Vanderbilt University Department of Economics Working Paper Series
Nashville, TN 37235-1819

Vanderbilt University Department of Economics Working Papers 16-00006

The Subgame Perfect Core

Parkash Chander
Jindal School of Government and Public Policy

Myrna Wooders
Vanderbilt University

Abstract

We propose a cooperative solution concept for games in extensive form that incorporates both cooperation and subgame perfection. This new concept, which we label the subgame-perfect core, is a refinement of the core of an extensive game in the same sense as the set of subgame-perfect Nash equilibria is a refinement of the set of Nash equilibria. Moreover, each subgame perfect core payoff vector can be obtained as a subgame-perfect Nash equilibrium payoff vector of a modified extensive game. We establish several additional properties of the subgame-perfect core and demonstrate its applicability by studying three applications: the centipede game, the two-player infinite bargaining game of alternating offers, and a dynamic game of climate change. In addition, we motivate and introduce a concept of subgame-perfect strong Nash equilibrium of an extensive game and show that it is coalition proof.

The authors are indebted to the Douglas Grey Fund for Research in Economics at Vanderbilt University for financial support. We are thankful to Philippe Jehiel and Francoise Forge for their helpful comments. We have also benefited from seminar presentations at Monash and University of Paris, Dauphine. url: www.parkashchander.com url: www.myrnawooders.com
Citation: Parkash Chander and Myrna Wooders, (2016) "The Subgame Perfect Core", *Vanderbilt University Department of Economics Working Papers*, VUECON-16-00006.
Contact: Parkash Chander - pchander@jgu.edu.in, Myrna Wooders - myrna.wooders@gmail.com.
Submitted: March 12, 2016. **Published:** March 13, 2016.
URL:http://www.accessecon.com/Pubs/VUECON/VUECON-16-00006.pdf

The Subgame-Perfect Core[1]

1. Introduction

We propose a cooperative solution concept for general extensive games that can improve upon subgame-perfect Nash equilibrium. This new concept, which we label the subgame-perfect core, is a refinement of the core of the extensive game in the same sense as the set of subgame-perfect Nash equilibria is a refinement of the set of Nash equilibria. From its definition and properties, our concept is a cooperative analog of the non-cooperative subgame-perfect Nash equilibrium. Moreover, the subgame-perfect core has "perfection" properties; each subgame-perfect core payoff vector can be obtained as a subgame-perfect Nash equilibrium payoff vector of a modified extensive game that differs from the original game only in terms of the distribution of players' payoffs at a terminal node. In this paper, we restrict ourselves to extensive games of perfect information with transferable utility. Consequently, each decision node determines a subgame, and terminal payoffs can be added and distributed among the players in the game.

Arguably, the most well-known approach to defining the core of a non-cooperative game is Aumann (1961), which addresses *strategic games*. That paper proposes two ways to define a characteristic function, both of which attribute a worth or payoff to each coalition.[2] With a characteristic function in place, the *core* of the strategic game is the core of the characteristic function game. In contrast, we address *extensive form games*. Defining a core concept for an extensive form game creates new challenges.

Given an extensive game, we assume that coalitions may form in any subgame. A coalition, however, can consist only of *active* players, that is, those players who still have decisions to make in that subgame. Additionally, coalition members can agree only upon actions still to be taken in the subgame. Thus, at any subgame, the past is finished and previous actions taken by players cannot be changed. Moreover, when a coalition forms in a subgame, a new game is

[1] The authors are indebted to the Douglas Grey Fund for Research in Economics at Vanderbilt University for financial support. We are thankful to Philippe Jehiel and Francoise Forge for their helpful comments. We have also benefited from seminar presentations at Monash and University of Paris, Dauphine.

[2] The so-called α and β approaches.

created in which members of the coalition act as a single player. Since utility is transferable, the total payoff to the members of any coalition can be divided among the members of the coalition. Note that the largest payoff to a coalition can be achieved by the grand coalition, consisting of the total player set. Roughly speaking, the subgame-perfect core is the set of payoff vectors that can be achieved by a distribution to the players of the grand coalition's payoff with the property that neither a singleton player nor a coalition of players can profitably deviate at any decision node in the game.

We stress that two notions are fundamental to our concept of the subgame-perfect core. First, we assume that a coalition becomes a single player, precisely as in the derivation of a characteristic function from a strategic game in Aumann (1961). Given a game in extensive form with player set N, when a coalition S forms, a new game, to be called the induced game, is created with player set $\{i \in N\backslash S, S\}$. However, since we treat extensive form games, the payoffs that S may be able to achieve can vary along a path of the game tree. Another fundamental idea is that, at any decision node of the extensive game, only those players who still have decisions to make can form coalitions and only they can coordinate their decisions from that point onwards.[3] While in this paper we treat mainly games of perfect information, these two ideas are formalized below and can be applied to games of both perfect and imperfect information.

Like the core of a characteristic function game, the subgame-perfect core of an extensive game may be empty. The emptiness of the subgame-perfect core, however, carries an important message about the game. In this case, a subgame-perfect Nash equilibrium remains the predicted outcome, despite the fact that cooperation is allowed.

Further discussion

In the treatment of cooperation within coalitions in a non-cooperative game a question arises: What is the response of the players in the complementary set given a deviation by a coalition? Because we can think of a coalition as simply a player in a game derived from the original game,

[3] Since the player set of the original game does not include any player who has no decisions to make in the game, we treat a subgame analogously by excluding players who no longer have decisions to make in the subgame. Not doing so would be inconsistent with the very notion of subgame perfection.

it is appealing to take the players in the complement as singletons, especially since that, as will be shown, leads to a core concept that in addition to its other properties nicely relates to both subgame-perfect strong and coalition-proof Nash equilibria, in which the players in the complement are also taken to be singletons.[4] Yet our approach goes beyond that. As we will demonstrate and explain below, our approach can be applied even if the players in the complement are not assumed to be singletons.

Another question that arises is whether binding agreements can be made. We assume that binding agreements can be made over distribution of the total payoff at *each* terminal node of the game. Thus these agreements resemble contingent contracts; the agreement that comes into effect depends on the terminal node reached. We characterize such contingent contracts and show that for a contract to be credible the distribution of total payoff at terminal nodes with highest total payoff for the grand coalition must be a subgame-perfect core payoff vector.

The subgame-perfect core takes into account interactions of coalitions in a fashion analogous to how Nash and subgame-perfect Nash equilibria (SPNE) take into account interactions of players. More specifically, a payoff vector belongs to the subgame-perfect core if (a) there is a history that leads to a terminal node for which the payoff vector is feasible and, (b) at any decision node in the history, no active coalition can improve upon its part of the payoff vector by deviating, where the payoff that a (deviating) coalition can obtain is assumed to be equal to its highest SPNE payoff in the induced game with origin at the decision node.

Relationships to other solution concepts

We show that an extensive game, like a strategic game, has a characteristic function form. This result is significant because it implies that concepts and ideas from the vast literature on games in characteristic form can now be extended and applied to games in extensive form. One immediate implication is that a non-empty subgame-perfect core exists *if and only if* the characteristic function game derived from the extensive game is balanced (Bondareva, 1963 and

[4] Chander (2007) shows that forming singletons is a subgame-perfect Nash equilibrium strategy of the players in the complement in an infinitely repeated game of coalition formation, which implies that the players in the complement may indeed have incentives to form singletons.

Shapely, 1967). Furthermore, as we will show, the fact that the characteristic function is derived from an extensive game generates additional subtleties and applications.

We also show that the subgame-perfect core is a refinement of the cores of a family of subgames and introduce an alternative equivalent concept as well as a weaker version of the subgame-perfect core. The weaker subgame-perfect core (WSPC) is generally larger. As will be shown, the WSPC is a complementary, rather than an alternative, concept.

We demonstrate applicability of the subgame-perfect core to three classes of games: the centipede game, the two-player infinite bargaining game of alternating offers, and a dynamic game of climate change. First, the application to the centipede game illustrates that in some games there may be no conflict between cooperation and non-cooperation. Second, the application to the bargaining game establishes an equivalence relationship between the subgame-perfect core and the static axiomatic Nash solution for the bilateral bargaining problem. Third, the application to the dynamic game of climate change shows that the subgame-perfect core can be applied also to games of imperfect information.

The centipede game (Rosenthal, 1981) has been at the center of the debate concerning the SPNE concept (e.g. Binmore, 1996 *and* Aumann, 1996) and has been often used to motivate the extensive form trembling-hand perfect Nash equilibrium (Selten, 1975). The inefficiency of the unique SPNE of this game has been also tested in experiments in game theory (McKelvey and Palfrey, 1992). We show that the subgame-perfect core of a centipede game may be non-empty and consist of a unique payoff vector that, unlike the SPNE, is efficient.

The two-player infinite bargaining game of alternating offers (Rubinstein, 1982) is *foundational* to the unification of the static axiomatic and the dynamic non-cooperative approaches to bargaining. As is well known, this game admits a unique SPNE. We show that the subgame-perfect core of this game is non-empty and equivalent to the unique SPNE payoff vector. This equivalence, as will be shown, is independent of the patience, i.e. the discount factors, of the players. However, Binmore, Rubinstein and Wolinsky (1986) show that if the players are patient, the unique SPNE payoff vector of Rubinstein's game is equal to the Nash bargaining solution. It follows that if players are patient, the unique subgame-perfect core payoff vector is also equal to the Nash bargaining solution, since it is then equal to the unique SPNE payoff vector in Rubinstein's game.

We introduce, as a byproduct of the conceptual framework introduced in this paper, a concept of subgame-perfect strong Nash equilibrium (SPSNE) for an *arbitrary* extensive game.[5] We show that the subgame-perfect core is generally a weaker concept than SPSNE, but also that the two may be equal in some games. In fact, we show that in Rubinstein's game, the unique SPNE is also a SPSNE which implies equivalence between the subgame-perfect core, the SPNE, the SPSNE, and the axiomatic Nash solution if players are patient. In order to justify the SPSNE as a convincing extension of the strong Nash equilibrium for strategic games, we show that, just as in strategic games, a SPSNE is a subgame-perfect coalition-proof Nash equilibrium.

Our work is related to research that seeks to unify cooperative and non-cooperative game theory through the melding of cooperative game theoretical solutions with non-cooperative Nash equilibria, the so-called "Nash Program". Numerous papers have contributed to this program including Perry and Reny (1994), Pérez-Castrillo (1994), Compte and Jehiel (2010), and Lehrer and Scarsini (2013), for example. In contrast to our work, these papers start with a cooperative game and a notion of the core and then propose a non-cooperative procedure to implement the core. While our paper makes a contribution to the Nash program our approach is entirely different; we start with an extensive game and the notion of subgame perfectness. We show that each payoff vector belonging to the subgame-perfect core of an extensive game is a SPNE payoff vector of another extensive game that differs from the original extensive game *only* with respect to players' payoffs at a terminal node with highest payoff for the grand coalition.

