Modelling of cooperative behavior in multistage game of renewable resource extraction[1]

Denis Kuzyutin[1,2], Nadezhda Smirnova[2,1]
and Elizaveta Tantlevskaya[1]

[1]Saint Petersburg State University

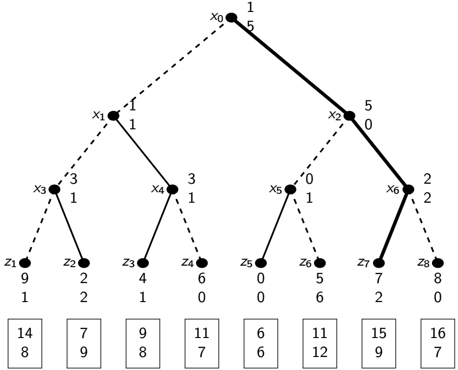
[2]HSE University

MSBC-2022
Vilnius, September 21–23, 2022

[1]The reported study was funded by RSF according to the research project 22-28-01221 (https://rscf.ru/project/22-28-01221/).

Motivating Example. Extensive game with terminal payoffs: S-P Core

$\bar{\omega} = (\bar{x}_0, \bar{x}_2, \bar{x}_6, z_7)$ is a cooperative path, (p_1, p_2) - is a distribution of the cooperative payoff, S is active coalition at $\bar{x}_t$:
$\gamma(S; \bar{x}_t) \leqslant \sum_{i \in S} p_i$.



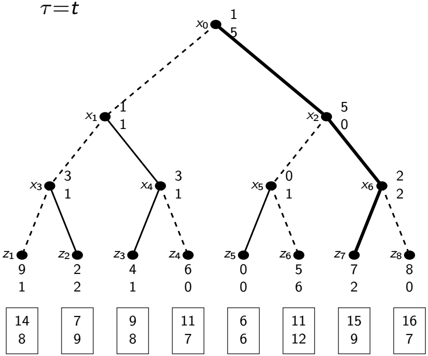
The S-P Core (Chander and Wooders, 2020) is empty since

$$p_1 + p_2 = 24, \ \gamma(\{1\}; \bar{x}_0) = 14 \leqslant p_1, \ \gamma(\{2\}; \bar{x}_2) = 12 \leqslant p_2.$$

Motivating Example. Extensive game with payoffs at each node

$$\beta_i(\bar{x}_\tau), \tau = 0, \dots, T \text{ — PDP for } (p_1, p_2)\text{: } \sum_{\tau=0}^{T} \beta_i(\bar{x}_\tau) = p_i, \ i \in N.$$

$$\gamma(S; \bar{x}_t) \leqslant \sum_{\tau=t}^{T} \beta_S(\bar{x}_\tau), \quad \bar{x}_t \in \bar{\omega}, \quad t = 0, \dots, T-1.$$



The β-S-P Core (Kuzyutin et al., 2021) is non-empty.

$$\widetilde{\beta}_1 = 15 + \varepsilon, \quad \widetilde{\beta}_2 = 9 - \varepsilon, \quad -1 \leqslant \varepsilon \leqslant 1.$$

Highlights

- We consider a competitive model of renewable resource extraction (RRE) as a multistage game with feedback information structure, where each player's stage payoff is *log* of the current extraction level (Lehvari and Mirman, 1980).
- We focus on the finite-horizon game when the players value differently the resource residual stock after the extraction process ends (the only source of asymmetry in our model).
- We explore the payoff distribution procedure (PDP) based approach (Petrosyan and Danilov, 1978) to reach and implement the cooperative agreement.

Highlights

- To derive non-cooperative and cooperative feedback strategies we use standard Dynamic Programming technique.
- We extend the novel β- subgame perfect core (β-S-P Core) concept to the class of multistage games.
- We introduce a refinement of the β-S-P Core and specify the quasi proportional PDP meeting several advantageous properties.
- Lastly, we provide a numerical example of the two-person multistage game of RRE.

The contributions of the paper is twofold:

- we derive analytical solution for specific finite-horizon multistage game of RRE with asymmetric players;
- we extend the β-S-P Core concept to n-person multistage games and provide an algorithm to construct quasi proportional PDP which belongs to non-empty β-S-P Core of a two-player game.

