FACULTY OF LAW, ECONOMICS AND FINANCE

CREA Discussion Paper 2014-26

Center for Research in Economics and Management
University of Luxembourg

Stability and Cooperative Solution in Stochastic Games

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Elena M. Parilina, Saint Petersburg State University
Alessandro Tampieri, CREA, Université de Luxembourg

November, 2014

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie
L-1511 Luxembourg

UNIVERSITÉ DU LUXEMBOURG

Stability and Cooperative Solution in Stochastic Games*

Elena M. Parilina† and Alessandro Tampieri‡

November 28, 2014

Abstract

Cooperative game theory is effective in explaining many economic interactions, such as risk-sharing agreements or the enforcing role of social norms. In a stochastic environment, the analysis of these issues is generalised by taking into account the presence of shocks. The paper finds the conditions of dynamic stability for cooperative stochastic games. Principles of dynamic stability include three conditions: subgame consistency, strategic stability and irrational-behaviour-proof of the cooperative agreement.

JEL codes: C71, C73.

Keywords: cooperative stochastic game, stationary strategies, time consistency, subgame consistency, strategic stability, irrational-behaviour-proof.

*We are grateful to Stefania Bortolotti, Andreas Irmen, Anastasia Litina, Arsen Palestini, Leon Petrosyan, Pierre Picard, Giuseppe Pignataro and Benteng Zou for helpful comments. A previous version of this paper was presented at Game Theory and Management Conference (St. Petersburg, 2011). The usual disclaimer applies.

†Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, Universitetskii prospekt 35, Petergof, Saint-Petersburg, Russia 198504; e.parilina@spbu.ru.

‡Faculty of Law, Economics and Finance, University of Luxembourg, Avenue de la Faïencerie 162a, L - 1511 Luxembourg; tamp79@gmail.com.

1 Introduction

In this paper we establish the conditions of stable cooperation in stochastic games. The starting point is the analysis of Petrosyan and Zenkevich (2009), who considers three principles of stable cooperation for dynamic games: "time consistency", "strategic stability" and "irrational-behaviour-proof". We apply the same principle to stochastic games.[1]

In particular, the paper considers stochastic games in stationary strategies with the finite number of states, any of which could be realised at every game stage. With this class of stochastic games, the claim of subgame consistency lays into the cooperative agreement.[2] In other words, consistency of cooperative agreement should take place in every position (state) of the game. Subgame consistency of the cooperative agreement let players be expecting the allocation from the same optimality principle in every stochastic subgame. The condition of strategic stability (Grauer and Petrosyan, 2002) ensures the existence of a Nash equilibrium in the regularised game with the payoffs that players expect to receive as a result of the cooperative agreement. We build up the regularisation of the game on the basis of the initial stochastic game with the use of the "payoff distribution procedure" (Petrosyan and Danilov, 1979). Finally, the irrational-behaviour-proof condition (Yeung, 2006) ensures that, if some player (group of players) deviates from the agreement and players play individually from this stage to the end, then they will receive not less than if each player plays individually during the whole game.

The definition of stochastic game was introduced by Shapley (1953a). Afterwards, the analysis of stochastic games developed in several strands of the economic literature (See Amir 2001 and 2003 for discussions). Kirman and Sobel (1974) model a dynamic oligopoly with random variable demand as a stochastic game. The concept of correlated equilibrium (Aumann, 1987) is adopted for the class of stochastic games in stationary strategies with finite set of players (Solan and Vieille, 2002). Stochastic games with infinite horizon allows to model the dynamic conflicts with uncertainties in the form of "shocks" with existing ergodic Markov equilibria for stochastic overlapping generations (Duffie *et al.*, 1994). Still in the overlapping generations framework, Messner and Polborn (2003) find the existence

[1] Petrosyan (1977) introduced the concept of dynamic consistency for differential games. This condition appeared to be relevant also for stochastic games (Petrosyan, 2006).

[2] The analogous concept of dynamic consistency in stochastic games is called "subgame consistency". The reason is the following. Unlike dynamic games, characterised by continuous time, stochastic games exhibit discrete times, and players may change their behaviour in the states which are the initial points of the subgames.

of a cooperative equilibrium. Finally, the existence of stationary Markov perfect equilibria in risk sensitive stochastic games was proved by Jaskiewicz and Nowak (2014). Stochastic games have been applied also in analysing tax evasion (Raghavan, 2006), and in the field of telecommunication system modeling (Parilina, 2010, Altman *et al.*, 2003).

The paper contributes to the literature on stochastic games by proposing the conditions of stability in a cooperative agreement. This result is very general and can be implemented in several real-world situations, and in many applications of the economic literature. For instance, a stochastic game allows to consider shocks in the environment (such as social shocks) whereas cooperative games have been already adopted to represents many economic issues, like risk-sharing agreements and the enforcing role of social norms.[3] Given the general purpose of this work, we do not enter into the details of a specific issue, by leaving the application of this setting to future research.

The remainder of the paper is organised as follows. Section 2 explains some applications of our framework. We begin by showing examples in which a stochastic framework is an effective analysis tool. The presence of social shocks justifies the use of stochastic games, in which states of the world randomly change. Applications of cooperative games are, for instance, risk-sharing agreements and social norms. Some stylised facts in which social shocks and social norms interact are also presented. Section 3 introduces the model, while Section 4 describes the cooperative equilibrium. Section 5 proves the conditions for stability of the cooperative solution, while Section 6 shows an example of stability. Section 7 concludes. All proofs can be found in the appendix.

2 Applications of stochastic cooperative games

2.1 Social shocks as a stochastic environment

In the real world, many situations, interactions and events are well modelled through stochastic games. A stochastic-game framework is particularly effective to interpret individuals' decisions in the presence of social shocks. Social shocks emerge as an abrupt discrete change (a "jump") in some aspect of the society. Examples of social shocks are:

- Economic crises influencing lifestyle: macroeconomic shocks may influence an individual's endowment through education and health, income or differences in labour

[3] Cooperative game theory is employed also in the analysis of inequality indexes. See Charpentier and Mussard (2011) for a discussion and comparisons between indexes.

market opportunities. Individual endowments, in turn, are known to be an important variable in explaining the formation of beliefs and attitude towards institutions (Di Tella, Galiani, and Schargrodsky 2007).

- Introducing legislative norms that affect social behaviour, such as banning tobacco in public places, anticorruption measures or tax-evasion campaigns.
- Some periods of the year with specific religious implication, inducing people to fasting, like Ramadan for Muslims or Lent for Christians, or to eating more than usual, like Christmas, Easter or Diwali (Chen, 2014).

The presence of social shocks involves changes in the game environment. Hence, a purely repeated game model cannot be employed to correctly represents these scenario. In particular, in a stochastic game, the presence of (positive or negative) social shocks depends on the state of the game.

2.2 Cooperative solution

Cooperative game theory is based on the idea that agents choose a joint strategy for mutual benefit. For example, profit-seeking firms in this framework are coalitions that manifest cooperative solutions just as cooperative organisations are. In this section we explore some scenarios in which a cooperative solution can be applied. We focus on two situations: risk-sharing arrangements and enforcing social norms.

2.2.1 Risk-sharing arrangements

A cooperative game-theoretic setting is useful to describe the formation of risk sharing arrangements. As a response to the large fluctuations in their income, individuals from developing countries often enter into informal insurance or quasi-credit arrangements. In order to be self enforcing, the expected net benefits from participating in the agreement must be at any point in time larger than the one time gain from defection.

The literature on risk-sharing without commitment in rural societies started with the suggestions of Posner (1980) and Kimball (1988) that schemes of mutual insurance with limited commitment were possible. Coate and Ravaillon (1993) characterised mutual insurance arrangements with a restriction to stationary transfers for a symmetric two-household model. Another strand of literature investigates efficient dynamic contracts in the absence of commitment (Kocherlakota, 1996, Kletzer and Wright, 2000, Ligon *et al.*, 2002). In an important paper, Genicot and Ray (2003) study informal insurance within communities, explicitly recognising the possibility that subgroups of individuals may destabilise insurance arrangements that are robust not only to single person deviations but also to potential deviations by subgroups.

Of course, the presence of social shocks is the very reason for which risk-sharing agreements emerge. Hence the implementation of stochastic games to these issues seems particularly natural.

2.2.2 Social norms

Like non-cooperative game theory, cooperative game theory is best thought as a rational action hypothesis, but the concept of rationality is different. Indeed neither of these rational action hypotheses corresponds well with experimental evidence (McKelvey and Palfrey, 1992, Fehr and Fischbacher, 2004, Fehr and Gachter, 2002, and Hoffman *et al.*, 1998, *inter alia*). However, when the self-interest hypothesis considers the presence of social norms, the experimental evidence can largely be accounted for (McCain, 2008).

