Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 187 (2020) 105017

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Subgame-perfect cooperation in an extensive game ☆

Parkash Chander [a,*], Myrna Wooders [b]

[a] *University of Pittsburgh, United States of America*
[b] *Vanderbilt University, United States of America*

Received 10 November 2018; final version received 21 November 2019; accepted 16 February 2020
Available online 5 March 2020

Abstract

We propose a new solution concept for games in extensive form that incorporates both cooperation and subgame perfection. From its definition and properties, the new solution concept, named 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. To further characterize the subgame-perfect core, we show that each subgame-perfect core payoff vector can be implemented as a non-cooperative solution, as it is a subgame-perfect Nash equilibrium payoff vector of an extensive form game that is closely related to the original game. We also motivate and introduce a related concept of subgame-perfect strong Nash equilibrium of an extensive game that is coalition proof.

JEL classification: C71-73; C79

Keywords: Extensive game; Subgame perfection; Coalitional game; Centipede game; Bargaining game; Coalition-proof Nash equilibrium

☆ The authors are indebted to the NSF Grant 422-422-3491 and the Douglas Grey Fund for Research in Economics at Vanderbilt University for financial support. We are thankful to multiple colleagues for their helpful comments, especially Philippe Jehiel, Francoise Forge, and Eko Riyanto. We are also thankful to an Editor and to two anonymous referees for their comments which have helped us greatly improve the paper. We have benefited from seminar presentations at the University of Arizona, the Sorbonne Summer Workshop on Economic Theory, Monash University, University of Paris (Dauphine), Nanyang Technological University, University of Pittsburgh, and the 5th World Congress of the Game Theory Society held in Maastricht.

* Corresponding author.
E-mail addresses: parchander@gmail.com (P. Chander), myrna.wooders@vanderbilt.edu (M. Wooders).
URLs: http://www.parkashchander.com (P. Chander), https://as.vanderbilt.edu/econ/bio/myrna-wooders (M. Wooders).

https://doi.org/10.1016/j.jet.2020.105017
0022-0531/

1. Introduction

We propose and characterize a new solution concept, the *subgame-perfect core* of an *extensive form game* with transferable utility.

Arguably, the most well-known approach to defining the core of a non-cooperative game is given in Aumann (1959), which addresses *strategic games*.[1] Aumann proposes two ways[2] to derive a worth function for coalitions in a strategic game.[3] With a worth function in place, he defines the *core* of the strategic game as the core of the derived coalitional game. In contrast to Aumann, we address *extensive form games*. Deriving a worth function from an extensive form game creates new challenges: the payoff that a coalition can achieve, as will be seen, may change as the game unfolds along a history generated by a strategy profile; each decision node determines a subgame and the payoff that a coalition can achieve may not be the same in all subgames.

As we will show, the subgame-perfect core is a refinement of the core of an extensive game in the same sense as the subgame-perfect Nash equilibrium is a refinement of Nash equilibrium. Thus, the subgame-perfect core is a cooperative analog of the non-cooperative subgame-perfect Nash equilibrium.

A coalition in a subgame can consist only of players who are *active* in the subgame, that is, players who still have decisions to make in the subgame.[4] Additionally, coalition members can coordinate only on those actions that are yet to be taken by them in the subgame. Thus, in any subgame, the past is finished and previous actions taken by the players or coalitions cannot be changed – bygones are bygones. Moreover, when a coalition forms, an induced game is created in which members of the coalition act as one single player and the remaining players act as singleton coalitions. The sequential rationality of subgame perfection dictates that each coalition must act rationally in each subgame. Thus, the subgame-perfect core takes into account interactions of coalitions in a fashion that is analogous to how Nash and subgame-perfect Nash equilibrium (SPNE) take into account interactions of individual players. More precisely, as utilities are assumed to be transferable, a payoff vector belongs to the subgame-perfect core of an extensive game if

- (a) there is a history that leads to a terminal node for which the payoff vector is feasible
- (b) no coalition of active players can improve upon its part of the payoff vector by deviating at any decision node along the history, where the payoff that a (deviating) coalition can achieve at any decision node in the history is equal to the highest SPNE payoff of the coalition in the induced game with origin at the decision node.

The motivation for part (b) of the definition comes from the fact, as illustrated by an example below, that the highest SPNE payoff that a coalition can achieve, may vary and be higher (or lower) as the game unfolds along a history.

To further characterize the subgame-perfect core, we show that each subgame-perfect core payoff vector can be implemented as a non-cooperative solution in that each subgame-perfect

---

[1] Also known as “normal form games”.

[2] The so-called α and β approaches.

[3] Recall that a “worth function” assigns a total payoff to each coalition: the sum of the payoffs to individual members of the coalition.

[4] Since player set of the original game *only* includes players who have decisions to make in the game, we treat a subgame analogously by including *only* those players who have decisions to make in the subgame.

core payoff vector is a SPNE payoff vector of an extensive game obtained by modifying the distribution of payoffs at certain terminal nodes of the original game. One interpretation of this result is that if maximizing social welfare requires the players to cooperate and coordinate their actions, then they can be incentivized to do so.

The paper also opens the way to further research. We note in particular that on the way to proving our main results we drive a coalitional game from an extensive game. Given the coalitional game, other cooperative game-theoretic solutions can be applied to the extensive game, such as the Shapley value, which exists and which assigns to each player a payoff that is equal to her expected marginal contribution to coalitions that include her and is thus viewed as fair. Investigation of such issues, however, is beyond the scope of this paper.

Apart from its theoretical interest, the motivation for the subgame-perfect core comes from the fact that there are many real-world situations in which players may cooperate to obtain higher payoffs compared to their payoffs in a SPNE of the game in the absence of cooperation. One such case is that of climate change formulated as a dynamic game (Chander, 2017). Nearly two hundred countries have signed a forward-looking cooperative agreement on climate change, known as the Paris Agreement. Since the countries are sovereign, no country can be forced to sign an agreement against her wishes. Thus, an agreement must make each country better-off compared to the status quo, i.e. the SPNE of the game, and no country or group of countries (such as developing or developed countries) should have incentive to withdraw from the agreement at any date. As will be shown, the subgame-perfect core payoff vectors in a general extensive game have these properties.

1.1. Relationships to other solution concepts

As a byproduct of the conceptual framework developed in this paper for defining the subgame-perfect core, we propose a related concept of subgame-perfect strong Nash equilibrium (SPSNE) for an *extensive game*.[5] To justify the SPSNE as a convincing extension of the familiar strong Nash equilibrium for *strategic games* (Aumann, 1959), we show that every SPSNE, just as in strategic games, is also a subgame-perfect coalition-proof Nash equilibrium (Bernheim et al., 1987). As an application of the SPSNE, we show that in the two-player infinite bargaining game of alternating offers (Rubinstein, 1982), the unique SPNE is actually a SPSNE and the subgame-perfect core consists of the unique SPNE/SPSNE payoff vector. Furthermore, if the players are patient, not only there is equivalence between the subgame-perfect core and the SPNE/SPSNE payoff vector but also between the subgame-perfect core and the Nash bargaining solution, because, if the players are patient, then, as Binmore et al. (1986) show, the unique SPNE payoff vector is equivalent to the Nash bargaining solution.

Our work introduces a new approach to the melding of coalitional game theoretic solutions with Nash equilibrium, 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 coalitional game and a notion of the core and then propose a non-cooperative procedure to implement the core. Our approach is different. We start with an extensive form game and the notion of subgame perfection and show that each subgame-perfect core payoff vector can be

[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.

implemented as a SPNE payoff vector of an extensive form game that is closely related to the original game.

Our work adds to the literature on core concepts for 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 et al. (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, our paper proposes a core concept for a *general* extensive form game that satisfies subgame perfection and that can be applied to a variety of specific dynamic models.

Our work differs from the interesting literature on the core of sequences of *coalitional* games; see, for example, Kranich et al. (2005), Habis and Herings (2010), and Predtetchinski et al. (2006). In contrast to our paper, these studies start from a coalitional game as the primitive and do not consider subgame perfectness; instead they place rules on admissible deviations. We conjecture that investigation of the subgame-perfect core of the sort of dynamic games considered in these papers would be a fruitful line of research, but beyond the scope of this paper.

1.2. 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 interprets a subgame-perfect core payoff vector as a subgame-perfect Nash equilibrium payoff vector of a closely related extensive form game. Section 4 derives a coalitional game from an extensive game and establishes additional properties and interpretations of the subgame-perfect core. Section 5 motivates and introduces the concept of a SPSNE for a general extensive game. Section 6 discusses an application of the subgame-perfect core to a dynamic game of climate change. Section 7 makes concluding remarks that further address the significance of this work and future directions for research. The proofs for the results in the paper are gathered in the Appendix.

2. The framework and a motivating example

To introduce the subgame-perfect core in simplest terms, we restrict ourselves primarily to finite and perfect information extensive games,[6] but note later in subsection 3.1 that the subgame-perfect core can be defined and applied to a more general class of extensive games.

We denote a finite and perfect information extensive game with transferable utilities by $\Gamma = (N, K, P, u)$ where $N = \{1, \dots, n\}$ is the set of players and K is the game tree with origin at o. Let Z denote the set of terminal nodes of the game tree K and X denote the set of non-terminal nodes, i.e. the 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 $u\colon Z \to R^n$ where $u_i(z)$ denotes the payoff of player i at terminal node z. A history leading to a terminal node z is the path that connects the origin of the game tree K to the terminal node z. For now, we do not state players' strategy sets explicitly.

[6] To avoid any ambiguity, we follow, throughout this paper, the definition of a finite and perfect information game in Osborne and Rubinstein (1994: Chapter 6).

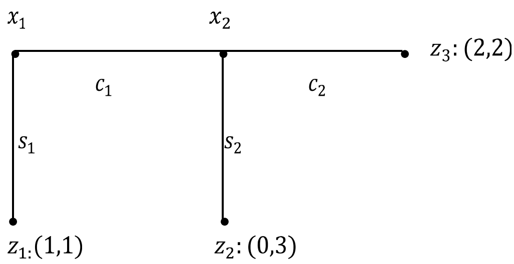

Fig. 1. A centipede game with two players and induced games.

2.1. The induced extensive games

Given a finite and perfect information 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 induced game has $n - s + 1$ players where s is the cardinality of S);
- $K^S = K$, i.e., the game tree is same as the game tree of the original 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 $X_S = \cup_{j \in S} X_j$ is the set of decision nodes of members of coalition 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 $i \in N \setminus S$.

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 fully coordinate their decisions in all subgames. It is worth noting that if S is a singleton coalition, then $\Gamma^S = \Gamma$. Induced games for extensive games that are not finite and perfect information are defined similarly.

