to (a).

We claim there is a sequence $\langle \lambda_k \rangle$ of continuous elements of Λ with $\lambda_k(z) \to \lambda_2(z)$ for each $z \in [0, \bar{n}]$. Indeed, let $D = \{z_i \colon i \in \mathbb{N}\}$ be a countable dense subset of $[0, \bar{n}]$ with $\bar{n} \in D$. For each $i, n \in \mathbb{N}$ choose a continuous element $\lambda'_{i,n}$ in Λ such that $\lambda'_{i,n} \geq \lambda_2$ and $\lambda'_{i,n}(z_i) < \lambda_2(z_i) + 1/(n+1)$, using the fact that λ_2, being in Λ, is non-decreasing and right-continuous. Re-index the family $\langle \lambda'_{i,n} \rangle_{i,n \in \mathbb{N}}$ as a sequence $\langle \lambda'_j \rangle_{j \in \mathbb{N}}$. For each $k \in \mathbb{N}$, let λ_k be the pointwise infimum of the set $\{\lambda'_j \colon j \leq k\}$. Then each λ_k is still continuous and non-decreasing, hence in Λ, and we have $\lambda_k(z_i) \to \lambda_2(z_i)$ for each $z_i \in D$. By the choice of D, and because λ_2 is right-continuous and each λ_k is non-decreasing and satisfies $\lambda_k \geq \lambda_2$, it follows that $\lambda_k(z) \to \lambda_2(z)$ for each $z \in [0, \bar{n}]$.

Now by Lebesgue's dominated convergence theorem and the fact that $\lambda_2 \in S \subseteq B$, there is a sequence $\langle \alpha_k \rangle$ of real numbers, with $\alpha_k \to 1$, such that $\alpha_k \lambda_k \in B$ for each k. Another invocation of Lebesgue's dominated convergence theorem shows that $v(\alpha_k\lambda_k, \gamma) \to v(\lambda_2, \gamma)$, so $\alpha_k\lambda_k$ satisfies (a) if k is large enough. Finally, once more using Lebesgue's dominated convergence theorem, we have $\rho(\lambda_2, \alpha_k\lambda_k) \to 0$ by definition of ρ, so the fact that λ_2 satisfies (b) implies that (b) is also satisfied by $\alpha_k\lambda_k$ if k is large enough. Thus, for large k, $\alpha_k\lambda_k$ is in B and satisfies all of (a)-(c) above. □

References

Aliprantis, C. & Border, K. (2006), *Infinite Dimensional Analysis*, 3rd edn, Springer, Berlin.

Bagh, A. & Jofre, A. (2006), 'Reciprocal upper semicontinuity and better reply secure games: A comment', *Econometrica* **74**, 1715–1721.

Balder, E. (1988), 'Generalized equilibrium results for games with incomplete information', *Mathematics of Operations Research* **13**, 265–276.

Balder, E. J. (1989), 'On Prohorov's theorem for transition probabilities', *Sém. Anal. Convexe Montpellier* **19**, 9.1–9.11.

Balder, E. J. (2000), 'Lectures on Young measure theory and its applications in economics', *Rend. Istit. Mat. Univ. Trieste Vol. XXXI, Suppl. 1, 1–69* .

Balder, E. J. (2002), 'A unifying pair of Cournot-Nash equilibrium existence results', *Journal of Economic Theory* **102**, 437–470.

Balder, E. J. (2011), 'An equilibrium closure result for discontinuous games', *Economic Theory* **48**, 47–65.

Barelli, P. & Duggan, J. (2012), Extremal choice equilibrium. University of Rochester.

Barelli, P. & Meneghel, I. (2013), 'A note on the equilibrium existence problem in discontinuous games', *Econometrica* **81**, 813–824.

Barelli, P. & Soza, I. (2009), On the existence of Nash equilibria in discontinuous and qualitative games. University of Rochester.

Bich, P. (2009), 'Existence of pure Nash equilibria in discontinuous and non quasiconcave games', *International Journal of Game Theory* **38**, 395–410.

Bich, P. & Laraki, R. (2012), A unified approach to equilibrium existence in discontinuous strategic games. Paris School of Economics.

Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.

Carmona, G. (2009), 'An existence result for discontinuous games', *Journal of Economic Theory* **144**, 1333–1340.

Carmona, G. (2011), 'Understanding some recent existence results for discontinuous games', *Economic Theory* **48**, 31–45.

Carmona, G. (2012), Reducible equilibrium properties: Comments on recent existence results. University of Surrey.

Carmona, G. & Podczeck, K. (2009), 'On the existence of pure strategy Nash equilibria in large games', *Journal of Economic Theory* **144**, 1300–1319.

Castaing, C. & Valadier, M. (1977), *Convex Analysis and Measurable Multifunctions*, Vol. 580 of *Lect. Notes Math*, Springer-Verlag, New York.

Chari, V. & Kehoe, P. (1990), 'Sustainable plans', *Journal of Political Economy* **98**, 783–802.