In many social contexts, the presence of social norms allows to reach a cooperative solution to the game. If a cooperative solution is seen as a common strategy to improve the payoffs to the members of the group, then social norms supporting the common strategy would be part of a "cooperative solution" of the game. Social norms work as are commitments for the players, and to the extent that people are rational in the sense of cooperative game theory they will be carried out.

An analysis going in this direction is McCain (2007), who develops a game of effort determination in a cooperative enterprise, in which productive efficiency requires an effort commitment by each individual that is greater than the commitment the person would choose on the basis of pure self-interest in a non-cooperative solution of the game. McCain (2007) supposes that there is a norm of effort commitment, and suppose that the norm is the effort commitment required for efficient production. Reciprocity motives lead to one or the other of two "solutions": in one, everyone obeys the norm, production is efficient, and there is no retaliation. The other replicates the noncooperative solution: everyone chooses the effort commitment on the basis of self-interest, without regard to social norms. López-Pérez (2008) explicitly introduces social norms in games, assuming that they shape some of the players' utility and beliefs. People feel bad by deviating from a binding norm, and they are sensitive to their deviation relative to how much the others deviate. He then studies how social norms and emotions affect cooperation, coordination, and punishment in a variety of games.

2.3 Social shocks and social norms: some facts

In this section we propose some stylised facts in the real world where social shocks and social norms interact. The idea is to show how the presence of social shocks can modify an individual's perception of his or her social behaviour. An example of interaction between social shocks and social norms is the relevance of macroeconomic shocks on the determination of attitudes toward the state, and ultimately different welfare systems. The welfare system established after the Great Depression was a radical break from the strong sense of individualism and self-reliance characterising American society. Many European countries also moved from partial or selective provision of social services to relatively comprehensive coverage of the population during the same period. The shock in turn affected the social beliefs and norms: individuals who experienced a recession when young believe that success in life depends more on luck than effort, support more government redistribution and tend to vote for left parties (Giuliano and Spilimbergo, 2012).

Another example of interaction between social norms and social shocks is the introduction of anti-corruption measures (Litina and Palivos, 2013). Anti-corruption campaigns aim at increasing awareness and sensitiveness of the public. One first attempt was the so-called San Fan (the Three Anti), being initiated in China in 1951. The three "antis" promoted were anti-corruption, anti-waste and anti-bureaucracy. All citizens and members of the Chinese Communist Party were mobilised to inspect and report corrupt activities. Offenders were widely exposed and in some instances severely punished (Spence, 1991).

In some cases, the presence of social norms in turn lead to social shocks. This is the case of gifts in social ceremonies, such as weddings and funerals, in some communities (Chen, 2014). Presenting gifts can be a financial burden in years with frequent social events. Ceremonies are often very costly, and on average cost more than twice the gifts received or several times of average income (Chen *et al.*, 2011). Because the poor often lack the necessary resources, they are forced to cut back on basic consumption in order to afford a gift to attend social festivals. For instance, in South Africa and Ghana, poor families often spend so profusely on funerals that they skimp on food for months afterwards (The Economist, 24/05/2007, Case *et al.*, 2013). Families in rural China are unable to afford a refrigerator yet spent on gift they can barely cover (The Economist, 28/11/2013). Costly entertainment, social festivals and malnutrition are ubiquitous in many developing countries (Banerjee and Duflo, 2007).

Finally, the introduction of anti-tax evasion campaigns. In order to increase peer pressure on tax evaders, governments often make publicly available the tax filling reports or publicise cases of tax evasion. For instance, in Finland, Norway and Sweden personal income tax fillings are publicly available (see Lenter *et al.*, 2003). In India a tax amnesty took place in 1997. This brought a substantial increase in revenue, mostly because the state hired marketing companies that used moral suasion to increase tax compliance (Torgler, 2004).

3 The model

Consider a stochastic game with n individuals, denoted by i and whose set is $N = \{1, \ldots, n\}$. Time is discrete and denoted by $j = 1, ...t$. In each stage of the game, individuals (also called "players") perform some actions. We denote the action profile as $x^j = (x_1^j, \ldots, x_n^j)$.

One of the finite number of states is realised at each stage of the stochastic game. The state of the stochastic game is determined as in a simultaneous normal form game of n players. Formally, the set of states is denoted by $\{\Gamma^j\}_{j=1}^t$, where $\Gamma^j = \langle N, X_1^j, \ldots, X_n^j, K_1^j, \ldots, K_n^j \rangle$ is state j, and X_i^j is the finite set of pure strategies of player $i \in N$ in Γ^j. The number of players N is the same for all Γ^j, $j = 1, \ldots, t$.

Each state corresponds to the presence of (positive or negative) shocks of different size. This will be reflected on the players' payoff. A player i's payoff function in state Γ^j, $j = 1, \ldots, t$ is denoted as

$$K_i^j(x_1^j, \ldots, x_n^j) = K_i^j(x^j).$$

Since the game is stochastic, each state is realised according to a certain distribution. If in the previous stage with state Γ^j, the action profile $x^j = (x_1^j, \ldots, x_n^j)$ has been realised, then the probability that state Γ^k is realised is $p(j, k; x^j)$, $j, k = 1, \ldots, t$. Note that $p(j, k; x^j) \geqslant 0$ and $\sum_{k=1}^t p(j, k; x^j) = 1$ for each $x^j \in X^j = \prod_{i \in N} X_i^j$ and for any $j, k = 1, \ldots, t$. We denote the vector of the initial distribution on states $\Gamma^1, \ldots, \Gamma^t$ as $\pi^0 = (\pi_1^0, \ldots, \pi_t^0)$, where π_j^0 is the probability that state Γ^j is realised in the first stage of the game, $\sum_{j=1}^t \pi_j^0 = 1$.

Along the paper, we will focus on stationary strategies. We denote the set of player i's stationary strategies as $\Xi_i = \{\eta_i\}$. This implies that, in each stage, the player's strategy in every state from set $\left\{\Gamma^1, \dots, \Gamma^t\right\}$ depends only on which state is realised at this stage, that is $\eta_i : \Gamma^j \longmapsto x_i^j \in X_i^j, j = 1, \dots, t$.

Begin by describing the non-cooperative game.

Definition 1 *A finite stochastic game in stationary strategies is defined as*

$$G = \left\langle N, \left\{\Gamma^j\right\}_{j=1}^t, \{\Xi_i\}_{i \in N}, \delta, \pi^0, \left\{p(j,k;x^j)\right\}_{j=\overline{1,t}, k=\overline{1,t}, x^j \in \prod_{i=1}^n X_i^j} \right\rangle. \tag{1}$$

Definition 2 *A finite stochastic subgame in stationary strategies G^j, $j = 1, \dots, t$, beginning with state Γ^j is a stochastic game* (1) *with vector $\pi^0 = (0, \dots, 0, 1, 0, \dots, 0)$ (with 1 in the j^{th} component).*

Note that a player i's stationary strategy in G is the player i's stationary strategy in any subgame $G^1, \dots, G^t$. A player's indirect utility function is then defined as the expectation of the player's payoff in stochastic game G. Let $E_i^j(\eta)$ be the expected payoff of player i in G^j when a strategy profile η is realised in the stochastic game G (subgame G^j), whose vectorial form is $E_i(\eta) = (E_i^1(\eta), \dots, E_i^t(\eta))$.

Hence a player i's indirect utility function in the subgame G^j has the shape of the following recurrent equation:

$$E_i^j(\eta) = K_i^j(x^j) + \delta \sum_{k=1}^t p(j,k;x^j) E_i^k(\eta), \tag{2}$$

$$\text{s.t. } \eta(\Gamma^j) = x^j, i.e. \eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot)),$$

where $\delta \in (0,1)$ is the discount factor, and $\eta_i(\Gamma^j) = x_i^j \in X_i^j$, $x^j = (x_1^j, \dots, x_n^j)$ for each $j = 1, \dots, t$, $i \in N$. The fact that the stochastic game G is considered in the class of stationary strategies and the set of states $\{\Gamma^1, \dots, \Gamma^t\}$ is finite allows us to consider only t of subgames $G^1, \dots, G^t$ beginning with the states $\Gamma^1, \dots, \Gamma^t$. Hereinafter, let $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$ be the stationary strategy profile such as $\eta_i(\Gamma^j) = x_i^j \in X_i^j$ where $j = 1, \dots, t$, $i \in N$. We restrict our attention to the set of player i's pure stationary strategies in stochastic game G. Denote it as $\widetilde{\Xi}_i$.