The Model. Non-cooperative behavior

Let $x(t)$ be a measure of the resource at time $t = 0, 1, \ldots, T$ (current stock level or state variable), while $u_j(t)$ denote player j's extraction level in that period (control variable).

$$H_j(\cdot) = \sum_{\tau=0}^{T-1} \delta^\tau \ln u_j(\tau) + K_j \, \delta^T \ln x(T), \qquad (1)$$

where $\delta \in (0,1)$ is a discount factor, and $K_j > 0$ is a parameter that specifies the player j's valuation of the resource residual stock.

$$x(t+1) = \alpha \cdot x(t) - \sum_{j=1}^{n} u_j(t, x(t)), \quad x(0) = x_0. \qquad (2)$$

$G^0(n, x_0, T)$ – multistage n-player game starting at time instant $t = 0$ with discrete dynamics (2), objective functions (1) and feedback information structure.

Each intermediate state $x(t),\ t = 0, \ldots, T-1$ determines a subgame $G^t(n, x(t), T)$ starting at time instant $\tau = t$ and initial state $x(t)$ with the subgame objective functions

$$H_j^t(\cdot) = \sum_{\tau=t}^{T-1} \delta^{\tau-t} \ln u_j(\tau) + K_j\ \delta^{T-t} \ln x(T), \quad j = 1, \ldots, n. \quad (3)$$

We employ the DP technique to determine the feedback-equilibrium strategies in multistage game $G^0(n, x(0), T)$.
The (present-valued) value function for player j in the subgame $G^t(n, x(t), T)$ satisfies

$$V_j(t,x) = \max_{u_j}\{\ln u_j + \delta \cdot V_j(t+1, \ \ \alpha x - u_j - \sum_{i \neq j} u_i^{SPE}(t,x))\}, \quad (4)$$

$$V_j(T,x) = K_j \cdot \ln x(T). \quad (5)$$

We guess the log-linear form of the value functions

$$V_j(t,x) = A_j(t) \ln x + B_j(t), \quad t = 0, 1, \ldots, T. \quad (6)$$

Prop. 1. A multistage finite-horizon game $G^0(n = 2, x(0), T)$ possesses a unique SPE

$$u_j(x) = \alpha \frac{A_{-j}(t+1)}{\varphi(t+1)} \cdot x, \quad j = 1, 2; \quad t = 0, \dots, T-1, \qquad (7)$$

where $\varphi(t+1) = A_j(t+1) + A_{-j}(t+1) + \delta A_j(t+1) A_{-j}(t+1)$, while coefficients $A_j(t)$ satisfy the recurrence formula

$$A_j(t) = 1 + \delta A_j(t+1), \quad A_j(T) = K_j. \qquad (8)$$

The SPE state trajectory is

$$x(t+1) = \frac{\alpha \delta A_j(t+1) \cdot A_{-j}(t+1)}{\varphi(t+1)} \cdot x(t), \quad t = 0, \dots, T-1. \qquad (9)$$

The value functions (6) represent the SPE payoffs in the subgame $G^t(n=2, x(t), T), t = 0, \dots, T-1$, while coefficients $B_j(t)$ satisfy the recurrence formula

$$B_j(t) = \Phi_j(\alpha, \delta, A_j(t+1), A_{-j}(t+1), B_j(t+1)) = \ln \frac{\alpha A_{-j}(t+1)}{\varphi(t+1)} +$$

$$+\delta[A_j(t+1) \cdot \ln \frac{\alpha\delta A_j(t+1) \cdot A_{-j}(t+1)}{\varphi(t+1)} + B_j(t+1)], \; B_j(T) = 0. \tag{10}$$

Remark 1. One can prove that multistage finite-horizon game $G^0(n, x(0), T)$, $n > 2$, still possesses a unique SPE, and moreover, the feedback equilibrium strategies $u_j(x)$ are proportional to x.

