MAGYAR TUDOMÁNYOS AKADÉMIA
Közgazdaság- és Regionális Tudományi Kutatóközpont

KRTK
CERS-HAS

Research Centre for Economic and Regional Studies
HUNGARIAN ACADEMY OF SCIENCES

MŰHELYTANULMÁNYOK DISCUSSION PAPERS

MT-DP – 2012/26

Stationary Consistent Equilibrium Coalition Structures Constitute the Recursive Core

LÁSZLÓ Á. KÓCZY

INSTITUTE OF ECONOMICS, RESEARCH CENTRE FOR ECONOMIC AND REGIONAL STUDIES,
HUNGARIAN ACADEMY OF SCIENCES BUDAPEST, 2012

Discussion papers
MT-DP – 2012/26

Institute of Economics, Research Centre for Economic and Regional Studies,
Hungarian Academy of Sciences

Stationary Consistent Equilibrium Coalition Structures Constitute the Recursive Core

Author:

László Á. Kóczy
senior research fellow
Institute of Economics
Research Centre for Economic and Regional Studies
Hungarian Academy of Sciences
Keleti Faculty of Business and Management, Óbuda University
Email: koczy@krtk.mta.hu

August 2012

ISBN 978-615-5243-28-8
ISSN 1785 377X

Stationary Consistent Equilibrium Coalition Structures Constitute the Recursive Core

László Á. Kóczy

Abstract

We study coalitional games where the coalitional payoffs depend on the entire coalition structure. We introduce a noncooperative, sequential coalition formation model and show that the set of equilibrium outcomes coincides with the recursive core, a generalisation of the core to such games. In order to extend past results limited to totally recursive-balanced partition function form games we introduce subgame-consistency that requires perfectness in relevant subgames only, while some unreached subgames are ignored. Due to the externalities, the profitability of deviations depends on the partition formed by the remaining players: the stability of core payoff configurations is ensured by a combination of the pessimism of players going for certain profits only and the assumption that players base their stationary strategies on a made-up history punishing some of the possible deviators - and getting this sometimes right.

Keywords: partition function, externalities, implementation, recursive core, stationary perfect equilibrium, time consistent equilibrium

JEL classification: C71, C72

Acknowledgement:

The author thanks Lars Ehlers for his comments and the Hungarian Academy of Sciences for the financial support (LP-004/2010).

Stacionárius konzisztens egyensúlyok által előállított koalícióstruktúrák alkotják a rekurzív magot

Kóczy Á. László

Összefoglaló

Olyan koalíciós játékokat vizsgálunk, amelyekben a koalíciók kifizetése függ a teljes koalíciósstruktúrától. Egy nonkooperatív, szekvenciális koalícióformációs modellt vezetünk be és igazoljuk, hogy az egyensúlyi kimenetek halmaza egybeesik a rekurzív maggal, a mag partíciós játékokra alkalmazható általánosításával. A korábbi, kizárólag tökéletesen rekurzív kiegyensúlyozott játékokra alkalmazható eredmény kiterjesztése érdekében bevezetjük a részjáték-konzisztencia fogalmát, ami a tökéletességet csak a releváns részjátékokban vizsgálja, míg a többi részjátékot figyelmen kívül hagyjuk. Az externáliák miatt az elhajlások nyereségessége függ a maradék játékosok által létrehozott partíciótól: a magbéli kifizetés-konfigurációk stabilitását garantálja egyrészt a kizárólag a biztos profitra törekvő játékosok borúlátása, másrészt az a feltétel, mely szerint a játékosok stacionárius stratégiájukat egy kitalált történelem alapján választják meg és e szerint büntetik a vélt deviánsokat, alkalmanként ráhibázva a valódi történelemre.

Tárgyszavak: partíciós függvény, externáliák, implementáció, rekurzív mag, stacionárius tökéletes egyensúly, idő-konzisztens egyensúly

JEL kódok: C71, C72

Stationary consistent equilibrium coalition structures constitute the recursive core*

László Á. Kóczy†

Abstract

We study coalitional games where the coalitional payoffs depend on the entire coalition structure. We introduce a noncooperative, sequential coalition formation model and show that the set of equilibrium outcomes coincides with the recursive core, a generalisation of the core to such games. In order to extend past results limited to totally recursive-balanced partition function form games we introduce subgame-consistency that requires perfectness in relevant subgames only, while some unreached subgames are ignored. Due to the externalities, the profitability of deviations depends on the partition formed by the remaining players: the stability of core payoff configurations is ensured by a combination of the pessimism of players going for certain profits only and the assumption that players base their stationary

*The author thanks Lars Ehlers for his comments and the Hungarian Academy of Sciences for the financial support (LP-004/2010).

†Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budaörsi út 45., H-1112 Budapest and Keleti Faculty of Business and Management, Óbuda University; koczy@krtk.mta.hu

strategies on a made-up history punishing some of the possible deviators – and getting this sometimes right.