Now, define the matrix of transition probabilities $t \times t$ in G under the realization of $\eta(\cdot)$ as:

$$\Pi(\eta) = \begin{pmatrix} p(1,1;x^1) & \dots & p(1,t;x^1) \\ p(2,1;x^2) & \dots & p(2,t;x^2) \\ \dots & \dots & \dots \\ p(t,1;x^t) & \dots & p(t,t;x^t) \end{pmatrix}, \tag{3}$$

where the element of the j^{th} row and the j'^{th} column is the probability to transit from state j^{th} to state j'^{th}.

Given (3), we can rewrite equation (2) in matrix form as follows:

$$E_i(\eta) = K_i(\eta) + \delta\Pi(\eta)E_i(\eta), \tag{4}$$

where $K_i(\eta) = (K_i^1(x^1), \dots, K_i^t(x^t))$, and $K_i^j(x^j)$ is the player i's payoff in state Γ^j conditional to $x^j \in X^j$ is realized in this state. Equation (4) is equivalent to:

$$E_i(\eta) = (\mathbb{I} - \delta\Pi(\eta))^{-1} K_i(\eta), \tag{5}$$

where $\mathbb{I}$ is an identity $t \times t$ matrix. A technical consideration is the following.

Lemma 1 *Matrix $(\mathbb{I} - \delta\Pi(\eta))^{-1}$ always exists for $\delta \in (0,1)$.*

The expected payoff of player i is then:

$$\overline{E}_i(\eta) = \pi^0 E_i(\eta). \tag{6}$$

4 Cooperative solution

Suppose now that players decide to cooperate by forming a coalition, denoted by $S \subset N$. Then the pure strategy profile maximising the sum of the expected players' payoffs in G is denoted as $\overline{\eta}(\cdot) = (\overline{\eta}_1(\cdot), \dots, \overline{\eta}_n(\cdot))$, where

$$\max_{\eta \in \prod_{i \in N} \widetilde{\Xi}_i} \sum_{i \in N} \overline{E}_i(\eta) = \sum_{i \in N} \overline{E}_i(\overline{\eta}). \tag{7}$$

The problem depicted in (7) may have more than one decision. We call the strategy profile $\overline{\eta}(\cdot)$ satisfying (7) as cooperative decision. In order to form a coalition, all players should decide not only what strategies to use to maximise the joint payoff but also how to allocate it. In cooperative game theory, the value function of the entire coalition is given for any coalition S and it is called the "characteristic function". The standard way to proceed is determining the characteristic function which gives the value of guaranteed payoff for each coalition. Indeed, the problem is to allocate the payoff of coalition among members in a "fare" way. The allocation needs to satisfy the following conditions: (i) any player should obtain no less than she may get by non-cooperative play ("individual rationality condition") and (ii) the sum of allocations corresponds to the characteristic function ("group rationality condition"). The aim of this section is to determine the characteristic function for any coalition S.

The coalitional form of the non-cooperative game described in the previous section is given by the pair $\langle N, V \rangle$, where N is the set of players and V is a real-valued function, called the characteristic function of the game, defined on the set 2^N (the set of all subsets of N). The value $V(S)$ is a real number for each coalition $S \subset N$, which may be interpreted as the value of coalition S when its members play together as a unit.

The characteristic function satisfies two properties:

Assumption 1 $V(\emptyset) = 0$.

Assumption 2 *For any disjoint coalitions* $S, T \subset N$, *and* $S \cap T = \emptyset$, *then* $V(S) + V(T) \leqslant V(S \cup T)$.

Assumption 1 implies that, if the coalition set is empty, then the weight of the coalition is nil. Assumption 2 implies "super-additivity", and says that the value of two disjoint coalitions is at least as great when they play together as when they act non-cooperatively. If this condition is not satisfied, it may be explained as coalition $S \cup T$ is not profitable, thus it will not be formed. The assumption of super-additivity is not needed and it is often omitted in cooperative game theory, because in real life there are a lot of motivations to consider both profitable and non-profitable coalitions.[4]

We define the characteristic function $\overline{V}(S)$ through the characteristic function $V^j(S)$ of stochastic subgames G^j, $j = 1, \ldots, t$, as follows:

$$\overline{V}(S) = \pi^0 V(S) \tag{8}$$

[4]As Aumann and Dreze (1974, p. 233) note, there are arguments for superadditivity that are quite persuasive, but, as they also note, superadditivity is quite problematic in some economic applications.

for any coalition $S \subset N$, where $V(S) = (V^1(S), \ldots, V^t(S))$, and $V^j(S)$ is the value of the characteristic function of the stochastic subgame G^j derived for S.

Begin the analysis by examining a coalition formed by the entire number of players, $S = N$. The Bellman equation of the characteristic function $V(N)$ represents the expected payoff of the entire population discounted over time, and it can be written as follows:

$$V(N) = \max_{\eta \in \prod\limits_{i \in N} \widetilde{\Xi}_i} \left[\sum_{i \in N} K_i(\eta) + \delta \Pi(\eta) V(N) \right] = \sum_{i \in N} K_i(\overline{\eta}) + \delta \Pi(\overline{\eta}) V(N),$$

where $\overline{\eta}(\cdot)$ is the pure strategy profile satisfying condition (7). From equation (5), we can infer the matrix form of $V(N)$:

$$V(N) = (\mathbb{I} - \delta \Pi(\overline{\eta}))^{-1} \sum_{i \in N} K_i(\overline{\eta}). \tag{9}$$

One way to determine the value of the characteristic function $V^j(S)$ for this coalition, $j = 1,\ldots,t$, for each subgame G^j, is to assume the "worst" situation for coalition S. This occurs when some players cooperate in coalition $N \backslash S$ against S. Since the two coalitions S and $N \backslash S$ play against each other, the payoff of coalition S is the loss of $N \backslash S$. Hence we introduce an auxiliary zero-sum stochastic game G^j_S where a coalition $S \subset N$ plays as a maximising player and coalition $N \backslash S$ plays as a minimising player. Define the value of function $V^j(S)$ for subgame G^j as a lower value of antagonistic stochastic game G^j_S in pure stationary strategies:[5]

$$V^j(S) = \max_{\eta_S} \min_{\eta_{N \backslash S}} \sum_{i \in S} E^j_i(\eta_S, \eta_{N \backslash S}), \tag{10}$$

where

- $(\eta_S(\cdot), \eta_{N \backslash S}(\cdot))$ is a strategy profile in pure stationary strategies;
- $\eta_S(\cdot) = (\eta_{i_1}(\cdot), \ldots, \eta_{i_k}(\cdot))$ is a vector of stationary strategies of players $i_1, \ldots, i_k \in S$, $S = i_1 \bigcup \ldots \bigcup i_k$, $\eta_S(\cdot) \in \prod\limits_{j=1}^{k} \widetilde{\Xi}_{i_j}$, where $\prod\limits_{j=1}^{k} \widetilde{\Xi}_{i_j}$ is the set of pure stationary strategies of coalition S;
- $\eta_{N \backslash S}(\cdot)$ is a vector of stationary strategies of players $i_{k+1}, \ldots, i_n \in N \backslash S$, $i_{k+1} \bigcup \ldots \bigcup i_n = N \backslash S$,

[5] In fact, the lower value of matrix game.

where $\prod_{j=k+1}^{n} \widetilde{\Xi}_{i_j}$ is the set of pure stationary strategies of coalition $N \backslash S$.

Finally, for $S = \emptyset$, the value of the characteristic function is:

$$V^j(\emptyset) = 0. \tag{11}$$

Simple algebra shows that the characteristic functions $\overline{V}(S)$ and $V^j(S)$ determined by (8) and (9)-(11), respectively, are super-additive. We are now in a position to define the cooperative version of the game and the subgame described in definitions (1) and (2).

Definition 3 *A cooperative stochastic subgame G^j_{co} is a set $\langle N, V^j(\cdot)\rangle$, where N is the set of players, and $V^j : 2^N \longrightarrow R$ is the characteristic function calculated by* (9) – (11)*.*

Definition 4 *A cooperative stochastic game G_{co} is a set $\langle N, \overline{V}(\cdot)\rangle$, where N is the set of players and $\overline{V} : 2^N \longrightarrow R$ is the characteristic function calculated by* (8)*.*

The next definitions display the allocation rule of the game within the characteristic function.

Definition 5 *An allocation in the subgame G^j_{co} $(j = 1, \ldots, t)$ is the vector $\alpha^j = (\alpha^j_1, \ldots, \alpha^j_n)$ satisfying the two following conditions:*

1. $\sum_{i \in N} \alpha^j_i = V^j(N)$,
2. $\alpha^j_i \geqslant V^j(\{i\})$ *for any* $i \in N$,

The set of allocations in cooperative subgame G^j_{co} is denoted as A^j, $j = 1, \ldots, t$.