Cooperative behavior. β-S-P Core

Given nonempty coalition $S \subset N$, the induced multistage game $G_S^0(n - |S| + 1, x(0), T)$ describes the case when coalition S becomes a new player, i.e. all the players in S fully coordinate their strategies to maximize the total payoff of S

$$H_S(\cdot) = \sum_{\tau=0}^{T-1} \delta^\tau \ln \sum_{j \in S} u_j(\tau) + \sum_{j \in S} K_j \cdot \delta^T \ln x(T). \qquad (11)$$

Denote by $\gamma(S, t, x)$ the SPE payoff of coalition S in the induced subgame $G_S^t(n - |S| + 1, x(t), T)$, $t = 0, \ldots, T-1$.

Consider the fully cooperative solution when all the players cooperate to reach the maximal total payoff

$$H_N(\cdot) = \sum_{\tau=0}^{T-1} \delta^\tau \ln u(\tau) + K \cdot \delta^T \ln x(T), \qquad (12)$$

where

$$u(\tau) = \sum_{j \in N} u_j(\tau), \quad K = \sum_{j \in N} K_j.$$

Again we suppose the log-linear form of the value function:

$$V(t, x) = A(t) \cdot \ln x + B(t), \quad t = 0, 1, \ldots, T. \qquad (13)$$

Prop. 2. A multistage finite-horizon game $G^0(n=2, x(0), T)$ possesses a cooperative solution

$$u(x) = \frac{\alpha}{1+\delta A(t+1)} \cdot x, \quad t = 0, \dots, T-1, \tag{14}$$

while coefficients $A(t)$ satisfy recurrence formula

$$A(t) = 1 + \delta A(t+1), \quad A(T) = K. \tag{15}$$

The cooperative state trajectory is

$$x(t+1) = \frac{\alpha\delta A(t+1)}{1+\delta A(t+1)} \cdot x(t), \quad t = 0, \dots, T-1. \tag{16}$$

The value function (13) determines the cooperative payoff in the subgame $G^t(n=2, x(t), T)$, $t = 0, \dots, T-1$, while coefficients $B(t)$ are given by the recurrence formula

$$B(t) = \ln \frac{\alpha}{1+\delta A(t+1)} + \delta A(t+1) \ln \frac{\alpha\delta A(t+1)}{1+\delta A(t+1)} + \delta B(t+1), B(T) = 0. \tag{17}$$

Let $\bar{\omega} = (x(0) = \bar{x}(0), \ldots, \bar{x}(t), \ldots, \bar{x}(T))$ denote a cooperative state trajectory (16). A vector $(p_1^t, \ldots, p_n^t)$ such that

$$\sum_{i \in N} p_i^t = V(t, \bar{x}(t)) \tag{18}$$

specifies a possible sharing rule to distribute the total cooperative (subgame) payoff between the players and could be considered as a cooperative solution for the subgame $G^t(n, \bar{x}(t), T)$.

Def. 1. Vectors $\beta_i(\bar{\omega}) = (\beta_i(\bar{x}(\tau)))$, $\tau = 0, \ldots, T$; $i = 1, \ldots, n$, is a Payoff Distribution Procedure (PDP) for the cooperative solution $(p_1^0, \ldots, p_n^0)$ if

$$p_i^0 = \sum_{\tau=0}^{T} \delta^\tau \beta_i(\bar{x}(\tau)), \tag{19}$$

while $\beta_i(\bar{x}(\tau))$ denotes the actual current payment that the i-th player should get at time τ when the players use PDP β under cooperative scenario.

We adopt in the paper the following assumption about the players non-cooperative behavior if a cooperative agreement is broken down at some intermediate time instant $t = 0, \ldots, T-1$, because of some coalition S deviation from cooperative scenario:

- all the players $j \in N \setminus S$ form singletons and switch (immediately and till the end of game) to non-cooperative (that is, SPE) behavior scheme in the subgame $G^t(n, \bar{x}(t), T)$.

Then the maximal payoff a coalition S could expect in $G^t(n, \bar{x}(t), T)$ in case of its deviation equals to $\gamma(S, t, \bar{x}(t))$ instead of $\sum_{\tau=t}^{T} \delta^{\tau-t} \beta_S(\bar{x}(\tau))$.

