Contributions to Game Theory and Management, VIII, 347–360

On Subgame Consistent Solution for NTU Cooperative Stochastic Dynamic Games*

David W.K. Yeung[1] **and Leon A. Petrosyan**[2]

[1] *Corresponding author. Phone +852-28065150 Fax +852-28068044*
Center of Game Theory, St Petersburg State University, Russia;
SRS Consortium for Advanced Study in Dynamic Cooperative Games,
Shue Yan University, Hong Kong
E-mail: dwkyeung@hkbu.edu.hk

[2] *Faculty of Applied Mathematics-Control Processes,*
St Petersburg State University, Russia
E-mail: Spbuoasis7@petrlink.ru

Abstract In cooperative dynamic games a stringent condition – subgame consistency – is required for a dynamically stable solution. In particular, a cooperative solution is subgame consistent if the optimality principle agreed upon at the outset remains in effect in any subgame starting at a later stage with a state brought about by prior optimal behavior. Hence the players do not have incentives to deviate from the previously adopted optimal behavior. Yeung and Petrosyan (2015) provided subgame consistent solutions in cooperative dynamic games with non-transferable payoffs/utility (NTU) using a variable payoffs weights scheme is analyzed. This paper extends their analysis to a stochastic dynamic framework. A solution mechanism for characterizing subgame consistent solutions is derived. The use of a variable payoff weights scheme allows the derivation of subgame consistent solutions under a wide range of optimality principles.

Keywords: stochastic dynamic games, subgame consistent cooperative solution, variable payoff weights.

1. Introduction

Cooperative games suggest the possibility of enhancing the participants' well-being in situations involving strategic interactions. One of the ways to uphold sustainability of a cooperation scheme is to maintain the condition of subgame consistency. In particular, a cooperative solution is subgame consistent if the optimality principle agreed upon at the outset remains in effect in any subgame starting at a later time with a state brought about by prior optimal behavior. Subgame consistent solutions for differential games and dynamic games with transferable payoffs under deterministic and stochastic dynamics can be found in Yeung and Petrosyan (2004, 2010). The use of transfer payments plays an important role in achieving subgame consistency in games with transferrable payoffs.

In NTU cooperative dynamic games, the inapplicability of transfer payments makes the design of cooperative schemes much more difficult. The number of studies in cooperative dynamic games with non-transferrable payoffs/utility (NTU) is much less than that of cooperative dynamic games with transferrable payoffs. Leitmann (1974), Dockner and Jorgensen (1984), Hamalainen et al (1986), Yeung and Petrosyan (2005 and 2006), de-Paz et al (2013), and Marin-Solano (2014) studied continuous-time cooperative differential games with non-transferable payoffs.

* This work was supported by the Saint-Petersburg State University No.9.38.245.2014

Sorger (2006) presented a recursive Nash bargaining solution for a discrete-time NTU cooperative dynamic game involving a productive asset. In NTU cooperative dynamic games, the inapplicability of transfer payments makes the derivation of subgame consistent solutions extremely strenuous. Yeung and Petrosyan (2005) presented subgame consistent solution in cooperative stochastic differential games with non-transferable payoffs under a constant weight scheme. However, the result is confined to a specific class of games under a very restrictive set of optimality principles. Yeung and Petrosyan (2015) provides an effective way in achieving subgame consistency using variable weights under a wide range of optimality principles in cooperative dynamic games.

This article extends Yeung and Petrosyan's analysis on subgame consistent solution in NTU cooperative dynamic games to NTU cooperative stochastic dynamic games. A mechanism for the derivation of subgame consistent solution in NTU cooperative stochastic dynamic games using a variable payoff weights scheme is presented. The game formulation and mathematical preliminaries are given in Section 2. Subgame consistency for NTU cooperative stochastic dynamic games under variable payoff weights are defined and examined in Section 3. In Section 4 a theorem is established to characterize and derive subgame consistent solutions. An illustration in public goods provision is given in Section 5. Concluding remarks are provided in Section 6.

2. Game Formulation and Mathematical Preliminaries

Consider the general $T-$stage $n-$person nonzero-sum discrete-time stochastic dynamic game with initial state x_1^0. The state space of the game is $X \in R^m$ and the state dynamics of the game is characterized by the stochastic difference equation:

$$x_{k+1} = f_k(x_k, u_k^1, u_k^2, \cdots, u_k^n) + G_k(x_k)\theta_k, \tag{2.1}$$
$$\text{for } k \in \{1, 2, \cdots, T\} \equiv \kappa \text{ and } x_1 = x_1^0,$$

where $u_k^i \in R^{m_i}$ is the control vector of player i at stage k, and $x_k \in X$ is the state of the game and θ_k is a set of independent random variable. The payoff that player i seeks to maximize is

$$E_{\theta_1,\theta_2,\cdots,\theta_T}\left\{ \sum_{\zeta=1}^{T} g_\zeta^i[x_\zeta, u_\zeta^1, u_\zeta^2, \cdots, u_\zeta^n, x_{\zeta+1}] + q^i(x_{T+1}) \right\}, \tag{2.2}$$
$$\text{for } i \in \{1, 2, \cdots, n\} \equiv N,$$

where $q^i(x_{T+1})$ is the terminal payoff that player i will received in stage $T+1$, and $E_{\theta_1,\theta_2,\cdots,\theta_T}$ is the expectation operation with respect to the statistics of θ_1,θ_2,$\cdots$, θ_T.

The payoffs of the players are not transferable. Using the standard HJB equations approach for solving stochastic dynamic games a feedback Nash equilibrium of the game can be characterzied (see Basar and Olsder (1999)). Let $\{\phi_k^i(x)$, for $k \in \kappa$ and $i \in N\}$ denote a set of strategies that provides a feedback Nash equilibrium solution to the game (2.1)-(2.2), and $\{V^i(k, x)$, for $k \in \kappa$ and $i \in N\}$ denote the value functions yielding the payoff to player i over the stages from k to $T+1$. Since the analysis is on cooperative schemes for improving the non-cooperative outcomes in NTU stochastic dynamic games, we would consider games with non-cooperative Nash equilibrium outcomes.