Definition 6 *An allocation in the game G_{co} is the vector $\overline{\alpha} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_n)$, where $\overline{\alpha}_i = \pi^0 \alpha_i$, $\alpha_i = (\alpha^1_i, \ldots, \alpha^t_i)$, and $(\alpha^j_1, \ldots, \alpha^j_n) = \alpha^j \in I^j$. The set of allocations in cooperative stochastic game G_{co} is denoted as $\overline{I}$.*

Suppose that the set of allocations in any subgame G^j_{co}, $j = 1, \ldots, t$, is non-empty. Therefore the set of allocations in the cooperative stochastic game G_{co} is also non-empty.

5 Principles of stable cooperation

In the theory of cooperative games, the solution of a cooperative game is generated by a set of optimality principles. Examples are the Shapley value (1953b), the Von Neumann-Morgenstern solution (1944) and the Nash bargaining solution (1953). For stochastic games, additional conditions are required to ensure stable cooperation. The derivation of the stability conditions is the main result of the present paper and is developed in this section. In particular, the principles of dynamic stability include: *subgame consistency*, *strategic stability* and *irrational-behaviour-proof* of the cooperative agreement. Each condition is analysed separately.

5.1 Subgame consistency

Subgame consistency of the cooperative agreement allows players to expect their allocation coming from the same optimality principle in every stochastic subgame. This concept deserves a detailed explanation. Players initiate cooperation before the game and choose an allocation rule. During the game they realise the cooperative strategy profile by maximising their total payoff. In any subgame beginning from a certain state, a player is able to calculate her expected payoff for the remainder of the game. This is given by the difference between the initial expected payoff according to the allocation and the sum of the payoffs in each stage that the player obtained by the current state. This difference may not be equal to her transfer for this subgame, since this has been chosen using the same allocation mechanism as in the beginning of the game. In other words, if players are paid according to their stage payoff functions, then the allocation rule changes over time. Therefore, in order to keep the same allocation mechanism during the whole game, players should agree to determine the transfer mechanism (payoff distribution procedure). For instance, if players choose the Shapley value at the beginning of the game as an allocation mechanism, then subgame consistency guarantees that in any stage of the game the vector of remaining parts of players' payoff is the Shapley value calculating for the subgame beginning from this stage. Finally, note that subgame consistency is not a problem in the cooperative/repeated game is deterministic. Indeed, the players' payoffs are equal to their payoffs in the cooperative strategy-profile, according to the allocation rule. This is because there is no need to reallocate the payoffs in each stage.

Suppose that players cooperate in the stochastic game and, for every subgame G^j_{co}, choose allocation $\alpha^j = (\alpha^j_1, \dots, \alpha^j_n) \in I^j$. The problem is how to realise transfers to the players at each stage of the stochastic game for getting the expected payoff α_i^j for player i in stochastic subgame G^j. If players receive transfers according to their payoff functions in the states, then they can hardly expect to get the payoff based on the allocation rule chosen at the beginning of the game. In order to solve this, we propose a procedure of transferring the players' total payoff in every state realised in the stochastic game process. Petrosyan (1979) introduced this method for differential games. In what follows, we adapt the Petrosyan (1979)'s method to stochastic games.

There are two principles of derive the transfers to the players in the dynamic game adapted to the theory of stochastic games:

1. The sum of transfers to players in every state is equal to the sum of players' payoffs in the strategy profile realised in this state according to the cooperative decision $\overline{\eta}(\cdot)$.
2. The expected sum of transfers to player i in each subgame G^j is equal to the i^{th} component of an allocation in the subgame G_{co}^j that players have chosen before the beginning of the game.

Since in the stochastic game (1) the number of subgames is equal to the number of possible states, we need a vector $\beta_i = (\beta_i^1, \ldots, \beta_i^t)$ for every $i \in N$, where β_i^j is the transfer to player i in state Γ^j, $j = 1, \ldots, t$. If these transfers satisfy the two above principles, then we call then as *payoff distribution procedure* (from now on, PDP) (Petrosyan 1979, and Petrosyan and Baranova, 2006).

The first principle is equivalent to the following equation:

$$\sum_{i \in N} \beta_i^j = \sum_{i \in N} K_i^j(\overline{x}^j), \tag{12}$$

where $\overline{x}^j$ is an action profile realized under cooperative decision $\overline{\eta}(\cdot)$ in state Γ^j, $j = 1, \ldots, t$. The second principle is satisfied according to the expected total payoff of player i, denoted as B_i, with new transfer β_i^j in state Γ^j, $j = 1, \ldots, t$. The recurrent equation of B_i is given by:

$$B_i^j = \beta_i^j + \delta \sum_{k=1}^{t} p(j, k; x^j) B_i^k,$$

or in vectorial form:

$$B_i = \beta_i + \delta \Pi(\overline{\eta}) B_i, \tag{13}$$

where $B_i = (B_i^1, \dots, B_i^t)$. Equation (13) is equivalent to the following:

$$B_i = (\mathbb{I} - \delta\Pi(\eta))^{-1}\, \beta_i. \tag{14}$$

Given the second principle of PDP and equation (14) we obtain:

$$\alpha_i = (\mathbb{I} - \delta\Pi(\eta))^{-1}\, \beta_i, \tag{15}$$

where $\alpha_i = (\alpha_i^1, \dots, \alpha_i^t)$, $(\alpha_1^j, \dots, \alpha_n^j) = \alpha^j \in I^j$. Equation (15) can be rewritten equivalently as:

$$\beta_i = (\mathbb{I} - \delta\Pi(\overline{\eta}))\alpha_i. \tag{16}$$

It is easy to show that β_i from (16) satisfies (12). Since $\sum_{i\in N} \beta_i^j$ is equal to $(\mathbb{I}-\delta\Pi(\overline{\eta})) \sum_{i\in N} \alpha_i = (\mathbb{I}-\delta\Pi(\overline{\eta}))V(N)$, and $V(N)$ from (9), then equation (12) holds. Finally, note that equation (16) equals the following:

$$\alpha_i = \beta_i + \delta\Pi(\overline{\eta})\alpha_i. \tag{17}$$

The second item in the right part of equation (17) is the expected value of the transfers from the next stage onwards. Suppose that the allocation for each subgame is chosen from the same optimality principle that has been chosen by the players at the beginning of the game. Obviously, if players maintain the cooperative strategy $\overline{\eta}(\cdot)$, then the expected payoff of player i with new transfers is equal to the expected value of the correspondent allocation component in the cooperative stochastic game G_{co}.

For every allocation $\overline{\alpha} = (\overline{\alpha}_1, \dots, \overline{\alpha}_n)$, where $\overline{\alpha}_i = \pi^0\alpha_i$, $\alpha_i = (\alpha_i^1, \dots, \alpha_i^t)$, $(\alpha_1^j, \dots, \alpha_n^j) = \alpha^j \in I^j$ we can determine the regularisation of stochastic game G by the following definition.

Definition 7 *An "α-regularisation" of a stochastic game G (subgame G^j) is a non-cooperative stochastic game G_α (subgame G_α^j, $j = 1, \dots, t$) if, for any player $i \in N$ in state Γ^j, the payoff function $K_i^{\alpha,j}(x^j)$ is defined as follows:*

$$K_i^{\alpha,j}(x^j) = \begin{cases} \beta_i^j, & \text{if } x^j = \overline{x}^j; \\ K_i^j(x^j), & \text{if } x^j \neq \overline{x}^j, \end{cases} \tag{18}$$

where the PDP $\beta = (\beta_1, \dots, \beta_n)$ is found from (16)*.*

The procedure of regularisation of the stochastic game G (subgame G^j) suggests a consistent method of transfers in every state. The payoff distribution procedure $\beta_i^1, \ldots, \beta_i^t$ in states $\Gamma^1, \ldots, \Gamma^t$ ensures that a player i receives the same expected payoff in game G_α (G_α^j), as she planned to receive in the cooperative stochastic game G_{co} (G_{co}^j). Moreover the expected payoff from future transfers is in line with the same optimality principle chosen by a player at the beginning of the game. Thus subgame consistency of the chosen cooperative agreement holds.