Our work also differs from another interesting literature that considers cores of sequences of *characteristic function* games; see, for example, Kranich, Perea, and Peters (2005), Habis and Herings (2010), and Predtetchinski, Herings, and Perea (2006). In contrast to our paper, these studies start from characteristic function games as the primitive and do not consider subgame perfectness; instead they place rules on admissible deviations. We conjecture that investigation

[5] Rubinstein (1980) introduces a strong perfect equilibrium for a "super" game. But a concept of a subgame-perfect strong Nash equilibrium for a *general* extensive game is apparently missing in the literature.

of the subgame-perfect core of the sort of dynamic games considered in these papers would be a fruitful line of research but it is beyond the scope of this paper.

This work contributes to our knowledge on core concepts in dynamic games, a field that has long attracted the interest of economists. Notably Gale (1978) explores the issue of time consistency in the Arrow-Debreu model with dated commodities and introduces the sequential core which consists of allocations that cannot be improved upon by anyone at any date. Similarly, Forges, Mertens, and Vohra (2002) propose the ex-ante incentive compatible core. Becker and Chakrabarti (1995) propose the recursive core as the set of allocations such that no coalition can improve upon its consumption stream at any time. In contrast, this paper proposes a core concept for a *general* extensive game that satisfies subgame-perfection and can be, as shown below, applied to a variety of dynamic games.

Organization of the paper

The paper is organized as follows. Section 2 introduces notation and a motivating example. Section 3 introduces the definition of the subgame-perfect core and applies it to the centipede and Rubinstein's infinite bargaining games. This section also introduces and characterizes credible contracts and establishes equivalence between the subgame-perfect core, the SPNE of Rubinstein's infinite bargaining game, and the axiomatic Nash solution for the two-person bargaining problem. Section 4 establishes several additional properties and interpretations, introduces an alternative equivalent definition and a weaker notion of the subgame-perfect core. Section 5 motivates and introduces the concept of SPSNE for a general extensive game. Section 6 presents the application to a dynamic game of climate change. Section 7 makes concluding remarks that further address the significance of this research and future directions for research.

2. The framework and a motivating example

We denote an extensive game of perfect information by $\Gamma = (N, K, P, u)$, where $N = \{1, \dots, n\}$ is the player set and K is the game tree with origin denoted by 0. Let Z denote the set of terminal nodes of game tree K and let X denote the set of non-terminal nodes, i.e., the set of decision nodes. The player partition of X is given by $P = \{X_1, \dots, X_n\}$ where X_i is the set of all decision nodes of player $i \in N$. The payoff function is $u: Z \to R^n$ where $u_i(z)$ denotes the payoff of player i at terminal node z. A payoff vector $p = (p_1, \dots, p_n)$ is *feasible* for terminal node $z \in Z$ if $\sum_{i \in N} p_i = \sum_{i \in N} u_i(z)$. The strategy sets need not be explicitly stated for now.

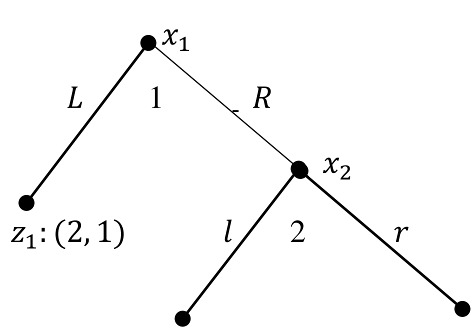
2.1 *The induced extensive games*

Given an extensive game $\Gamma = (N, K, P, u)$ and a coalition $S \subset N$, the *induced* extensive game $\Gamma^S = (N^S, K^S, P^S, u^S)$ is defined as follows:

- $N^S = \{S, (i)_{i \in N\setminus S}\}$: the player set wherein coalition S and all $i \in N\setminus S$ are the players (thus the game has $n - s + 1$ players);
- $K^S = K$: the game tree, equal to the game tree of the game with player set N (thus the set of decision and terminal nodes remain X and Z, respectively)
- $P^S = \{X_S, (X_i)_{i \in N\setminus S}\}$: the player partition of X where $X_S = \cup_{j \in S} X_j$ (thus the decision nodes of the coalition player S are all the decision nodes of members of S)
- $u^S = (u_S, (u_i)_{i \in N\setminus S})$: the profile of payoff functions of the players in N^S, where for all $z \in Z, u_S(z) = \sum_{j \in S} u_j(z)$ is the payoff function of S and $u_i(z)$ is the payoff function of $i \in N\setminus S$.

Notice that if S is a singleton coalition then $\Gamma^S = \Gamma$. For each $S \subset N$, the induced game $\Gamma^S = (N^S, K^S, P^S, u^S)$ represents the situation in which the players in S form a coalition to coordinate their decisions at all their decision nodes. Example 1, below, illustrates the definitions so far.

Example 1 Let Γ denote the extensive game depicted in Fig.1. Then, x_1 is the origin of the game tree $K, N = \{1,2\}$: the set of players, $Z = \{z_1, z_2, z_3\}$: the set of terminal nodes, $X = \{x_1, x_2\}$: the set of decision nodes, $P = \{\{x_1\}, \{x_2\}\}$: the player partition, and the payoff function $u: Z \to R^2$ is given by $u(z_1) = (2,1), u(z_2) = (4,2)$, and $u(z_3) = (1,3)$.



z_2: (4, 2) z_3: (1, 3)

Fig. 1

The induced extensive game Γ^N where players 1 and 2 form a coalition to coordinate their decisions in all their decision nodes is depicted in Fig. 2. The game tree is the same, but now we have a one-player game with player set$\{N\}$. So $P^N = \{\{x_1, x_2\}\}$, $u_N(z_1) = 3, u_N(z_2) = 6$, and $u_N(z_3) = 4$. Notice that each strategy of player N in game Γ^N generates a history of game Γ.

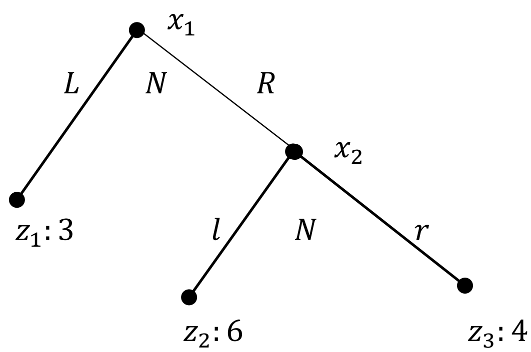


Fig. 2

2.2 *Defining achievable coalitional payoffs*

To derive the characteristic function of an extensive game we need to define the payoff that each coalition can achieve. We do so for every subgame and then explain by an example why this is necessary. Thus we first define the induced game with origin x for each $x \in X$.

Given a decision node $x \in X$, let Γ_x denote the subgame with origin at x. Since the origin of Γ is denoted by 0, $\Gamma_0 = \Gamma$ and if $x \neq 0$, then Γ_x is a proper subgame of Γ. It may be noted that the player set of a proper subgame Γ_x may be smaller than the set N (though not necessarily). A player is *active* in subgame Γ_x if some decision node in Γ_x is a decision node of the player. Similarly, a coalition is *active* in subgame Γ_x if *all* its members are active in the subgame Γ_x. Let S be an active coalition in subgame Γ_x. Then, the induced game Γ_x^S is defined from Γ_x in exactly the same way as the induced game Γ^S is defined from Γ. Clearly, $\Gamma_0^S = \Gamma^S$. Since Γ, by assumption, is a game of perfect information, so is each game Γ_x^S, $x \in X$ and S an active coalition in Γ_x. In what follows, it will be often convenient to refer to "a coalition that is active in the subgame with origin at x" simply as "an active coalition at x".

A SPNE of an extensive game induces a Nash equilibrium in each subgame of the extensive game. Therefore, for each coalition S which is active at x, a SPNE strategy of S in the game Γ_x^S prescribes a play that is optimal for S from point x onwards, given the optimal strategies of the remaining active players. Thus, a SPNE payoff of a coalition S in the induced game Γ_x^S is a payoff that S can achieve if the game reaches node x.

The subgame-perfect Nash equilibria of the family of extensive games Γ_x^S, $x \in X$ and S an active coalition in Γ_x, determine the payoffs that coalition S can achieve at each decision node x of the game Γ. If the induced game Γ_x^S has more than one SPNE, then a SPNE with highest payoff for coalition S is selected. Choosing the highest SPNE payoff leads to a core concept which is independent of which SPNE may be actually played.

We return to Example 1 to illustrate the additional definitions introduced. Since the game Γ in Example 1 has only two players, $\Gamma_{x_1}^{\{1\}} = \Gamma_{x_1}^{\{2\}} = \Gamma$. The SPNE payoff of coalition $\{1\}$ in the induced game $\Gamma_{x_1}^{\{1\}}$ is 2 and its SPNE strategy is L. Similarly, the SPNE payoff of $\{2\}$ in the induced game $\Gamma_{x_1}^{\{2\}}$ is 1 and its SPNE strategy is rR ($\equiv\ r$ if 1 plays R).

The SPNE payoff of player N in the single player game $\Gamma^N (= \Gamma_{x_1}^N)$ in Fig. 2 is 6 and its SPNE strategy is $(R, lR) (\equiv R; l$ if N plays $R)$. Notice that the SPNE strategy (R, lR) of coalition N is not compatible with the SPNE strategies L and rR of coalitions $\{1\}$ and $\{2\}$, respectively.

Changing coalitional payoffs

The need for defining the payoff of each coalition at each decision node of an extensive game arises from the fact that coalitional payoffs may change as the game unfolds along a history. This important fact can be explained in terms of Example 1.

If players 1 and 2 form a coalition, the payoff of the coalition is 6, as implied by the SPNE of $\Gamma_{x_1}^N$. If coalition $\{1\}$ decides to deviate from the SPNE strategy (R, lR) of N in the beginning of the game, its resulting payoff is 2, as implied by the SPNE of $\Gamma_{x_1}^{\{1\}}$. Similarly, if $\{2\}$ decides to deviate in the beginning of the game, its resulting payoff is 1, as implied by the SPNE of $\Gamma_{x_1}^{\{2\}}$. In sum, the coalitions $\{1\}, \{2\}$, and N, which are active at x_1, can obtain payoffs of 2, 1, and 6, respectively. Thus, none of them can improve upon a payoff vector (p_1, p_2) such that $p_1 \geq 2, p_2 \geq 1, p_1 + p_2 = 6$. E. g., if the payoff vector (3.5, 2.5), then no coalition can obtain a higher payoff by deviating from the grand coalition's strategy (R, lR) in the *beginning* of the game.