2.2. An example

The centipede game (Rosenthal, 1981) has been at the center of the debate concerning the SPNE concept (see e.g. Binmore, 1996 *and* Aumann, 1996).[7] The "inefficiency" of the unique SPNE of this game has been often tested in experiments (McKelvey and Palfrey, 1992). In this game, there are two players, labeled 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 2 dollars each. This game is depicted in Fig. 1.

The node x_1 is the origin of the game tree K, the set $N = \{1, 2\}$ is the set of players, the set $Z = \{z_1, z_2, z_3\}$ is the set of terminal nodes, the set $X = \{x_1, x_2\}$ is the set of decision nodes, the set $P = \{\{x_1\}, \{x_2\}\}$ is the player partition, and the function u; $Z \to R^2$ where $u(z_1) = (1, 1)$, $u(z_2) = (0, 3)$, and $u(z_3) = (2, 2)$ is the payoff function. The game has three terminal histories,

[7] This game has also been used to motivate the extensive form trembling-hand perfect Nash equilibrium (Selten, 1975).

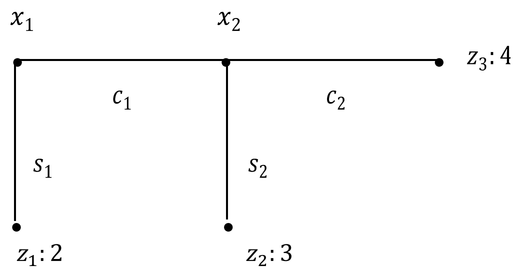

Fig. 2. The game tree when both players form a coalition.

i.e., three paths each connecting the origin x_1 of the game tree K to one of the terminal nodes z_1, z_2, or z_3. Since the game has two players, it has only three induced games: $\Gamma^{\{1\}}$, $\Gamma^{\{2\}}$, and $\Gamma^{\{1,2\}}$. By definition, $\Gamma^{\{1\}} = \Gamma^{\{2\}} = \Gamma$. The induced game $\Gamma^{\{1,2\}}(= \Gamma^N)$ when players 1 and 2 form a coalition to fully coordinate their actions in all subgames, is depicted in Fig. 2. The game tree is the same, but now we have a one-player game with player set $\{N\}$. So $P^N = \{\{x_1, x_2\}\}$, $u_N(z_1) = 2$, $u_N(z_2) = 3$, and $u_N(z_3) = 4$. The game Γ^N also has three terminal histories, which are the same as in the original game Γ.

2.3. *Defining achievable coalitional payoffs*

We define the payoff that each coalition can achieve in each subgame and illustrate the same by means of a centipede game. For this, we first define the *induced* subgame with origin at x for each $x \in X$.

Given a decision node $x \in X$, let Γ_x denote the subgame with origin at x. If the origin of Γ is o, then $\Gamma_o = \Gamma$ and for any $x \neq o$, the game Γ_x is a proper subgame of Γ. It may be noted that the player set of a proper subgame Γ_x can be smaller than the set N. 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 a coalition which is active 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, $\Gamma_o^S = \Gamma^S$ and if Γ is a finite and perfect information game, then so is each induced game $\Gamma_x^S, x \in X$ and S an active coalition in Γ_x. Therefore, if Γ is a finite and perfect information game, then each induced game $\Gamma_x^S, x \in X$ and S an active coalition in Γ_x, admits a SPNE, because, as is well-known, every finite and perfect information game admits a SPNE.

In what follows, it will often be convenient to refer to "a coalition which is active in the subgame with origin at x" simply as "a coalition active at x". It is worth noting that the family of induced games $\Gamma_x^S, x \in X$ and S an active coalition in Γ_x, includes all subgames of the original game Γ, as $\Gamma_x^{\{i\}} = \Gamma_x$ for each i and $x \in X$.

Since a SPNE of an extensive game, by definition, induces a Nash equilibrium in each subgame of the extensive game, a SPNE strategy of coalition S in the induced game Γ_x^S (S active at x), prescribes a play that is optimal for S from node x onwards, given the equilibrium strategies of the remaining active individual players. Thus, a SPNE payoff of coalition S in the induced game Γ_x^S is a payoff that S can achieve, if the game reaches node x, without cooperation of the players outside S.

In summary, the SPNE of the family of extensive games $\Gamma_x^S, x \in X$ and S an active coalition in Γ_x, determine the payoffs that coalition S can achieve at each decision node x of extensive game Γ without cooperation of the remaining players. If the induced game Γ_x^S has more than one SPNE, then a SPNE with highest payoff for coalition S is selected as the achievable payoff of coalition S at x. Selecting the highest SPNE payoff of S in the induced game Γ_x^S as the achievable payoff of S at x leads to a core concept which is independent of which SPNE of the induced game Γ_x^S may be actually played. However, we choose the highest SPNE payoff only for the sake of concreteness and because of lack of refinement procedures that can select a unique SPNE. If there were a refinement procedure to select among the SPNEs, our approach could be applied. As will be clear, the subgame-perfect core defined below, by selecting the highest SPNE payoffs as the achievable payoffs, is a subset of any other core that may be similarly defined by selecting among the SPNE payoffs that are not necessarily equal to the highest SPNE payoffs. Another aspect of our definition of achievable payoffs is that if a coalition deviates, then the remaining players form singletons – they do not form a coalition of their own. We make this assumption because (i) it 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 also the remaining players are assumed to form singletons, and (ii) assuming that the remaining players form a coalition of their own or some other coalition structure is no less arbitrary then assuming that they form singletons.[8] Yet our approach goes beyond that. As we will explain in the concluding section, our approach can also be applied if the remaining players are assumed to form a coalition of their own or some other coalition structure.

To illustrate the additional definitions just introduced, we return to the centipede game in Fig. 1. The SPNE payoff of coalition $\{1\}$ in the induced game $\Gamma_{x_1}^{\{1\}}$ is 1 dollar and its SPNE strategy is s_1. Similarly, the SPNE payoff of coalition $\{2\}$ in the induced game $\Gamma_{x_1}^{\{2\}}$ is 1 dollar and its SPNE strategy is s_2c_1 (= play s_2 if 1 plays c_1). The SPNE payoff of player N in the single player game $\Gamma_{x_1}^N$ ($= \Gamma^N$) is 4 dollars and its SPNE strategy is (c_1, c_2c_1) (= play c_1; play c_2 if N plays c_1). It is noteworthy that the SPNE strategy (c_1, c_2c_1) of coalition N is not the same as the SPNE strategies s_1 and s_2c_1 of coalitions $\{1\}$ and $\{2\}$, respectively. Thus, the history generated by the SPNE strategy of coalition N is not the same as the history generated by the SPNE strategy of either coalition $\{1\}$ or $\{2\}$.

2.3.1. *Varying coalitional payoffs*

The need to define the SPNE payoffs of each active coalition at *each* decision node of an extensive game arises from the fact that the SPNE payoffs may vary and be higher or lower[9] as the game unfolds along a history, as seen, for instance, in the centipede game in Fig. 1.

The SPNE payoff of the grand coalition N is 4 dollars, as implied by the unique SPNE of $\Gamma_{x_1}^N = \Gamma^N$. If coalition $\{1\}$ decides to deviate from the SPNE strategy of N in the beginning of the game, i.e. at x_1, its achievable payoff is 1, as implied by the (unique) SPNE of $\Gamma_{x_1}^{\{1\}}$. Similarly, if $\{2\}$ decides to deviate in the beginning of the game, its achievable payoff is 1, as implied by the SPNE of $\Gamma_{x_1}^{\{2\}}$. In sum, all three coalitions $\{1\}$, $\{2\}$ and N, which are active at x_1, can achieve payoffs of 1, 1, and 4, respectively. But if N follows its SPNE strategy (c_1, c_2c_1) in the game $\Gamma_{x_1}^N$ to achieve its highest payoff of 4, the game would reach the decision node x_2. Coalition $\{2\}$

8 This is the so-called γ-core assumption. Chander (2007; 2018a; 2018b) justifies this assumption in the context of a *strategic* game by showing that, in an intuitive game of coalition formation, forming singletons is a SPNE strategy of the remaining players, i.e. the players in the complement of a deviating coalition indeed have incentives to form singletons. Chander (2017) provides a justification for this assumption in the context of a specific *dynamic* game.

9 However, as will be clear below, the lower payoffs are not binding. Only the higher payoffs matter.

is active in the subgame with origin at x_2 and can achieve a higher payoff of 3 (> 1) by taking action s_2.

The above analysis of the centipede game 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 that may be different from their SPNE strategies. For instance, as noted above, coalition $\{2\}$ can achieve a payoff of only 1 by deviating from the SPNE strategy of N in the subgame with origin at x_1, but a payoff of 3 by deviating in the subgame with origin at x_2. This is possible, despite the fact that coalition $\{2\}$ follows its SPNE strategy in the induced game $\Gamma_{x_1}^{\{2\}}$, because x_2 is not reached in the history generated by the SPNE of the induced game $\Gamma_{x_1}^{\{2\}}$. More generally, this is possible because a SPNE strategy of a coalition (e.g. $\{1, 2\}$) is not necessarily a SPNE strategy of a proper subcoalition (e.g. $\{2\}$).

In summary, as the centipede game in Fig. 1 illustrates, the payoff achievable by a coalition may be higher (or lower) as a game unfolds along the history generated by a strategy profile. This can be seen even more starkly by considering a centipede game with more "legs", as below:

Coalitions $\{1\}$ and N are active in all but the last subgame with origin at x_{18}. The SPNE payoff of N is 20 in all subgames in which it is active. The SPNE payoff of coalition $\{1\}$ in the subgame with origin at x_i is $(i + 1)/1$ if i is odd. Therefore, the SPNE payoff of $\{1\}$ is higher in every successive alternating subgame as the game unfolds along the longest terminal history of the game. Similarly, for coalition $\{2\}$. To conclude, any meaningful core concept for an extensive game must take account of the rising bargaining power of the coalitions as the game unfolds along a history of the game.

3. The subgame-perfect core

We need some additional notions. A payoff vector $(p_1, \ldots, p_n)$ is *feasible* if $\sum_{i \in N} p_i = u_N(z)$ for some terminal node z.[10] We shall refer to a history leading to a terminal node for which a payoff vector $(p_1, \ldots, p_n)$ is feasible as a *history leading to the feasible payoff vector* $(p_1, \ldots, p_n)$. It may be noted that there can be more than one history leading to a feasible payoff vector $(p_1, \ldots, p_n)$. Indeed, this is the case if and only if $u_N(z) = u_N(z')$ for two or more terminal nodes z and z'. Given a finite and perfect information game Γ and the family of induced games Γ_x^S, let $w^\gamma(S; x)$ denote the highest SPNE payoff of coalition S in the induced subgame Γ_x^S (by definition of induced subgame Γ_x^S, coalition S must be active in the subgame Γ_x).