Dasgupta, P. & Maskin, E. (1986), 'The existence of equilibrium in discontinuous economic games, i: Theory', *Review of Economics Studies* **53**, 1–26.

de Castro, L. (2011), 'Equilibria existence and approximation of regular discontinuous games', *Economic Theory* **48**, 67–85.

Diestel, J. & Uhl, J. J. (1977), *Vector Measures*, American Mathematical Society, Providence.

Engelking, R. (1989), *General Topology*, Vol. 6 of *Sigma Series in Pure Mathematics*, Heldermann Verlag, Berlin.

Farhi, E., Sleet, C., Werning, I. & Yeltekin, S. (2011), Nonlinear capital taxation without commitment. Harvard University, Carnegie Mellon University and MIT.

Fremlin, D. H. (2008), *Measure Theory*, Vol. 5: Set-Theoretic Measure Theory, Torres Fremlin, Colchester.

Golosov, M., Kocherlakota, N. & Tsyvinski, A. (2003), 'Optimal indirect and capital taxation', *Review of Economic Studies* **70**, 569–587.

Karni, E. & Levin, D. (1994), 'Social attributes and strategic equilibrium: A restaurant pricing game', *Journal of Political Economy* **102**, 822–840.

Keisler, H. J. & Sun, Y. (2009), 'Why saturated probability spaces are necessary', *Advances in Mathematics* **221**, 1584–1607.

Khan, M. (1986), 'Equilibrium points of nonatomic games over a Banach space', *Transactions of the American Mathematical Society* **293**, 737–749.

Khan, M., Rath, K. & Sun, Y. (1997), 'On the existence of pure strategy equilibria in games with a continuum of players', *Journal of Economic Theory* **76**, 13–46.

Khan, M. & Sun, Y. (1999), 'Non-cooperative games on hyperfinite Loeb spaces', *Journal of Mathematical Economics* **31**, 455–492.

Kim, T. & Yannelis, N. (1997), 'Existence of equilibrium in Bayesian games with infinitely many players', *Journal of Economic Theory* **77**, 330–353.

Kocherlakota, N. (2010), *The New Dynamic Public Finance*, Princeton University Press, Princeton.

Mailath, G. & Samuelson, L. (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press, Oxford.

Mandelkern, M. (1990), 'On the uniform continuity of Tietze extensions', *Archiv der Mathematik* **55**, 387–388.

Martins da Rocha, V. & Topuzu, M. (2008), 'Cournot-Nash equilibria in continuum games with non-ordered preferences', *Journal of Economic Theory* **140**, 314–327.

Mas-Colell, A. (1984), 'On a theorem by Schmeidler', *Journal of Mathematical Economics* **13**, 201–206.

McLennan, A., Monteiro, P. & Tourky, R. (2011), 'Games with discontinuous payoffs: a strengthening of Reny's existence theorem', *Econometrica* **79**, 1643–1664.

Milgrom, P. & Weber, R. (1985), 'Distributional strategies for games with incomplete information', *Mathematics of Operations Research* **10**, 619–632.

Mirrlees, J. (1971), 'An exploration in the theory of optimum income taxation', *Review of Economic Studies* **38**, 175–208.

Nessah, R. (2011), 'Generalized weak transfer continuity and Nash equilibrium', *Journal of Mathematical Economics* **47**, 659–662.

Páscoa, M. (1993), 'Noncooperative equilibrium and Chamberlinian monopolistic competition', *Journal of Economic Theory* **60**, 335–353.

Podczeck, K. (2009), 'On purification of measure-valued maps', *Economic Theory* **38**, 399–418.

Prokopovych, P. (2011), 'On equilibrium existence in payoff secure games', *Economic Theory* **48**, 5–16.

Prokopovych, P. (2013), 'The single deviation property in games with discontinuous payoffs', *Economic Theory* **53**, 383–402.

Rath, K. (1996), 'Existence and upper hemicontinuity of equilibrium distributions of anonymous games with discontinuous payoffs', *Journal of Mathematical Economics* **26**, 305–324.

Reny, P. (1999), 'On the existence of pure and mixed strategy equilibria in discontinuous games', *Econometrica* **67**, 1029–1056.

Reny, P. (2009), Further results on the existence of Nash equilibria in discontinuous games. University of Chicago.

Reny, P. (2011), 'Strategic approximations of discontinuous games', *Economic Theory* **48**, 17–29.

Reny, P. (2013), Nash equilibrium in discontinuous games. University of Chicago.

Roughgarden, T. & Tardos, E. (2004), 'Bounding the inefficiency of equilibria in nonatomic congestion games', *Games and Economic Behavior* **47**, 389–403.

Sabourian, H. (1990), 'Anonymous repeated games with a large number of players and random outcomes', *Journal of Economic Theory* **51**, 92–110.

Schmeidler, D. (1973), 'Equilibrium points of nonatomic games', *Journal of Statistical Physics* **4**, 295–300.

Schwartz, L. (1973), *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press, Oxford.

Yannelis, N. C. (2009), 'Debreu's social equilibrium theorem with asymmetric information and a continuum of agents', *Economic Theory* **38**, 419–432.