Yet, we claim that the strategy profile (R, lR) and the payoff vector (3.5, 2.5) are not a sensible prediction of the game. That is because if the strategy profile (R, lR) is followed, the game would reach node x_2. As a result, the strategy profile (R, lR) and the payoff vector (3.5, 2.5) should also be immune to deviations by all active coalitions at x_2. However, it is not. The only active coalition at x_2 is $\{2\}$ and it can obtain a higher payoff of 3 (> 2.5) by taking action r once the game reaches x_2. Thus, the strategy profile (R, lR) is not immune to deviations by all active coalitions along the history generated by it.

The above analysis of Example 1 demonstrates that the relative bargaining power of coalitions following their SPNE strategies may change as the game unfolds along the history generated by a strategy profile. For instance, coalition $\{2\}$ can obtain a payoff of only 1 by deviating from the SPNE strategy (R, lR) of N at x_1, but a payoff of 3 by deviating at x_2. Despite the fact that coalition $\{2\}$ follows a SPNE strategy in the induced game $\Gamma_{x_1}^{\{2\}}$, this is possible because x_2 is not reached in the history generated by the SPNE of the induced game $\Gamma_{x_1}^{\{2\}}$. In more general terms, it is possible because a SPNE strategy of a coalition (e.g. $\{1,2\}$ in Example 1) is not necessarily a SPNE strategy of a proper subcoalition (e.g. $\{2\}$ in Example 1).

In summary, as Example 1 illustrates, the payoff achievable by a coalition following its SPNE strategies may change as the game unfolds along the history generated by a strategy profile. A core concept that takes account of this fact implies a possibly smaller core.

3. The subgame-perfect core

We need some additional definitions. A payoff vector $(p_1, \ldots, p_n)$ is *feasible* for a strategy profile if $\sum_{i \in N} p_i = u_N(z)$, where z is the terminal node of the history generated by the strategy profile.

By a history *leading* to a payoff vector $(p_1, \dots, p_n)$ we mean a history with a terminal node z such that $u_N(z) = \sum_{i \in N} p_i$. Given an extensive game Γ and the resulting family of the induced (extensive) games Γ_x and Γ_x^S, let $w^\gamma(S; x)$ denote the highest SPNE payoff of coalition S in the game Γ_x^S.[6]

Given the set of SPNE payoffs $w^\gamma(S; x)$, $x \in X$ and S, an active coalition at x, the subgame-perfect core of the extensive game Γ consists of payoff vectors with the property that no coalition can improve upon its payoff by deviating not only at the origin but also at any decision node along the histories leading to the terminal nodes for which the payoff vectors are feasible.[7] It may be noted that the history generated by any strategy profile begins at the origin of game Γ and all coalitions including coalition N are active at the origin.

Definition 1 The subgame-perfect core of an extensive game Γ is the set of payoff vectors $(p_1, \dots, p_n)$ such that $w^\gamma(S; x) \leq \sum_{i \in S} p_i$ for all decision nodes x along the histories leading to the payoff vector $(p_1, \dots, p_n)$ and all coalitions $S \subset N$ that are active at x.

Let $z^* \in Z$ be a terminal node such that $u_N(z^*) \geq u_N(z)$ for all $z \in Z$. Such a terminal node exists if the extensive game Γ is finite or if the strategy sets are compact and the payoff functions are continuous. Definition 1 implies that the subgame-perfect core of the extensive game Γ must be a subset of the set of feasible payoff vectors $(p_1, \dots, p_n)$ such that $\sum_{i \in N} p_i = u_N(z^*)$. That is because the origin of the extensive game Γ is a decision node along every history of the game and coalition N is active at the origin. Thus, $u_N(z^*) = w^\gamma(N; 0)$ and there are no other feasible payoff vectors $(p_1, \dots, p_n)$ such that $w^\gamma(N; 0) \leq \sum_{i \in N} p_i$.

Definition 1 takes into account the possibility that the terminal node at which the total payoff $u_N(z)$ is highest may not be unique and that the payoffs that coalitions can obtain along the nodes of different histories leading to different terminal nodes with highest total payoff may be different. As we will discuss, this implies a concept that may be considered “too strong”. Therefore, we also introduce the concept of the weak subgame-perfect core, which is weaker in

[6] By definition of Γ_x^S, coalition S must be active in the subgame Γ_x.

[7] As Example 1 demonstrates, the SPNE payoffs that a coalition can obtain as the game unfolds along the history generated by a strategy profile may be higher.

the sense that the *weak* subgame-perfect core is non-empty if the subgame-perfect core is, but the converse is not true.

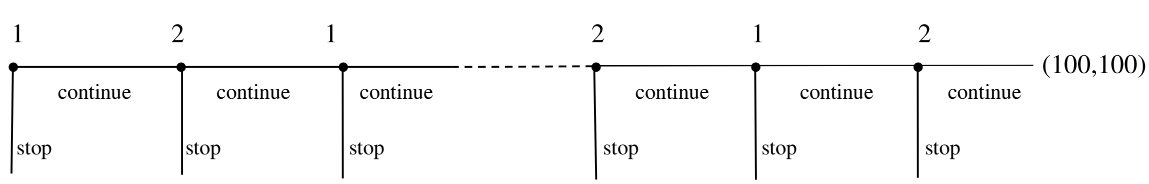
Let $Z^* \subset Z$ be such that if $z^* \in Z^*$, then $u_N(z^*) \geq u_N(z)$ for all $z \in Z$. Let $X(z^*)$ denote the set of decision nodes along the history leading to the terminal node z^*. Let $X^* = \cup X(z^*)$ where the union is taken over all $z^* \in Z^*$.

Definition 1 implies that the subgame-perfect core of an extensive game Γ consists of payoff vectors from the set $\{ (p_1, \dots, p_n): \sum_{i \in N} p_i = u_N(z^*)\}, z^* \in Z^*$, which are immune to deviations by all coalitions that are active at the decision nodes in the set X^*. Since the origin $0 \in X^*$ and all coalitions are active at the origin, the payoff vectors must additionally satisfy at least $w^\gamma(S; 0) \leq \sum_{i \in S} p_i$ for all $S \subset N$. In particular, $w^\gamma(\{i\}; 0) \leq p_i$ for each $i \in N$. Thus, in our model of cooperation, the players are at least as well-off as they would be if the game were played non-cooperatively without any cooperation. In other words, no player stands to lose by cooperating in the manner assumed in this paper. The same property cannot be shown to hold, if the players in the complement of a deviating coalition do not form singletons.

3.1 *The centipede and the infinite bargaining games*

We illustrate the subgame-perfect core and some of its properties by applying it to two well-known examples of extensive games.

Example 2 *The centipede game* (Rosenthal, 1981): There are two players, 1 and 2. The players have 1 dollar each in the beginning of the game. When a player says "continue", 1 dollar is taken by a regulator from her pile and 2 dollars are put in her opponent's pile. As soon as either player says "stop", play is terminated, and each player receives the money currently in her pile. The play *also* stops if both players' piles reach 100 dollars each. The extensive form of this game is depicted in Fig. 3; the first number in each pair is the payoff of player 1 and the second number is the payoff of player 2.



(1,1) (0,3) (2,2) (97,100) (99,99) (98,101)

Fig. 3

This centipede game admits a unique SPNE and the equilibrium strategy for each player is to choose "stop" whenever it is her turn to move. In this SPNE equilibrium the payoffs are $ 1 to each player while $100 each is possible. For this reason the centipede game has often appeared in the debates concerning the SPNE concept. Several experimental studies have demonstrated that SPNE is rarely observed. Instead, players regularly show partial cooperation: playing “continue” for several moves before eventually choosing “stop”. It is rare for players to cooperate through the whole game. For examples see McKelvey and Palfrey (1992) and Nagel and Tang (1998).

In the centipede game in Fig. 3, the total payoff of the grand coalition of 1 and 2 is highest at the last terminal node and equal to 200 dollars. Therefore, the set Z^* is a singleton. Player 2 is active in all subgames and player 1 and the grand coalition of 1 and 2 in all but the last subgame. The payoffs $w^\gamma(\{1\},x)$ of player 1, $w^\gamma(\{2\},x)$ of player 2, and $w^\gamma(\{1,2\},x)$ of the grand coalition of 1 and 2 are well-defined in each subgame in which they are active. The set X^* of decision nodes along the history leading to the unique terminal node at which the payoff of the grand coalition is highest includes all decision nodes of the game. The payoff $w^\gamma(\{1\},x), x \in X^*$, of player 1 is highest in the last subgame in which it is active and equal to 99 dollars. Similarly, the payoff $w^\gamma(\{2\},x), x \in X^*$, of player 2 is highest in the last subgame in which it is active and equal to 101 dollars. These calculations imply that the subgame-perfect core of the game is non-empty and consists of the unique payoff vector (99,101). Unlike the SPNE payoff vector, this payoff vector is “efficient” and not bad for either of the players or the grand coalition.[8]

There is one more important property of the subgame-perfect core worth noting. Consider a modified centipede game which is identical to the original game in Fig. 3 except that the payoffs

[8] However, it may be noted that the subgame-perfect core of a centipede game may be empty and, thus, no cooperation may be possible. In this case a SPNE of the game is a predicted outcome.

at the last terminal node have been replaced with the subgame-perfect core payoffs $(99,101)$. It is easily verified that the subgame-perfect core payoff vector $(99,101)$ is a SPNE outcome of the so-modified game. This means that each subgame-perfect core payoff vector (a cooperative solution) can be supported as a SPNE outcome (a non-cooperative solution) of an extensive game that differs from the original game only in terms of players' payoffs at just one terminal node.[9] We show below (Proposition 1) that this property holds not just for the centipede game but for every extensive game of perfect information. This also leads us to conjecture that the subgame-perfect core of a *general* two-player extensive game is non-empty if it admits a SPNE with an "efficient" outcome. Before we confirm the conjecture (see Proposition 6 below), it is worth checking whether the conjecture holds for another well-known two-player game which admits an SPNE with an efficient outcome.