Given the set of coalitional payoffs $w^\gamma(S; x)$, $x \in X$ and S an active coalition at x, the subgame-perfect core of an extensive game is the set of all feasible payoff vectors such that no coalition (including the grand coalition N) can achieve a payoff that is higher than its total payoff in any feasible payoff vector in the set by deviating at any decision node along the histories leading to the payoff vector.

Definition 1. The subgame-perfect core of a finite and perfect information game Γ is the set of all feasible payoff vectors $(p_1, \ldots, p_n)$ such that, for all decision nodes x along any history leading to a payoff vector $(p_1, \ldots, p_n)$ in the set and all coalitions $S \subset N$ that are active at x, the coalitional payoff $w^\gamma(S; x) \le \sum_{i \in S} p_i$.[11]

[10] It may be noted that a feasible payoff vector is not necessarily "efficient".

[11] The symbol "$\subset$" denotes weak inclusion unless stated otherwise.

Since every history of game Γ begins at origin o of Γ and coalition N is active at origin o, Definition 1 implies that the subgame-perfect core of Γ must be a subset of the set of feasible payoff vectors $(p_1, \ldots, p_n)$ such that $\sum_{i \in N} p_i \geq w^\gamma(N; o)$. Let $z^* \in Z$ be a terminal node such that $u_N(z^*) \geq u_N(z)$ for all $z \in Z$.[12] Then, $w^\gamma(N; o) = u_N(z^*)$ and, therefore, $\sum_{i \in N} p_i = w^\gamma(N; o) = u_N(z^*)$, because there is no history leading to a payoff vector $(p_1, \ldots, p_n)$ such that $\sum_{i \in N} p_i > w^\gamma(N; o) = u_N(z^*) \geq u_N(z)$ for all $z \in Z$, but there is a history or histories leading to a payoff vector $(p_1, \ldots, p_n)$ such that $\sum_{i \in N} p_i = w^\gamma(N; o) = u_N(z^*) \geq u_N(z)$ for all $z \in Z$. In summary, the subgame-perfect core of Γ is a subset of the set of feasible payoff vectors $(p_1, \ldots, p_n)$ that are "efficient" in the sense that $\sum_{i \in N} p_i = u_N(z^*) \geq u_N(z)$ for all $z \in Z$.

Definition 1 does not rule out the possibility that the terminal node at which the total payoff $u_N(z)$ of the grand coalition is highest may not be unique. Accordingly, Definition 1 takes into account the fact that the SPNE payoffs that coalitions can achieve at nodes along different histories leading to different terminal nodes with highest payoff for the grand coalition N may be different. It is necessary to include *all* histories leading to terminal nodes with highest payoff for the grand coalition N because each of them is generated by a SPNE of the induced game Γ^N and ignoring any one of them would be arbitrary. But, as we will discuss, including *all* histories leading to terminal nodes with highest payoff for N implies a core concept that may be considered "strong". For now, we note some general properties of the subgame-perfect core as per Definition 1.

Let $Z^* \subset Z$ denote the set of all terminal nodes $z^* \in Z$ such that $u_N(z^*) \geq u_N(z)$ for all $z \in Z$. Let $X(z^*)$ denote the set of decision nodes along the history leading to the terminal node $z^* \in Z^*$. Let $X^* = \cup_{z^* \in Z^*} X(z^*)$.

Definition 1 implies that the subgame-perfect core payoff vectors are efficient and such that no coalition, active at any decision node in the set X*, can achieve a payoff that is higher than its total payoff in any subgame-perfect core payoff vector. Since the origin $o \in X^*$ and every singleton coalition is active at origin o, every subgame-perfect core payoff vector $(p_1, \ldots, p_n)$ must be such that $w^\gamma(\{i\}; o) \leq p_i$ for each singleton coalition $\{i\} \subset N$. By definition, $w^\gamma(\{i\}; o)$ is equal to the SPNE payoff of player i in the induced game $\Gamma^{\{i\}} = \Gamma$. Thus, every subgame-perfect core payoff vector is such that the payoff of each player i is at least as high as the payoff she would obtain if she were to deviate at the origin o and play her individual SPNE strategy. In other words, no player stands to lose by cooperating (rather than non-cooperating) with the other players and obtain her subgame-perfect core payoff. This property of the subgame-perfect core payoffs comes from our assumption that the players in the complement of a deviating coalition form singletons, as then, and only then, the payoff that a singleton coalition can obtain by deviating at the origin is necessarily equal to its SPNE payoff. The payoff of a singleton coalition may not be equal to its SPNE payoff and can be lower if the player in the complement form one or more coalitions and the game exhibits "negative externalities".

3.1. A more general class of games

We have restricted ourselves to finite and perfect information games to enable us to motivate and present our new concept in simplest terms. As every finite game of perfect information is known to admit a SPNE, we can simply presume the existence of a SPNE for each induced game

[12] The existence of such a terminal node is ensured because the extensive game Γ, by assumption, is finite.

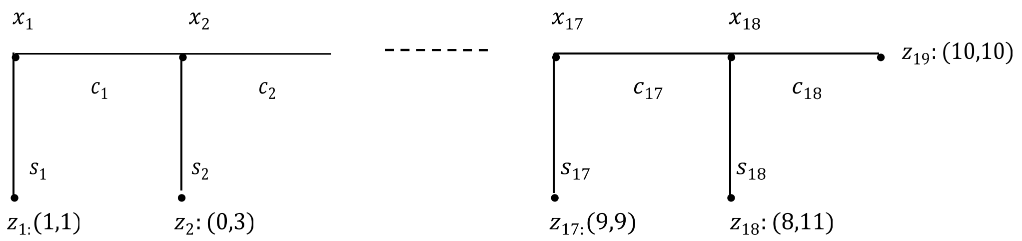

Fig. 3. A centipede game with rising subgame-perfect Nash equilibrium payoffs of coalitions.

Γ_x^S, $x \in X$, and, therefore, of the highest coalitional payoffs $w^\gamma(S; x)$, $x \in X$, which are used in the definition of the subgame-perfect core.

However, the subgame-perfect core can be defined for, and is applicable to, a wider class of games, because, as is well-known, a SPNE may also exist for each induced game that is neither finite nor perfect information. One such class of games, discussed below, are dynamic games of climate change, which are neither finite nor perfect information. Another example is Rubinstein's infinite bargaining game, which is perfect information but not finite.[13] Yet, in both these classes of games each induced game admits a unique SPNE, which guaranties existence of payoffs $w^\gamma(S; x)$, $x \in X$ and S an active coalition at x. Other potential applications of the subgame-perfect core include dynamic models with public goods, dynamic models of oligopoly, the chain store games, and repeated games in which a strategic game is played in each period.

More generally, the subgame-perfect core can be defined for, and is applicable to, any situation that can be modeled as an extensive game in which (i) each induced game Γ_x^S, $x \in X^*$, admits a SPNE, and (ii) either the number of SPNEs is finite or each player's strategy set is compact and its payoff is a continuous function of the strategies. In case of (ii), each achievable coalitional payoff $w^\gamma(S; x)$, $x \in X^*$, can be taken to be equal to the supremum of SPNE payoffs of coalition S in the induced game Γ_x^S, $x \in X^*$.

3.2. An illustration of the subgame perfect core

It is worth illustrating the subgame-perfect core by means of the centipede game in Fig. 3. In this game, the total payoff of coalition N is highest at the last terminal node and equal to 20 dollars. Therefore, the set Z^* of terminal nodes with highest payoff for the grand coalition is a singleton. Player 2 is active in all subgames and player 1 and the grand coalition N are active in all but the last subgame. The payoffs $w^\gamma(\{1\}; x)$ of player 1, $w^\gamma(\{2\}; x)$ of player 2, and $w^\gamma(N; x)$, where $N = \{1, 2\}$, are well-defined for each decision node x at 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 N is highest includes all decision nodes of the game. The payoff $w^\gamma(\{1\}; x)$, $x \in X^*$, of player 1 is highest at the last decision node x_{17} at which it is active and is equal to 9 dollars. Similarly, the payoff $w^\gamma(\{2\}, x)$, $x \in X^*$, of player 2 is highest in the last decision node x_{18} at which it is active and equal to 11 dollars. These achievable payoffs imply that the subgame-perfect core of this game is non-empty and consists of the unique payoff vector (9, 11). This payoff vector is "efficient" and cannot be rejected by any coalition in any subgame in which the coalition is active.

[13] In this game, as will be seen, the set X^* of decision nodes along the histories leading to the highest payoff for the grand coalition is finite, even though the game is not finite and both X and Z include infinitely many nodes.

However, it may be noted that the subgame-perfect core may not *generally* consist of a *unique* payoff vector. In centipede games with non-empty cores, increasing the total payoff at the final or penultimate terminal node will lead to a larger core (i.e., one with larger measure), while decreasing it may lead to an empty core. For instance, the subgame-perfect core consists of infinitely many payoff vectors if the payoffs at the final terminal node z_{19} in the centipede game in Fig. 3 are (15, 15) instead of (10, 10), but is empty if the players' payoffs at the final terminal node are $(9, 9)$ instead of $(10, 10)$.[14]

The centipede game in Fig. 3 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 each even though $10 for each is possible. For this reason, the centipede game has often appeared in the debates concerning the SPNE concept. Several experimental studies concerning the centipede game have demonstrated that the SPNE is rarely observed. Instead, players regularly show partial cooperation: playing "continue" for several moves before eventually choosing "stop". It is rare for the players to cooperate through the whole game. See McKelvey and Palfrey (1992) and Nagel and Tang (1998) among others for experimental evidence in support of this phenomenon.

The subgame-perfect core of the centipede game in Fig. 3 has an additional important property: Consider a modified centipede game which is identical to the original game except that the payoffs at the last terminal node c_{19} have been replaced with the subgame-perfect core payoffs $(9, 11)$. It is easily verified, by backward induction, that the subgame-perfect core payoff vector $(9, 11)$ is a SPNE payoff vector of the so-modified game. This means that each subgame perfect core payoff vector (a cooperative solution) can be supported as a SPNE payoff vector (a non-cooperative solution) of an extensive game that differs from the original game only in terms of players' payoffs at just the terminal node with the highest payoff for the grand coalition.[15] We show (Proposition 1) below that this property holds not just for the centipede game but for every finite and perfect information game. This also leads us to conjecture that the subgame perfect core of a *general* two-player perfect information extensive game is non-empty if the game admits a SPNE with an efficient outcome. We confirm this conjecture as Proposition 6 below.