2.1. Cooperation Under Constant Weights

To enhance their payoffs the players would consider formulating a cooperative scheme. In particular, the players agree to cooperate and enhance their payoffs according to an agreed-upon optimality principle. Since payoffs are non-transferable the payoffs of individual players are directly determined by the optimal cooperative strategies adopted. Pareto efficient cooperative strategies can be derived from the maximization of the expected weighted sum of payoffs of the players under a set of agreed-upon payoff weights. See Leitmann (1974), Dockner and Jorgensen (1984), Hamalainen et al (1986) Yeung and Petrosyan (2005 and 2015) and Yeung et al (2007) for examples.

To establish the optimization foundation of the variable weights schme we consider first the case in which the players adopt a vector of constnt payoff weights $\alpha = (\alpha^1, \alpha^2, \cdots, \alpha^n)$ in all stages, where $\sum_{j=1}^{n} \alpha^j = 1$. Conditional upon the vector of weights α, the players' optimal cooperative strategies can be generated by solving the stochastic dynamic programming problem of maximizing the expected weighted sum of payoffs :

$$E_{\theta_1,\theta_2,\cdots,\theta_T} \left\{ \sum_{j=1}^{n} \left[\sum_{k=1}^{T} \alpha^j g_k^j(x_k, u_k^1, u_k^2, \cdots, u_k^n) + \alpha^j q^j(x_{T+1}) \right] \right\} \tag{2.3}$$

subject to (2.1).

An optimal solution to the problem (2.1) and (2.3) can be characterized by the technique of stochastic dynamic programming. Let $\{\psi_k^{(\alpha)i}(x)$, for $k \in \kappa$ and $i \in N\}$ denote the set of cooperative strategies that provides an optimal solution to the problem (2.1) and (2.3), and let $W^{(\alpha)}(k, x)$, for $k \in K$, denote the expected weighted sum of cooperative payoffs over the stages from k to $T + 1$. Substituting the optimal control $\{\psi_k^{(\alpha)i}(x)$, for $k \in \kappa$ and $i \in N\}$ into (2.1), one can obtain the cooperative trajectory $x_k^{(\alpha)} \in X_k^{(\alpha)}$, for $k \in \kappa$.

2.2. Individual Payoffs and Individual Rationality

Given that all players are adopting the cooperative strategies the payoff of player i under cooperation can be obtained as:

$$W^{(\alpha)i}(t, x) = E_{\theta_t,\theta_{t+1},\cdots,\theta_T} \left\{ \sum_{k=t}^{T} g_k^i[x_k^{(\alpha)}, \psi_k^{(\alpha)1}(x_k^{(\alpha)}), \psi_k^{(\alpha)2}(x_k^{(\alpha)}), \cdots \right.$$

$$\left. \cdots, \psi_k^{(\alpha)n}(x_k^{(\alpha)})] + q^i(x_{T+1}^{(\alpha)}) \,\middle|\, x_t^{(\alpha)} = x \right\}, \tag{2.4}$$

for $i \in N$ and $t \in \kappa$.

To allow the derivation of the functions $W^{(\alpha)i}(t, K)$ in a more direct way we follow the analysis in Yeung (2013) and characterize individual player's payoffs under cooperation by the following Theorem.

Theorem 2.1. *The payoff of player i over the stages from t to $T+1$ can be characterized as the value function $W^{(\alpha)i}(t,x)$ satisfying the following recursive system of equations:*

$$W^{(\alpha)i}(T+1,x) = q^i(x_{T+1}),$$
$$W^{(\alpha)i}(t,x) = E_{\theta_t}\Big\{ \; g_t^i[x,\psi_t^{(\alpha)1}(x),\psi_t^{(\alpha)2}(x),\cdots,\psi_t^{(\alpha)n}(x)]$$
$$+ W^{(\alpha)i}[t+1, f_t(t,\psi_t^{(\alpha)1}(x),\psi_t^{(\alpha)2}(x),\cdots,\psi_t^{(\alpha)n}(x)) + G(x)\theta_t] \; \Big\},$$
$$\textit{for} \quad i\in N \quad \textit{and} \quad t\in\kappa. \quad (2.5)$$

Proof. $W^{(\alpha)i}(t,x)$ in (2.4) can be expressed as:

$$W^{(\alpha)i}(t,x) = E_{\theta_t,\theta_{t+1},\cdots,\theta_T}\Big\{ \; g_t^i[x,\psi_t^{(\alpha)1}(x),\psi_t^{(\alpha)2}(x),\cdots,\psi_t^{(\alpha)n}(x)]$$
$$+ \sum_{k=t+1}^{T} g_k^i[x_k^{(\alpha)},\psi_k^{(\alpha)1}(x_k^{(\alpha)}),\psi_k^{(\alpha)2}(x_k^{(\alpha)}),\cdots,\psi_k^{(\alpha)n}(x_k^{(\alpha)})] + q^i(x_{T+1}^{(\alpha)}) \; \Big\}. \quad (2.6)$$

Invoking (2.4) again, we have:

$$W^{(\alpha)i}(t+1,x_{t+1}^{(\alpha)}) =$$
$$E_{\theta_{t+1},\theta_{t+2},\cdots,\theta_T}\Big\{ \; \sum_{k=t+1}^{T} g_k^i[x_k^{(\alpha)},\psi_k^{(\alpha)1}(x_k^{(\alpha)}),\psi_k^{(\alpha)2}(x_k^{(\alpha)}),\cdots\cdots,\psi_k^{(\alpha)n}(x_k^{(\alpha)})]$$
$$+ q^i(x_{T+1}^{(\alpha)}) \; \Big\}, \quad (2.7)$$

Using (2.6) and (2.7), one can obtain (2.5). □

For individual rationality to be maintained throughout all the stages $t\in\kappa$, it is required that:

$$W^{(\alpha)i}(t,x_t^{(\alpha)}) \geq V^i(t,x_t^{(\alpha)}), \quad \text{for} \quad i\in N \quad \text{and} \quad t\in\kappa. \quad (2.8)$$

Let the set of weights α that satisfies (2.8) be denoted by Λ. If Λ is not an empty set, a vector $\hat{\alpha} = (\hat{\alpha}^1,\hat{\alpha}^2,\cdots\;\cdots,\hat{\alpha}^n)\in\Lambda$ agreed upon by all players would yield a cooperative solution which satisfies both individual rationality and Pareto optimality throughout the cooperation duration.

Remark 2.1. The pros of the constant payoff weights scheme is that full Pareto efficiency is satisfied in the sense that there does not exist any strategy path which would enhance the cooperative payoff of a player without lowering the cooperative payoff of at least one of the other players in all stages.