Def. 2. A PDP $\beta = (\beta_i(\bar{x}(\tau)))$, $i = 1, \dots, n$; $\tau = 0, \dots, T$, belongs to the β-Subgame-Perfect Core (β-S-P Core) of the multistage finite-horizon game $G^0(n, x(0), T)$ if for each nonempty coalition $S \subset N$ and each intermediate time instant $t = 0, \dots, T-1$ the following inequality holds

$$\sum_{\tau=t}^{T} \delta^{\tau-t} \beta_S(\bar{x}(\tau)) \geqslant \gamma(S, t, \bar{x}(t)). \tag{20}$$

Prop. 3. β-Subgame-Perfect Core of a multistage finite-horizon game $G^0(2, x(0), T)$ is non-empty.

Remark 2. The $n \times (T+1)$ components $\beta_j(\bar{x}_\tau)$ of the PDP β from β-S-P Core have to satisfy a system of non-strict linear inequalities (20) and linear equations (19). Hence, a non-empty β-S-P Core for multistage finite-horizon game $G^0(n, x(0), T)$ is a convex closed polytope Δ in $R^{n\times(T+1)}$.

Def. 3. A payoff distribution procedure β satisfies the strict balance constraints if

$$\begin{aligned} \sum_{j\in N} \beta_j(\bar{x}(\tau)) &= \ln \bar{u}(\bar{x}(\tau)), \quad \tau = 0, \ldots, T-1; \\ \sum_{j\in N} \beta_j(\bar{x}(T)) &= K \cdot \ln \bar{x}(T). \end{aligned} \tag{21}$$

How to choose a unique PDP from β-S-P Core?

If we apply maxmin RBC approach to multistage finite-horizon game $G^0(n, x(0), T)$ and focus on the case when $\gamma(\{i\}, 0, \bar{x}(0)) > 0, \; i \in N$, we need to solve the following optimization problem

$$\max_{\beta \in \Delta} \min_{i \in N} \frac{p_i^0 - \gamma(\{i\}, 0, \bar{x}(0))}{\gamma(\{i\}, 0, \bar{x}(0))} \tag{22}$$

and then distribute each player i's cooperative payoff $p_i^0 = \sum_{\tau=0}^{T} \delta^\tau \beta_i(\bar{x}(\tau))$ along the cooperative trajectory in such a way that PDP β meets (20) and (21).

Maxmin RBC for two-player game

Remark 3. Note that for two-player game problem (22) takes the following simple form

$$\frac{p_1^0 - \gamma(\{1\}, 0, \bar{x}(0))}{\gamma(\{1\}, 0, \bar{x}(0))} = \frac{p_2^0 - \gamma(\{2\}, 0, \bar{x}(0))}{\gamma(\{2\}, 0, \bar{x}(0))}. \tag{23}$$

Algorithm (quasi proportional PDP from β-S-P Core)

1. Using Prop. 2 find a cooperative trajectory $\bar{\omega} = (\bar{x}(0), \bar{x}(1), \ldots, \bar{x}(T-1), \bar{x}(T))$ and corresponding sequence of cooperative extraction levels $\bar{u}(\bar{x}(t)), \quad t = 0, \ldots, T-1$.
2. Calculate $\gamma(\{j\}, t, \bar{x}(t)), \quad t = 0, \ldots, T-1; \quad j = 1, 2$, using (6), (8) and (10) in accordance with Prop. 1.
3. Solve (23) and (18) to obtain p_1^0 and p_2^0.

Algorithm (quasi proportional PDP from β-S-P Core)

4. Using strict balance constraints (21) and inequalities (20) write the system of double inequalities for $\sum\limits_{\tau=t}^{T} \delta^{\tau-t}\beta_1(\bar{x}(\tau)), \quad t = T-1, T-2, \ldots, 1$ in the form:

$$\begin{cases} c_1^{T-1} \leqslant \beta_1(\bar{x}(T-1)) + \delta \cdot \beta_1(\bar{x}(T)) \leqslant C_1^{T-1} \\ \vdots \\ c_1^{t} \leqslant \beta_1(\bar{x}(t)) + \delta \cdot \beta_1(\bar{x}(t+1)) + \ldots + \delta^{T-t} \cdot \beta_1(\bar{x}(T)) \leqslant C_1^{t} \\ \vdots \\ c_1^{1} \leqslant \sum\limits_{\tau=1}^{T} \delta^{\tau-1} \cdot \beta_1(\bar{x}(\tau)) \leqslant C_1^{1} \end{cases} \tag{24}$$

where $c_1^t \leqslant C_1^t$, for all $t = 1, \ldots, T-1$.