3.3. A non-cooperative implementation of the subgame-perfect core

We define a *strategic-transform* of an extensive game: (a) If the subgame-perfect core of an extensive game is non-empty, a strategic-transform of the extensive game is a modified extensive game in which the players' payoffs at all terminal nodes with highest total payoff for the grand coalition have been replaced by a subgame-perfect core payoff vector; (b) if the subgame-perfect core is empty, a strategic-transform of the extensive game is same as the original extensive game.

Thus, every finite and perfect information extensive game – whether its subgame perfect core is non-empty or not – has a strategic-transform. We show that if the subgame-perfect core of an

[14] It may also be noted that the subgame-perfect core may become smaller, but not empty, if the payoffs at the final terminal node z_{19} are (14, 14) instead of (15, 15) or $(13, 13)$ instead of $(14, 14)$ and so on. The subgame-perfect core eventually reaches the "tipping point" and becomes empty if the payoffs at the final terminal node z_{19} are reduced below $(10, 10)$, say to $(9.9, 9.9)$.

[15] Given that in the modified centipede game, unlike the original game, there is no conflict between maximization of individual and group payoffs, it would be interesting to conduct experiments to check whether in such centipede games also, which have not been considered in experimental studies so far, the players would cooperate only partially.

extensive game is non-empty, each subgame-perfect core payoff vector is a SPNE payoff vector of a strategic-transform of the extensive game.

Proposition 1. *If the subgame-perfect core of a finite and perfect information game* Γ *is non-empty, each subgame-perfect core payoff vector is a SPNE payoff vector of a strategic-transform* Γ^* *of* Γ.

The proof of Proposition 1 is in the Appendix. As an application to the infinite bargaining game below demonstrates, the game need not necessarily be finite for Proposition 1 to hold.

Proposition 1 justifies the subgame-perfect core as a solution of a non-cooperative game as it shows that the subgame-perfect core payoff vectors (a cooperative solution concept) can be implemented as a SPNE (a non-cooperative solution concept) of a strategic transform of the original game. This means if maximizing social welfare (i.e. the total payoff) requires the players to cooperate and coordinate their actions, then they can be incentivized to take those actions in a self-enforcing manner.

A more "applied" message of Proposition 1 is that, if the subgame-perfect core is non-empty, then credible cooperation is possible and each subgame-perfect core payoff vector is a solution of the game in the sense that it is a SPNE payoff vector of a strategic-transform of the game. But if the subgame-perfect core is empty, then no credible cooperation is possible and a SPNE payoff vector is a solution of the game, as it is a SPNE payoff vector of a strategic-transform of the game. In other words, every finite and perfect information extensive form game has a solution irrespective of whether its subgame-perfect core is empty or non-empty. This solution is efficient if the subgame-perfect core is empty, but may not be efficient if the core is empty.

4. Additional characterization of the subgame-perfect core

Aumann (1959) defines the core of a *strategic game* by converting the strategic game into a coalitional game. In this section, we extend Aumann's approach to *extensive* games by converting an extensive form game into a coalitional game. However, the extension is not straightforward. First, as already noted, dealing with an extensive game creates new challenges as the payoff that a coalition can achieve may not be the same in every subgame. Second, Aumann (1959) assumes that the players in the complement of a deviating coalition max-min or min-max the payoff of the deviating coalition by adopting strategies that are least favorable to the deviating coalition.[16] Whereas our approach is to assume that the players in the complement of a deviating coalition adopt their individually best reply strategies that satisfy subgame-perfection.[17] For each $S \subset N$, let

$$w^\gamma(S) = \sup_{x \in X^*} w^\gamma(S; x),$$

where the supremum is taken only over those decision nodes $x \in X^*$ at which coalition S is active. However, it may be noted that the origin $o \in X^*$ and *every* coalition $S \subset N$ is active at

[16] See Avrachenkov et al. (2013) for an application of Aumann's max-min approach to a dynamic game.

[17] Chander (2007) shows that to max-min or min-max the payoff of a deviating coalition, the players in the complement of the deviating coalition may not follow even their dominant strategies. See Chander (2007; 2018a; 2018b) for a detailed discussion of the difference between our approach and Aumann's max-min or min-max approach in the context of a strategic game.

least at origin o. Therefore, the function w^γ is defined for *every* coalition $S \subset N$. For this reason, we shall refer to the real valued function $w^\gamma(S)$, $S \subset N$, as the *worth function* of the extensive game Γ.

Proposition 2. *Every finite and perfect information extensive game Γ has a coalitional game form, to be denoted by (w^γ, N), and the subgame-perfect core of the extensive game Γ is equal to the core of the coalitional game (w^γ, N).*

The proof of Proposition 2 is in the Appendix. The proposition implies that the subgame-perfect core of an extensive game is non-empty *if and only if* its coalitional game form (w^γ, N) is balanced (Bondareva, 1963 and Shapley, 1967).

Remark. Proposition 2 has important implications as it associates a coalitional game with an extensive game. The literature on extensive games has been growing separately from the literature on coalitional games. Proposition 2 bridges the two literatures and opens the way for application of concepts and ideas from the vast literature on coalitional games to extensive games. To illustrate, every coalitional game, as is well-known, has a Shapley value (Shapley, 1953). Thus, by Proposition 2, every finite and perfect information extensive game has a Shapley value. In other words, every finite and perfect information extensive game has a single-valued solution that maximizes social welfare and is "fair". Another application is that a finite and perfect information extensive game has a non-empty subgame-perfect core if its coalitional game form is convex (Shapley, 1971).

We establish some additional properties of the subgame-perfect core. To this end, notice that if coalition S is active in subgame Γ_x, then so is every coalition $S' \subset S$. Therefore, for each node $x \in X$ and active coalition S in Γ_x, $w^\gamma(S'; x)$, $S' \subset S$, satisfies the standard definition of a worth function of a coalitional game with player set S. Because N is an active coalition at the origin o of Γ, we interpret the core of the coalitional game with worth function $w^\gamma(S; o)$, $S \subset N$, as the *core* (in contrast to the subgame-perfect core, which, by Proposition 2, is the core of the coalitional game with worth function $w^\gamma(S)$, $S \subset N$) of the extensive game Γ. Similarly, for each $x \in X$, we define the core of the coalitional game with worth function $w^\gamma(S'; x)$, $S' \subset S$, where S is the largest coalition that is active in the subgame Γ_x, as the *core* of Γ_x. We show that the *subgame-perfect core* of an extensive game Γ is a refinement of the *core* of Γ. To show this, we interpret the subgame-perfect core of Γ as a refinement of the cores of the family of subgames $\Gamma_x, x \in X^*$.

Notice that the family of subgames $\Gamma_x, x \in X^*$ includes at least one subgame in which all n players are active, namely, the subgame $\Gamma(= \Gamma_o)$. But in some games, this family of subgames may include more than one subgame in which all n players are active: for example, in the centipede game with two players in Fig. 3, the set X* includes all decision nodes and both (i.e., all n) players are active in all but one subgame in the family $\Gamma_x, x \in X^*$. The *core* of the centipede game in Fig. 3 consists of the set of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 20$, $p_1 \geq 1$, and $p_2 \geq 1$, because $w^\gamma(\{1,2\}; x_1) = 20$, $w^\gamma(\{1\}; x_1) = 1$, and $w^\gamma(\{2\}; x_1) = 1$. Whereas the *subgame-perfect core* of this game, as noted in Section 3, consists of the unique payoff vector (9, 11), because $w^\gamma(\{1,2\}) = \sup_{x \in X^* \setminus x_{18}} w^\gamma(\{1,2\}; x) = 20$, $w^\gamma(\{1\}) = \sup_{x \in X^* \setminus x_{18}} w^\gamma(\{1\}; x) = 9$, and $w^\gamma(2) = \sup_{x \in X^*} w^\gamma(\{2\}; x) = 11$, and, therefore, the subgame-perfect core consists of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 20$, $p_1 \geq 9$, and $p_2 \geq 11$. Thus, the subgame-perfect core is a subset of the *core* of the game.

Proposition 3. *The subgame-perfect core of a finite and perfect information game Γ with n players is a subset of the intersection of the cores of those subgames in the family Γ_x, $x \in X^*$, in which all n players are active. If all n players are active in all subgames in the family Γ_x, $x \in X^*$, then it is exactly equal to the intersection.*

The proof of Proposition 3 is in the Appendix. On reflection, Proposition 3 really shows that the subgame-perfect core is a *subset* of the intersection of the cores of subgames in a family of subgames and is exactly *equal* to the intersection if all n players are active in all subgames in the family.[18] Proposition 3 also implies that the subgame-perfect core is non-empty only if the core of every n-player game in the family of subgames Γ_x, $x \in X^*$, is non-empty.

The centipede game Γ in Fig. 3 illustrates the first part of Proposition 3. In this game, $X^* = \{x_1, \dots, x_{18}\}$. Both players are active in all but the subgame $\Gamma_{x_{18}}$ in the family of subgames $\Gamma_x, x \in X^*$. The *core* of subgame Γ_{x_i}, $i \in \{1, \dots, 17\}$, consists of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 20$, $p_1 \geq \frac{i+1}{2}$, and $p_2 \geq \frac{i+1}{2}$, if i is odd, or $p_1 + p_2 = 20$, $p_1 \geq \frac{i-2}{2}$, and $p_2 \geq \frac{i+4}{2}$, if i is even. Therefore, the intersection of the cores of the subgames in which all players are active (i.e. the subgames Γ_{x_i}, $i = 1, \dots, 17$) is the set $p_1 + p_2 = 20$, $p_1 \geq 9$, and $p_2 \geq 10$. Whereas the *subgame-perfect core* of Γ is a subset of this intersection as it consists of the unique payoff vector $(p_1, p_2) = (9, 11)$, because in the subgame $\Gamma_{x_{18}}$ in the family of subgames $\Gamma_x, x \in X^*$, only player 2 is active, and can obtain a SPNE payoff of 11.