Example 3 *The two-player infinite bargaining game of alternating offers* (Rubinstein, 1982): Two players, 1 and 2, bargain to split 1 dollar. The rules are as follows: The game, to be denoted by Γ, begins in period 1 in which player 1 makes an offer of a split (a real number between 0 and 1) to player 2, which player 2 either accepts or rejects. Acceptance by player 2 ends the game and the proposed split is immediately implemented. If player 2 rejects, nothing happens until period 2. In period 2, the players' roles are reversed with player 2 making an offer of split to player 1 and player 1 then accepting or rejecting it. The bargaining can potentially go on forever. If that indeed happens, both players get zero. Each player i "discounts" the future using the discount factor $\delta_i \in (0\ 1)$. That is, a dollar received by player i in period t is worth only δ_i^{t-1} in period 1 dollars. Rubinstein (1982) shows that this game admits a unique SPNE, in which

- Player 1 always offers $p^* = (p_1^*, p_2^*)$ and accepts an offer if and only if $q_1 \geq q_1^*$
- Player 2 always offers $q^* = (q_1^*, q_2^*)$ and accepts a proposal p if and only if $p_2 \geq p_2^*$,

where

[9] Given that in the modified centipede game there is no conflict between a player's self-interest and mutual benefit, it would be interesting to conduct experiments to check whether the players would cooperate during the entire modified game.

$$p^* = \left(\frac{1-\delta_2}{1-\delta_1\delta_2}, \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2}\right)$$

$$q^* = \left(\frac{\delta_1(1-\delta_2)}{1-\delta_1\delta_2}, \frac{1-\delta_1}{1-\delta_1\delta_2}\right).$$

This equilibrium strategy profile implies an outcome in which player 1 offers p^* at the start of the game, and player 2 accepts it immediately. Therefore, p^* is the unique SPNE payoff vector and, given that there are only two players, the unique SPNE is also the unique SPNE in both the induced games $\Gamma^{\{1\}}$ and $\Gamma^{\{2\}}$. Hence, $w^\gamma(\{1\};0) = p_1^*, w^\gamma(\{2\};0) = p_2^*$, and $w^\gamma(\{1,2\};0) = 1$. Since each subgame has exactly the same structure as the original game and the future payoffs are discounted, $X^* = \{0\}$,[10] and, therefore, $w^\gamma(\{1\}) = w^\gamma(\{1\};0) = p_1^*, w^\gamma(\{2\}) = w^\gamma(\{2\};0) = p_2^*$, and $w^\gamma(\{1,2\}) = w^\gamma(\{1,2\};0) = 1$. Since $p_1^* + p_2^* = 1$, the subgame-perfect core is non-empty and consists of the unique SPNE payoff vector p^* independently of the values of the discount factors δ_1 and δ_2.

This shows that in Rubinstein's two-player infinite bargaining game, like the modification of the centipede game discussed above, there is no conflict between a player's self-interest and mutual benefit. Furthermore, Binmore, Rubinstein, and Wolinsky (1986) show that if the players are patient, the SPNE payoff vector p^* is also the axiomatic Nash solution of the bilateral bargaining game (Nash, 1950). This means that if the players are patient, the equivalence between the subgame-perfect core and the SPNE for the two-play er infinite bargaining game of alternating offers, as established above, also implies equivalence between the subgame-perfect core and the axiomatic Nash solution.[11]

3.2 *A non-cooperative interpretation of the subgame-perfect core*

We show that each subgame-perfect core payoff vector can be supported as a SPNE payoff vector of a modified game that is identical to the original game except that the individual payoffs

[10] This confirms the point made above that even for games with infinite horizon the set X^* may contain only a finite number of nodes. The set X^*includes all decision node in the case of the finite centipede game, but only one node -- the origin -- in the case of the infinite bargaining game.

[11] That is, $\delta_1, \delta_2 \to 1$ in our notation.

at a terminal node with highest total payoff for the grand coalition have been replaced by the subgame-perfect core payoff vector.

We shall henceforth refer to a SPNE that leads to a terminal node with highest total payoff for the grand coalition as an *efficient* SPNE. Also, we define a *strategic-transform* of an extensive game as follows: (a) If the subgame-perfect core is empty, a strategic transform is identical to the original game except that the players' payoffs at a terminal node with highest total payoff for the grand coalition are replaced by a subgame-perfect core payoff vector; (b) if the subgame-perfect core is empty the strategic–transform is identical to the original game.

Note that with this definition, every extensive game has a strategic-transform. We show that each subgame-perfect core payoff vector of an extensive game is a SPNE payoff vector of a strategic-transform of the extensive game.

Proposition 1 Given an extensive game Γ with a non-empty subgame-perfect core, each subgame-perfect core payoff vector of the game is a SPNE payoff vector of a strategic-transform of the game.

Proof: Let z^* be a terminal node of Γ with highest total payoff for the grand coalition and let $X(z^*) \equiv \{x_1^*, \dots, x_K^*\}$ and $(a_1^*, \dots, a_K^*)$ denote (respectively) the decision nodes and actual actions taken along the history leading to the terminal node z^*. Let i_k^* denote the player who moves and takes action a_k^* at the decision node x_k^*.

Given a subgame-perfect core payoff vector $(p_1^*, \dots, p_n^*)$, let Γ^* denote the strategic-transform of Γ obtained by replacing the individual payoffs at the terminal node z^* with the subgame-perfect core payoff vector $(p_1^*, \dots, p_n^*)$. By definition, $\sum_{i=1}^{n} p_i^* = \sum_{i=1}^{n} u_i(z^*)$.

We prove by backward induction that $(p_1^*, \dots, p_n^*)$ is a SPNE payoff vector in the strategic-transform Γ^*. As above, let $w^\gamma(S, x)$, S active at x, denote the highest SPNE payoff of coalition S in the induced subgame Γ_x^S. We start by determining the optimal actions for moves at the final decision nodes in the game tree. Since play at these nodes involves no further strategic interactions among the players, the determination of optimal behavior at these decision nodes involves a single-person decision problem. At the particular final decision node x_K^*, by definition of a SPNE of Γ, $w^{\gamma}(\{i_K^*\}, x_K^*) \geq u_{i_K^*}(z)$ at all terminal nodes z of subgame $\Gamma_{x_K^*}$. Since $(p_1^*, \dots, p_n^*)$ is a subgame-perfect core payoff vector and x_K^* is a node along the history leading to the terminal node z^*, it follows that $p_{i_K^*}^* \geq w^{\gamma}(\{i_K^*\}, x_K^*) \geq u_{i_K^*}(z)$ at all terminal nodes z of the subgame $\Gamma_{x_K^*}$. Hence, a_K^* is an optimal action of player i_K^* in the subgame $\Gamma_{x_K^*}^*$ of the strategic-transform Γ^*. Consider next the reduced games with origins at next-to-last decision nodes and determine the optimal decisions to be taken there by players who correctly anticipate the actions that will be taken at the final decision nodes. The reduced forms of the subgames $\Gamma_{x_{K-1}^*}^*$ and $\Gamma_{x_{K-1}^*}$, obtained after applying backward induction at the final decision nodes, also involve a single-person decision problem. Since x_{K-1}^* is a decision node of player i_{K-1}^*, by definition of a SPNE of $\Gamma_{x_{K-1}^*}$, we have $w^{\gamma}(\{i_{K-1}^*\}, x_{K-1}^*) \geq u_{i_{K-1}^*}(z)$ at all terminal nodes of the reduced form of the subgame $\Gamma_{x_{K-1}^*}$. Since $(p_1^*, \dots, p_n^*)$ is a subgame-perfect core payoff vector and x_{K-1}^* is a node along the history leading to the terminal node z^*, it follows, by definition of a subgame-perfect core payoff vector, that $p_{i_{K-1}^*}^* \geq w^{\gamma}(\{i_{K-1}^*\}, x_{K-1}^*) \geq u_{i_{K-1}^*}(z)$ at all terminal nodes z of the reduced form of the subgame $\Gamma_{x_{K-1}^*}^*$ and a_{K-1}^* is an optimal action of player i_{K-1}^*. Thus, continuing backward through the game tree will eventually lead to a SPNE of the game Γ^* with payoffs $(p_1^*, \dots, p_n^*)$. ■

Proposition 1 justifies the subgame-perfect core, when non-empty, as a non-cooperative solution concept satisfying the perfection property of the SPNE. Furthermore, from our definition of a strategic-transform of a game, each SPNE of an extensive game with an empty subgame-perfect core is a solution of the extensive game. Then, given that the strategic-transform of a game with an empty subgame-perfect core is identical to the original game, the proposition implies that an extensive game always has a solution and every solution is a SPNE of a strategic transform of the game irrespective of whether the subgame-perfect core of the game is empty. The more "applied" message of the above proposition is that, if the subgame-perfect core is non-empty, then cooperation in an extensive game is credible and the subgame-perfect core determines the set of reasonable predictions. But if the subgame-perfect core is empty, we cannot expect that a cooperative solution will be reached and in this case a SPNE remains the predicted outcome, despite the fact that cooperation is allowed.

3.3 *The subgame-perfect core payoff vectors as credible contracts*

We now interpret subgame-perfect core payoff vectors as credible contracts. A *contract* in an extensive game is an assignment of payoffs to each player at each terminal node of the game. Thus, a contract in an extensive game induces a contract in every subgame. We assume that any coalition can cancel a contract at any point of the game and take actions to maximize its own payoff, and a contract is binding *if and only if* it has not been cancelled.[12] A contract is *credible* in a subgame if no coalition of players who are active in the subgame has incentive to cancel the contract at the beginning of the subgame and take actions to obtain a higher payoff for the coalition. A contract is credible, if it is credible in each subgame. Examples 1 and 2 above further clarify this concept.

In Example 1, consider the contract that assigns payoffs equal to (2,1) at z_1, (3.4,2.6) at z_2, and (1.5, 2.5) at z_3.[13] This contract is not credible, since in the subgame with origin at x_2, player 2 can cancel the contract at the beginning of the subgame and take action r to obtain a payoff of 3 (>2.6). By contrast, if the contract were to assign payoffs equal to (2,1) at $z_1$and (1.5, 2.5) at z_3, but (2.9,3.1) at z_2, then neither player 1, nor player 2, nor the coalition of players 1 and 2 will have incentive to cancel the contract at the beginning of any subgame in which they are active. Clearly, only those contracts that assign payoffs equal to those in a subgame-perfect core payoff vector at terminal node z_2 (the terminal node with highest total payoff for the grand coalition) are credible.

In the case of the centipede game in Fig. 3, the payoffs at the last terminal node under any credible contract must be equal to those in the unique subgame-perfect core payoff vector (99,101). That is so because any other contract that assigns different payoffs will be either cancelled by player 2, who can play "stop" to obtain a payoff of 101 in the subgame with origin

[12] In other words, a contract is binding if and only if it has not been cancelled.

[13] We are intellectually indebted to Philippe Jehiel for counter posing this specific example of a contract which eventually led us to propose and characterize the concept of a credible contract in an extensive game.

at his last decision node, or by player 1, who can play "stop" to obtain a payoff of 99 in the subgame with origin at his last decision node.[14]

Proposition 2 If a contract is credible in an extensive game Γ, then the payoffs assigned to the players at each terminal node with highest payoff for the grand coalition must be equal to those in a subgame-perfect core payoff vector.