The cons of the constant payoff weights scheme include the inflexibility in accommodating the preferences of the players according to the initial cooperative agreement and the high possibility of the non-existence of the set of weights Λ that satisfies individual rationality throughout the cooperation duration. □

In the cited literature on NTU cooperative dynamic games in Section 1 only Sorger (2006), Marin-Solano (2014) and Yeung and Petrosyan (2015) adopted a variable payoff weights scheme.

3. Subgame Consistent Cooperative Solution

Now, we proceed to consider subgame consistent solutions in NTU cooperative stochastic dynamic games. A salient property of a subgame consistent solution is that the agreed-upon optimality principle remains in effect at each stage of the game and hence the players do not possess incentives to deviate from the solution plan. Let $\Gamma(t, x_t)$ denote the cooperative game in which the objective of player i is

$$E_{\theta_t,\theta_{t+1},\cdots,\theta_T}\left\{\sum_{k=t}^{T} g_k^i(x_k, u_k^1, u_k^2, \cdots, u_k^n) + q^i(x_{T+1})\right\}, \quad \text{for} \quad i \in N, \tag{3.1}$$

and the state dynamics is

$$x_{k+1} = f_k(x_k, u_k^1, u_k^2, \cdots, u_k^n) + G_k(x_k)\theta_k, \tag{3.2}$$

for $k \in \{t, t+1, \cdots, T\}$ and the state at stage t is x_t.

Let the agreed-upon optimality principle be denoted by $P(t, x_t)$. For subgame consistency to be maintained the agreed-upon optimality principle $P(t, x_t)$ must be satisfied in the subgame $\Gamma(t, x_t)$ for $t \in \{1, 2, \cdots, T\}$. Hence, when the game proceeds to any stage t, the agreed-upon solution policy remains effective. Examples of optimality principles $P(t, x_t)$ include criteria like the Nash bargaining solution, cooperative gains proportional to non-cooperatives payoffs and the mid-value of feasible payoff weights.

A time-invariant weights scheme is usually hardly applicable for the derivation of a subgame consistent solution in general. As stated in Remark 2.1, the set Λ which satisfies individual rationality throughout the game duration is often empty. In general, typtical optimality principles in classical game theory could not be maintained as the game proceeds under a time-invariant payoff weights cooperative scheme.

To derive a set of subgame consistent strategies in a cooperative solution with optimality principle $P(t, x_t)$ a variable payoff weight scheme has to be adopted. In particular, at each stage $t \in \kappa$ the players would adopt a vector of payoff weights $\hat{\alpha}_t = (\hat{\alpha}_t^1, \hat{\alpha}_t^2, \cdots, \hat{\alpha}_t^n)$ for $\sum_{j=1}^{n} \hat{\alpha}_t^j = 1$ which satisfies the agreed-upon optimality principle. The chosen set of weights $\hat{\alpha}_t = (\hat{\alpha}_t^1, \hat{\alpha}_t^2, \cdots, \hat{\alpha}_t^n)$ must lead to the satisfaction of the optimality principle $P(t, x_t)$ in the subgame $\Gamma(t, x_t)$ for $t \in \{1, 2, \cdots T\}$.

4. Derivation of Subgame Consistent Cooperative Strategies

To derive the optimal cooperative strategies in a subgame consistent solution for NTU cooperative stochastic dynamic games with variable payoff weights we invoke the principle of backward induction and begin with the final stage of the cooperative game.

4.1. Optimal Cooperative Strategies in Ending Stages

Consider first the last operation stage, that is stage T, with the state $x_T = x \in X$. The players will select a set of payoff weight $\alpha_T = (\alpha_T^1, \alpha_T^2, \cdots, \alpha_T^n)$ which satisfies the optimality principl $P(T,x)$. The players' optimal cooperative strategies can be generated by solving the stochastic dynamic programming problem of maximizing the weighted sum of their payoffs

$$E_{\theta_T}\left\{ \sum_{j=1}^{n}\left[\alpha_T^j g_T^j(x_T,u_T^1,u_T^2,\cdots,u_T^n)+\alpha_T^j q^j(x_{T+1}) \right] \right\} \tag{4.1}$$

subject to

$$x_{T+1}=f_T(x_T,u_T^1,u_T^2,\cdots,u_T^n)+G_T(x_T)\theta_T, \quad x_T=x. \tag{4.2}$$

Let $\{u_T^i = \psi_T^{(\alpha_T)i}$, for $i\in N\}$ denote the optimal cooperative stratgies in stage T that solves the stochastic dynamic programming problem (4.1)-(4.2). When all players are adopting the cooperative strategies the payoff of player i under cooperation covering stages T and $T+1$ can be obtained as:

$$W^{(\alpha_T)i}(T,x)=E_{\theta_T}\left\{ g_T^i[x,\psi_T^{(\alpha_T)1}(x),\psi_T^{(\alpha_T)2}(x),\cdots,\psi_T^{(\alpha_T)n}(x)]+q^i(x_{T+1}^{(\alpha_T)}) \right\},$$
$$\text{for} \quad i\in N. \tag{4.3}$$

Invoking Theorem 2.1 one can characterize player i's payoff $W^{(\alpha_T)i}(T,x)$ by the following equations

$$\begin{aligned} &W^{(\alpha_T)i}(T+1,x)=q^i(x),\\ &W^{(\alpha_T)i}(T,x)=E_{\theta_T}\Big\{ g_T^i[x,\psi_T^{(\alpha_T)1}(x),\psi_T^{(\alpha_T)2}(x),\cdots,\psi_T^{(\alpha_T)n}(x)]\\ &+W^{(\alpha_T)i}[T+1,f_T(x,\psi_T^{(\alpha_T)1}(x),\psi_T^{(\alpha_T)2}(x),\cdots,\psi_T^{(\alpha_T)n}(x))+G_T(x)\theta_T] \Big\} \end{aligned}$$
$$\text{for} \quad i\in N. \tag{4.4}$$

For individual rationality to be maintained, it is required that:

$$W^{(\alpha_T)i}(T,x)\geq V^i(T,x), \quad \text{for} \quad i\in N. \tag{4.5}$$

Since the maximization problem (4.1)-(4.2) with payoff weight α_T yields a Pareto optimal cooperative solution and the non-cooperative outcome is (in general) sub-optimal there always exists a set of weights that satisfies (4.5). We use Λ_T to denote the set of weights α_T that satisfies (4.5). Then we use $\hat{\alpha}_T=(\hat{\alpha}_T^1,\hat{\alpha}_T^2,\cdots,\hat{\alpha}_T^n)\in\Lambda_T$ to denote the payoff weights in stage T that leads to the satisfaction of the optimality principle $P(T,x)$.