Our definition of the worth function of an extensive game does not rule out the possibility that the worth function $w^\gamma(S)$, $S \subset N$, of an extensive game Γ may be equal to the worth function $w^\gamma(S; x)$, $S \subset N$, of some subgame Γ_x, $x \in X^*$, in which all n players are active.[19] Indeed, if the family of subgames Γ_x, $x \in X^*$, includes a subgame in which all n players are active, say Γ_{x^*}, such that for each $S \subset N$, $w^\gamma(S) = w^\gamma(S; x^*)$, then the subgame-perfect core of Γ and the core of the subgame Γ_{x^*} are equal, because their worth functions are equal. If $x^* \neq o$, the subgame-perfect core of Γ could be smaller than the core of Γ (i.e. the core of the coalitional game with worth function $w^\gamma(S; o)$, $S \subset N$). But if $x^* = o$, then no refinement of the core occurs as the game unfolds along the nodes in the set X^* and the *subgame-perfect core* of Γ is equal to the *core* of Γ. This is indeed so in Rubinstein bargaining game we discuss next.

4.1. The two-player infinite bargaining game of alternating offers (Rubinstein, 1982)

Rubinstein 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; 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; player 2 makes an offer of a split to player 1 and player 1 then accepts or rejects. The bargaining can potentially

[18] See Chander (2017) for a n-player game in which all n players are active in all subgames. However, a simple two-player game (with imperfect information) in which both players are active in all subgames is given in Fig. 4 in the Appendix. As Proposition 3 shows, the subgame-perfect core of this game is exactly equal to the intersection of the cores of all three subgames in the family.

[19] If the number of active players is less than n, then the function $w^\gamma(S; x)$ is not defined for all coalitions $S \subset N$, and thus is not comparable to the worth function $w^\gamma(S)$, $S \subset N$.

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 $p_1 \geq p_1^*$,

where

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

This equilibrium strategy profile implies an outcome in which player 1 offers p^* at the start of the game, and player 2 immediately accepts. 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, $w^\gamma(\{1\}; o) = p_1^*$, $w^\gamma(\{2\}; o) = p_2^*$, and $w^\gamma(\{1,2\}; o) = 1$, where o is the origin of the game. Since each subgame has exactly the same structure as the original game and the future payoffs are discounted, $X^* = \{o\}$, and, therefore, $w^\gamma(\{1\}) = w^\gamma(\{1\}; o) = p_1^*$, $w^\gamma(\{2\}) = w^\gamma(\{2\}; o) = p_2^*$, and $w^\gamma(\{1,2\}) = w^\gamma(\{1,2\}; o) = 1$. Since $p_1^* + p_2^* = 1$, the subgame-perfect core is non-empty and consists of the unique SPNE payoff vector p^*, whatever be the values of the discount factors δ_1 and δ_2.

It may be noted that Rubinstein game is not a finite game. Yet we could prove the existence of a non-empty subgame-perfect core because each subgame in the family Γ_x, $x \in X^*$, admits a unique SPNE.

This shows that in Rubinstein two-player infinite bargaining game, like the modified centipede game discussed above, there is no conflict between maximization of individual and group payoffs. Furthermore, Binmore et al. (1986) show that if the players are patient (i.e., $\delta_1, \delta_2 \to 1$ in our notation), the SPNE payoff vector p^* is also the 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-player infinite bargaining game of alternating offers, as established above, also implies equivalence between the subgame-perfect core and the Nash bargaining solution.

4.2. A further characterization of the subgame-perfect core

We now introduce an alternative but equivalent definition of the subgame-perfect core that leads to its further characterization. Recall the definitions of the sets Z^* and $X(z^*)$, $z^* \in Z^*$, in the third paragraph following Definition 1. For each $z^* \in Z^*$ and $S \subset N$, let

$$w_{z^*}^\gamma(S) = \sup_{x \in X(z^*)} w^\gamma(S; x),$$

where the supremum is taken only over those nodes $x \in X(z^*)$ at which S is active. Because the origin of the game $o \in X(z^*)$ for every $z^* \in Z^*$, each coalition S is active at least at one node $x \in X(z^*)$. Moreover, $w_{z^*}^\gamma(N) = u_N(z^*)$. Therefore, for each $z^* \in Z^*$, the function $w_{z^*}^\gamma(S)$, $S \subset N$, is a worth function of a coalitional game. Accordingly, we shall refer to $w_{z^*}^\gamma(S)$, $S \subset N$, as the worth function for the coalitional game *corresponding to the terminal node* z^*. Thus, there are as many coalitional games associated with an extensive game as the number of terminal nodes in the set Z^*. We shall refer to the core of a coalitional game corresponding to a terminal node z^* as the subgame-perfect core corresponding to the terminal node z^*.

Proposition 4. *The subgame-perfect core of a finite and perfect information game* Γ *is equal to the intersection of the cores of all coalitional games with worth functions* $w^{\gamma}_{z^*}$, $z^* \in Z^*$.

The proof of Proposition 4 is in the Appendix. The proposition 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 at which the total payoff of the grand coalition is highest. However, this refinement, like many others in game theory, though intuitive, reveals that the concept of subgame-perfect core is "strong" in the sense that the subgame-perfect core of an extensive game is non-empty *only if* the subgame-perfect core corresponding to each terminal node with highest total payoff for the grand coalition is non-empty.

Since we regard the subgame-perfect core as a 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, we can define a weaker concept of a subgame-perfect core of an extensive game as the intersection of only those subgame-perfect cores corresponding to terminal nodes with highest total payoff for the grand coalition that are non-empty. However, in most applications there is no difference between the subgame-perfect core and the so-defined weaker notion of the subgame-perfect core because 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 *every* terminal node $z^* \in Z^*$ is non-empty. Moreover, the so-defined weaker subgame-perfect core has the same properties as the subgame-perfect core except that it is less likely to be empty. In essence, the weaker subgame-perfect core is the subgame-perfect core of a modified extensive game in which the payoffs of the players at all terminal nodes in the set Z^* for which the corresponding subgame-perfect core is empty have been reduced by arbitrary small amounts. Importantly, the weaker subgame-perfect core is not an alternative, but a complementary concept that differs from the subgame-perfect core and can be a useful concept if the subgame-perfect core is empty. Our analysis so far would remain unchanged if we were to switch to the weaker subgame-perfect core instead.

5. Other concepts of cooperation in extensive games

Aumann (1959) introduces a concept of strong Nash equilibrium for a *strategic game* which allows coalitional deviations. Similarly, Bernheim et al. (1987) introduce a concept of a coalition-proof Nash equilibrium for games in both strategic and extensive forms. Using our approach regarding how coalitions may interact in an extensive game, we now introduce a concept of subgame-perfect strong Nash equilibrium (SPSNE) for an *extensive game* and then study how the SPSNE and the subgame-perfect coalition-proof Nash equilibrium (SPCPNE) à la Bernheim et al. (1987) are related to the subgame-perfect core.

5.1. The subgame-perfect strong Nash equilibrium

As we are familiar with the definition of a SPCPNE in terms of players' strategies, it is convenient to define a SPSNE of an extensive game also in terms of players' strategies. Moreover, unlike the subgame-perfect core, SPSNE and SPCPNE are strategy profiles rather than payoff vectors. Thus, both SPSNE and SPCPNE are to be defined in the space of strategies rather than payoff vectors. Let T_i denote the strategy set of player i in an extensive game.[20] Then, $T = T_1 \times \cdots \times T_n$ is the set of strategy profiles, $t = (t_1, \ldots, t_n) \in T$ is a strategy profile, and $u_1 = (t_1, \ldots, t_n)$ is the payoff function of player i.[21]

Definition 2. Given an extensive game Γ, a strategy profile $\bar{t} = (\bar{t}_1, \ldots, \bar{t}_n) \in T$ is a SPSNE of Γ, if $\bar{t}$ is a SPNE of each induced game Γ^S, $S \subset N$.

Requiring the *same* strategy profile $\bar{t}$ to be a SPNE in *each* induced game Γ^S, $S \subset N$, is the key to the definition of a SPSNE, as the same strategy profile may not generally be a SPNE in every induced game. For example, in the centipede game in Fig. 1, the strategy profile (s_1, s_2c_1) is a SPNE in both induced games $\Gamma^{\{1\}}$ and $\Gamma^{\{2\}}$ but not in the induced game $\Gamma^{\{1,2\}}$. In contrast, the strategy profile (c_1, c_2c_1) is a SPNE in all three (modified) induced games $\Gamma^{*\{1\}}$, $\Gamma^{*\{2\}}$, and $\Gamma^{*\{1,2\}}$ of the strategic transform Γ^*, which is identical to the centipede game in Fig. 1 except that the payoffs at terminal node z_3 are $(1, 3)$ instead of $(2, 2)$. Thus, the strategy profile (c_1, c_2c_1) is a SPSNE of game Γ^* but not of game Γ.

We first note some basic conceptual similarities between the strong Nash equilibrium of a *strategic game* and the SPSNE of an *extensive game*. In strategic games, a strong Nash equilibrium of a game is also a Nash equilibrium of the game. (i) Similarly, in extensive games, a SPSNE of an extensive game is also a SPNE of the extensive game, because a SPSNE of an extensive game Γ, by definition, is a SPNE in every induced game including any induced game $\Gamma^{\{i\}} = \Gamma$, $i \in N$. (ii) If an extensive game is a single-stage (imperfect information) game and equivalent to a strategic game, a SPSNE of the extensive game reduces to a strong Nash equilibrium of the strategic game. In strategic games, a strong Nash equilibrium of a strategic game is efficient. (iii) In extensive games, a SPSNE is efficient, because a SPSNE of an extensive game Γ, by definition, is a SPNE of the induced game Γ^N and, therefore, the SPSNE maximizes the total payoff of the grand coalition N.

A SPSNE satisfies subgame-perfection in the sense that the restriction of a SPSNE of an extensive game to any subgame is a SPSNE of the subgame. This is because a SPSNE of an extensive game Γ, by definition, is a SPNE of every induced game Γ^S and, therefore, its restriction to any subgame Γ_x^S, where S is an active coalition in the subgame Γ_x, is a SPNE in Γ_x^S. Thus, the restriction of a SPSNE to any subgame Γ_x is a SPSNE of Γ_x.

If a two-player extensive game Γ admits an SPNE, say $\bar{t}$, and $\bar{t}$ is efficient, i.e. there exists no strategy profile $t \in T$ such that $\sum_{i \in N} u_i(t) > \sum_{i \in N} u_i(\bar{t})$, then $\bar{t}$ is a SPNE of all three induced games $\Gamma^{\{1\}}$, $\Gamma^{\{2\}}$, and $\Gamma^{\{1,2\}}$ and, therefore, the SPNE is actually a SPSNE. Thus, the unique SPNE in Rubinstein game is actually a SPSNE.