Subject classification: C71, C72

Keywords and phrases: partition function, externalities, implementation, recursive core, stationary perfect equilibrium, time consistent equilibrium

1 Introduction

Throughout its history the theory of coalitional games has mostly focussed on the study of games with *orthogonal coalitions*, that is, coalitions, which can be studied independently of each other. The most obvious example is the commonest form of a TU-game with a characteristic function that assigns a payoff to a coalition disregarding other players and coalitions. When we look at the usual interpretations of coalitions, be those trading blocks (Yi, 1996), trusts (Bloch, 1995) or international environmental agreements (Funaki and Yamato, 1999; Eyckmans and Tulkens, 2003), the orthogonality assumption is difficult to maintain; we believe it is the exception rather than the rule that coalitions can be studied independently of each other.

Since the seminal paper of Thrall and Lucas (1963) introducing the partition function form numerous cooperative approaches and solution concepts have been proposed to solve games with externalities, but in the absence of an implementation by non-cooperative equilibria these remain interesting heuristics (Chander and Tulkens, 1995; Ray and Vohra, 1997; Hyndman and Ray, 2007). For games with orthogonal coalitions the implementation of cooperative solution concepts, such as the core has an extensive literature (Chatterjee et al., 1993; Lagunoff, 1994; Perry and Reny, 1994), but these results do not directly generalise to games with externalities. In this do-main Huang and Sjöström (2006) and Kóczy (2009) have provided partial results that are limited to games with non-empty cores in all subgames, or, in terms of sequential coalition formation games: to games with stationary perfect equilibria. It turns out that perfectness is a very demanding condition and the implementation might fail even for simple TU games. We therefore introduce a generalisation, subgame-consistency, and show that the set of partitions formed under the resulting equilibria coincides with the recursive core.

Subgame-consistency is a weaker concept than subgame-perfectness, but more demanding than time-consistency (Kydland and Prescott, 1977). While subgame-perfectness requires all subgames to be perfect, time-consistency insists only on the consistency of playing the equilibrium. Subgame consistency insists on the perfectness of most subgames, only ignoring subgames that are not relevant for studying the stability of the equilibrium play. Subgame-perfect equilibria are therefore also subgame-consistent and subgame-consistent equilibria are also time-consistent. Moreover stationary perfect equilibria are stationary consistent. For more on the relation of subgame-perfect and time-consistent strategies see Fershtman (1989) and Asilis (1995).

The structure of the paper is as follows. After this introduction a long second section follows introducing both the cooperative and noncooperative theories to solve games in partition function form, we introduce the notation and simple terminology we are going to use. We present the cooperative solution, namely the recursive core and similarly the noncooperative coalition formation game and its equilibria. A novel equilibrium concept, *subgame consistency* and the corresponding notion of *relevant subgame* are also introduced here. We state and prove our main result in the third section. The paper ends with a brief conclusion.

2 Preliminaries

Let N denote the set of players. Subsets are called *coalitions*. A partition $\mathcal{S}$ of S is a splitting of S into disjoint coalitions. $\Pi(S)$ denotes the set of partitions of S. In general we use capital and calligraphic letters to denote a set and its partition (the set of players N being an exception), indexed capital letters are elements of the partition. We write $i \in \mathcal{S}$ if there exists S_k such that $i \in S_k \in \mathcal{S}$ and if $i \in \mathcal{S}$ we write $\mathcal{S}(i)$ for the coalition embedded in $\mathcal{S}$ containing i.

The game (N, V) is given by a player set N and a *partition function* (Thrall and Lucas, 1963) $V : \Pi(N) \to (2^N \to \mathbb{R})$, where $V(S_i, \mathcal{S})$ denotes the payoff for coalition S_i embedded in partition $\mathcal{S}$. For vectors $x, y \in \mathbb{R}^N$ we write x_S for the restriction to the set S and $x_S > y_S$ if $x_i \geq y_i$ for all $i \in S \subset N$ and there exists $j \in S$ such that $x_j > y_j$.

Due to the externalities in the partition function form game the standard solution concepts do not work; we consider outcomes instead of imputations and the recursive core (Kóczy, 2007), a generalisation of the core, as the solution concept.

The pair $\omega = (x, \mathcal{P})$ consisting of a payoff vector $x \in \mathbb{R}^N$ and a partition $\mathcal{P} \in \Pi(N)$ is a *payoff configuration* (or *outcome*) if $\sum_{i \in S} x_i = V(P_i, \mathcal{P})$ for all $P_i \in \mathcal{P}$. The set of outcomes of game (N, V) is denoted $\Omega(N, V)$.

Let $S \subsetneq N$ and $\overline{\mathcal{S}}$ the partition of $\overline{S} = N \setminus S$. Then the *residual game* $S, V^{\overline{\mathcal{S}}}$ is the partition function game played over S such that $V^{\overline{\mathcal{S}}}(S_i, \mathcal{S}) = V(S_i, \mathcal{S} \cup \overline{\mathcal{S}})$ for all $S_i \subseteq S$ and $\mathcal{S} \in \Pi(S)$.

Definition 1 (Recursive core (Kóczy, 2007))**.** For a single-player game the recursive core is trivially defined. Now assume that the core $C(N, V)$ has been defined for all games with $|N| < k$ players. Then for an $|N|$-player game an outcome $(x, \mathcal{P})$ is dominated if there exists a coalition Q forming partition $\mathcal{Q}$ and an outcome $(y, \mathcal{Q} \cup \overline{\mathcal{Q}}) \in \Omega(N, V)$ such that $y_Q > x_Q$ and if $C(\overline{Q}, V^{\mathcal{Q}}) \neq \varnothing$ then $(y_{\overline{Q}}, \overline{\mathcal{Q}}) \in C(\overline{Q}, V^{\mathcal{Q}})$. The *core* $C(N, V)$ of (N, V) is the set of undominated outcomes.