5.2 Strategic stability

Suppose players come to an agreement, and find the cooperative strategy profile that maximises the expected total payoff in the whole game. Also, suppose that, if a player deviates in any state s, then other players use the grim-trigger strategy from the next stage to punish the deviating player. The strategic stability condition ensures that the one-shot payoff in a certain state of the deviating player plus her max-min payoff in the rest of the game is not larger than playing the cooperative strategy in the whole game. In other words, this condition allows the existence of a Nash equilibrium in the regularised game with the same payoffs that players expect to receive within the cooperative agreement. Along the section, we build up the regularisation of the game on the basis of the initial stochastic game with the use of payoff distribution procedure (Petrosyan and Danilov, 1979). Comparing our approach with the standard analysis of deterministic (repeated) games, the condition of strategic stability corresponds to the existence of Nash equilibrium in grim-trigger strategies. The difference is that, in our setting, players first find the payoff distribution procedure by achieving subgame consistency.

Begin by introducing some additional notations. Denote the set $\Gamma(k)$ as the state realised at stage k of the stochastic game G, $\Gamma(k) \in \{\Gamma^1, \ldots, \Gamma^t\}$, and $x(k)$ as the strategy profile realised in state $\Gamma(k)$. Denote as $G_\alpha^{\Gamma(k)}$ as the the subgame of the stochastic game G_α beginning from state $\Gamma(k)$. Finally, denote the history of stage k as $h(k)$, which is the sequence $((\Gamma(1), x(1)),(\Gamma(2), x(2)),\ldots,(\Gamma(k-1), x(k-1)))$, and let T be $\{(\Gamma^1, \overline{x}^1), (\Gamma^2, \overline{x}^2), \ldots, (\Gamma^t, \overline{x}^t)\}$. Then, G and G_α are stochastic games with perfect information if at each stage k ($k = 1, 2, \ldots$) all players from N know state $\Gamma(k)$ and the history of stage k.

Definition 8 *A strong transferable equilibrium in the regularised game G_α is a behaviour strategy profile $\varphi^*(\cdot) = (\varphi_1^*(\cdot),\ldots,\varphi_n^*(\cdot))$ if, for any coalition $S \subset N$, $S \neq \emptyset$,*

- *the inequality*

$$\sum_{i\in S}\overline{E}_i^{\alpha}(\varphi^*) \geqslant \sum_{i\in S}\overline{E}_i^{\alpha}(\varphi^* \parallel \varphi_S) \tag{19}$$

  is true for any behaviour strategy of coalition S*:* $\varphi_S(\cdot) = \{\varphi_i(\cdot)\}_{i\in S} \in \prod_{i\in S}\Phi_i$;

- $\overline{E}_i^{\alpha}(\cdot)$ *is the expected payoff of player* i *in* α*-regularisation of stochastic game* G*.*

The definition of strong transferable equilibrium allows us to introduce the following theorem.

Theorem 3 *If in an* α*-regularisation of stochastic game* G *such that* $\overline{\alpha} = \pi^0\alpha$*, the following inequality holds for any coalition* $S \subset N$*,* $S \neq \emptyset$*:*

$$\sum_{i\in S}\beta_i \geqslant (\mathbb{I} - \delta\Pi(\overline{\eta}))F(S), \tag{20}$$

where

$$F(S) = (F^1(S), \ldots, F^t(S)),$$

$$F^j(S) = \max_{\substack{x_S^j \in \prod_{i\in S} X_i^j \\ x_S^j \neq \overline{x}_S^j}} \left\{ \sum_{i\in S} K_i^j(\overline{x}^j \parallel x_S^j) + \delta\sum_{l=1}^{t} p(j,l;\overline{x}^j \parallel x_S^j)V^l\,(S) \right\},$$

then, in the regularised game G_α *there exists a strong transferable equilibrium with payoffs* $(\overline{\alpha}_1,\ldots,\overline{\alpha}_n)$*.*

From Theorem 3 follows

Corollary 1 *Suppose an* α*-regularisation of the game* G *for any player* $i \in N$*. if the next inequality takes place:*

$$\beta_i \geqslant (\mathbb{I} - \delta\Pi(\overline{\eta}))W_i,$$

where

$$W_i = (W_i^1, \ldots, W_i^t),$$

$$W_i^j = \max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \overline{x}_i^j}} \left\{ K_i^j(\overline{x}^j \parallel x_i^j) + \delta\sum_{l=1}^{t} p(j,l;\overline{x}^j \parallel x_i^j)V^l\,(\{i\}) \right\},$$

then there exists a Nash equilibrium with payoffs $(\overline{\alpha}_1, \ldots, \overline{\alpha}_n)$*.*

Theorem 3 and Corollary 1 show the conditions under which a cooperative equilibrium exists and is stable. This in turn ensures the existence of strategic stability.

5.3 Irrational-behaviour-proof

One of the most commonly used assumptions to handle deviation of players from the cooperative path is that cooperation will break down and players would revert to non-cooperative behaviour if deviations occur. Given that time-consistent imputations satisfy group and individual rationalities throughout the cooperative trajectory, no rational players will deviate from the cooperative path. However, in reality deviations may appear for various reasons. For instance, a player may use 'irrational' acts to extort additional gains if later circumstances allow. Refusal of other players to yield to his extortion would result in the dissolution of the cooperative scheme. Thus in this case, deviation would appear as "irrational behaviour".[6]

Yeung (2006) proposes a condition under which, even if an irrational behaviour emerges in the game, a player would still be performing better under the cooperative scheme. In particular, in order to protect players against losses due to irrational behaviour, it is necessary that the following equation holds:

$$V(\{i\}) \leqslant E_i^{\alpha,[1,k]} + \delta^k \Pi^k(\overline{\eta}) V(\{i\}), \text{ for every } i \in N \text{ and any } k = 1, 2, \ldots, \tag{21}$$

where $E_i^{\alpha,[1,k]}$ is the expectation of player i's payoff at the first k stages of the regularised game G_α.

Suppose that, before the beginning of a stage, players know if the cooperation has broken up or not, so that information is not delayed. In the left hand side of inequality (21) $V(\{i\})$ is the value of the characteristic function $V(\{i\}) = (V^1(\{i\}), \ldots, V^t(\{i\}))$ derived for player i where $V^j(\{i\})$ is the value of the characteristic function of player i in the subgame G^j. In the right hand side of inequality (21), the first term is equal to the expected value of player i's payoff if, in the first k stages, players play the cooperative strategy $\overline{\eta}(\cdot)$. The second term is the expected payoff of player i from stage $k+1$ if she plays non-cooperatively at $k+1$ onwards.

Proposition 1 *In a stochastic game G_α a sufficient condition of irrational-behaviour-proof*

[6] Note that it is possible to formulate an analogous condition for deterministic repeated games.

is that, for any $i \in N$*:*

$$(\mathbb{I} - \delta\Pi(\overline{\eta}))(\alpha_i - V(\{i\})) \geqslant 0, \tag{22}$$

where $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^t)$ *and* α_i^j *is the* i^{th} *component of allocation* $\alpha^j \in I^j$*.*

6 An example

Consider the following stochastic version of the game "Battle of the sexes" with two players ("1", woman, "2", man). Let the set of states be $\{\Gamma^1, \Gamma^2\}$, where $\Gamma^j = \langle N, X_1^j, X_2^j, K_1^j, K_2^j \rangle$, $j = 1, 2$, and $X_i^j = \{B_j, F_j\}$ is the set of actions of player $i = 1, 2$. Strategy B_j means "to go to the ballet", D_j means "to go to the football game". For state Γ^1, players' payoffs are:

$$\begin{array}{c} \\ B_1 \\ F_1 \end{array} \begin{array}{c} \begin{array}{cc} B_1 & F_1 \end{array} \\ \begin{pmatrix} (3,1) & (0,0) \\ (0,0) & (1,3) \end{pmatrix}, \end{array}$$

whereas for state Γ^2 players' payoffs are:

$$\begin{array}{c} \\ B_2 \\ F_2 \end{array} \begin{array}{c} \begin{array}{cc} B_2 & F_2 \end{array} \\ \begin{pmatrix} (4,2) & (0,0) \\ (1,0) & (3,4) \end{pmatrix}. \end{array}$$

If we consider state Γ_1 as a one-shot game, then the game has two pure Nash equilibria (B_1, B_1) and (F_1, F_1). Equilibrium (B_1, B_1) is preferred by player 1, and (F_1, F_1) is preferred by player 2. Players are assumed to be symmetric in this state, meaning that the payoff of player 1 (player 2) in (B_1, B_1) equals the payoff of player 2 (player 1) in (F_1, F_1).

Let us consider the one-shot game of state Γ_2 (with larger payoffs), where there also exist two pure Nash equilibria (B_1, B_1) and (F_1, F_1). However, in this game players are asymmetric, since the payoff of player 2 in (B_1, B_1) does not equal the payoff of player 1 in (F_1, F_1). The max-min payoffs of player 1 and 2 in this state are 1 and 0. This implies that player 1 may ensure payoff 1 even if player 2 plays against player 1. This fact influences the values of characteristic function calculated for the cooperative stochastic game.