Now we proceed to cooperative scheme in the second to last stage. Given that the payoff of player i in stage T is $W^{(\hat{\alpha}_T)i}(T,x)$, his payoff covering stages $T-1$ to $T+1$ can be expressed as:

$$E_{\theta_{T-1}}\left\{ g_{T-1}^i(x_{T-1},u_{T-1}^1,u_{T-1}^2,\cdots,u_{T-1}^n)+W^{(\hat{\alpha}_T)i}(T,x_T) \right\},$$
$$\text{for} \quad i\in N. \tag{4.6}$$

In this stage the players will select payoff weights $\alpha_{T-1} = (\alpha^1_{T-1}, \alpha^2_{T-1}, \cdots, \alpha^n_{T-1})$ which satisfy optimality principle $\Gamma(T-1, x)$. The players' optimal cooperative strategies $\{u^i_{T-1} = \psi^{(\alpha_{T-1})i}_{T-1}$, for $i \in N\}$ in stage $T-1$ can be generated by solving the stochastic dynamic programming problem of maximizing

$$E_{\theta_{T-1}} \left\{ \sum_{j=1}^{n} \alpha^j_{T-1} \left[g^j_{T-1}(x_{T-1}, u^1_{T-1}, u^2_{T-1}, \cdots, u^n_{T-1}) + W^{(\hat{\alpha}_T)j}(T, x_T) \right] \right\} \tag{4.7}$$

subject to

$$x_T = f_{T-1}(x_{T-1}, u^1_{T-1}, u^2_{T-1}, \cdots, u^n_{T-1}) + G_{T-1}(x_{T-1})\theta_{T-1}, \quad x_{T-1} = x. \tag{4.8}$$

Invoking Theorem 2.1 one can characterize the payoff of player i under cooperation covering the stages $T-1$ to $T+1$ by:

$$W^{(\alpha_{T-1})i}(T, x) = W^{(\hat{\alpha}_T)i}(T, x_T),$$

$$W^{(\alpha_{T-1})i}(T-1, x) = E_{\theta_{T-1}} \Big\{ g^i_{T-1}[x, \psi^{(\alpha_{T-1})1}_{T-1}(x), \psi^{(\alpha_{T-1})2}_{T-1}(x), \cdots, \psi^{(\alpha_{T-1})n}_{T-1}(x)]$$

$$+W^{(\alpha_{T-1})i}[T, f_{T-1}(x, \psi^{(\alpha_{T-1})1}_{T-1}(x), \psi^{(\alpha_{T-1})2}_{T-1}(x), \cdots, \psi^{(\alpha_{T-1})n}_{T-1}(x)) + G_{T-1}(x)\theta_{T-1}] \Big\},$$

$$\text{for} \quad i \in N. \tag{4.9}$$

For individual rationality to be maintained, it is required that:

$$W^{(\alpha_{T-1})i}(T-1, x) \geq V^i(T-1, x), \quad \text{for} \quad i \in N. \tag{4.10}$$

We use Λ_{T-1} to denote the set of weights α_{T-1} that satisfies (4.10). We use the vector $\hat{\alpha}_{T-1} = (\hat{\alpha}^1_{T-1}, \hat{\alpha}^2_{T-1}, \cdots \cdots, \hat{\alpha}^n_{T-1}) \in \Lambda_{T-1}$ to denote the set of payoff weights that leads to satisfaction of the optimality principle $\Gamma(T-1, x)$.

4.2. Optimal Cooperative Strategies in Preceding Stages

Now we proceed to characterize the cooperative scheme in stage $k \in \{1, 2, \cdots, T-2\}$. Following the analysis in Section 4.1, the players will select a set of weights $\alpha_k = (\alpha^1_k, \alpha^2_k, \cdots, \alpha^n_k)$ which satisfies the optimality principle $P(k, x)$. The players' optimal cooperative strategies $\{u^i_k = \psi^{(\alpha_k)i}_k$, for $i \in N\}$ in stage k can be generated by solving the following stochastic dynamic programming problem of maximizing

$$E_{\theta_k} \left\{ \sum_{j=1}^{n} \alpha^j_k \left[g^j_k(x_k, u^1_k, u^2_k, \cdots, u^n_k) + W^{(\hat{\alpha}_{k+1})j}(k+1, x_{k+1}) \right] \right\}, \tag{4.11}$$

subject to

$$x_{k+1} = f_k(x_k, u^1_k, u^2_k, \cdots, u^n_k) + \theta_k, \quad x_k = x. \tag{4.12}$$

Invoking Theorem 2.1 the payoff of player i under cooperation can be characterized by the following equations

$$
\begin{aligned}
&W^{(\alpha_k)i}(k+1,x) = W^{(\hat{\alpha}_{k+1})i}(k+1,x),\\
&W^{(\alpha_k)i}(k,x) = E_{\theta_k}\Big\{\ g_k^i[x,\psi_k^{(\alpha_k)1}(x),\psi_k^{(\alpha_k)2}(x),\cdots,\psi_k^{(\alpha_k)n}(x)]\\
&+W^{(\alpha_k)i}[k+1,f_k(x,\psi_k^{(\alpha_k)1}(x),\psi_k^{(\alpha_k)2}(x),\cdots,\psi_k^{(\alpha_k)n}(x))+G_k(x)\theta_k]\ \Big\},\\
&\text{for}\quad i\in N. \qquad (4.13)
\end{aligned}
$$

For individual rationality to be maintained in stage k, it is required that:

$$
W^{(\alpha_k)i}(k,x) \geq V^i(k,x), \quad \text{for} \quad i \in N. \qquad (4.14)
$$

We use Λ_k to denote the set of weights α_k that satisfies (4.14). Again, we use $\hat{\alpha}_k = (\hat{\alpha}_k^1, \hat{\alpha}_k^2, \cdots, \hat{\alpha}_k^n) \in \Lambda_k$ to denote the set of payoff weights that leads to the satisfaction of the optimal principle $P(k,x)$, for $k \in \kappa$.