Huang and Sjöström's (2003) r-core coincides on a broad class of games that does not, however include the standard TU-games without externalities. Ray's (2007) standard equilibrium, defined for for symmetric partition function form games has a similar recursive structure. For an interpretation and the discussion of the properties of the recursive core see Kóczy (2007, 2009); Huang and Sjöström (2010).

3 Sequential coalition formation

The sequential coalition formation game we define is similar to Bloch's (1996) and Perry and Reny's (1994). First a player proposes the formation of some coalitions. The offer specifies not only who should be the members of these coalitions, but also how the coalitional payoffs will be shared. If all involved players accept the offer, the coalitions form and leave the game. When the offer is rejected, a new proposal is made and so on, until all players exit.

3.1 The game

Consider a game (N, V) with a player set N and partition function V. Time t is continuous, but we assume that there is always a open time interval between two actions: there is time to respond. Let $Q^t \subseteq N$ denote those who have already quit the game by time t, forming partition $\mathcal{Q}^t$. Player i

can make proposals

$$P_i^t = \left\{ (\mathcal{P}^t, w^t) \middle| \mathcal{P}^t \in \Pi(P^t), P^t \subseteq \overline{Q^t}, P^t \ni i, w^t \in \mathbb{R}^{P^t}, \forall P_k^t \in \mathcal{P}^t \sum_{j \in P_k^t} w_j^t = 1 \right\}$$

the current proposer is i^t making the proposal $p^t = (\mathcal{P}^t, w^t)$ to the players in P^t, already accepted by the players in $A^t \subseteq P^t$ (we assume $i^t \in A^t$) have already accepted the proposal. At t a player i can

1. accept the proposal p^t if $i \in P^t$,
2. make a new proposal, or
3. do nothing.

The strategy σ_i of a player i specifies a complete protocol of actions for all times and contingencies during the game and σ_i^t the action at t. Let σ denote the strategy profile collecting the strategies of all players.

A state s^t is a snapshot of the game and is given by a tuple $(\mathcal{Q}^t, i^t, p^t, A^t)$. *History* h is a complete record of actions: we only need a small part of this information. Technically let $t_1, t_2, ...$ be points in time such that $s^{t_k} = s^t \neq s^{t_{k+1}}$ for all $t_k < t < t_{k+1}$. Then $h = \{s^{t_k}\}_k$. Let h^t be a truncation of history up to time t. There is a one-to-one correspondence between history truncations and decision nodes in the game tree. We will identify subgames by truncated histories. Let $\sigma|_{h^t}$ denote the restriction of strategy σ to the subgame h^t.

There is no guarantee that the game ends, that all players quit. We therefore specify the payoffs for coalitions Q_k embedded in an incomplete partition $\mathcal{Q}$. We take a conservative approach: players only receive their guaranteed payoff, the lowest value the coalition can obtain in a partition embedding the coalition structure of departed players:

$$V(Q_k, \mathcal{Q}) = \begin{cases} \min_{\mathcal{P} \supset \mathcal{Q}} V(Q_k, \mathcal{P}) & Q_k \in \mathcal{Q} \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Given a strategy profile σ the payoff of player i is $x_i(\sigma)$.

3.2 Equilibria

Now that we have specified the available strategies (actions), the resulting payoffs (incentives) we can focus on the outcomes of the coalition formation game. We hope to answer two questions simultaneously: (i) which coalitions will form (ii) how are coalitional payoffs distributed.

Recall that players are conservative and only go for certain profits: If different beliefs lead to different subsequent actions from the other players, a deviation may or may not be profitable under all such scenarios.

Definition 2. The strategy profile σ^* is a subgame-perfect equilibrium if for all $h \in \mathcal{H}$, time t, $i \in N$, strategies σ_i the corresponding restrictions $\sigma^*|_{h^t}$ and $\sigma_i|_{h^t}$ to the subgame at h^t we have

$$x_i(\sigma^*|_{h^t}) \geq x_i(\sigma_i|_{h^t}, \sigma^*_{-i}|_{h^t}). \tag{3.2}$$

The set of perfect equilibria may be too inclusive (see Muthoo (1990, 1995); Perry and Reny (1994); Osborne and Rubinstein (1990) for a discussion of folk-theorem-like results) so we focus on stationary strategies.

Definition 3. A strategy σ is *stationary* if it does not depend on time. Formally: if for all h and t_1, t_2 with $h^{t_1} = h^{t_2}$ we have $\sigma|_{h^{t_1}} = \sigma|_{h^{t_2}}$.

Definition 4. A *stationary perfect equilibrium* σ^* is a strategy profile that is both subgame-perfect and stationary.

For games with nonempty residual cores the set of stationary perfect equilibrium partitions coincide with the recursive core (Kóczy, 2009). This equivalence result predicts that games containing empty residual cores do not have stationary perfect equilibria.