Let the transition probabilities from states Γ^1 and Γ^2 be

$$\begin{pmatrix} (0.3, 0.7) & (0.5, 0.5) \\ (0.5, 0.5) & (0.7, 0.3) \end{pmatrix}, \quad \begin{pmatrix} (0.7, 0.3) & (0.5, 0.5) \\ (0.5, 0.5) & (0.7, 0.3) \end{pmatrix},$$

where the elements of the matrix consist of the transition probabilities from state Γ^j to states Γ^1, Γ^2 respectively. Let $\delta = 0.99$ be the discount factor and $\pi^0 = (0.5, 0.5)$ be the vector of the initial distribution on the set of states.

Begin by determining the cooperative form G_{co} of the stochastic game G. Firstly, we compute the cooperative decision $\overline{\eta} = (\overline{\eta}_1, \overline{\eta}_2)$ in stationary strategies using (6) and (7). We obtain the unique stationary strategy $\overline{\eta}_1 = (B_1, B_1)$ and $\overline{\eta}_2 = (F_2, F_2)$, for player 1 and 2, respectively. Secondly, we find the values of the characteristic function $V(\cdot) = (V^1(\cdot), V^2(\cdot))$ for all possible coalitions using (9)-(11):

$$V(\emptyset) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; V(\{1\}) = \begin{pmatrix} 49.5 \\ 50.5 \end{pmatrix}, \; V(\{2\}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; V(\{1,2\}) = \begin{pmatrix} 548.9 \\ 551.1 \end{pmatrix}.$$

Using (8), we then calculate the values of the characteristic function $\overline{V}(\cdot)$ for all possible coalitions:

$$\overline{V}(\emptyset) = 0.00, \; \overline{V}(\{1\}) = 50, \; \overline{V}(\{2\}) = 0, \; \overline{V}(\{1,2\}) = 550.$$

Thus, we determine the cooperative stochastic subgame G^j_{co} as the set $\langle N, V^j(\cdot)\rangle$, $j = 1, 2$, and the cooperative stochastic game G_{co} as the set $\langle N, \overline{V}(\cdot)\rangle$. Before to verify the stability of cooperation, we suppose that players choose the Shapley value as the allocation of their total payoff in the cooperative stochastic game G_{co} and in all subgames G^j_{co}, $j = 1, 2$. The Shapley values calculated for subgames are:

$$\alpha_1 = \begin{pmatrix} 299.2 \\ 300.8 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 249.7 \\ 250.3 \end{pmatrix},$$

where $\alpha_i = (\alpha_i^1, \alpha_i^2)$, and α_i^j is the i^{th} component of the Shapley value of subgame G^j_{co} using the characteristic function $V^j(\cdot)$, $j = 1, 2$, $i \in N$. Then taking into account the vector of initial distribution π^0, we are able to determine the allocation $\overline{\alpha}$ in G_{co} by Definition 6:

$$\overline{\alpha} = (\overline{\alpha}_1, \overline{\alpha}_2) = (300, 250).$$

We are now in a position to verify the principles of stable cooperation. To satisfy the principle of subgame consistency we should calculate the PDP for the allocation $\overline{\alpha}$ equals to $\pi^0\alpha$ using (16):

$$\beta_1 = \begin{pmatrix} 1.9 \\ 4.1 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} 2.1 \\ 2.9 \end{pmatrix}.$$

The components of PDP are not all equal to the corresponding payoffs of the players in the states. Hence, if players obtain their payoffs from the initial rule of the game, they cannot receive the expected chosen allocation. This in turn implies subgame inconsistency of the cooperative agreement. We then realise the α-regularisation of the initial stochastic game G using PDP and Definition 7. The players' payoffs in state Γ^1 are:

$$\begin{pmatrix} (1.9, 2.1) & (0,0) \\ (0,0) & (1,3) \end{pmatrix},$$

while in state Γ^2 the payoffs are:

$$\begin{pmatrix} (4,2) & (0,0) \\ (1,0) & (4.1, 2.9) \end{pmatrix}.$$

In state Γ_1 both players go to ballet and player 1 gives 1.1 from her payoff 7 to player 2, in state Γ_2 both players go to football game and player 2 gives (4-2.9)= 1.1 from his payoff to player 1 to obtain a subgame consistent Shapley value.

We now check for strategic stability of cooperative agreement. This implies to verify inequality (20) from Theorem 3. Compute $F(S)$ for all $S \subset N$ and obtain that

$$F(\{1\}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0.99 \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} V^1(\{1\}) \\ V^2(\{1\}) \end{pmatrix} = \begin{pmatrix} 49.5 \\ 50.5 \end{pmatrix},$$

$$F(\{2\}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 0.99 \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} V^1(\{2\}) \\ V^2(\{2\}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

Inequality (20) is equivalent to:

$$\beta_1 = \begin{pmatrix} 1.9 \\ 4.1 \end{pmatrix} \geqslant \begin{pmatrix} -0.198 \\ 1.198 \end{pmatrix} = (\mathbb{I} - \delta\Pi(\overline{\eta}))F(\{1\}),$$

$$\beta_2 = \begin{pmatrix} 2.1 \\ 2.9 \end{pmatrix} \geqslant \begin{pmatrix} 0 \\ 0 \end{pmatrix} = (\mathbb{I} - \delta\Pi(\overline{\eta}))F(\{2\}).$$

Hence strategic stability is verified.

Finally, we verify the irrational-behaviour-proof condition. The sufficient condition from Proposition 1 since:

$$(\mathbb{I} - \delta\Pi(\overline{\eta}))(\alpha_1 - V(\{1\})) = (\mathbb{I} - \delta\Pi(\overline{\eta}))(\alpha_2 - V(\{2\}))$$

$$= \begin{pmatrix} 0.703 & -0.693 \\ -0.693 & 0.703 \end{pmatrix} \begin{pmatrix} 249.7 \\ 250.3 \end{pmatrix} = \begin{pmatrix} 2.1 \\ 2.9 \end{pmatrix} \geqslant 0.$$

By computing the α-regularisation of the initial game G and checking the principles of stable cooperation, they all are satisfied.

7 Concluding remarks

In this paper we have examined the conditions of dynamic stability for cooperative stochastic games. Principles of dynamic stability include three conditions: subgame consistency, strategic stability and irrational-behaviour-proof of the cooperative agreement. This approach to obtain stable cooperation has been left general purposely, and can be applied hereinafter in specific contexts in which stochastic cooperative games are an effective framework of analysis. These developments are left for future research.

References

[1] Altman, E., El-Azouzi, R. and Jimenez, T. 2003. Slotted Aloha as a stochastic game with partial information. *Proceedings of Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks*. Sophia Antipolis, France.

[2] Amir, R. 2001. Stochastic games in economics and related fields: an overview. *CORE working paper* **60**.

[3] Amir, R. 2003. Stochastic games in economics: The lattice-theoretic approach. In A. Neyman and S. Sorin (eds.) *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences **570**: 443-453.

[4] Aumann, R. 1987. Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica* **55**: 1-18.

[5] Aumann, R.J., and Dreze, J.H., 1974. Cooperative games with coalition structure. *International Journal of Game Theory* **3**: 217-237.

[6] Banerjee A., and Duflo E. 2007. The economic lives of the poor. *Journal of Economic Perspectives* **21**: 141-168.

[7] Bellman, R. 1957. *Dynamic programming*. Princeton University Press, Princeton, NJ.

[8] Case C., Garrib A., Menendez A., and Olgiati A. 2013 Paying the piper: The high cost of funerals in South Africa. *Economic Development and Cultural Change* **62**: 1-20.

[9] Charpentier, A. and Mussard, S. 2011. Income inequality games. *Journal of Economic Inequality* **9**: 529-554.

[10] Chen X. 2014. Fetus, Fasting, and Festival: The Persistent Effects of in Utero Social Shocks. *IZA working Paper* **8494**.

[11] Chen X., Kanbur R., and Zhang X. 2011 Peer effect, risk-pooling and status seeking: What explains gift spending escalation in rural China? *CEPR Working Paper* **8777**.

[12] Duffie, D., Geanakoplos, J., Mas-Colell, A. and McLennan, A. 1994. Stationary Markov Equilibria. *Econometrica* **62:** 745-781.

[13] Dutta, P. 1995. A folk theorem for stochastic games. *Journal of Economic Theory* **66**: 1-32.

[14] Giuliano, P., and Spilimbergo, A. 2014. Growing up in a recession. *Review of Economic Studies* **81**: 787-817.