Algorithm (quasi proportional PDP from β-S-P Core)

5. Denote by μ the first player's part of the total cooperative payoff $\frac{p_1^0}{p_1^0+p_2^0}$. Then $\frac{p_2^0}{p_1^0+p_2^0} = 1-\mu$.
   Accept $\beta_1(\bar{x}(T)) = \mu \cdot K \cdot \ln \bar{x}(T)$.
6. Solve (24) in series assuming that in each subgame $G^t(2, \bar{x}(t), T)$, $t = T-1, T-2, \dots, 1$, player 1 receives exactly part μ of the admissible range $(C_1^t - c_1^t)$ of the subgame payment. Namely,

$$\begin{cases} \beta_1(\bar{x}(T-1)) = c_1^{T-1} + \mu(C_1^{T-1} - c_1^{T-1}) - \delta\beta_1(\bar{x}(T)) \\ \vdots \\ \beta_1(\bar{x}(1)) = c_1^1 + \mu(C_1^1 - c_1^1) - \sum_{\tau=2}^{T} \delta^{\tau-1}\beta_1(\bar{x}(\tau)) \end{cases} . \quad (25)$$

Algorithm (quasi proportional PDP from β-S-P Core)

7. Take

$$\beta_1(\bar{x}(0)) = p_1^0 - \sum_{\tau=1}^{T} \delta^\tau \beta_1(\bar{x}(\tau)). \tag{26}$$

8. Calculate $\beta_2(\bar{x}(t))$, $t = 0, \ldots, T$, from (21).

Remark 4. Payoff distribution procedure β specified above satisfies the following properties:

- it belongs to the β-S-P Core of multistage game $G^0(2, \bar{x}(0), T)$;
- the resulting cooperative solution (p_1^0, p_2^0) maximizes the relative benefit from cooperation (22) of the least winning player;
- it meets the strict balance constraints (21);
- PDP β implements a reasonable and subgame-consistent sharing rule in a sense that in each intermediate state $\bar{x}(t))$, $t = 1, \ldots, T$, the first player receives the same share of the current range $(C_1^t - c_1^t)$ of the admissible subgame $G^t(2, \bar{x}(t), T)$ payment $\sum_{\tau=2}^{T} \delta^{\tau-1} \beta_1(\bar{x}(\tau))$ as he is expected to obtain in the whole game $G^0(2, \bar{x}(0), T)$ in accordance with the cooperative solution (p_1^0, p_2^0).

Numerical example

Parameters values: $T = 2$ $(t = 0, 1, 2)$, $\alpha = 1.5$, $\delta = 0.95$, $K_1 = 1$, $K_2 = 0.5$, $K = K_1 + K_2 = 1.5$. The first player is more concerned about the resource residual stock (and hence, he is more interested in cooperation).

Table: Table 1. SPE strategies and cooperative strategy

t	u_1^{SPE}	u_2^{SPE}	u^{Coop}
0	$0.3593 \cdot x_0$	$0.475 \cdot x_0$	$0.454 \cdot x_0$
1	$0.2528 \cdot x_0$	$0.5056 \cdot x_0$	$0.647 \cdot x_0$

The relative values of these strategies at time instants $t = 0$ and $t = 1$ (current extraction levels divided by $\bar{x}_0$) are presented in Fig. 1 and connected by a dashed lines for visual clarity.