Bloch (1996) presents a 3-player example, where player 1 would like to form a coalition with 2, 2 with 3, 3 with 1. This game does not have stationary-perfect equilibria. Since residual games are also partition function form games, the smallest residual game for which the corresponding subgame of the sequential game has no stationary strategies has an empty core. By a sufficiently large payoff for the grand coalition the core of the original game is nevertheless empty. Perfectness only holds globally, that is, if the tiniest subgame fails to have stationary perfect equilibria this failure extends to the entire game. On the other hand, just as the recursive core may be non-empty even if the game has empty residual cores, with a weaker concept of perfection we may retain an essentially perfect behaviour in the corresponding sequential coalition formation games, too.

Time-consistency (Kydland and Prescott, 1977) merely requires the *equilibrium strategy* to be consistent or revision-free and is therefore unaffected by empty cores in other elsewhere. This is a promising direction for weakening subgame perfectness as most subgames are never reached anyway. *Subgame-consistency*, that we introduce below is the *aurea mediocritas*, the golden mean: the perfectness/consistency criterion is only required in *relevant* subgames. What is relevant?

Definition 5. For a strategy profile σ a subgame at t is *relevant* if

1. it is the original game ($t = 0$),
2. it can be reached via an elementary profitable deviation from σ, or

3. it is a relevant subgame of a relevant subgame.

Let $\sigma|_h$ denote the truncation of σ to the subgame corresponding to h.

Definition 6. The strategy profile σ^* is a *subgame-consistent equilibrium* if for all relevant subgames h^t, $i \in N$, strategies σ_i the corresponding restrictions $\sigma^*|_{h^t}$ and $\sigma_i|_{h^t}$ to h^t we have

$$x_i(\sigma^*|_{h^t}) \geq x_i(\sigma_i|_{h^t}, \sigma^*_{-i}|_{h^t}). \tag{3.3}$$

For an equilibrium strategy profile, this requires checking the equilibrium path only: Since no profitable deviation exist, other subgames need not be checked. Were there profitable deviations they would have to be supported by a strategy that is subgame perfect along that strategy.

Clearly, subgame perfect equilibrium strategies are also subgame-consistent.

3.3 Alternative histories

Consider a game with nonempty recursive-core. For a general implementation result we want to define an equilibrium strategy profile that produces core outcomes. In Subsection 3.2 we have solved the issue of empty residual cores, here we describe and solve another issue. Before we move to the example let us clarify that in a sequential, noncooperative coalition formation game players do not "respond" to deviations by punishment, and players will especially not revert to non quitting the game. Instead, players will have strategies to the same effect and those are these strategies that we try to mend to salvage the equilibrium.

We assume that players obtain the payoff x when the equilibrium strategy is played. Suppose now that players find themselves in an off-equilibrium state with two deviant coalitions (see Figure 1): A and B. In case A deviates

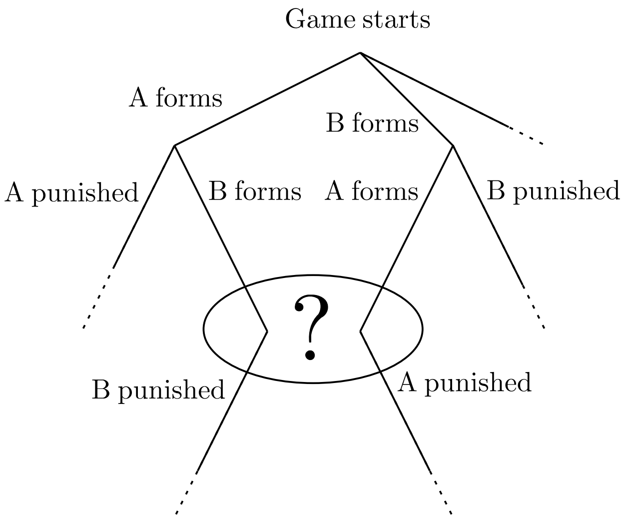

Figure 1: Stationary strategies cannot react to different histories.

first $\overline{A}$ (including B) should stop this deviation producing payoff-vector y^A such that $y^A_A < x_A$. When B deviates, too, the remaining $\overline{A \cup B}$ should also choose an action to get a payoff z^B such that $z^B_B < y^A_B$. Consequently B does not deviate, y^A forms which is bad for A, hence A does not deviate and the equilibrium is preserved.

What if $\overline{A \cup B}$ think B deviated first? They want response y^B, which A did not comply with thus A must be punished by z^A, where $y^B_B < x_B$ and $z^A_A < y^B_A$. If $z^A_B > y^B_B$ and $z^A_A > x_A$ the response does not work. It seems it is essential for $\overline{A \cup B}$ to know the correct history: When looking at the sequential game we must explicitly determine which equilibrium strategy responds to a particular history, which equilibrium strategy can render the same deviations non-profitable. When looking at stationary strategies, history is masked from the players, who only see the current state of the game. At some point players may see a state that is very different from the equilibrium play. What is the equilibrium that prevents such deviations?