Proposition 5. *Let Γ be a perfect information extensive game such that each induced game Γ^S, $S \subset N$, admits a unique SPNE. Then, if Γ admits a SPSNE, the SPSNE is unique and the*

[20] For perfect information games, a strategy of player i is a function that assigns an action to each decision node of player i. For imperfect information games, it is a function that assigns an action to each information set of player i. See Osborne and Rubinstein (1994: pp. 94 and 203) for a formal definition of a player's strategy in an extensive game.

[21] In terms of earlier notation $u_i(t_1, \ldots, t_n) \equiv u_i(z)$, where z is the terminal node of the history generated by the strategy profile $(t_1, \ldots, t_n)$.

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.*

The proof of Proposition 5 is in the Appendix. Since, as was noted above, a finite extensive game of perfect information admits a non-empty subgame-perfect core if and only if its coalitional game representation is balanced, Proposition 5 implies that in a general class of extensive games the subgame-perfect core is a weaker concept than the SPSNE in the sense that the *necessary and sufficient* condition for the existence of a non-empty subgame-perfect core is not generally sufficient for the existence of a SPSNE. For example, the centipede game in Fig. 3 admits a non-empty subgame-perfect core, but has no SPSNE.

The proof of Proposition 5 also illustrates the point made above that the worth functions $w^{\gamma}(S)$ and $w^{\gamma}(S; x)$, $x \in X^*$, S active at x, may be closely related, and may even be equal for some $x \in X^*$, if the extensive game has additional structure.

Proposition 6. *If a two-player extensive game of perfect information admits a unique SPNE and the SPNE is efficient, the subgame-perfect core is non-empty and consists of the unique SPNE payoff vector.*

The proof of Proposition 6 is in the Appendix.

5.2. The subgame-perfect strong and the coalition-proof Nash equilibria

We now establish an additional conceptual similarity between the strong Nash equilibrium of a strategic game and a SPSNE of an extensive game. 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 above to qualify as a convincing extension of the strong Nash equilibrium concept for strategic games to extensive games, it needs to be shown that a SPSNE of an *extensive game* is indeed a SPCPNE of the extensive game.

We recall the definition of a SPCPNE of an extensive game as in Bernheim et al. (1987); for which we need some additional notation. Given a strategy profile $t = (t_1, \ldots, t_n) \in T$ and a subset $S \subset N$, let $t_S \equiv (t_i)_{i \in S}, t_{-S} \equiv (t_j)_{j \in N \backslash S}$, and $(t_S, t_{-S}) \equiv t$. Let $T_S \equiv \times_{i \in S} T_i$. Then, the strategy set of coalition S in the induced game Γ^S is T_S and the strategy set of each player j in $N \backslash S$ is T_j. Let $\Gamma / \bar{t}_{-S}$ denote the extensive game induced on subgroup S in which the strategy of each player $j \in N \backslash S$ has been fixed at $\bar{t}_j \in T_j$, the number of players in $\Gamma / \bar{t}_{-S}$ is equal to the cardinality of S, the strategy set of a player $i \in S$ is T_i, and the payoff function of each player $i \in S$ is $\bar{u}_i = u_i(t_S, \bar{t}_{-S})$, where $t_S \in T_S$ is a strategy profile of the players in S. By definition of a restricted game $\Gamma / \bar{t}_{-S}$, two straightforward implications are that if $\bar{t} = (\bar{t}_S, \bar{t}_{-S})$ is a SPNE (or SPSNE) of Γ, then for each $S \subset N, \bar{t}_S$ is a SPNE (or SPSNE resp.) of the restricted game $\Gamma / \bar{t}_{-S}$. Finally, we define the *number of stages* of an extensive game as the maximum number of nested proper subgames.

Definition 3. (1) In a single-player, single-stage extensive game Γ, $\bar{t} \in T$ is a SPCPNE if and only if $\bar{t}$ is a SPNE of Γ.

(2) Let (n, m) $\geq (1, 1)$. Assume that SPCPNE has been defined for all extensive games with fewer than n players and m or fewer stages:

(a) For any game Γ with n players and m stages, a strategy profile $\bar{t} \in T$ is *perfectly self-enforcing* if, for every proper subset $S \subset \{1, \ldots, n\}$, $\bar{t}_S$ is a SPCPNE of the game $\Gamma/\bar{t}_{-S}$, and if the restriction of $\bar{t}$ to any proper subgame is a SPCPNE of the subgame.

(b) For an extensive game Γ with n players and m stages, $\bar{t} \in T$ is a SPCPNE if it is perfectly self-enforcing and if there does not exist another perfectly self-enforcing strategy profile $t \in T$ such that $\sum_{i \in N} u_i(t) > \sum_{i \in N} u_i(\bar{t})$.

Proposition 7. *Every SPSNE of a perfect-information extensive game Γ with m stages is a SPCPNE of Γ, but the converse is not true.*

The proof of Proposition 7 is in the Appendix.

6. Additional applications

For the sake of a simple and transparent exposition of our solution concept for extensive games, we have restricted ourselves to *finite and perfect information* games. However, as already noted in section 3.1, the concept of subgame-perfect core is of wider applicability. Here we briefly discuss an application of the subgame-perfect core to a specific n-player dynamic game that is neither finite nor a perfect information game. More specifically, Chander (2017) formulates climate change as a dynamic game and shows that the subgame-perfect core of the dynamic game is non-empty. Non-emptiness of the subgame-perfect core means that climate change can be tackled by cooperation, as a cooperative agreement to tackle climate change has indeed been negotiated and signed by close to 200 countries. An empty subgame-perfect core would have implied that no stable cooperation to tackle climate change was possible and the status quo, which is best described by a SPNE of the dynamic game of climate change, cannot be improved upon, unless the countries were willing to agree to some fairness criteria such as the one implicit in the Shapley value.

As Chander (2017) notes, the dynamic game of climate change is similar to the dynamic game with a public good in Marx and Matthews (2000), and the subgame-perfect core of the dynamic game of climate change is comparable to the efficient Bayesian equilibria sustained by trigger strategies of the dynamic game with a public good. The dynamic game of climate change is also related to the dynamic game in Harstad (2012), who, unlike our approach, assumes that countries can write contracts that commit them to a profile of strategies of the dynamic game.

7. Concluding remarks

This paper brings together two of the most important solution concepts in game theory: the subgame-perfect Nash equilibrium of an extensive game and the core of a coalitional game. A link between the two has been apparently missing in the literature. Thus, the paper opens the door for further applications of concepts and solutions for coalitional games 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 are assumed to 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 these authors address strategic rather than extensive games, subgame perfection plays no role in their framework. Maskin (2003) proposes a core concept for partition function games in which if a coalition deviates, the remaining players form a coalition of their own. Our approach can be used to extend the basic 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 subgame Γ_x^S consists of $\{S', S \backslash S'\}$. As in the case of subgame-perfect core, the highest SPNE payoff of this induced subgame is the highest payoff that coalition S' can obtain in the induced game.[22] Defined thusly, the properties of Maskin's core for an extensive game with perfect information can be explored further. However, it is clear from the definition itself that Maskin's core payoff vectors, unlike the subgame-perfect core payoff vectors, cannot be related to subgame-perfect strong and coalition-proof Nash equilibria and are not in general Pareto improvements over the status quo that is best described by an SPNE.

Our analysis shows that extensive form games can be divided into three broad categories: (a) those in which the subgame-perfect core is empty, (b) those in which the subgame-perfect core is non-empty, but no subgame-perfect core payoff vector is a SPNE payoff vector, and (c) those in which the subgame-perfect core is non-empty and a subgame-perfect core payoff vector is also a SPNE payoff vector. In category (a) games, no stable cooperation is possible and no SPNE can be Pareto improved upon by cooperation. In category (b) games, cooperation is possible and a SPNE can be Pareto improved upon by cooperation, but there could be a conflict between individual and group incentives. In category (c) games, not only cooperation is possible, but also there is no conflict between individual and group incentives and every subgame-perfect core payoff vector is also a SPNE payoff vector. It was also shown that every category (b) game can be transformed into a category (c) game by suitably changing the distribution of payoffs at a terminal node with highest payoff for the grand coalition.

Finally, every finite and perfect information extensive game, as shown, has a coalitional game form and, therefore, a Shapley value. Thus, every finite and perfect information extensive game has a solution that is efficient and fair in the sense of Shapley (1953). It would be interesting to do separate experiments on all three categories of centipede games and compare their Shapley values with the outcomes obtained in the experiments and the degree of cooperation observed in extensive games in each category.

We showed that every subgame-perfect core payoff vector can be implemented as a non-cooperative solution of an extensive form game that is closely related to the original extensive form game. In other words, we showed how incentives can be designed such that the players will choose those actions that maximize social welfare. The subgame-perfect core payoff vectors can also be interpreted as state dependent contracts that are binding if an only if a pre-specified state (i.e. a terminal node) occurs. This type of contracts may be necessary in environments in which actions of the players cannot be observed, but the state that may occur can be observed and verified.

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 any subgame induced by the partition.

[22] The α- or β-core of an extensive game can be defined similarly by assuming that in each induced subgame the complementary coalition of a deviating coalition of active players chooses strategies that max-min or min-max the payoff of the deviating coalition. However, the max-min or min-max strategies may not be SPNE strategies and, thus, not credible.

Our approach differs from that in Chander (2007; 2018a) and others in that we consider extensive form 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 coalitional actions are considered through the equilibrium notion – the subgame-perfect Nash equilibrium.

Appendix A

Proof of Proposition 1. We prove the proposition for the case in which the terminal node with the highest payoff for the grand coalition is unique and then note that the proof also holds for the case when it is not. Let z^* denote the terminal node of Γ with highest total payoff for the grand coalition and let $(x_1^*, \ldots, x_K^*)$ denote the decision nodes – in descending order – along the history leading to the terminal node z^*. Thus, x_1^* is the origin of Γ, x_2^* is the immediate successor of x_1^*, and x_K^* is the immediate predecessor of z^*. Let a_k^* denote the actual action taken at decision node x_k^* and i_k^* be the player required (by the grand coalition) to move and take action a_k^* at decision node x_k^*, $k = 1, \ldots, K$.