[15] Jaskiewicz, A., Nowak, A.S. 2014. Stationary Markov perfect equilibria in risk sensitive stochastic overlapping generations models *Journal of Economic Theory*, **151**: 411-447.

[16] Kirman, A.P. and Sobel, M.J. 1974 Dynamic oligopoly with inventories. *Econometrica* **42**: 279-287.

[17] Knuth, D. E. 1997. *The Art of Computer Programming* **1**. Addison-Wesley.

[18] Lenter, D., Shackelford, D. and Slemrod, J. 2003. Public disclosure of corporate tax return information: Accounting, economics, and legal perspectives. *National Tax Journal* **56**: 803-830.

[19] Litina, A. and Palivos, T. 2013. Corruption and tax evasion: reflections on greek tragedy. Working paper.

[20] López-Pérez, R. 2008 Introducing social norms in game theory. in A. Innocenti and P. Sbriglia (eds.), *Games, Rationality and Behavior*, Palgrave Mcmillan.

[21] McCain, R., 2007. Cooperation and effort, reciprocity and mutual supervision in worker cooperatives. In: Novkovic, S. (Ed.), *Advances in the Economic Analysis of Participatory and Labor-Managed Firms*. JAI Press, Greenwich, CT.

[22] McCain, R. A. 2008. Cooperative games and cooperative organizations. *Journal of Socio-Economics* **37**: 2155-2167.

[23] Messner, M. and Polborn, M.K. 2003. Cooperation in Stochastic OLG games. *Journal of Economic Theory* **108**: 152-168.

[24] Parilina, E. M. 2010. Cooperative data transmission game in wireless network. *Upravlenie Bol'simi Sistemami* **31**: 93-110 (in Russian).

[25] Petrosyan, L.A. (1977). Stability of the solutions in differential games with several players. *Vestnik Leningrad. Univ. Mat. Mekh. Astronom* **19**: 46-52 (in Russian).

[26] Petrosjan, L.A. 2006. Cooperative Stochastic Games. *Advances in dynamic games, Annals of the International Society of Dynamic Games* **8**: 139-146.

[27] Petrosjan, L. A. and Baranova, E. M. (2006). Cooperative Stochastic Games in Stationary Strategies. *Game Theory and Applications* **11**: 7-17.

[28] Petrosyan, L. A., Baranova, E. M., and Shevkoplyas, E. V. 2004. Multistage cooperative games with random duration. Optimal control and differential games. *Proceedings of Institute of mathematics and mechanics* **10**: 116-130 (in Russian).

[29] Petrosjan, L. A. and Danilov, N. N. (1979). Stability of the solutions in nonantagonistic di erential games with transferable payoffs. *Vestnik Leningrad. Univ. Mat. Mekh. Astronom.* **1**: 52-59 .

[30] Petrosyan, L.A. and Zenkevich, N.A. (2009) Conditions for Sustainable Cooperation *Contributions to Game Theory and Management* **2**: 344-355.

[31] Raghavan, T. E. S. 2006. A Stochastic Game Model of Tax Evasion. *Advances in dynamic games, Annals of the International Society of Dynamic Games* **8**: 397-420.

[32] Shapley, L. S. 1953a. Stochastic Games. *Proceedings of National Academy of Sciences of the USA* **39**: 1095-1100.

[33] Shapley, L. S. 1953b. A Value for n-person Games. Contributions to the Theory of Games-II, by H.W. Kuhn and A.W. Tucker (eds) *Annals of Mathematical Studies* **28**. Princeton University Press: 307-317.

[34] Solan, E. and Vieille N. 2002. Correlated Equilibrium in Stochastic Games *Games and Economic Behavior*, **38: 362-399**.

[35] Spence, J., 1991. *The Search for Modern China*. WW Norton &Company.

[36] The Economist. Can Ghanaians afford such splendid funerals? May 24; 2007.

[37] The Economist. Two weddings, two funerals, no fridge. November 28; 2013.

[38] Torgler, B., 2004. Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance* **5**: 235-253.

[39] Yeung, D.W.K. 2006. An irrational-behavior-proof condition in cooperative differential games. *International Game Theory Review* **8**: 739-744.

[40] Yeung D. W.K., Petrosyan L.A. 2012 *Subgame Consistent Economic Optimization*. New York: Birkhauser.

8 Appendix

8.1 Proof of Lemma 1

Let λ be an eigenvalue of $\Pi(\eta) = \{p(i,j;x^i)\}_{\substack{i=1,\dots,t\\ j=1,\dots,t}}$. This must satisfy the equality $\Pi(\eta)\Lambda = \lambda\Lambda$ for some eigenvector Λ. Let k be such that $|\Lambda_j| \leqslant |\Lambda_k|$ for any $j \in 1,\dots,t$. The k^{th} component of each side of equation $\Pi(\eta)\Lambda = \lambda\Lambda$ gives $\sum\limits_{j=1}^{t} p(k,j;x^k)\Lambda_j = \lambda\Lambda_k$. Therefore, we have

$$|\lambda\Lambda_k| \leqslant |\lambda|\cdot|\Lambda_k| = |\sum_{j=1}^{t} p(k,j;x^k)\Lambda_j| \leqslant \sum_{j=1}^{t} p(k,j;x^k)|\Lambda_j| \leqslant \sum_{j=1}^{t} p(k,j;x^k)|\Lambda_k| = |\Lambda_k|,$$

yielding $|\lambda| \leqslant 1$. Hence the eigenvalues of the stochastic matrix $\Pi(\eta)$ are in the interval $[-1,1]$.

For the existence of matrix $(\mathbb{I} - \delta\Pi(\eta))^{-1}$ it is necessary and sufficient that the determinant of matrix $\left(\Pi(\eta) - \frac{1}{\delta}\mathbb{I}\right)$ is non-negative, implying that $\frac{1}{\delta}$ must not be an eigenvalue. Indeed, $\delta \in (0,1) \Longrightarrow \frac{1}{\delta} < 1$, hence it cannot be the eigenvalue of $\Pi(\eta)$.

8.2 Proof of Theorem 3

Consider the behaviour strategy profile $\widehat{\varphi}(\cdot) = (\widehat{\varphi}_1(\cdot),\dots,\widehat{\varphi}_n(\cdot))$ in the game G_α:

$$\widehat{\varphi}_i(h(k)) = \begin{cases} \overline{x}_i^j, & \text{if } \Gamma(k) = \Gamma^j, j = \overline{1,t}, h(k) \subset T; \\ \widehat{x}_i^j(S), & \text{if } \Gamma(k) = \Gamma^j,\ j = \overline{1,t},\ \exists\ l \in [1, k-1], \\ & S \subset N,\ i \notin S\text{: } h(l) \subset T, \text{ and} \\ & (\Gamma(l), x(l)) \notin T, \text{ but} \\ & (\Gamma(l), (x(l) \parallel \overline{x}_S(l)) \in T; \\ \text{anyone} & \text{otherwise,} \end{cases} \tag{23}$$

where $\widehat{x}_i^j(S)$ is the player i's pure strategy in state Γ^j which with strategies $x_p^j(S)$, $p \neq i$, $p \notin S$ forms the strategy of coalition $\{N \backslash S\}$ in the antagonistic game against coalition S in the subgame $G^{\Gamma(j)}$.

The proof of the theorem follows the folk theorem for stochastic games (Dutta, 1995) using the structure of strategy (23). We prove that $\widehat{\varphi}(\cdot) = (\widehat{\varphi}_1(\cdot),\dots,\widehat{\varphi}_n(\cdot))$ determined in (23) is a strong transferable equilibrium in the stochastic game G_α. Given definition (23) and provided that all players do not deviate from the cooperative strategy $\overline{\eta}(\cdot)$, the expected payoff of coalition S in the subgame G_α^j, $j = 1, \ldots, t$, is:

$$E_S^j(\widehat{\varphi}(\cdot)) = \sum_{i \in S} E_i^j(\widehat{\varphi}(\cdot)) = \sum_{i \in S} E_i^j(\overline{\eta}(\cdot)).$$

Let $E_S(\widehat{\varphi}(\cdot))$ be equal to the vector $(E_S^1(\widehat{\varphi}(\cdot)), \ldots, E_S^t(\widehat{\varphi}(\cdot)))$. Then for any coalition $S \subset N$, $S \neq \emptyset$ the next equality holds:

$$E_S(\widehat{\varphi}) = (\mathbb{I} - \delta \Pi(\overline{\eta}))^{-1} \sum_{i \in S} \beta_i. \tag{24}$$