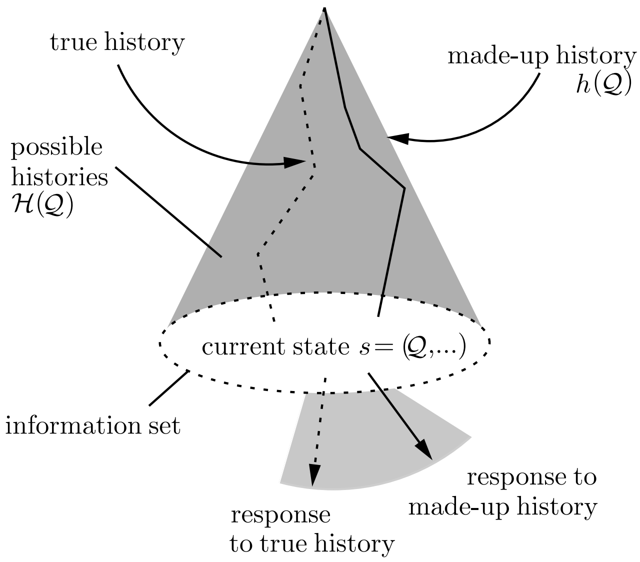

Figure 2: Stationary decision with possible histories

Although players do not know the history, given the current state s they can reconstruct one of the possible histories $h(s)$ satisfying $s^{h^t(s)} = s$ for some t. This history $h(s)$ provides a plausible, but not necessarily true explanation of the current state. Let $\mathcal{H}(s) = \left\{h \,\middle|\, \exists t: \; s^{h^t} = s\right\}$ be the set of plausible histories to the current state s.

Unfortunately in a stationary process these possible histories only provide a temporary explanation as they get forgotten, too. We assume that h is arbitrarily regenerated when the partition $\mathcal{Q}^t$ changes, that is, each time some players quit the game. It is common knowledge among the remaining players, who choose their strategies treating h as the true history, but taking into account that future deviations may generate a new alternative history. "History" is not preserved, subsequent alternative histories are totally unrelated, it may well be that the current state will not have happened at all. Let $\mathcal{H}(\sigma)$ denote the set of all possible histories happening as σ is played (and there are no deviations). Finally $\mathcal{H}(\sigma, s) \subseteq \mathcal{H}(\sigma) \cap \mathcal{H}(s)$ denotes the set of possible histories passing through s when σ is played after s is reached. Let us stress that many of the $h \in \mathcal{H}(\sigma, s)$ will believe that s did not exist at all, but at s players are aware of their limited rationality in the future.

The coalitional payoffs are not related to history and are thus unaffected, but what happens to the payoff of the individual players? In the non-stationary game players can calculate the payoffs by just looking at the strategies. With stationarity we must take the subsequent updating of the history into account. The subsequent development of the game therefore depends on the current state s – in particular, the partition $\mathcal{Q}$ – and the strategy restricted to this subgame. Since we focus on stationary strategies it is sufficient to say that these strategies are restricted to a subgame s. Then the subsequent development of history depends on some strategy $\sigma|_s$ restricted to s. Let $x(\sigma, h) \in \mathbb{R}^N$ denote the vector of payoffs in case σ is played along the history h. Then the payoff players can expect is

$$x(\sigma, s) = \min_{h \in \mathcal{H}(\sigma, s)} x(\sigma, h). \tag{3.4}$$

Note the pessimism of the players. When uncertain about the subsequent development of the game, they assume that the remaining players will fabricate histories that are the least favourable to them. While subgame perfectnes can be formulated with these expectations, too, the resulting equilibria are different in general. Since the additional "information" comes from the past, the concepts of stationarity and stationary equilibria are not affected, the stationary equilibria remain the same and the recursive core equivalence result remains valid. Likewise, subgame consistency can be redefined in this environment, but we first clarify what is a relevant subgame.

Definition 7. For a strategy profile σ a subgame at h^t is *relevant* if

1. it is the original game ($\mathcal{Q}^{h^t} = \varnothing$),

2. there exists an elementary deviation σ' producing $\mathcal{Q}'$ such that

$$x_i(\sigma, s^{h^t}) < x_i(\sigma', \mathcal{Q}', \varnothing), \tag{3.5}$$

3. or it is a relevant subgame of a relevant subgame.

The subgame-consistency is accordingly modified replacing h by an arbitrary (compatible) history $h(s)$ in Inequality 3.3 and payoffs are now given by Equation 3.4 and are conditional on the current s via the different updates of made-up history.

$$x_i(\sigma^*|_{h^t(s)}, s) \geq x_i(\sigma_i|_{h^t(s)}, \sigma^*_{-i}|_{h^t(s)}, s). \tag{3.6}$$

The condition becomes clear now: it has implications not so much for the present, but for the reactions of the remaining players.

Let σ be a stationary strategy and $\sigma|_s$ its restriction to a state s.

A *stationary consistent equilibrium* σ^* is a strategy profile that is both subgame-consistent and stationary, that is, if for all relevant subgames corresponding to some s we have

$$x_i(\sigma^*|_s, s) \geq x_i(\sigma_i|_s, \sigma^*_{-i}|_s, s). \tag{3.7}$$

We denote the set of stationary consistent equilibria by $\mathrm{SCE}(N, V)$ and outcomes resulting from playing such equilibrium strategies by $\Omega^*(N, V)$.