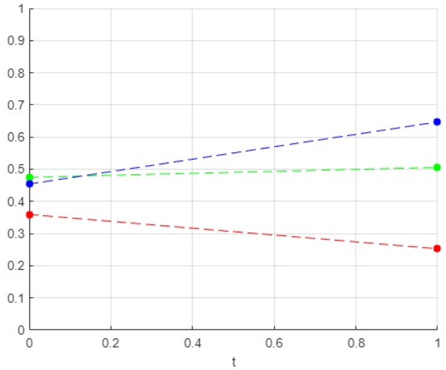


Figure: Fig. 1. SPE strategies – for the first player (red), for the second player (green), cooperative strategy (blue)

Table: Table 2. SPE trajectory and cooperative trajectory

t	x^{SPE}/x_0	x^{Coop}/x_0
0	1	1
1	0.6656	1.046
2	0.2401	0.922

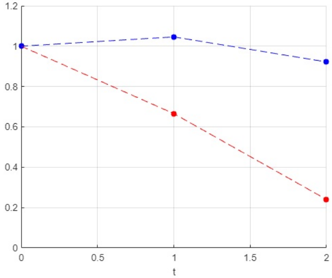


Figure: Fig. 2. SPE (red) and cooperative (blue) state trajectories

To compare the sum of the players' SPE payoffs (6) and the cooperative payoff (13) in the whole game $G^0(2, \bar{x}(0), T)$ and in the subgames $G^t(2, \bar{x}(t), T), t = 1, 2$, along the cooperative trajectory we fix initial state $x_0 = e^{1.5} \approx 4.4817$. The results are presented in Table 3.

Table: Sum of the players' SPE payoffs versus cooperative payoff along cooperative trajectory

t	$V_1(t, \bar{x}(t)) + V_2(t, \bar{x}(t))$	$V(t, \bar{x}(t))$
0	2.227	3.642
1	2.595	3.086
2	2.128	2.128

Following the algorithm and using (18) and (23) we receive conditions on p_1^0, p_2^0:

$$\frac{p_1^0 - 0.66}{0.66} = \frac{p_2^0 - 1.56}{1.56}, \; p_1^0 + p_2^0 = 3.64,$$

from where we get $p_1^0 = 1.08, \; p_2^0 = 2.56$. Then, system of inequalities (24) takes the following form:

$$1.076 \leqslant \beta_1(\bar{x}(1)) + 0.95 \cdot \beta_1(\bar{x}(2)) \leqslant 1.567.$$

Coefficient μ is equal to 0,297 for this game. Further, using (25), (26) and the strict balance constraints (21) we obtain quasi proportional PDP from β-S-P Core:

	$\beta_i(\bar{x}(0))$	$\beta_i(\bar{x}(1))$	$\beta_i(\bar{x}(2))$
$i = 1$	-0.079	0.621	0.632
$i = 2$	0.789	0.443	1.496

Note that a negative payment to some player in accordance with (26) can only arise in the initial state of a multistage game (when the players just enter into a cooperative agreement).

Breton, M., Keoula, M. Y.: A great fish war model with asymmetric players. Ecological Economics **97**, 209–223 (2014)

Chander, P., Wooders, M.: Subgame-perfect cooperation in an extensive game. Journal of Economic Theory **187**, 105017 (2020)

Kuzyutin, D., Smirnova, N., Skorodumova, Y.: Implementation of subgame-perfect cooperative agreement in an extensive-form game. In: Petrosyan, L.A., Zenkevich, N.A. (eds.) Contrib. to Game Theory and Manag., pp. 257–272. St. Petersburg State Univ., St. Petersburg, Russia (2021).

Mazalov, V.V., Rettieva, A.N.: Asymmetry in a cooperative bioresource management problem. In: Game-Theoretic Models in Mathematical Ecology, pp. 113–152. Nova Science Publishers (2015).

Petrosyan, L.A., Danilov, N.N.: Stability of solutions in non-zero sum differential games with transferable payoffs. Astronomy **1**, 52–59 (1979)

Petrosyan, L., Kuzyutin, D.: Games in Extensive Form: Optimality and Stability. Saint Petersburg University Press, St. Petersburg, Russia (2000), (in Russian)

Rettieva, A. N.: A discrete-time bioresource management problem with asymmetric players. Automation and Remote Control **75**(9), 1665–1676 (2014)

Rettieva, A.: Cooperation in bioresource management problems. In: Petrosyan, L.A., Zenkevich, N.A. (eds) Contributions to Game Theory and Management, vol. 10, pp. 245—286. St. Petersburg State Univ., St. Petersburg, Russia (2017)

Thank you for the attention!