Consider then the profile of strategies $(\widehat{\varphi}(\cdot) \parallel \varphi_S(\cdot))$, $S \subset N$, $S \neq \emptyset$, when some coalition S deviates from strategy $\widehat{\varphi}_S(\cdot)$. Let k be such that there exists a number $l \in [1, k-1]$ such that $h(l) \subset T$, $(\Gamma(l), x(l)) \notin T$, but $(\Gamma(l), (x(l) \parallel \overline{x}_S(l))) \in T$. Without loss of generality, we simplify $\Gamma(k) = \Gamma^j$. We are now able to determine the total payoff of the coalition S in the game G_α with strategies profile $(\widehat{\varphi}(\cdot) \parallel \varphi_S(\cdot))$ by

$$\sum_{i \in S} \overline{E}_i^\alpha(\widehat{\varphi} \parallel \varphi_S) = \pi^0 \sum_{i \in S} E_i^\alpha(\widehat{\varphi} \parallel \varphi_S),$$

where

$$\sum_{i \in S} E_i^\alpha(\widehat{\varphi} \parallel \varphi_S) = \sum_{i \in S} E_i^{\alpha,[1,k-1]}(\widehat{\varphi} \parallel \varphi_S) \tag{25}$$
$$+ \delta^{k-1} \Pi^{k-1}(\widehat{\varphi} \parallel \varphi_S) \sum_{i \in S} E_i^{\alpha,[k,\infty)}(\widehat{\varphi} \parallel \varphi_S).$$

The first term in the right hand side of equation (25) is the expected payoff of coalition S in the first $k-1$ stages of the game G_α, the second term is the expected payoff of coalition S in the subgame of G_α beginning from stage k, where $E_i^{\alpha,[k,\infty)}(\widehat{\varphi} \parallel \varphi_S)$ is the vector $(E_i^{\alpha,1}(\widehat{\varphi} \parallel \varphi_S), \ldots, E_i^{\alpha,t}(\widehat{\varphi} \parallel \varphi_S))$, with $E_i^{\alpha,j}(\widehat{\varphi} \parallel \varphi_S)$ being the player i's expected payoff in the regularised subgame G_α^j beginning at state Γ^j. Since there were no deviations of any coalition from the cooperative decision $\overline{\eta}(\cdot)$ up to stage $k-1$, the following equalities hold:

$$\sum_{i \in S} E_i^{\alpha,[1,k-1]}(\widehat{\varphi} \parallel \varphi_S) = \sum_{i \in S} E_i^{\alpha,[1,k-1]}(\overline{\eta}),$$

$$\Pi^{k-1}(\widehat{\varphi} \parallel \varphi_S) = \Pi^{k-1}(\overline{\eta}).$$

We now find the expected payoff of coalition S in the subgame G_α^j beginning with stage k and state $\Gamma(k)$ is equal to Γ^j. The following formula takes place:

$$\sum_{i \in S} E_i^{\alpha,j}(\widehat{\varphi} \parallel \varphi_S) = \sum_{i \in S} K_i^j(\overline{x}^j \parallel x_S^j) + \delta \sum_{l=1}^{t} p(j,l;\overline{x}^j \parallel x_S^j) V^l(S). \tag{26}$$

Players from coalition $N \setminus S$ punish coalition S by playing the "antagonistic" game from stage $k+1$, according to the definition of strategy profile $\widehat{\varphi}(\cdot)$. In (26), the value of the characteristic function $V^j(S)$ is determined by (10). Since the expected payoffs of coalition S in the strategy profiles $\widehat{\varphi}(\cdot)$ and $(\widehat{\varphi}(\cdot) \parallel \varphi_S(\cdot))$ do not change up to stage $k-1$, then a deviation may increase coalition S's payoff only at the expenses of the expected payoff in subgame G_α^j, $j = 1, \ldots, t$. In particular, coalition S in the strategy profile $(\widehat{\varphi}(\cdot) \parallel \varphi_S(\cdot))$ ensures the following expected payoff from stage k:

$$\max_{\substack{x_S^j \in \prod_{i \in S} X_i^j \\ x_S^j \neq \overline{x}_S^j}} \left\{ \sum_{i \in S} K_i^j(\overline{x}^j \parallel x_S^j) + \delta \sum_{l=1}^{t} p(j,l;\overline{x}^j \parallel x_S^j) V^l(S) \right\}. \tag{27}$$

According to the definition of PDP, the expected payoff of coalition S in the regularised subgame G_α^j with profile of strategies $\widehat{\varphi}(\cdot)$ can be found from:

$$\sum_{i \in S} E_i^\alpha(\widehat{\varphi}) = (\mathbb{I} - \delta \Pi(\overline{\eta}))^{-1} \sum_{i \in S} \beta_i, \tag{28}$$

where $E_i^\alpha(\widehat{\varphi}(\cdot)) = (E_i^{\alpha,1}(\widehat{\varphi}(\cdot)), \ldots, E_i^{\alpha,t}(\widehat{\varphi}(\cdot))$. Taking into account inequality (20) from (27), (28) and the above discussion we get

$$E_S^\alpha(\widehat{\varphi}(\cdot)) \geqslant E_S^\alpha(\widehat{\varphi}(\cdot) \parallel \varphi_S(\cdot)).$$

Thus the behaviour strategy profile (23) is the strong transferable equilibrium in the α-regularisation of game G. The expected payoff of player i in the game G_α with profile of strategies $\widehat{\varphi}(\cdot)$ is equal to $\overline{\alpha}_i$, where $\overline{\alpha}_i = \pi^0 \alpha_i$, while $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^t)$ consists of i^{th} components of allocations α^1, …, α^t derived from the cooperative subgames G^1, …, G^t accordingly.

8.3 Proof of Corollary 1

Condition (20) for coalition $S = \{i\}$ takes the form:

$$\beta_i \geqslant (\mathbb{I} - \delta\Pi(\overline{\eta}))F(\{i\}), \tag{29}$$

where

$$F(\{i\}) = (F^1(\{i\}), \ldots, F^t(\{i\})),$$
$$F^j(\{i\}) = \max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \overline{x}_i^j}} \left\{ K_i^j(\overline{x}^j \parallel x_i^j) + \delta \sum_{l=1}^{t} p(j, l; \overline{x}^j \parallel x_i^j) V^l(\{i\}) \right\}.$$

Note that $F(\{i\}) = W_i$ for any $i \in N$. Inequality (29) can be rewritten as

$$(\mathbb{I} - \delta\Pi(\overline{\eta}))^{-1}\beta_i \geqslant W_i. \tag{30}$$

The left hand side of equation (30) is equal to

$$E_i^{\alpha}(\widehat{\varphi}(\cdot)) \geqslant W_i \geqslant E_i^{\alpha}(\widehat{\varphi}(\cdot) \parallel \varphi_i(\cdot)),$$

for any $i \in N$. This is the definition of Nash equilibrium for a behaviour strategy profile $\widehat{\varphi}(\cdot)$ in which a deviation is punished using strategies of form (23).

8.4 Proof of Proposition 1

In what follows, we show that condition (22) is sufficient for inequality (21) for any $k = 1, 2, \ldots$. The proof is based on the mathematical induction method (Knuth, 1997). First, we rewrite inequality (21) for $k = 1$. Then we transform inequality (22) by considering definition α_i and using PDP (15). We get

$$V(\{i\}) \leqslant \beta_i + \delta\Pi(\overline{\eta})V(\{i\}). \tag{31}$$

Suppose that (22) implies (21) for $k = l$. Rewriting (21) for $k = l$ yields:

$$V(\{i\}) \leqslant \beta_i + \ldots + \delta^{l-1}\Pi^{l-1}(\overline{\eta})\beta_i + \delta^l\Pi^l(\overline{\eta})V(\{i\}). \tag{32}$$

We adopt the same procedure for $k = l + 1$. Inequality (21) for $k = l + 1$ is:

$$V(\{i\}) \leqslant \beta_i + \ldots + \delta^l \Pi^l(\overline{\eta})\beta_i + \delta^{l+1}\Pi^{l+1}(\overline{\eta})V(\{i\}). \qquad (33)$$

At this point we need to prove that, if (22) holds, then (21) holds for $k = l + 1$. After the transformation the right hand side of (33) is:

$$\beta_i + \delta\Pi(\overline{\eta}) \left\{ \beta_i + \delta\Pi(\overline{\eta})\beta_i + \ldots + \delta^{l-1}\Pi^{l-1}(\overline{\eta})\beta_i + \delta^l \Pi^l(\overline{\eta})V(\{i\}) \right\}.$$

Taking into account (32), the expression in brackets is not less than $V(\{i\})$. Therefore the right part of (33) is not less than $\beta_i + \delta\Pi(\overline{\eta})V(\{i\})$. From equation (16) and (22), we get (21) for $k = l + 1$, which proves the proposition.