4 Results

Theorem 1. *Let (N, V) be a partition function form game. Then its recursive core $C(N, V)$ coincides with the set $\Omega^*(N, V)$ of outcomes supported by stationary consistent equilibrium strategy profiles.*

The rest of this section is devoted to the inductive proof of this theorem. As the proof is long, we break it into a number of propositions and finally present a summary of these results. The first proposition requires no proof:

Proposition 2. *Let* $(\{1\}, V)$ *be a trivial, single-player partition function form game. Then* $C(\{1\}, V) = \Omega^*(\{1\}, V)$*.*

Now *assume* that Theorem 1 holds for all games with less than k players. In the following we extend it to games with k players. In order to show $\Omega^*(N, V) = C(N, V)$, first we show $\Omega^*(N, V) \subseteq C(N, V)$ then $\Omega^*(N, V) \supseteq C(N, V)$.

Lemma 3. *If Theorem 1 holds for all games with up to* $k-1$ *players,* $\Omega^*(N, V) \subseteq C(N, V)$ *for all* k*-player games.*

Proof. If $\Omega^*(N, V) = \varnothing$ the result is trivial, otherwise there exists a SCE σ producing $\omega(\sigma, h) = (x(\sigma, h), \mathcal{P}(\sigma, h)) \in \Omega^*(N, V)$ for some sequence of possible histories $h \in \mathcal{H}(\sigma, \varnothing)$. In particular, we assume that $\omega(\sigma, h) \notin C(N, V)$ and prove contradiction.

By this assumption there exists a profitable deviation $\mathcal{D}$ by some set D of players. By this we really mean a deviation by a single player $i \in D$ that results in the departure of $\mathcal{D}$, that is, in a state s' with $\mathcal{Q}^{s'} = \mathcal{D}$, where, without loss of generality, we assume that the deviation occurs at s when no other players have yet left the game. The induced subgame has fewer players so the inductive assumption can be applied. In the sequential game the deviation at h^t is expressed by the strategy profile $\sigma' = (\sigma_i|_{-h^t}, \sigma'_i|_{h^t}, \sigma_{-i})$ differing only for i and only in the subgame h^t. We discuss three cases.

Case 1. For the strategy profile σ' the subgame at $\mathcal{D}$ is not relevant. Then for all σ_{-i} there exists $i \in D$ and $h' \in \mathcal{H}(\sigma', s')$ such that $x_i(\sigma', h') <$

$x_i(\sigma, h)$ – thus the deviation cannot be profitable in the cooperative game; contradiction.

Case 2. The subgame is relevant, the core of the corresponding residual subgame is empty. Then $V(D, \mathcal{D} \cup \overline{\mathcal{D}}) > \sum_{i \in S} x_i(\sigma, h)$ for all $\overline{\mathcal{D}} \in \Pi(\overline{D})$. As

$$V(D, \mathcal{D} \cup \overline{\mathcal{D}}) = \min_{h' \in \mathcal{H}(\sigma', \mathcal{D})} \sum_{i \in S} x_i(\sigma', h'),$$

a player i in D should immediately propose to form $\mathcal{D}$. By subgame consistency all in D will accept. Therefore σ is *not* a stationary consistent equilibrium, moreover the outcome $\omega(\sigma, h)$ cannot be supported by other equilibria either. Contradiction.

Case 3. The induced subgame is relevant and the core of the corresponding residual subgame is not empty. Since σ is a SCE its restriction $\sigma|_s$ to this relevant subgame s is stationary consistent, too. Moreover the deviation from σ to form $\mathcal{D}$ is not profitable, therefore

$$x_D(\sigma|_s, s) \geq x_D(\sigma'|_{s'}, s') \tag{4.1}$$

On the other hand, by the inductive assumption,

$$\omega(\sigma'|_{s'}, s') \in C(\overline{D}, V^{\mathcal{D}}). \tag{4.2}$$

This, however, implies that the deviation $\mathcal{D}$ is not profitable in the cooperative game; contradiction.

We have discussed all cases, and found the assumptions contradicting. Therefore $\omega(\sigma, H) \in C(N, V)$. □

Punishment strategy Before we move on to our next lemma, we specify a "response" to each deviation that turns these deviations unprofitable. In the recursive core a deviation is only profitable if it represents an improvement in the payoffs for all residual assumptions. In a game (N, V) the core is nonempty if for all outcomes $(x, \mathcal{P}) \in C(N, V)$ and for all deviations $\mathcal{D}$ there exists an outcome $(y_{\overline{D}}, \overline{\mathcal{D}}) \in \Omega(\overline{D}, V^{\mathcal{D}})$ such that

1. there exists $D_1 \in \mathcal{D}$ such that $\sum_{i \in D_1} x_i \geq V(D_1, \mathcal{D} \cup \overline{\mathcal{D}})$ and
2. $(y_{\overline{D}}, \overline{\mathcal{D}}) \in C(\overline{D}, V^{\mathcal{D}})$ if $C(\overline{D}, V^{\mathcal{D}}) \neq \varnothing$.

Generally, for a residual game $(\overline{S}, V^{\mathcal{S}})$ the response to $\mathcal{D}$ is denoted as $(x_{\overline{S}}(\mathcal{D}), \overline{\mathcal{S}}(\mathcal{D}))$.