4.3. Subgame Consistent Solution: A Mathematical Theorem

A theorem characterizing a subgame consistent solution of the cooperative stochastic dynamic game (2.1)-(2.2) with the optimality principle $P(k,x_k)$ can be obtained as follows.

Theorem 4.1. *A set of payoff weights $\{\hat{\alpha}_k = (\hat{\alpha}_k^1, \hat{\alpha}_k^2, \cdots, \hat{\alpha}_k^n)$, for $k \in \kappa\}$ and a set of strategies $\{\psi_k^{(\hat{\alpha}_k)i}(x)$, for $k \in \kappa$ and $i \in N\}$ provides a subgame consistent solution to the cooperative stochastic dynamic game (2.1)-(2.2) with optimality principle $P(k,x)$ if there exist functions $W^{(\hat{\alpha}_k)}(k,x)$ and $W^{(\hat{\alpha}_k)i}(k,x)$, for $i \in N$, $k \in \kappa$, which satisfy the following recursive relations:*

$$
\begin{aligned}
&W^{(\hat{\alpha}_{T+1})i}(T+1,x) = q^i(x_{T+1}),\\
&W^{(\hat{\alpha}_k)}(k,x) = \max_{u_k^1,u_k^2,\cdots,u_k^n}\Big\{\ E_{\theta_k}\Big[\ \sum_{j=1}^{n}\hat{\alpha}^j g_k^j(x,u_k^1,u_k^2,\cdots,u_k^n)\\
&+\sum_{j=1}^{n}\hat{\alpha}_k^j W^{(\hat{\alpha}_{k+1})j}[k+1,f_k(x,u_k^1,u_k^2,\cdots,u_k^n)+G_k(x)\theta_k]\ \Big]\ \Big\};\\
&W^{(\hat{\alpha}_k)i}(k,x) = E_{\theta_k}\Big\{\ g_k^j(x,\psi_k^{(\hat{\alpha}_k)1}(x),\psi_k^{(\hat{\alpha}_k)2}(x),\cdots,\psi_k^{(\hat{\alpha}_k)n}(x))\\
&+W^{(\hat{\alpha}_{k+1})i}[k+1,f_k(x,\psi_k^{(\hat{\alpha}_k)1}(x),\psi_k^{(\hat{\alpha}_k)2}(x),\cdots,\psi_k^{(\hat{\alpha}_k)n}(x))+G_k(x)\theta_k]\ \Big\},
\end{aligned}
$$

for $i \in N$ and $k \in \kappa$; and the optimality principle $P(k,x)$ in all stages $k \in \kappa$. (4.15)

Proof. Follow the analysis from equation (4.1) to equation (4.13) in Sections 4.1 and 4.2. □

In the case when the agreed-upon optimality principle requires the proportion each player's cooperative payoff to his non-cooperative payoff being equal, the optimality principle $P(k,x)$ in Theorem 4.1 becomes

$$\frac{W^{(\hat{\alpha}_k)i}(k,x)}{V^i(k,x)} = \frac{W^{(\hat{\alpha}_k)j}(k,x)}{V^j(k,x)}, \quad \text{for} \quad i,j \in N \quad \text{and} \quad k \in \kappa.$$

If the optimality principle requires the satisfaction of the Nash bargaining solution, $P(k,x)$ becomes

$$\hat{\alpha}_k = \arg\max_{\alpha_k} \left\{ \prod_{j=1}^{n} [W^{(\alpha_k)j}(k,x) - V^j(k,x)] \right\};$$

for $k \in \kappa$.

In the two-player case, the optimality principle $P(k,x)$ may require the chosen payoff weights $\hat{\alpha}_k = \{\hat{\alpha}_k^1, \hat{\alpha}_k^2\}$ to be the mid-value of the maximum and minimum of the payoff weight α_k^1 and that of the payoff weights α_k^2 in the set Λ.

Remark 4.1. The subgame consistent solution presented in Theorem 4.1 is conditional Pareto efficient in the sense that the solution is a Pareto efficient outcome satsifying the condition that the agreed-upon optimality principle is maintained in all stages.

Remark 4.2. A subgame consistent solution is fully Pareto efficient only if the optimality principle $P(t,x)$ requires the choice of a set of time-invariant payoff weights.

Only very restrictive optimality principles in specific game structures would yield subgame consistent time-invariant weights (see Yeung and Petrosyan (2005)). Since full Pareto efficiency is of less importance than reaching a cooperative solution, achieving the latter at the expense of the former is a practical way out.

5. An Illustration in Public Capital Build-up

Consider a stochastic version of the example in Yeung and Petrosyan (2015) in which there are 2 asymmetric agents. These agents receive benefits from an existing public capital stock x_t at each stage $t \in \{1, 2, \cdots, 4\}$. The accumulation dynamics of the public capital stock is governed by the stochastic difference equation:

$$x_{k+1} = x_k + \sum_{j=1}^{2} u_k^j - \delta\, x_k + \theta_k x_k, \quad x_1 = x_1^0, \quad \text{for} \quad t \in \{1, 2, 3\}, \tag{5.1}$$

where u_k^i is the physical amount of investment in the public good and δ is the rate of obsolescence and θ_k is a random variable affecting the rate of obsolescence with range $\{\theta_k^1, \theta_k^2, \theta_k^3, \theta_k^4\}$ and corresponding probabilities $\{\lambda_k^1, \lambda_k^2, \lambda_k^3, \lambda_k^4\}$. In particular, $(\delta - \theta_k^y)$ is non-negative and not greater than one for $y \in \{1, 2, 3, 4\}$ and $t \in \{1, 2, 3\}$.

The objective of agent $i \in \{1, 2\}$ is to maximize the payoff:

$$E_{\theta_1,\theta_2,\theta_3} \left\{ \sum_{k=1}^{3} [a_k^i x_k - c_k^i (u_k^i)^2](1+r)^{-(k-1)} + (q^i x_4 + m^i)(1+r)^{-3} \right\}, \tag{5.2}$$

subject to the dynamics (5.1), where a_k^i, c_k^i, r, q^i and m^i are positive model parameters.