Proof: We prove the proposition by contradiction. First, suppose contrary to the assertion that a vector of payoffs $(p_1, \dots, p_n)$ assigned to the players at a terminal node $z^*$with highest total payoff for the grand coalition is not a subgame-perfect core payoff vector. Accordingly, there exists a decision node x along the history leading to z^*, i.e. $x \in X^*$, and a coalition S of players who are active in the subgame with origin at x such that $w^\gamma(S, x) > \sum_{i \in S} p_i$. That, however, contradicts the supposition that the contract is credible because it implies that there is a subgame in which a coalition of players who are active in the subgame can cancel the contract and obtain a higher payoff. ■

Given the focus on the subgame-perfect core in this paper, Proposition 2 does not characterize the entire credible contract. Notice, nonetheless, that a credible contract induces a credible contract in each subgame by definition. It follows from Proposition 2 that the payoffs at some of the other terminal nodes must be equal to those in a payoff vector belonging to the subgame-perfect core of a subgame. However, a credible contract for a proper subgame may never come into effect, except by mistake.

The form of contracting proposed above is relevant for environments in which the terminal node (state) that has been reached in a play of the game is verifiable and a mechanism to enforce terminal node-dependent (i.e. state-dependent) contracts is in place whereas actions taken are not verifiable by the enforcement authorities. Because most contracts in real life are of this type, existing institutions and mechanisms are sufficient for their enforcement.[15] In fact, forward

[14] In operational terms, the unique subgame-perfect core payoff vector (99,101) requires player 1 to transfer 1 dollar to player 2 unless the contract has been cancelled. Despite that, player 1 has no incentive to cancel the contract.

[15] This type of contracts fit well the original definition of a contingent contract in Debreu (1959:100) and further elaborated in Shaffer (1984) who notes that “The only risk faced by those who buy these contracts is the possibility

looking players may write in the beginning of the game a contract for *each* terminal node that is binding *if and only if* a play of the game reaches the terminal node for which the contract is written.[16] Since a terminal node is reached in a play of the game if and only if the players take a certain set of actions (indeed, a unique set of actions), a contract is binding if and only if the players take those actions. Put differently, players (including coalition of players) are free to not fulfil a contract by deviating from the implicitly required actions and prevent the play from reaching the terminal node for which the contract is written and binding.

4. Additional properties of the subgame-perfect core

We now establish some additional properties of the subgame-perfect core. To that end, we may notice that if coalition S is active in the subgame Γ_x, then so is every coalition $S' \subset S$. Therefore, for each $x \in X$, $w^\gamma(S;x)$ satisfies the standard definition of a characteristic function. Accordingly, we interpret the core of the characteristic function game $w^\gamma(S;0)$ as the *core of the extensive game* Γ, and, for each $x \in X$, the core of the characteristic function game $w^\gamma(S;x)$ as the *core of the subgame* Γ_x. We show below that the subgame-perfect core is a refinement of the core of the extensive game Γ.

As an illustration of this property, we may notice that the core of the centipede game in Fig. 3 consists of the set of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 200, p_1 \geq 1$, and $p_2 \geq 1$. In contrast, the subgame-perfect core consists of the unique payoff vector (99,101) in the set. The question remains as to whether the subgame-perfect core of an extensive game, like the core of a strategic game, is also the core of a characteristic function game. To see that, for each $S \subset N$, let

$$w^\gamma(S) = \sup_x w^\gamma(S;x)$$

that a state of the world occurs in which delivery is not promised in the contract; if those circumstances occur under which delivery is promised, then delivery is (correctly) foreseen to occur with certainty."

[16] We can interpret the collection of contracts, one for each terminal node, as a "grand contract" and a contract for a terminal node as a "clause" in the grand contract which may never come into effect unless it is for a terminal node at which grand coalition's payoff is highest.

where the supremum is taken over all nodes $x \in X^*$ at which S is active. We shall refer to the function $w^\gamma(S)$ as the *characteristic function form* of the extensive game Γ.[17]

Proposition 3 The subgame-perfect core of an extensive game Γ is equal to the core of the characteristic function game w^γ.

Proof: Let $(p_1, \dots, p_n)$ be a payoff vector that belongs to the core of the characteristic function game w^γ. Then, $\sum_{i \in N} p_i = w^\gamma(N)$ and $\sum_{i \in S} p_i \geq w^\gamma(S),\ S \subset N$. By definition of the characteristic function w^γ, for each $S \subset N, w^\gamma(S) \geq w^\gamma(S; x)$ at each decision node $x \in X^*$ at which S is active and $w^\gamma(N) = u_N(z^*)$, for all $z^* \in Z^*$. The inequalities imply that for each $S \subset N, \sum_{i \in S} p_i \geq w^\gamma(S) \geq w^\gamma(S; x)$ at each $x \in X^*$ at which S is active and the equalities imply that $\sum_{i \in N} p_i = u_N(z^*)$ for all $z^* \in Z^*$, that is $(p_1, \dots, p_n)$ is a feasible payoff vector for all terminal nodes z^* with the highest payoff for coalition N. Hence, $(p_1, \dots, p_n)$ meets all conditions for a payoff vector to be in the subgame-perfect core of Γ.

Conversely, let $(p_1, \dots, p_n)$ be a payoff vector in the subgame-perfect core of the extensive game Γ, then for each $S \subset N,\ w^\gamma(S; x) \leq \sum_{i \in S} p_i$ at each decision node x along the history generated by any strategy profile for which the payoff vector $(p_1, \dots, p_n)$ is feasible. Since the origin 0 is a decision node of the history generated by any strategy profile and coalition N is active at the origin, $\sum_{i \in N} p_i \geq w^\gamma(N; 0)$. Furthermore, since $(p_1, \dots, p_n)$ is a feasible payoff vector, $\sum_{i \in N} p_i = w^\gamma(N, 0) = w^\gamma(N)$. Accordingly, $\sum_{i \in N} p_i = w^\gamma(N)$ and $(p_1, \dots, p_n)$ is a feasible payoff vector for any history of the game leading to a $z^* \in Z^*$. Therefore, for each $S \subset N,\ w^\gamma(S; x) \leq \sum_{i \in S} p_i$ at each $x \in X^*$. Thus, $\sum_{i \in S} p_i \geq w^\gamma(S)$ for each $S \subset N$, and the payoff vector $(p_1, \dots, p_n)$ is in the core of the characteristic function game w^γ. This proves that the core of the characteristic function game w^γ is equal to the subgame-perfect core of the extensive game Γ. ■

Since the proposition shows that an extensive game can be converted into a characteristic function game, it extends the standard approach for strategic games (Aumann, 1961) to extensive

[17] In many games, including games of infinite horizon as in Example 3, the set X^* may be finite and, thus, the supremum a maximum.

games. This means that the concepts and ideas from the vast literature on characteristic function games can now be applied to games in extensive form. For instance, it is now possible to define and calculate the Shapley value of a centipede game and compare it with the actual outcomes of the game obtained in experiments.

In the following proposition, we interpret the subgame-perfect core in terms of the cores of the family of subgames $\Gamma_x, x \in X^*$ and show that the former is a refinement of the latter. Note that the family of subgames $\Gamma_x, x \in X^*$, includes at least one subgame in which all n players are active, namely, Γ (=Γ_0). But in some games the family of subgames may include many more: For example, in the centipede game the set X^* includes all decision nodes and both players are active in all but one subgame.

Proposition 4 The subgame-perfect core of an extensive game Γ with n players is a subset of the intersection of the cores of subgames in the family Γ_x, $x \in X^*$ with n active players. If all n players are active in all subgames in the family $\Gamma_x, x \in X^*$, then it is equal to the intersection and, therefore, non-empty if and only if the core of every subgame in the family is non-empty.

Proof: Let $(p_1, \dots, p_n)$ be a payoff vector in the subgame-perfect core. Then, for each coalition $S \subset N$, $w^\gamma(S) \le \sum_{i \in S} p_i$ and $w^\gamma(S;x) \le w^\gamma(S)$ for all $x \in X^*$ at which coalition S is active. Therefore, for each $x \in X^*$, $w^\gamma(S;x) \le \sum_{i \in S} p_i$ for all coalitions S which are active at x. Furthermore, if $\Gamma_x, x \in X^*$, is a game with n players, then x is a node in the set X^*at which coalition N is active. Therefore, $w^\gamma(N;x) = w^\gamma(N) = \sum_{i \in N} p_i$. This proves that $(p_1, \dots, p_n)$ belongs to the core of each subgame with n players in the family $\Gamma_x, x \in X^*$. However, if $(p_1, \dots, p_n)$ is a payoff vector in the subgame-perfect core, then, by definition, it must satisfy the constraints $w^\gamma(S;x) \le \sum_{i \in S} p_i$ *also* at nodes $x \in X^*$ at which not all n players are active. Therefore, the set of payoff vectors in the subgame-perfect core may be a strict subset of the intersection of the cores of subgames with n players in the family $\Gamma_x, x \in X^*$, as indeed is the case in Example 1.

If all n players are active in all games in the family Γ_x, $x \in X^*$, then coalition N is active in each Γ_x, $x \in X^*$, and the set of decision nodes along any history generated by any strategy profile that maximizes the payoff of coalition N in Γ_x, $x \in X^*$, is a subset of the set X^*.

Therefore, $w^\gamma(N,x) = w^\gamma(N,0)$ for all $x \in X^*$. This implies that if $(p_1, \dots, p_n)$ belongs to the core Γ_x, then $w^\gamma(N;x) = w^\gamma(N) = \sum_{i \in N} p_i$. Furthermore, if $(p_1, \dots, p_n)$ belongs to the cores of all Γ_x, $x \in X^*$, then for each coalition S which is active at x, $w^\gamma(S;x) \leq \sum_{i \in S} p_i$ for all $x \in X^*$. Given that $w^\gamma(S) = \sup_x w^\gamma(S;x)$, this implies that there is no coalition S for which $w^\gamma(S) > \sum_{i \in S} p_i$ at some decision node $x \in X^*$. This proves that if all n players are active in all games in the family Γ_x, $x \in X^*$, then a payoff vector $(p_1, \dots, p_n)$ that belongs to the intersection of the cores of the games in the family also belongs to the subgame-perfect core. It was shown above that if $(p_1, \dots, p_n)$ belongs to the subgame-perfect core then it also belongs to intersection of all games with n active players in the family Γ_x, $x \in X^*$. ■

On reflection, Proposition 4 really shows that the subgame-perfect core is a subset of the intersection of the *subgame-perfect* cores of subgames in a family and equal to the intersection if all players are active in all subgames in the family.

The game in Example 1 illustrates the first part of Proposition 4. Both players are active in only one subgame and the core of this subgame consists of vectors (p_1, p_2) such that $p_1 + p_2 = 6, p_1 \geq 2$, and $p_2 \geq 1$, but the *subgame-perfect* core is the smaller set $p_1 + p_2 = 6, p_1 \geq 2$, and $p_2 \geq 3$, since in one of the subgames in the family only player 2 is active and can obtain a payoff of 3.