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

We prove by backward induction that $(p_1^*, \ldots, p_n^*)$ is a SPNE payoff vector in the strategic-transform Γ^*. We start with subgames $\Gamma^*_{x_K^*}$ and $\Gamma_{x_K^*}$. Since $(p_1^*, \ldots, p_n^*)$ is a subgame perfect core payoff vector and x_K^* is a node along the history leading to the terminal node z^*, it follows, by definition of a subgame perfect core payoff vector, that $p^*_{i_K^*} \geq u_{i_K^*}(z)$ at all terminal nodes z of the subgame $\Gamma_{x_K^*}$. Because players' payoffs at all but one terminal node of the subgame $\Gamma^*_{x_K^*}$ are the same as in the subgame $\Gamma_{x_K^*}$ and are different and equal to $(p_1^*, \ldots, p_n^*)$ only at the terminal node z^* with $p^*_{i_K^*} \geq u_{i_K^*}(z)$ at all terminal nodes z of $\Gamma_{x_K^*}$, action a_K^* is optimal for player i_K^* also in $\Gamma^*_{x_K^*}$ and players' SPNE payoffs in $\Gamma^*_{x_K^*}$ are equal to $(p_1^*, \ldots, p_n^*)$. Consider next the subgame $\Gamma^*_{x_{K-1}^*}$ and its subgames Γ^*_x. Then each subgame Γ^*_x is identical to subgame Γ_x if $x \neq x_K^*$ is an immediate successor of x_{K-1}^*. Thus, the SPNE and players' SPNE payoffs in these subgames are equal. The reduced forms of subgames $\Gamma^*_{x_{K-1}^*}$ and $\Gamma_{x_{K-1}^*}$ obtained by applying backward induction to each subgame Γ^*_x/Γ_x where $x \neq x_K^*$ is an immediate successor of x_{K-1}^*, are single-player decision problems of player i_{K-1}^*. 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 $(p_1^*, \ldots, p_n^*)$ 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 $\Gamma_{x_{K-1}^*}$ Since players' payoffs at all but one terminal node of the reduced form of $\Gamma^*_{x_{K-1}^*}$ are the same as in the reduced form of $\Gamma_{x_{K-1}^*}$ and are different and equal to $(p_1^*, \ldots, p_n^*)$ only at one of the terminal nodes of the reduced form of $\Gamma^*_{x_{K-1}^*}$, it follows that $p^*_{i_{K-1}^*} \geq u_{i_{K-1}^*}(z)$ at all terminal nodes z of the reduced form of the subgame $\Gamma^*_{x_{K-1}^*}$. Therefore, $a^*_{i_{K-1}^*}$ is an optimal action of player i_{K-1}^* in the reduced form of $\Gamma^*_{x_{K-1}^*}$ and players' SPNE payoffs in $\Gamma^*_{x_{K-1}^*}$ are equal to $(p_1^*, \ldots, p_n^*)$.

By continued application of backward induction, we eventually obtain a SPNE of the subgame $\Gamma^*_{x_2^*}$ with SPNE payoffs equal to $(p_1^*, \ldots, p_n^*)$. Similarly, by backward induction, we solve each subgame Γ^*_x/Γ_x where $x \neq x_2^*$ is an immediate successor of x_1^*; and obtain a reduced form of the strategic transform Γ^* that is single-player decision problem of i_1^*. In this reduced form of Γ^*, the payoffs of the players at one of the terminal nodes are equal to $(p_1^*, \ldots, p_n^*)$ and at all other terminal nodes they are the same as their SPNE payoffs in the subgames Γ^*_x/Γ_x where $x \neq x_2^*$ is an immediate successor of x_1^*. Since x_2^* is a decision node along the history leading the terminal node z^*, we have, as above, $p^*_{i_2^*} \geq w^\gamma(\{i_2^*\}, x_2^*) \geq u_{i_2^*}(z)$ at all terminal nodes z of the reduced form of the subgame $\Gamma^*_{x_2^*}$. Therefore, $a^*_{i_1^*}$ is an optimal action of player i_1^* in the reduced form of Γ^* and players' SPNE payoffs in Γ^* are equal to $(p_1^*, \ldots, p_n^*)$.

Finally, if the terminal node with highest payoff for the grand coalition is not unique, we can modify the original game by replacing players' payoffs at every terminal node with highest payoff for the grand coalition with a subgame-perfect core payoff vector. By applying backward induction in the same way as in the above part of the proof, it is seen that the subgame-perfect core payoff vector is a SPNE payoff vector in the modified game. □

Proof of Proposition 2. Let $(p_1, \ldots, p_n)$ 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* above 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* above imply that $\sum_{i \in N} p_i = u_N(z^*)$ for all $z^* \in Z^*$ that is $(p_1, \ldots, p_n)$ is a feasible payoff vector for all terminal nodes z^* with the highest payoff for coalition N. Hence, $(p_1, \ldots, p_n)$ meets all conditions for a payoff vector to be in the subgame-perfect core of Γ.

Conversely, let $(p_1, \ldots, p_n)$ 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 $(p_1, \ldots, p_n)$ 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 $(p_1, \ldots, p_n)$ 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 $(p_1, \ldots, p_n)$ 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 $(p_1, \ldots, p_n)$ 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. □

Proof of Proposition 3. Let $(p_1, \ldots, p_n)$ be a payoff vector in the subgame-perfect core. Then, for each coalition $S \subset N$, $w^\gamma(S) \leq \sum_{i \in S} p_i$ and $w^\gamma(S; x) \leq w^\gamma(S)$ for all $x \in X^*$ at which coalition S is active. Therefore, for each $x \in X^*$, $w^\gamma(S; x) \leq \sum_{i \in S} p_i$ 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 $(p_1, \ldots, p_n)$ belongs to the core of each subgame with n players in the family $\Gamma_x, x \in X^*$. However, if $(p_1, \ldots, p_n)$ is a payoff vector in the subgame-perfect core, then, by definition, it must satisfy the constraints $w^\gamma(S; x) \leq \sum_{i \in S} p_i$ *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 $\Gamma_x, x \in X^*$, then coalition N is active in each $\Gamma_x, x \in X^*$ and the set of decision nodes along any history generated by any strategy profile

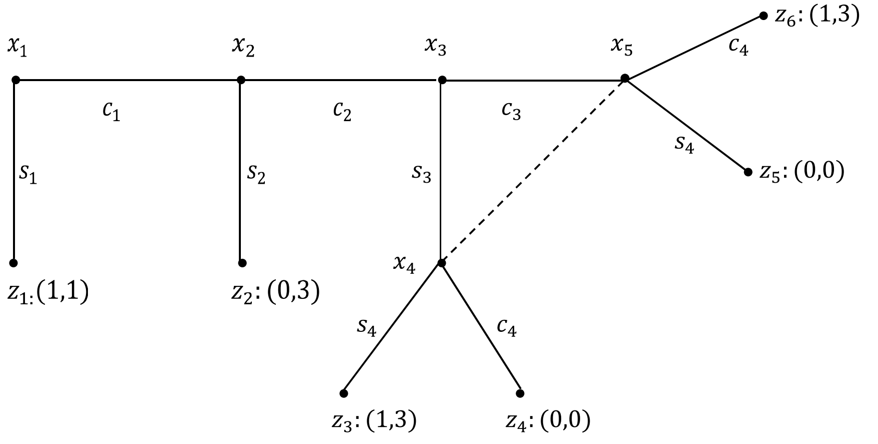

Fig. 4. A game in which all players are active in all subgames.

that maximizes the payoff of coalition N in $\Gamma_x, x \in X^*$, 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 $(p_1, \ldots, p_n)$ belongs to the core Γ_x, then $w^\gamma(N; x) = w^\gamma(N) = \sum_{i \in N} p_i$. Furthermore, if $(p_1, \ldots, p_n)$ belongs to the cores of all $\Gamma_x, x \in X^*$, 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 $\Gamma_x, x \in X^*$ then a payoff vector $(p_1, \ldots, p_n)$ 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 $(p_1, \ldots, p_n)$ belongs to the subgame-perfect core then it also belongs to intersection of all games with n active players in the family $\Gamma_x, x \in X^*$. □

A simple example of a game in which all players are active in all subgames: The two-player extensive game Γ in Fig. 4 below has three subgames: $\Gamma_{x_1}, \Gamma_{x_2}$, and Γ_{x_3}. Both players are active in all three subgames. The core of subgame Γ_{x_1}consists of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 4, p_1 \geq 1, p_2 \geq 1$, because $w^\gamma(x_1; \{1\}) = w^\gamma(x_1; \{2\}) = 1$ and $w^\gamma(x_1; \{1, 2\}) = 4$. The core of subgame Γ_{x_2} consists of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 4, p_1 \geq 0, p_2 \geq 3$, because $w^\gamma(x_2; \{1\}) = 0, w^\gamma(x_2; \{2\}) = 3$ and $w^\gamma(x_2; \{1, 2\}) = 4$. The core of subgame Γ_{x_3} consists of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 4, p_1 \geq 1, p_2 \geq 3$, because $w^\gamma(x_3; \{1\}) = 1, w^\gamma(x_3; \{2\}) = 3$ and $w^\gamma(x_3; \{1, 2\}) = 4$. Thus, the intersection of the cores of the three subgames is the set of payoff vectors (p_1, p_2) such that $p_1 + p_2 = 4, p_1 \geq 1, p_2 \geq 3$, and the subgame-perfect core of Γ is exactly equal to the intersection, because $w^\gamma(\{1\}) = 1, w^\gamma(\{2\}) = 3$ and $w^\gamma(\{1, 2\}) = 4$.

Proof of Proposition 4. By definition, $X^* = \cup X(z^*)$ where the union is taken over all $z^* \in Z^*$. Thus, by definitions of the worth functions w^γ and $w^\gamma_{z^*}$, we have for each coalition $S \subset N$, $w^\gamma(S) = \sup_{z^*} w^\gamma_{z^*}(S)$ where the supremum is taken over all $z^* \in Z^*$. Let $(p_1, \ldots, p_n)$ be a subgame-perfect core payoff vector of an extensive game Γ. Then, by Proposition 2, $(p_1, \ldots, p_n)$ belongs to the core of the coalitional game with worth function w^γ. Thus, for each $S \subset N$, we have $\sum_{i \in S} p_i \geq w^\gamma(S) = \sup_{z^*} w^\gamma_{z^*}(S)$ where the supremum is taken over all $z^* \in Z^*$. This implies, for each $S \subset N$, we have $\sum_{i \in S} p_i \geq w^{\gamma}_{z^*}(S)$ for each $z^* \in Z^*$. Thus, $(p_1, \ldots, p_n)$ belongs to the core of every coalitional game with worth function $w^{\gamma}_{z^*}$, $z^* \in Z^*$. Conversely, let $(p_1, \ldots, p_n)$ be a core payoff vector for every coalition game $w^{\gamma}_{z^*}, z^* \in Z^*$. Then, for each $S \subset N$, we have $\sum_{i \in S} p_i \geq w^{\gamma}_{z^*}(S)$ for every $z^* \in Z^*$. This means, for each $S \subset N$, we have $\sum_{i \in S} p_i \geq \sup_{z^*} w^{\gamma}_{z^*}(S) = w^{\gamma}(S)$. In view of Proposition 2, this implies that $(p_1, \ldots, p_n)$ is a subgame-perfect core payoff vector of Γ. Thus, the subgame perfect core of an extensive game is equal to the intersection of the subgame-perfect cores corresponding to the terminal nodes in the set Z*. □

Proof of Proposition 5. We first claim that if the extensive game Γ admits a SPSNE, then it must be unique. Because, if not, then at least one induced games Γ^S admits more than one SPNE (because a SPSNE, by definition, is a SPNE of each induced game Γ^S) which contradicts our supposition that each induced game Γ^S admits a unique SPNE. Let $\bar{t} \in T$ denote the unique SPSNE.