The payoffs of the agents are not transferable and the non-cooperative payoffs of agent i can be obtained as:

$$V^i(t,x) = [A_t^i x + C_t^i](1+r)^{-(t-1)}, \quad \text{for} \quad i \in \{1,2\} \quad \text{and} \quad t \in \{1,2,3\}, \tag{5.3}$$

where

$A_3^i = a_3^i + q^i(1-\delta+\sum_{\ell=1}^{4}\lambda_3^\ell\theta_3^\ell)(1+r)^{-1}$, and

$$C_3^i = -\frac{(q^i)^2(1+r)^{-2}}{4c_3^i} + \left[q^i \left(\sum_{j=1}^{2} \frac{q^j(1+r)^{-1}}{2c_3^j} \right) + m^i \right](1+r)^{-1},$$

$A_2^i = a_2^i + A_3^i(1-\delta+\sum_{\ell=1}^{4}\lambda_2^\ell\theta_2^\ell)(1+r)^{-1}$, and

$$C_2^i = -\frac{1}{4c_2^i}\left(A_3^i(1+r)^{-1} \right)^2 + \left[A_3^i \left(\sum_{j=1}^{2} \frac{A_3^j(1+r)^{-1}}{2c_2^j} \right) + C_3^i \right](1+r)^{-1} \Bigg\};$$

$A_1^i = a_1^i + A_2^i(1-\delta+\sum_{\ell=1}^{4}\lambda_1^\ell\theta_1^\ell)(1+r)^{-1}$, and

$$C_1^i = -\frac{1}{4c_1^i}\left(A_2^i(1+r)^{-1} \right)^2$$
$$+\left[A_2^i \left(\sum_{j=1}^{2} \frac{A_2^j(1+r)^{-1}}{2c_1^j} \right) + C_2^i \right](1+r)^{-1} \Bigg\}; \text{ for } i \in \{1,2\}.$$

5.1. Cooperative Solution

Now consider first the case when the agents agree to cooperate and maintain an optimality principle $P(t,x_t)$ requiring the adoption of the mid values of the maximum and minimum of the payoff weight α_t^i in the set Λ_t, for $i \in \{1,2\}$ and $t \in \{1,2,3\}$.

Invoking the technique of stochastic dynamic programming the value function $W^{(\alpha_3)}(3,x)$ in stage 3 can be obtained as:

$$W^{(\alpha_3)}(3,x) = [A_3^{(\alpha_3)}x + C_3^{(\alpha_3)}](1+r)^{-2}, \tag{5.4}$$

where

$A_3^{(\alpha_3)} = \sum_{j=1}^{2}\alpha_3^j \left[a_3^j + q^j(1-\delta+\sum_{\ell=1}^{4}\lambda_3^\ell\theta_3^\ell)(1+r)^{-1} \right]$, and

$$C_3^{(\alpha_3)} = -\sum_{j=1}^{2}\alpha_3^j \left[\frac{(1+r)^{-2}}{4\alpha_3^j c_3^j} \left(\sum_{\ell=1}^{2}\alpha_3^\ell q_3^\ell \right)^2 \right]$$

$$+\sum_{j=1}^{2}\alpha_3^j \left[q^j \left(\sum_{j=1}^{2} \frac{(1+r)^{-1}}{2\alpha_3^j c_3^j} \left(\sum_{\ell=1}^{2}\alpha_3^\ell q^\ell \right) \right) + m^j \right](1+r)^{-1} \Bigg\}.$$

Invoking Theorem 2.1, the payoff of player i under cooperation can be obtained as:

$$W^{(\alpha_3)i}(3,x) = [A_3^{(\alpha_3)i}x + C_3^{(\alpha_3)i}](1+r)^{-2}, \tag{5.5}$$

for $i \in \{1,2\}$,

where $A_3^{(\alpha_3)i} = \left[\; a_3^j + q^j(1-\delta+\sum_{\ell=1}^{4}\lambda_3^\ell\theta_3^\ell)(1+r)^{-1} \; \right]$, and

$$C_3^{(\alpha_3)i} = -\left[\; \frac{(1+r)^{-2}}{4\alpha_3^i c_3^i}\left(\; \sum_{\ell=1}^{2}\alpha_3^\ell q_3^\ell \; \right)^2 \; \right]$$

$$+\left[\; q^i\left(\; \sum_{j=1}^{2}\frac{(1+r)^{-1}}{2\alpha_3^j c_3^j}\left(\sum_{\ell=1}^{2}\alpha_3^\ell q^\ell\right) \; \right) + m^i \; \right](1+r)^{-1} \; \Bigg\}.$$

To identify the range of α_3 that satisfies individual rationality we examine the functions which gives the excess of agent i's cooperative over his non-cooperative payoff, that is

$$W^{(\alpha_3)i}(3,x) - V^i(3,x) = [C_3^{(\alpha_3)i} - C_3^i](1+r)^{-2}, \quad \text{for} \quad i \in \{1,2\}, \tag{5.6}$$

because $A_3^{(\alpha_3)i} = A_3^i$.

For individual rationality to be satisfied, it is required that $W^{(\alpha_3)i}(3,x) - V^i(3,x) \geq 0$ for $i \in \{1,2\}$. Using $\alpha_3^j = 1 - \alpha_3^i$ and upon rearranging terms $C_3^{(\alpha_3)i}$ can be expressed as:

$$\begin{aligned} C_3^{(\alpha_3)i} = q^i \Bigg[\; & \frac{(1+r)^{-2}}{2c_3^i}\left(\; \frac{\alpha_3^i q^i + (1-\alpha_3^i)q^j}{\alpha_3^1} \; \right) \\ & + \frac{(1+r)^{-2}}{2c_3^j}\left(\; \frac{\alpha_3^i q^i + (1-\alpha_3^i)q^j}{1-\alpha_3^i} \; \right) \; \Bigg] + m^i(1+r)^{-1} \\ & - \frac{(1+r)^{-2}}{4c_3^i}\left(\; \frac{\alpha_3^i q^i + (1-\alpha_3^i)q^j}{\alpha_3^i} \; \right)^2, \\ & \text{for} \quad i,j \in \{1,2\} \quad \text{and} \quad i \neq j. \end{aligned} \tag{5.7}$$

Differentiating $C_3^{(\alpha_3)i}$ with respect to α_3^i yields

$$\frac{\partial C_3^{(\alpha_3)i}}{\partial \alpha_3^i} = \frac{(1+r)^{-2}}{2c_3^j}\left(\; \frac{(q^i)^2}{(1-\alpha_3^i)^2} \; \right) + \frac{(1+r)^{-2}}{2c_3^i}\left(\; \frac{(1-\alpha_3^i)q^j}{\alpha_3^i} \; \right)\left(\; \frac{q^j}{(\alpha_3^i)^2} \; \right), \tag{5.8}$$

which is positive for $\alpha_3^i \in (0,1)$.