For each coalition $S \subset N$, the functions $w^\gamma(S;x)$, $x \in X^*$, determine the lower bounds on the characteristic function $w^\gamma(S)$. Since these lower bounds may be attained for different coalitions at different nodes x, the characteristic function $w^\gamma(S)$ may not inherit all "characteristics" of the family of functions $w^\gamma(S;x)$ unless the bounds are all attained at the same node $x \in X^*$.[18] Indeed, if the family of subgames Γ_x, $x \in X^*$, includes a game with n players, say Γ_{x^*} such that for each $S \subset N$, $w^\gamma(S) = w^\gamma(S;x^*)$, then the subgame-perfect core and core of the subgame Γ_{x^*} are equal, since their derived characteristic functions are equal. If $x^* \neq 0$, the subgame-perfect core is possibly smaller than the core of Γ. But if $x^* = 0$, then the subgame-perfect core is equal

[18] However, this is not peculiar to the subgame-perfect core of an extensive game. The same is also true for the SPNE of an extensive game which may not have the similar characteristics as the SPNE of proper subgames unless the game has additional structure in place.

to the core of Γ, since no refinement takes place as the game unfolds along the nodes in the set $x \in X^*$. That is indeed so in the case of Rubinstein's bargaining game discussed above, since $X^* = \{0\}$.

4.1 *The weak subgame-perfect* core

To motivate the weaker notion, we introduce first an alternative, but equivalent definition of the subgame-perfect core.

For each $z^* \in Z^*$, let $w^{\gamma}_{z^*}(S) = \max_{x} w^{\gamma}(S; x), S \subset N$, where the maximum is taken over all nodes $x \in X(z^*)$ at which S is active. Because the origin of the game $0 \in X(z^*)$, each coalition S is active at least at some $x \in X(z^*)$, and $w^{\gamma}_{z^*}(N) = u_N(z^*)$. We shall refer to the function $w^{\gamma}_{z^*}(S), S \subset N$, as the characteristic function corresponding to the terminal node z^*, and the core of the characteristic function game $w^{\gamma}_{z^*}$ as the subgame-perfect core corresponding to the terminal node $z^* \in Z^*$.[19] Knowing that $w^{\gamma}(S) = \max_{z^*} w^{\gamma}_{z^*}(S)$ where the maximum is taken over all $z^* \in Z^*$, the subgame-perfect core of an extensive game as proposed in Definition 1 is equal to the intersection of the subgame-perfect cores corresponding to the terminal nodes in the set Z^*.

This equivalence suggests an additional interpretation of the subgame-perfect core, pointedly that it is a refinement of the set of subgame-perfect cores corresponding to the terminal nodes with highest payoff for the grand coalition. However, this refinement, like many others in game theory, though intuitive implies a concept which is in a sense "too strong", since it implies that the subgame-perfect core is empty if the subgame-perfect core corresponding to *any* terminal node with highest payoff for the grand coalition is empty.[20] Since we regard the subgame-perfect

[19] Notice that the subgame-perfect core corresponding to a terminal node $z^* \in Z^*$ has the same properties as the subgame-perfect core for the extensive game. That is because the subgame-perfect core corresponding to a terminal node $z^* \in Z^*$ is the subgame-perfect core of a modified extensive game in which the payoffs of the players at all terminal nodes in Z^*, except z^*, have been reduced by arbitrary small amounts.

[20] Many selection procedures in game theory are motivated by intuitive criteria. However, they can sometimes lead to an empty solution set even though the game has a natural solution, e.g., a strategic game with a unique Nash equilibrium may have no strong Nash equilibrium.

core as the rule for the distribution of gains from coalitional choices, it makes sense to assume that the grand coalition will not choose a strategy that leads to a terminal node for which the corresponding subgame-perfect core is empty. Thus, a weaker concept results if the intersection is restricted only to the non-empty subgame-perfect cores corresponding to terminal nodes. We shall refer to the so-defined weaker notion as the *weak* subgame-perfect core (WSPC) of an extensive game. In most applications there is no difference between the two notions; either the set of terminal nodes Z^* with the highest payoff for the grand coalition is a singleton or the subgame-perfect core corresponding to each terminal node $z^* \in Z^*$ is non-empty. However, the WSPC may be non-empty in some instances in which the subgame-perfect core is not.

The WSPC has the same properties as the subgame-perfect core. In essence, the WSPC is the subgame-perfect core of a modified extensive game in which the payoffs of the players at every terminal node in the set Z^* for which the corresponding subgame-perfect core is empty have been reduced by arbitrary small amounts. Importantly, the WSPC is not really an alternative concept. In fact, it is a complementary concept that differs and can be useful *only* if the subgame-perfect core is empty. Our analysis above would remain unchanged even if we were to use the concept of the WSPC instead.

5. Other concepts of cooperation in extensive games

Aumann (1959) introduces a concept of strong Nash equilibrium for a *strategic game* which allows coaitional deviations. Similarly, Bernheim, Peleg, and Whinston (1987) introduce a concept of a coalition-proof Nash equilibrium for games in both strategic and extensive forms. In this section, we first introduce a concept of strong Nash equilibrium for an *extensive game* and then study how these two alternative concepts of cooperation are related to the subgame-perfect core.[21]

5.1 *The subgame-perfect strong Nash equilibrium*

[21] However, unlike the subgame-perfect core, there are no known sufficient conditions for the existence of the strong and the coalition-proof Nash equilibria. In fact, their existence can be proved in applications only by ad hoc methods.

Given an extensive game Γ, let $\Omega \equiv [\{u_i\}_{i\in N}, \{T_i\}_{i\in N}]$, where T_i is the strategy set of player i, denote its strategic form.[22] Then, $T = T_1 \times \cdots \times T_n$ becomes the set of strategy profiles, $t = (t_1, \dots, t_n) \in T$ is a strategy profile, and $u_i(t_1, \dots, t_n)$ is the payoff function of player i.
Given $t = (t_1, \dots, t_n) \in T$, let $t_S \equiv (t_i)_{i\in S}, t_{-S} \equiv (t_j)_{j\in N\backslash S}$, and $(t_S, t_{-S}) \equiv t = (t_1, \dots, t_n)$. Similarly, let $T_S \equiv \times_{i\in S} T_i$ and $T_{-S} \equiv \times_{i\in N\backslash S} T_i$. Finally, for each $\bar{t}_S \in T_S$, let $\Gamma/\bar{t}_{-S}$ denote the game restricted to the players in S by the strategies $\bar{t}_{-S}$ for the players in $N\backslash S$ and let $\Omega/\bar{t}_{-S} \equiv [\{\bar{u}_i\}_{i\in S}, \{T_i\}_{i\in S}]$, where $\bar{u}_i(t_S) = u_i(t_S, \bar{t}_{-S})$ for all $i \in S$ and $t_S \in T_S$, denote the corresponding restricted strategic game. The strategy set of coalition S in the induced game Γ^S is $T_S \equiv \times_{i\in S} T_i$.

Definition 2 Given an extensive game Γ, a strategy profile $\bar{t} = (\bar{t}_1, \dots, \bar{t}_n) \in T$ is a strong subgame perfect Nash equilibrium in Γ, if $(\bar{t}_S, \bar{t}_{-S}) = \bar{t}$ is a subgame perfect Nash equilibrium in every induced game Γ^S, $S \subset N$.

Unlike the subgame-perfect core, a subgame-perfect Nash equilibrium (SPSNE) in an extensive game requires the *same* strategy $\bar{t}$ to be a SPNE in every induced game Γ^S, $S \subset N$. Clearly, appropriate restrictions of the SPSNE strategy $\bar{t}$, by definition, are SPNE in every subgame of each induced game Γ^S, $S \subset N$, and, therefore, SPSNE in every subgame of Γ.[23] Notice that for each coalition $S \subset N$, if $\bar{t} = (\bar{t}_S, \bar{t}_{-S})$ is a SPSNE in the extensive game Γ, then it is also a SPSNE in every restricted game $\Gamma/\bar{t}_{-S}$. This suggests the following recursive but equivalent definition of SPSNE which, as will be seen below, is sometimes more convenient to use.

Definition 3 (1) In a single player extensive game Γ, $\bar{t} \in T$ is a SPSNE if and only if $\bar{t}$ is a SPNE in Γ. (2) Let $n > 1$ and assume that SPSNE has been defined for extensive games with fewer than n players. For any extensive game Γ with n players, $\bar{t} \in T$ is a SPSNE in Γ if for all proper

---

[22]See Osborne and Rubinstein (1994: 94) for an elegant definition of the strategic form of an extensive game. In terms of the earlier notation, $u_i(t_1, \dots, t_n) \equiv u_i(z)$ where z is the terminal node generated by the strategy profile $(t_1, \dots, t_n)$.

[23] For a game of perfect information Γ, the SPSNE strategy, by definition, induces a SPNE in every subgame of each induced game Γ_x^S, $S \subset N, x \in X$, and, therefore, a SPSNE in every subgame $\Gamma_x, x \in X$.

subsets $S \subset N$, $\bar{t}_S$ is a SPSNE in the restricted game $\Gamma/\bar{t}_{-S}$ and if there does not exist a strategy $t \in T$ such that $\sum_{i\in N} u_i(t) > \sum_{i\in N} u_i(\bar{t})$.

Proposition 5 Let Γ be an extensive game such that each induced game $\Gamma^S, S \subset N$, admits a unique SPNE. Then if Γ admits a SPSNE, the SPSNE is unique and the subgame-perfect core consists of the unique SPSNE payoff vector. But if Γ admits no SPSNE, the subgame-perfect core of Γ may still be non-empty.

Proof: We first prove that if the extensive game Γ admits a SPSNE, then it must be unique. If not, then some induced games Γ^S must admit more than one SPNE, which contradicts our supposition that each induced game Γ^S admits a unique SPNE. Therefore, let $\bar{t} \in T$ denote the unique SPSNE. Then, $\bar{t} = (\bar{t}_S, \bar{t}_{-S})$ is a unique SPNE in every induced game Γ^S, $S \subset N$, and, therefore, it also induces a unique SPNE in every subgame of each induced game Γ^S, $S \subset N$.

By supposition, Γ^N admits a unique SPNE, thus the terminal node with highest payoff for coalition N is unique. Since $\bar{t}$ is the unique SPNE of every induced game $\Gamma^S, S \subset N$, the SPNE of each induced game $\Gamma^S, S \subset N$, generates a history which is identical to the history leading to the terminal node with the highest payoff for coalition N. Let X^* denote the set of nodes along the history leading to the terminal node with the highest payoff for N. Then, for each $x \in X^*$, $w^\gamma(S;x) = w^\gamma(S;0) = \sum_{i\in S} u_i(\bar{t}), S \subset N$, since X^* is the set of nodes along the history generated by the unique SPNE of $\Gamma^S, S \subset N$. Thus, $w^\gamma(S) = w^\gamma(S;0), S \subset N$. Therefore, if $(p_1, \dots, p_n)$ belongs to the subgame-perfect core, then it must satisfy $\sum_{i\in N} p_i = w^\gamma(N) = \sum_{i\in N} u_i(\bar{t})$ and $\sum_{i\in S} p_i \geq w^\gamma(S) = \sum_{i\in S} u_i(\bar{t})$. Consequently, the SPSNE payoff vector $(u_1(\bar{t}), \dots, u_n(\bar{t}))$ is the unique subgame-perfect core payoff vector. This proves the first part of the proposition.