In the sequential game, primary deviations can be punished, but due to stationarity, with multiple departed coalitions finding the right punishment is difficult.

In the following we specify the punishment strategy to a deviation knowing that some other coalitions left, too. We assume that $\mathcal{Q}$ has already left the game, but $\tilde{\mathcal{Q}} \subseteq \mathcal{Q}$ was (or at least $\overline{Q}$ think it was) the last to exit. In the partition function form game $(\overline{Q \setminus \tilde{Q}}, V^{\mathcal{Q} \setminus \tilde{\mathcal{Q}}})$ the partition $\tilde{\mathcal{Q}}$, as a deviation, defines a residual game $(\overline{Q}, V^{\mathcal{Q}})$, where the response to $\tilde{\mathcal{Q}}$ is $(x_{\overline{Q}}(\tilde{\mathcal{Q}}), \overline{\mathcal{Q}}(\tilde{\mathcal{Q}}))$.

Lemma 4. *If Theorem 1 holds for all games with less than k players, then $\Omega^*(N, V) \supseteq C(N, V)$ for all k-player games (N, V).*

Proof. The proof is partly inspired by Bloch's (1996, Proposition 3.2), and is by construction. We show that if $(x^*, \mathcal{P}^*) \in C(N, V)$ there exists a stationary consistent strategy profile σ^* such that for all for all possible histories $h \in \mathcal{H}(\sigma^*, \varnothing)$ we have $\omega(\sigma^*, h) \in C(N, V)$.

In the following we define the stationary strategy σ_i^* for player i. Due to stationarity it is sufficient to specify the strategy for each triple $(\mathcal{Q}, \tilde{\mathcal{Q}}, p)$ consisting of the partition of players who have already quit, the subpartition consisting of the coalitions that left last according to the current made-up history and the current proposal $p = (\mathcal{T}, w)$. Then the strategy of player i is given as

$$\sigma_i^*(\mathcal{Q},\tilde{\mathcal{Q}},\mathcal{T},w) = \begin{cases} \text{accept} & \text{if } x_i(\sigma^*,\mathcal{Q}\cup\mathcal{T},\varnothing) > x_i(\sigma^*,\mathcal{Q},\varnothing) \\ \left(\mathcal{P}^*, \frac{x*}{|x^*|}\right) & \text{if } \mathcal{T}=\mathcal{Q}=\varnothing \\ \left(\overline{\mathcal{Q}}(\tilde{\mathcal{Q}}), \frac{x_{\overline{\mathcal{Q}}}(\tilde{\mathcal{Q}})}{\left|x_{\overline{\mathcal{Q}}}(\tilde{\mathcal{Q}})\right|}\right) & \text{if } \mathcal{T}=\varnothing, \text{ but } \mathcal{Q}\neq\varnothing \\ \text{wait} & \text{otherwise.} \end{cases} \tag{4.3}$$

In equilibrium $\mathcal{P}(\sigma^*) = \mathcal{P}^*$ and the strategy is stationary by construction so we only need to verify subgame-consistency. We show this by induction. As subgame-consistency holds for a trivial game we may assume that it holds for all games of size less than $|N|$.

Now consider game (N, V) and observe that if Q departed to form $\mathcal{Q}$ the subgame is simply a coalition formation game with less players. We discuss two cases based on the emptiness of the residual core.

1. If the residual core is not empty, the proposed strategy exhibits the same similarity property: in equilibrium the core partition is proposed and accepted, while residual cores form off-equilibrium.

The inductive assumption then ensures that the off-equilibrium path is subgame-consistent so we only need to check whether a deviation $\tilde{\mathcal{Q}}$ is ever accepted. This deviation corresponds to a deviation in the partition function game. Since $(x^*, \mathcal{P}^*) \in C(N, V)$, by the construction of $\left(\overline{\mathcal{Q}}(\tilde{\mathcal{Q}}), \frac{x_{\overline{\mathcal{Q}}}(\tilde{\mathcal{Q}})}{\left|x_{\overline{\mathcal{Q}}}(\tilde{\mathcal{Q}})\right|}\right)$ we know that for some history $h(\mathcal{Q})$ there exists a player in $\tilde{Q}$ for whom the deviation $\tilde{\mathcal{Q}}$ is not profitable. Given the pessimism of the players, this is sufficient to deter this player from accepting the proposal to deviate.

2. The emptiness of the residual core, by our assumption, implies that there are no stationary consistent equilibrium strategy profiles. In the absence of such strategy profiles the strategy σ^* will be abandoned and so

the players in $\tilde{Q}$ cannot predict the partition of $\overline{\mathcal{Q}}$ – in this case, by Expression 3.1, they, individually, expect the worst. As $\tilde{\mathcal{Q}}$ only forms if it is a profitable deviation, that is, only if $x_i(\sigma^*, h)$ is an improvement for all $h \in \mathcal{H}(\sigma^*, \mathcal{Q}, \varnothing)$. Since $(x^*, \mathcal{P}^*) \in C(N, V)$ this is not the case. This, implies that post-deviation subgame is not relevant. Also, the formation of $\mathcal{P}^*$ is unaffected by possible deviations in this subgame, meeting the first condition of subgame-consistency. □