One can readily observed that $\lim_{\alpha_3^i \to 0} C_3^{(\alpha_3)i} \to -\infty$ and $\lim_{\alpha_3^i \to 1} C_3^{(\alpha_3)i} \to \infty$. Since the cooperative solution is Pareto optimal and the non-cooperative outcome is (in general) suboptimal an $\underline{\alpha}_3^i \in (0,1)$ can be obtained such that

$$W^{(\underline{\alpha}_3^i, 1-\underline{\alpha}_3^i)i}(3,x) = V^i(3,x)$$

and yields agent i's minimum payoff weight value satisfying his own individual rationality. Similarly there exist an $\bar{\alpha}_3^i \in (0,1)$ such that

$$W^{(\bar{\alpha}_3^i, 1-\bar{\alpha}_3^i)j}(3,x) = V^j(3,x)$$

and yields agent i's maximum payoff weight value while maintaining agent j's individual rationality. According to the agreed-upon optimality principle $P(t,x_t)$, the cooperative weights in stage 3 is $\hat{\alpha}_3 = \left(\frac{\underline{\alpha}_3^i + \bar{\alpha}_3^i}{2}, 1 - \frac{\underline{\alpha}_3^i + \bar{\alpha}_3^i}{2}\right)$.

Now consider the stage 2 problem. We use $W^{(\hat{\alpha}_3)j}(3,x)$ for $j \in \{1,2\}$ to form the terminal payoff $\sum_{j=1}^{2} \alpha_2^j W^{(\hat{\alpha}_3)j}(3,x)$ for the cooperation scheme in stage 2. Following the analysis in stage 3, one can obtain

$$W^{(\alpha_2)}(2,x) = [A_2^{(\alpha_2)} x + C_2^{(\alpha_2)}](1+r)^{-1},$$

$$W^{(\alpha_2)i}(2,x) = [A_3^{(\alpha_3)i} x + C_3^{(\alpha_3)i}](1+r)^{-2}, \quad \text{for} \quad i \in \{1,2\},$$

where $A_2^{(\alpha_2)}$, $C_2^{(\alpha_2)}$, $A_3^{(\alpha_3)i}$ and $C_3^{(\alpha_3)i}$ are functions that depend on α_2.

One can readily verified that $A_t^{(\alpha_t)i} = A_t^i$ is independent of α_t and $C_t^{(\alpha_t)i}$ is strictly increasing in α_t^i and $C_t^{(\alpha_t)j}$ is strictly decreasing in α_t^i. Hence agent i's minimum payoff weight is $\underline{\alpha}_2^i \in (0,1)$ which leads to

$$W^{(\underline{\alpha}_2^i, 1-\underline{\alpha}_2^i)i}(2,x) = V^i(2,x),$$

and his maximum payoff weight is $\bar{\alpha}_2^i \in (0,1)$ which leads to

$$W^{(\bar{\alpha}_2^i, 1-\bar{\alpha}_2^i)j}(2,x) = V^j(2,x).$$

Invoking the agreed-upon optimality principle $P(t,x_t)$ the cooperative weights in stage 2 is $\hat{\alpha}_2 = \left(\frac{\underline{\alpha}_2^i + \bar{\alpha}_2^i}{2}, 1 - \frac{\underline{\alpha}_2^i + \bar{\alpha}_2^i}{2}\right)$.

Finally, following the analysis in stages 2 and 3, one can obtain the cooperative weights in stage 1 as $\hat{\alpha}_1 = \left(\frac{\underline{\alpha}_1^i + \bar{\alpha}_1^i}{2}, 1 - \frac{\underline{\alpha}_1^i + \bar{\alpha}_1^i}{2}\right)$.

Note that the parameter a_t^i in $C_t^{(\alpha_t)i}$ may be different at different stages of $t \in \{1,2,3\}$, therefore $\underline{\alpha}_t^i$ and $\bar{\alpha}_t^i$ cannot be the same. In the finite horizon, even if the parameter a_t^i is time-invariant, $\underline{\alpha}_t^i$ and $\bar{\alpha}_t^i$ would change as t changes because of the adjustments towards the terminal condition. Moreover, in general, there is no guarantee for the existence of a constant payoff weight such that the basic requirement of individual rationality is satisfied in all subsequent stages. An example is provided below.

Example 5.1. Consider the case in which $q^1 = 3$, $q^2 = 4$, $m^1 = 10$, $m^2 = 20$, $r = 0.05$, $\delta = 0.02$, $c_3^1 = 2$, $c_3^2 = 4$, $a_3^1 = 4$, $a_3^2 = 1$, $c_2^1 = 7$, $c_2^2 = 2$, $a_2^1 = 1$, $a_2^2 = 2$, $c_1^1 = 1$, $c_1^2 = 4$, $a_1^1 = 2$, $a_1^2 = 1$. In stage 1, a constant α_1^1 has to be between $\underline{\alpha}_1^1 = 0.435$ and $\bar{\alpha}_1^1 = 0.545$. In stage 2, a constant α_2^1 has to be between $\underline{\alpha}_2^1 = 0.33$ and $\bar{\alpha}_2^1 = 0.43$. In stage 3, α_3^1 has to be between $\underline{\alpha}_3^1 = 0.55$ and $\bar{\alpha}_3^1 = 0.655$. Therefore there does not exist a constant choice of α_t^1 for $t \in \{1,2,3\}$ such that individual rationality is satisfied in all the subsequent subgame stages.

5.2. Other Optimality Principles

In this section, we consider deriving subgame consistent solutions for the cooperative stochastic dynamic game (5.1)-(5.2) under two alternative optimality principles. Consider first the case where the agents agree with an optimality principle $P(t, x_t)$ that requires the excess of the players' cooperative payoffs over their respective non-cooperative payoffs satisfies the Nash bargaining solution. They would first search for an α_3 in stage 3 to maximize the Nash product

$$\prod_{j=1}^{2} [W^{(\alpha_3)j}(3, x) - V^j(3, x)].$$

Invoking (5.6), the issue becomes solving the problem

$$\max_{\alpha_3^i} \prod_{j=1}^{2} [C_3^{(\alpha_3)j} - C_3^j](1 + r)^{-2}, \tag{5.9}$$

in the range of $\alpha_3^i \in [\underline{\alpha}_3^i, \bar{\alpha}_3^i]$.