For the second part of the proposition, note that the centipede game in Fig.3 is a game in which every induced game admits a unique SPNE. Given that the unique SPNE of the game is not efficient, it is not a SPSNE. But, as seen, the game admits a non-empty subgame-perfect core. ■

Propositions 3 and 5 together imply that the subgame-perfect core of an extensive game is a weaker concept than SPSNE in the sense that the necessary and sufficient condition for the existence of a non-empty subgame-perfect core is not sufficient for the existence of a SPSNE. The proof for the proposition also illustrates the point made above that the characteristic functions $w^{\gamma}(S)$ and $w^{\gamma}(S;x), x \in X$ and S active at $x,$ may be closely related if the extensive game has additional structure. We note two implications of Proposition 5.

Proposition 6 If a two-player extensive game Γ admits a unique SPNE and the SPNE is efficient, then it admits a non-empty subgame–perfect core.

Proof: The unique SPNE of Γ is also a unique SPNE in both the induced games $\Gamma^{\{1\}}$ and $\Gamma^{\{2\}}$. Furthermore, given that the unique SPNE is efficient, by Definition 3, it is actually a unique SPSNE of Γ. Thus, by Proposition 5, the subgame-perfect core of Γ is non-empty and consists of the unique SPNE payoff vector. ■

As an application of this proposition, we may note that Rubinstein's two-player bargaining game admits a unique SPNE and the SPNE is efficient. Therefore, its subgame-perfect core, by Proposition 6, is non-empty. Incidentally, this implies equivalence between the subgame-perfect core, the Nash bargaining solution (if the players are patient), the SPSNE, and the SPNE of Rubinstein's two-player bargaining game.

5.2 *Subgame-perfect strong and coalition-proof Nash equilibria*

As is well-known, a strong Nash equilibrium of a *strategic* game is also a coalition-proof Nash equilibrium. Therefore, for the concept of SPSNE, introduced and compared with the subgame-perfect core, to qualify as a convincing extension of the strong Nash equilibrium for a strategic game, it must be shown that a SPSNE of an extensive game is also a coalition-proof subgame-perfect Nash equilibrium (SPCPNE) of the extensive game. For that we need to reproduce the definition of a SPCPNE.

Definition 4 (1) In a single player extensive game Γ, $\bar{t} \in T$ is a SPCPNE if and only if $\bar{t}$ is a SPNE of Γ. (2) Let $n > 1$ and assume that SPCPNE has been defined for extensive games with fewer than n players. For any extensive game Γ with n players, $\bar{t} \in T$ is *self-enforcing* if for all proper subsets $S \subset N$, $\bar{t}_S$ is a SPCPNE of the restricted game $\Gamma/\bar{t}_{-S}$. For an extensive game Γ with n players, $\bar{t} \in T$ is a SPCPNE if it is self-enforcing and if there does not exist another self-enforcing strategy $t \in T$ such that $\sum_{i \in N} u_i(t) > \sum_{i \in N} u_i(\bar{t})$. [24]

Observe that if $\bar{t}$ is a SPCPNE in Γ, then $\bar{t}$ is a SPNE in Γ and for every $S \subset N$, $\bar{t}_S$ is a SPCPNE in the restricted extensive game $\Gamma/\bar{t}_{-S}$. We show that, as in strategic games, a SPSNE in an extensive game is a SPCPNE. However, the converse is not true, since the unique SPNE in the centipede game in Fig. 3 is a SPCPNE, but not a SPSNE.

Proposition 7 Every SPSNE of an extensive game Γ is a SPCPNE in Γ.

Proof: The proof is by induction. The proposition is true for games with a single player. Suppose the proposition is true for games with k players $1 \le k < n$. We show that then it is also true for games with $k+1$ players. Let $\bar{t} \in T$ be a SPSNE of a game Γ with $k+1$ players. Then, as noted earlier, $\bar{t}_S$ is a SPSNE of $\Gamma/\bar{t}_{-S}$ for each proper subset S of players, and , by definition, there is no $t \in T$ such that $\sum_{i=1}^{k+1} u_i(t) > \sum_{i=1}^{k+1} u_i(\bar{t})$. Since the proposition is true for games with k players, $\bar{t}_S$ is a SPCPNE of $\Gamma/\bar{t}_{-S}$ for each proper subset S of players. Furthermore, there is no self-enforcing $t \in T$ such that $\sum_{i=1}^{k+1} u_i(t) > \sum_{i=1}^{k+1} u_i(\bar{t})$. Hence, $\bar{t} \in T$ is a SPCPNE of game Γ with $k+1$ players. ■

6. An application to a dynamic game of climate change

For the sake of a clear exposition, we have so far restricted ourselves to applications of the subgame-perfect core to games with two-players. To demonstrate its wider applicability, we now consider an n-player extensive game of imperfect information. For that we model climate change as a finite horizon dynamic game in discrete time in which each country/player chooses its level of economic activity, which generates benefits for the country as well as emissions that add to the existing stock of greenhouse gases (GHG). Any addition to the GHG stock spurs climate change and negatively affects the welfare of all countries in the current and future periods. The

[24] This definition is equivalent to the original definition of SPCPNE in Bernheim, Peleg, and Whinston (1987:10). We believe it to be slightly less cumbersome.

GHG stock evolves over time through additions due to emissions and depletions due to natural decay. We show that if players' payoff functions are linear in the GHG stock, the game admits a non-empty subgame-perfect core.

6.1 *The model*

There are n countries, indexed by $i = 1, \dots, n$. Time is discrete, indexed by $t = 1, \dots, T$, where T is finite but may approach infinity. The variables $x_{it} \geq 0$ and $y_{it} \geq 0$ denote the consumption and production, respectively, of a composite private good of country i in period t; x_{it} and y_{it} may differ because transfers between countries are permitted. The variables $e_{it} \geq 0$ and $z_t \geq 0$ denote, respectively, the amount of GHG emitted by country i and the GHG stock in period t. While x_{it}, y_{it}, and e_{it} are flow variables, z_t is a stock variable which evolves overtime according to the second equation below. The output of the private good and emissions of each country i are related according to the equation $y_{it} = g_i(e_{it})$ where $g_i(e_{it})$ is the benefit function. Each country i suffers damages from climate change and derives utility from the private good consumption in each period t according to the (utility) function $v_i(x_{it}, z_t) = x_{it} - d_i(z_t)$, where x_{it} is the private good consumption and $d_i(z_t)$ is the damage function. Thus, the model is similar to the classical model with one private and one public good and quasi-linear utility functions except that it is dynamic, the endowments of the private good are not exogenously fixed, and the public good is a public bad.

We assume that the benefit function, $g_i(e_{it})$, of each country i is strictly increasing and strictly concave, and the damage function, $d_i(z_t)$ is strictly increasing and convex or linear, i.e., $g_i'(e_{it}) > 0, g_i''(e_{it}) < 0, d_i'(z_t) > 0$, and $d_i''(z_t) \geq 0$. We assume that for all $z \geq 0$ and each country i, there exists an $e^0 > 0$ such that $g_i'(e^0) \leq d_i'(e^0 + z)$ and $\lim_{e_i \to 0} g_i'(e_i) > \sum_{j \in N} d_i'(z)$. This assumption implies that for all levels of the GHG stock z, the marginal benefit of emissions for each country i is smaller (larger) than its own marginal damages for large (small) enough emissions. The assumption ensures that each utility maximizing country i will choose its emissions e_{it} in any period t such that $0 < e_{it} < e^0, t = 1, \dots, T$.

Given an initial GHG stock $z_0 \geq 0$, a time-profile of consumption $(x_{1t}, \dots, x_{nt}; z_t)_{t=1}^T$ is *feasible* if there exists a time-profile of emissions $(e_{1t}, \dots, e_{nt})_{t=1}^T$ such that

$$\sum_{i=1}^{n} x_{it} = \sum_{i=1}^{n} g_i(e_{it}) = \sum_{i=1}^{n} y_{it}$$

$$z_t = (1-\delta)z_{t-1} + \sum_{i=1}^{n} e_{it}\, , t = 1, \dots, T.$$

Here $0 \le \delta < 1$ is the natural rate of decay of the GHG stock. Each feasible consumption time-profile $(x_{1t}, \dots, x_{nt}; z_t)_{t=1}^{T}$ uniquely generates an aggregate utility $\sum_{t=1}^{T} \beta^{t-1} v_i(x_{it}, z_t) = \sum_{t=1}^{T} \beta^{t-1} [x_{it} - d_i(z_t)]$ for each country i where $0 < \beta \le 1$ is the discount factor, assumed to be the same for all countries. In the optimal control literature, the GHG emissions $(e_{it})_{t=1}^{T}, i = 1, \dots, n$, are called *control variables* and the resulting GHG stocks $z_{t-1}, t = 1, \dots, T$, are called the *state variables*. Although the latter are not strategies in the dynamic game introduced below, they are generated by the former and appear in the payoff functions of the countries. In fact, they have the same role as decision nodes in a dynamic game.

6.2 *The dynamic game*

Given an initial stock $z_0 \ge 0$ and time periods $T > 1$, Γ_{z_0}denotes the dynamic game in which

- $N = \{i = 1, 2, \dots, n\}$ is the player set
- $E = E_1 \times E_2 \times \cdots \times E_n$, where $E_i = \{e_i \equiv (e_{it})_{t=1}^{T}: 0 \le e_{it} \le e^0\}$, is the set of all terminal histories,
- $u = (u_1, \dots, u_n)$ is the profile of payoff functions such that for each terminal history $e \equiv (e_1, \dots, e_n) \equiv ((e_{1t})_{t=1}^{T}, \dots, (e_{nt})_{t=1}^{T}) \in E$, $u_i(e) = \sum_{t=1}^{T} \beta^{t-1}[g_i(e_{it}) - d_i\,(z_t)]$, where $z_t = (1-\delta)z_{t-1} + \sum_{j \in N} e_{jt}\, , t = 1, \dots, T.$

We assume that the strategy of each player i is a function $e_i(z_{t-1}), 0 \le e_i(z_{t-1}) \le e^0, t = 1, \dots, T$ and the "statistic" z_{t-1} summarizes the history before the game reaches the state z_{t-1}. Accordingly, the set of all terminal histories after the game reaches a state z_{t-1} is $E_{1t} \times E_{2t} \times \cdots \times E_{nt}$, where $E_{it} = \{(e_{i\tau})_{\tau=t}^{T}: 0 \le e_{i\tau} \le e^0\}$. Thus, for each state z_{t-1}, we denote a subgame of the dynamic game Γ_{z_0} by $\Gamma_{z_{t-1}}, t = 1, \dots T$, in which each player i's strategy is a function $e_i(z_{t-1}), 0 \le e_i(z_{t-1}) \le e^0, t = 1, \dots, T$, and the payoff corresponding to each terminal history $((e_{1\tau})_{\tau=t}^{T}, \dots, (e_{n\tau})_{\tau=t}^{T})$ of the subgame is $\sum_{\tau=t}^{T} \beta^{\tau-1}[g_i(e_{i\tau}) - d_i\,(z_\tau)]$ with $z_\tau =$

$(1-\delta)z_{\tau-1}+\sum_{j\in N} e_{j\tau}$, $\tau = t, \ldots, T$. Notice that each subgame $\Gamma_{z_{t-1}}$depends only on z_{t-1} and not on the history before the game reaches the decision node z_{t-1}. Thus, each subgame $\Gamma_{z_{t-1}}$, $t = 1, \ldots T$, has essentially the same mathematical structure as the original game Γ_{z_0}.