Proof of Theorem 1. The proof is by induction. The result holds for trivial, single-player games. Assuming that the result holds for all $k-1$ player games, the result for k-player games is a corollary of Lemmata 3 & 4. □

5 Conclusion

Theorem 1 holds for arbitrary games in discrete partition function form, but of course it is most interesting for games where some of the residual cores are empty. When a proposal is made in a game without externalities the invited players do not even (need to) consider the residual game and therefore the emptiness of a residual core is not addressed. Huang and Sjöström (2006) and Kóczy (2009) focus on games where the residual cores are non-empty, in fact the r-core (Huang and Sjöström, 2003) is not even defined for games with empty residual cores. As already pointed out by Kóczy (2007) this is not only an enormous limitation given the number of conditions such games must satisfy (one for each residual game), but the definitions/results do not apply to some games without externalities and so they are not generalisations of the well-known results for TU-games. The present paper heals this deficiency.

If the concept is a generalisation of the core for TU-games, it is natural to ask how our game proceeds in the special case when the partition function form game at hand is actually a TU-game as it does not have externalities. In the absence of externalities, there is way to "punish" deviations as their payoff does not depend on the partition of the remaining players. Since there are no punishments, any strategy profile resulting in a core outcome in the remaining game will equally be a punishment strategy and therefore considering alternative histories does not really have a bite here.

Similarly, the expectation that the residual players will form a residual core outcome does not influence the decisions of active players: their payoff will not depend on the coalitions formed in the residual game. Perry and Reny (1994) study market games (Shapley and Shubik, 1969), which are totally-balanced. For such games the "residual core" is always non-empty and therefore *i)* the game always terminates with all players leaving the game and *ii)* there always exist stationary perfect equilibria. For such games subgame consistency and subgame perfectness coincide and so the two procedures provide the same implementation.

References

Asilis, C. M., 1995, A note on the equivalence of time consistency and subgame perfection in stochastic games, European Economic Review 39, 245–251.

Bloch, F., 1995, Endogenous structures of association in oligopolies, The RAND Journal of Economics 26, 537–556.

Bloch, F., 1996, Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division, Games and Economic Behavior 14, 90–123.

Chander, P. and H. Tulkens, 1995, A core-theoretic solution for the design of cooperative agreements on transfrontier pollution, International Tax and Public Finance 2, 279–293.

Chatterjee, K., B. Dutta, D. Ray, and K. Sengupta, 1993, A Noncooperative Theory of Coalitional Bargaining, Review of Economic Studies 60, 463–477.

Eyckmans, J. and H. Tulkens, 2003, Simulating coalitionally stable burden sharing agreements for the climate change problem, Resource and Energy Economics 25, 299–324.

Fershtman, C., 1989, Fixed rules and decision rules:: Time consistency and subgame perfection, Economics Letters 30, 191–194.

Funaki, Y. and T. Yamato, 1999, The core of an economy with a common pool resource: A partition function form approach, International Journal of Game Theory 28, 157–171.

Huang, C. and T. Sjöström, 2003, Consistent solutions for cooperative games with externalities, Games and Economic Behavior 43, 196–213.

Huang, C.-Y. and T. Sjöström, 2006, Implementation of the recursive core for partition function form games, Journal of Mathematical Economics 42, 771–793.

Huang, C.-Y. and T. Sjöström, 2010, The Recursive Core for Non-Superadditive Games, Games 1, 66–88.

Hyndman, K. and D. Ray, 2007, Coalition Formation with Binding Agreements, Review of Economic Studies 74, 1125–1147.

Kóczy, L. A., 2007, A recursive core for partition function form games, Theory and Decision 63, 41–51.

Kóczy, L. A., 2009, Sequential Coalition Formation and the Core in the Presence of Externalities, Games and Economic Behavior 66, 559–565.

Kydland, F. E. and E. C. Prescott, 1977, Rules rather than discretion: The inconsistency of optimal plans, Journal of Political Economy 85, 473–490.

Lagunoff, R. D., 1994, A Simple Noncooperative Core Story, Games and Economic Behavior 7, 54–61.

Muthoo, A., 1990, Bargaining without commitment, Games and Economic Behavior 2, 291–297.

Muthoo, A., 1995, A bargaining model with players' perceptions on the retractability of offers, Theory and Decision 38, 85–98.

Osborne, M. and A. Rubinstein, 1990, Bargaining and Markets 59, 50–77.

Perry, M. and P. J. Reny, 1994, A noncooperative View of Coalition Formation and the Core, Econometrica 62, 795–817.

Ray, D., 2007, A Game-Theoretic Perspective on Coalition Formation (Oxford University Press, Oxford).

Ray, D. and R. Vohra, 1997, Equilibrium Binding Agreements, Journal of Economic Theory 73, 30–78.

Shapley, L. S. and M. Shubik, 1969, On market games, Journal of Economic Theory 1, 9–25.

Thrall, R. M. and W. F. Lucas, 1963, N-Person Games in Partition Function Form, Naval Research Logistics Quarterly 10, 281–298.

Yi, S.-S., 1996, Endogenous formation of customs unions under imperfect competition: open regionalism is good, Journal of International Economics 41, 153–177.