Invoking the derivative property in (5.8), there exist an $\hat{\alpha}_3^i \in [\underline{\alpha}_3^i, \bar{\alpha}_3^i]$ that solves problem (5.9).

Then one can obtain $W^{(\hat{\alpha}_3)j}(3, x)$ for $j \in \{1, 2\}$ and used them to solve the cooperation scheme in stage 2. Repeating the above analysis, one can identify $\hat{\alpha}_2$ which yields the Nash bargaining solution in stage 2. Finally, in a similar manner, $\hat{\alpha}_1$ which yields the Nash bargaining solution in stage 1 can be obtained.

Now consider another optimality principle $P(t, x_t)$ which requires the proportion of each player's cooperative payoff to his non-cooperative payoff to be equal. In particular, a subgame consistent solution requires payoff weights $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$ leading to

$$\frac{W^{(\hat{\alpha}_t)1}(t, x_t)}{V^1(t, x_t)} = \frac{W^{(\hat{\alpha}_t)2}(t, x_t)}{V^2(t, x_t)}, \quad \text{for} \quad t \in \{1, 2, 3\},$$

along the cooperation trajectory.

Invoking the value functions $V^i(t, x)$ and $W^{(\hat{\alpha}_t)i}(t, x)$, for $i \in \{1, 2\}$ and $t \in \{1, 2, 3\}$, a subgame consistent solution to the problem can be obtained with the payoff weights $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$ which satisfy:

$$\frac{A_t^{(\hat{\alpha}_t)1} x_t + C_t^{(\hat{\alpha}_t)1}}{A_t^1 x_t + C_t^1} = \frac{A_t^{(\hat{\alpha}_t)2} x_t + C_t^{(\hat{\alpha}_t)2}}{A_t^2 x_t + C_t^2}, \quad \text{for} \quad t \in \{1, 2, 3\}, \tag{5.10}$$

and

$$x_{t+1} = x_t \sum_{j=1}^{2} \left[\frac{(1 + r)^{-1}}{2\hat{\alpha}_t^j c_t^i} \sum_{\ell=1}^{2} \hat{\alpha}_t^\ell A_{t+1}^{(\hat{\alpha}_{t+1})\ell} \right] - \delta x_t, \quad x_1 = x_1^0, \tag{5.11}$$

for $t \in \{12, 3\}$ and $A_4^{(\hat{\alpha}_4)\ell} = q^\ell$.

Again with $A_t^{(\alpha_t)i} = A_t^i$ being independent of α_t and $C_t^{(\alpha_t)i}$ being strictly increasing in α_t^i and $C_t^{(\alpha_t)j}$ being strictly decreasing in α_t^i for $\alpha_t^i \in [\underline{\alpha}_t^i, \bar{\alpha}_t^i]$, therefore one can readily identify payoff weights $\hat{\alpha}_t^i$ such that (5.10) is satisfied.

6. Concluding Remarks

This paper considers subgame consistent solutions in NTU cooperative stochastic dynamic games using variable payoff schemes. A mathematical theorem characterizing such a solution under different optimality principles is established. It resolves the problems of the lack of guarantee for the agreed-upon optimality principle to be maintained throughout the planning duration. The analysis contributes to the solving of subgame consistent solutions for NTU cooperative stochastic dynamic games under a wide range of optimality principles.

References

Basar, T. and G. J. Olsder (1999). *Dynamic Noncooperative Game Theory.* Philadelphia, SIAM.

de-Paz, A., J. Marn-Solano and J. Navas (2013). *Time-Consistent Equilibria in Common Access Resource Games with Asymmetric Players Under Partial Cooperation.* Environmental Modeling & Assessment, **18(2)**, 171–184.

Dockner, E. and S. Jørgensen (1984). *Cooperative and Non-cooperative Differential Games Solutions to an Investment and Pricing Problem.* The Journal of the International Operational Research Society, **35**, 731–739.

Hamalainen, R., A. Haurie and V. Kaitala (1986). *Equilibria and Threats in a Fishery Management Game.* Optimal Control Applications and Methods, **6**, 315–333.

Leitmann, G. (1974). *Cooperative and Non-Cooperative Many Players Differential Games.* New York, Springer-Verlag.

Marin-Solano, J. (2014). *Time-consistent Equilibria in a Differential Game Model with Time Inconsistent Preferences and Partial Cooperation*, forthcoming in J. Haunschmied, V. Veliov and S. Wrzaczek (eds.) Dynamic Games in Economics, Springer Verlag, New York, pp. 219–238.

Sorger, G. (2006). *Recursive Bargaining Over A Productive Asset.* Journal of Economic Dynamics and Control, **30**, 2637–2659

Yeung, D. W. K. (2013). *Nontransferable Individual Payoffs in Cooperative Stochastic Dynamic Games.* International Journal of Algebra, **7(13)**, 597–606.

Yeung, D. W. K. and L. A. Petrosyan (2004). *Subgame Consistent Cooperative Solutions in Stochastic Differential Games.* Journal of Optimization Theory Applications, **120**, 651–666.

Yeung, D. W. K. and L. A. Petrosyan (2005). *Subgame Consistent Solution of A Cooperative Stochastic Differential Game With Nontransferable Payoffs.* Journal of Optimization Theory and Applications, **124(3)**, 701–724.

Yeung, D. W. K. and L. Petrosyan (2006). *Cooperative Stochastic Differential Games*, New York: Springer-Verlag, 242 pp.

Yeung, D. W. K. and L. A. Petrosyan (2010). *Subgame Consistent Solutions for cooperative Stochastic Dynamic Games.* Journal of Optimization Theory and Applications, **145(3)**, 579–596.

Yeung, D. W. K. and L. A. Petrosyan (2015). *Subgame Consistent Cooperative Solution for NTU Dynamic Games via Variable Weights*, forthcoming in Automatica.

Yeung, D. W. K., L. A. Petrosyan, and P. M. Yeung (2007). *Subgame Consistent Solutions for a Class of Cooperative Stochastic Differential Games with Nontransferable Payoffs.* Annals of the International Society of Dynamic Games, **9**, 153–170.