As hypothesized, Γ^N admits a unique SPNE. Thus, the terminal node with highest payoff for coalition N is unique. Since each induced game, by hypothesis, admits a unique SPNE and $\bar{t}$ is the unique SPSNE, $\bar{t}$ is the unique SPNE of every induced game Γ^S, $S \subset N$. Therefore, the history generated by the unique SPNE $\bar{t}$ of each induced game Γ^S, $S \subset N$ is identical to the history leading to the unique terminal node with highest payoff for coalition N. Let X* denote the set of nodes along the history leading to the unique terminal node with highest payoff for N. Then, $o \in X^*$ and for each $x \in X^*$, $w^{\gamma}(S; x) = w^{\gamma}(S; o) = \sum_{i \in S} u_i(\bar{t})$, $S \subset N$, because X^* is the set of nodes along the history generated by the unique SPNE $\bar{t}$ of each Γ^S, $S \subset N$, and, therefore, the restriction of $\bar{t}$ to each induced subgame Γ^S_x, $x \in X^*$ is a unique SPNE of the subgame. Thus, $w^{\gamma}(S) = w^{\gamma}(S; o)$, $S \subset N$, and if $(p_1, \ldots, p_n)$ 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 $\sum_{i \in S} p_i \geq w^{\gamma}(S) = \sum_{i \in S} u_i(\bar{t})$. Hence, the unique SPSNE payoff vector $(u_i(\bar{t}), \ldots, u_n(\bar{t}))$ 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 each induced game admits a unique SPNE. But the game admits no SPSNE because there is no SPNE that is a SPNE in *every* induced game. In particular, the unique SPNE of the centipede game Γ in Fig. 3 is not a SPNE of the induced game $\Gamma^{\{1,2\}}$. However, as noted above, the game admits a non-empty subgame-perfect core. □

Proof of Proposition 6. The unique SPNE of Γ is also a unique SPNE in both the induced games $\Gamma^{\{1\}}$ and $\Gamma^{\{2\}}$, because $\Gamma^{\{1\}} = \Gamma^{\{2\}} = \Gamma$. Furthermore, since the unique SPNE is efficient, it is also a SPNE of the induced game $\Gamma^{\{1,2\}}$. Therefore, by definition of SPSNE, the unique SPNE of Γ is a SPSNE of Γ. By Proposition 5, the SPSNE is unique and the subgame-perfect core is non-empty and consists of the unique SPSNE/SPNE payoff vector. □

Proof of Proposition 7. The proof is by induction on the number of players and stages. The proposition is true for games with a single player and m or fewer stages. Suppose the proposition is true for games with k or fewer players, i.e. $1 \leq k < n$, and $h \leq m$ stages. We show that then it is also true for games with $k+1$ players and $h \leq m$ stages, i.e. a SPSNE of any extensive game with $k+1$ players and $h \leq m$ stages is a SPCPNE of the game.

Suppose $\bar{t} \in T$ is a SPSNE of a single-stage game Γ with $k+1$ players, then, by definition of a SPSNE, $\bar{t} = (\bar{t}_S, \bar{t}_{-S})$ is a SPNE of each induced game Γ^S, $S \subset \{1, \ldots, k+1\}$ and, therefore, $\bar{t}_S$ is a SPNE of each induced game in the restricted game $\Gamma/\bar{t}_{-S}$. Thus, $\bar{t}_S$ is a SPSNE of the restricted game $\Gamma/\bar{t}_{-S}$ for each *proper* subset S of players. Since, as hypothesized, the proposition is true for games with k or fewer players and m or fewer stages and $\bar{t}_S$ is a SPSNE of $\Gamma/\bar{t}_{-S}$ for each proper subset $S \subset \{1, \ldots, k+1\}$, the strategy $\bar{t}_S$ is a SPCPNE of $\Gamma/\bar{t}_{-S}$ for each *proper* subset $S \subset \{1, \ldots, k+1\}$. Therefore, $\bar{t}$ is perfectly self-enforcing in the single-stage game Γ with $k+1$ players. Furthermore, by hypothesis, $\bar{t}$ is a SPSNE of Γ and, therefore, a SPNE of the (single player) induced game Γ^N, i.e. there is no $t \in T$ such that $\sum_{i=1}^{k+1} u_i(t) > \sum_{i=1}^{k+1} u_i(\bar{t})$. Therefore, by part (2) of Definition 3, $\bar{t} \in T$ is a SPCPNE of the single-stage game Γ with $k+1$ players. Similarly, if Γ is a two-stage game with $k+1$ players and $\bar{t}$ is a SPSNE of Γ, the strategy $\bar{t}_S$ is a SPCPNE of $\Gamma/\bar{t}_{-S}$ for each proper subset $S \subset \{1, \ldots, k+1\}$, and restriction of $\bar{t}$ to any subgame of Γ is a SPCPNE of the subgame, because any subgame of a two-stage game Γ is a single-stage game and has at most $k+1$ players and, therefore, as already shown a SPCPNE of any subgame of Γ with $k+1$ or fewer players. Furthermore, there is no $t \in T$ such that $\sum_{i=1}^{k+1} u_i(t) > \sum_{i=1}^{k+1} u_i(\bar{t})$, because $\bar{t}$ is a SPSNE of Γ. Proceeding in this way, the proposition is true for any game with $k+1$ players and m or fewer stages, if it is true for any game with k players and m or fewer stages.

Lastly, the converse is not true, because the unique SPNE in the centipede game in Fig. 3 is a SPCPNE, but not a SPSNE, as the unique SPNE of this game is not a SPNE of the induced game Γ^N. □

References

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.

Aumann, R., 1996. A reply to Binmore. Games Econ. Behav. 17, 138–146.

Avrachenkov, K., Cottatellucci, L., Maggi, L., 2013. Cooperative Markov decision processes: time consistency, greedy players satisfaction, and cooperation maintenance. Int. J. Game Theory 42, 239–262.

Becker, R.A., Chakrabarti, S.K., 1995. The recursive core. Econometrica 63, 401–423.

Bernheim, B.D., Peleg, B., Whinston, M.D., 1987. Coalition-proof equilibria 1: concepts. J. Econ. Theory 42, 1–12.

Binmore, K., 1996. A note on backward induction. Games Econ. Behav. 17, 135–137.

Binmore, K., Rubinstein, A., Wolinsky, A., 1986. The Nash bargaining solution in economic modelling. Rand J. Econ. 17, 176–188.

Bondareva, O.N., 1963. Some applications of linear programming methods to the theory of cooperative games [in Russian]. Probl. Kibern. 10, 119–139.

Chander, P., 2007. The gamma-core and coalition formation. Int. J. Game Theory 2007, 539–556.

Chander, P., 2017. Subgame-perfect cooperative agreements in a dynamic game of climate change. J. Environ. Econ. Manag., 173–188.

Chander, P., 2018a. The core of a strategic game. B.E. J. Theor. Econ., 20170155. Available online.

Chander, P., 2018b. Game Theory and Climate Change. Columbia University Press, New York.

Compte, O., Jehiel, P., 2010. The coalitional Nash bargaining solution. Econometrica 78, 1593–1623.

Forges, F., Mertens, J.-F., Vohra, R., 2002. The ex-ante incentive compatible core in the absence of wealth effects. Econometrica 70, 1865–1892.

Gale, D., 1978. The core of a monetary economy without trust. J. Econ. Theory 19, 456–491.

Habis, H., Herings, P.J-J., 2010. A note on the weak sequential core of dynamic games. Int. Game Theory Rev. 12, 407–416.

Harstad, B., 2012. Climate contracts: a game of emissions, investments, negotiations, and renegotiations. Rev. Econ. Stud. 79, 1527–1557.

Kranich, L., Perea, A., Peters, H., 2005. Core concepts for dynamic TU games. Int. Game Theory Rev. 7, 43–61.

Lehrer, E., Scarsini, M., 2013. On the core of dynamic cooperative games. Dyn. Games Appl. 3, 359–373.

Marx, L.M., Matthews, S.A., 2000. Dynamic voluntary contribution to a public project. Rev. Econ. Stud. 67, 327–358.

Maskin, E., 2003. Bargaining, Coalitions and Externalities. Presidential Address to the Econometric Society. Institute for Advanced Study, Princeton.

McKelvey, R., Palfrey, T., 1992. An experimental study of the centipede game. Econometrica 69, 803–836.

Nagel, R., Tang, F.F., 1998. Experimental results on the centipede game in normal form: an investigation on learning. J. Math. Psychol. 42, 356–384.

Nash, J., 1950. The bargaining problem. Econometrica 18, 155–162.

Osborne, M., Rubinstein, A., 1994. A Course in Game Theory. MIT Press, Cambridge, Mass.

Pérez-Castrillo, D., 1994. Cooperative outcomes through non-cooperative games. Games Econ. Behav. 7, 428–440.

Perry, M., Reny, P.J., 1994. A non-cooperative view of coalition formation and the core. Econometrica 62, 795–817.

Predtetchinski, A., Herings, P.J-J., Perea, A., 2006. The weak sequential core for two-period economies. Int. J. Game Theory 34, 55–65.

Ray, D., Vohra, R., 1997. Equilibrium binding agreements. J. Econ. Theory 73, 30–78.

Rosenthal, R., 1981. Games of perfect information, predatory pricing, and the chain-store paradox. J. Econ. Theory 25, 92–100.

Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica 50, 97–109.

Rubinstein, A., 1980. Strong perfect equilibrium in supergames. Int. J. Game Theory 9, 1–12.

Selten, R., 1975. Re-examination of the perfectness concept for equilibrium points in extensive games. Int. J. Game Theory 4, 25–55.

Shapley, L.S., 1971. Core of convex games. Int. J. Game Theory 1, 11–26.

Shapley, L.S., 1967. On balanced sets and cores. Nav. Res. Logist. Q. 14, 453–460.

Shapley, L.S., 1953. A value for n-person games. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games, vol. II. Princeton University Press, Princeton, NJ, pp. 307–317.