6.3 *The subgame-perfect core*

We first specify the payoffs that a coalition can achieve in each subgame without cooperation of the other countries. To that end, given the dynamic game Γ_{z_0}, let $\Gamma_{z_0}^S, S \subset N$, denote the induced dynamic game in which coalition S acts as one single player. In other words, within the coalition the individual strategies are selected so as to maximize the sum of the total payoffs of its members, given the strategies of the non-members. Similarly, let $\Gamma_{z_{t-1}}^S$, $S \subset N$, denote an induced game of the subgame $\Gamma_{z_{t-1}}$, to be called an induced subgame. Then, we show that each induced subgame admits a SPNE. By definition, a SPNE of Γ_{z_0} is also a SPNE of each induced game $\Gamma_{z_0}^{\{i\}}, i = 1, \ldots, n$.

Let $w^\gamma(S, z_{t-1})$ denote the highest SPNE payoff of coalition S in the induced game $\Gamma_{z_{t-1}}^S$, $S \subset N$. Then, a feasible consumption time-profile $(x_{1t}, \ldots, x_{nt}; z_t)_{t=1}^T$ belongs to the subgame-perfect core of the dynamic game Γ_{z_0}, if $w^\gamma(S, z_{t-1}) \leq \sum_{i\in S}\sum_{\tau=t}^T \beta^{\tau-t}(x_{i\tau} - d_i(z_\tau))$ for each coalition $S \subset N$ and $t = 1, \ldots T$.

Proposition 8 Each induced subgame $\Gamma_{z_{t-1}}^S$, $S \subset N, z_{t-1} \geq 0, t = 1, \ldots, T$, admits a unique SPNE if the benefit functions $g_i, i = 1, \ldots n$ are strictly concave, the damage functions $d_i, i = 1, \ldots, n$ are strictly convex or linear and the third derivatives $g_i''' = d_i''' = 0$, $i = 1, \ldots n$. The unique subgame-perfect Nash equilibrium payoff $w^\gamma(S; z_{t-1}), S \subset N$, is a non-increasing and concave function of z_{t-1}.

Proposition 9 The dynamic game Γ_{z_0} admits a non-empty subgame-perfect core if the benefit functions g_i are strictly concave with $g_i''' = 0$, $i = 1, \ldots n$, and the damage functions $d_i, i = 1, \ldots, n$ are linear.

Proofs for these propositions may be found in Chander (2015: theorems 3 and 4). These propositions show that the concept of subgame-perfect core can be applied also to a general class of extensive games of imperfect information. Since the subgame-perfect core of the dynamic game is non-empty, it follows that climate change can be tackled by cooperation, as indeed a cooperative agreement on climate change has been recently signed by 196 countries-- known as the Paris Agreement. An empty subgame-perfect core would have implied that no cooperation was possible and a SPNE of the dynamic game was the more likely outcome.

7. Concluding remarks

This paper brings together two of the most important solution concepts in game theory: subgame-perfect Nash equilibrium of a non-cooperative game and the core of a cooperative game. A link between the two is apparently missing in the extant literature. It opens the door for applications of the concepts and ideas from cooperative game theory to extensive form games.

Our approach to define coalitional payoffs and subgame-perfect cooperation can be extended to the case in which if a coalition deviates, the remaining players may form one or more non-singleton coalitions. Papers taking this approach include Ray and Vohra (1997) and Maskin (2003). Ray and Vohra address the question of the properties that might be expected of binding agreements. Because the authors address strategic rather than extensive games, subgame perfection plays no role in their framework. Maskin proposes a core concept for partition function form games in which if a coalition deviates, the remaining players form a coalition of their own. Our approach can be used to extend the idea underlying Maskin's core to an extensive game. More specifically, the induced games will now have only two players: If S is the set of *all* active players at a decision node x, then for each $S' \subset S$, the player set of the induced game with origin at x consists of $\{S', S\backslash S'\}$. As in the case of subgame-perfect core, the highest SPNE payoff of this induced game is the highest payoff that coalition S' can obtain in the induced game. Defined thusly, the properties of Maskin's core for an extensive game with perfect information need to be explored further. Unlike the subgame-perfect core, it does not seem to have similar properties and cannot be related to subgame-perfect strong and coalition-proof Nash equilibria.

More generally, our approach can be used to derive a partition function from an extensive game: For each partition of the total player set, consider the induced game in which each coalition in the partition becomes a single player. Then, the worth of a coalition in a partition is equal to its highest SPNE payoff in the game induced by the partition. Since the player set in an extensive game, unlike a strategic game, may become smaller as the game unfolds along a history, the same partition of players may no longer be possible. As a result no refinement may take place and the so-derived partition function may just be equal to the partition function of the strategic form of the extensive game. Therefore, both the subgame-perfect core and Maskin's core of the so-derived partition function may be larger than those proposed in this paper.

Our approach differs from that in Chander (2007) and others in that we consider extensive games and subgame perfection. Our approach rests on two fundamental ideas discussed in the introduction: Coalitions become players and, at the origin of any subgame, only those players who still have decisions to make can become part of a coalition. Possibilities for coalition actions are taken into account through the equilibrium notion – in this, paper, the subgame-perfect core.

References

1. Aumann, R. (1959), "Acceptable points in general cooperative *n*-person games", in *Contributions to the Theory of Games* IV," Princeton University Press, Princeton, N. J.

2. Aumann, R. (1961), "The core of a cooperative game without side payments", *Transactions of the American Mathematical Society*, 98, 539-552.

3. Aumann, R. (1996), A reply to Binmore", *Games and Economic Behavior*, 17, 138-146.

4. Becker, R. A. and S. K. Chakrabarti (1995), "The recursive core", *Econometrica,* 63, 401-423.

5. Bernheim, B. D., B. Peleg, and M. D. Whinston (1987), "Coalition-proof equilibria 1: Concepts", *Journal of Economic Theory*, 42, 1-12.

6. Binmore, K., "A note on backward induction", *Games and Economic Behavior*, 17, 135-137.

7. Binmore, K., A, Rubinstein and A. Wolinsky (1986), "The Nash bargaining solution in economic modelling", *The Rand Journal of Economics*, 17, 176-188.

8. Bondareva, O. N. (1963), "Some applications of linear programming methods to the theory of cooperative games" [in Russian]. Problemy Kibernetiki, 10, 119-139.

9. Chander, P. (2015), “Subgame-perfect agreements in a dynamic game of climate change", Jindal Global University, typescript.

10. Chander, P. and M. Wooders (2010), “Subgame perfect cooperation in an extensive game”, Vanderbilt Working Paper No. 10-W08.

11. Chander, P. (2007), “The gamma-core and coalition formation”, *International Journal of Game Theory,* 2007: 539-556.

12. Compte, O. and P. Jehiel (2010), "The coalitional Nash bargaining solution", *Econometrica*, 78, 1593-1623.

13. Debreu, G. (1959), *Theory of Value*, Wiley: New York.

14. Forges, F., J. -F. Mertens and R. Vohra (2002), “The ex-ante incentive compatible core in the absence of wealth effects”, *Econometrica*, 70, 1865-1892.

15. Gale, D. (1978), “The core of a monetary economy without trust”, *Journal of Economic Theory*, 19, 456-491.

16. Gillies, D.B. (1953), “Discriminatory and bargaining solutions to a class of symmetric n-person games”, in H. W. Kuhn and A.W. Tucker (eds.) *Contributions to Theory of games* 2, pp. 325-342.

17. Habis, H. and P. J-J. Herings (2010), “A note on the weak sequential core of dynamic games”, *International Game Theory Review*, 12, 407–416.

18. Kranich, L., A. Perea, and H. Peters (2005), “Core concepts for dynamic TU games”, *International Game Theory Review* 7: 43–61.

19. Lehrer, E. and M. Scarsini (2013), “On the Core of Dynamic Cooperative Games”, *Dynamic Games and Applications*, 3, 359-373.

20. Maskin, E. (2003), “Bargaining, coalitions and externalities”, Presidential Address to the Econometric Society, Institute for Advanced Study, Princeton.

21. McKelvey, R. and T. Palfrey (1992), An experimental study of the centipede game", *Econmetrica*, 69, 803-836.

22. Nash, J. (1950), “The bargaining problem”, *Econometrica*, 18, 155-162.

23. Osborne, M. and A. Rubinstein (1994), *A Course in Game Theory*, MIT Press, Cambridge, Mass.

24. Pérez-Castrillo, D. (1994), “Cooperative outcomes through non-cooperative games”, *Games and Economic Behavior,* 7, 428-440.

25. Perry, M. and P. J. Reny (1994), “A non-cooperative view of coalition formation and the core”, *Econometrica*, 62, 795-817.

26. Predtetchinski, A., P. J-J. Herings, and A. Perea (2006), “The weak sequential core for two-period economies. *International Journal of Game Theory*, 34, 55-65.

27. Ray, D. and R. Vohra (1997), “Equilibrium Binding Agreements,” *Journal of Economic Theory*, 73, 30-78.

28. Rosenthal, R. (1981), "Games of perfect information, predatory pricing, and the chain-store paradox, *Journal of Economic Theory*, 25, 92-100.

29. Rubinstein, A. (1982), “Perfect equilibrium in a bargaining model”, *Econometrica*, 50, 97-109.

30. Rubinstein, A. (1980), “Strong perfect equilibrium in supergames”, *International Journal of Game Theory*, 9, 1-12.

31. Shaffer, S. (1984), “Production and contingent contracts: comment”, *Journal of Post Keynesian Economics*, 6, 634-636.

32. Selten, R. (1975), "Re-examination of the perfectness concept for equilibrium points in extensive games", *International Journal of Game Theory*, 4, 25-55.

33. Shapley, L. S. (1967), “On balanced sets and cores”, *Naval Research Logistics Quarterly*, 14, 453-460.