Nash Equilibrium Strategies in Discrete-Time Finite-Horizon Dynamic Games with Risk-and Effort-Averse Players

Dan Protopopescu[1]
Center for the Study of Organizations and Decisions in Economics
Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona, Spain
E-mail: *dan.protopopescu@uab.cat*

November 2009

Abstract

The objective of this paper is to re-examine the risk-and effort attitude in the context of strategic dynamic interactions stated as a discrete-time finite-horizon Nash game. The analysis is based on the assumption that players are endogenously risk-and effort-averse. Each player is characterized by distinct risk-and effort-aversion types that are unknown to his opponent. The goal of the game is the optimal risk-and effort-sharing between the players. It generally depends on the individual strategies adopted and, implicitly, on the the players' types or characteristics.

Keywords: Dynamic Nash game, optimal path, closed-loop control, endogenous risk-and effort-aversion, adaptive risk-and effort management, optimal risk-and effort-sharing.

JEL Classification: C71, C73, D81, D82.

1. Introduction

For reaching better results, very often in practice, the players are incited to cooperate. A non-cooperative solution is not necessarily satisfying. In the absence of cooperation, it is possible that the results are inferior to what could be achieved with coordinated behavior. An additional act of cooperation always brings a positive contribution.

A cooperative decision making approach is useful when dealing with a strategic complexity of the players. Although the players cooperate, they can remain independent. Moreover, a cooperative behavior may emerge in non-cooperative situations when the nature of the interaction is for a long term.

In the real world, there exist many situations when the players are engaged in a game containing both cooperative and non-cooperative elements. For example, they can form cooperative coalitions which interact in a non-cooperative fashion. Moreover, the players can exhibit variable degrees of cooperation. The form of the relationships between the players' actions and the resulting outputs may affect the level of cooperation.

Suppose that two players with potentially conflicting interests share the same dynamic environment. Its trajectory is influenced by their individual actions. Some pure conflicting interests are generally rare and in most economic applications there are both, common and conflicting interests present.

When players have common strategic objectives, it is expected that the quality and the speed with which decisions are undertaken are improved. Inappropriate decisions or uncorrelated objectives can generate unexpected outputs, with undesirable consequences on the equilibrium and stability of the game.

Nash equilibrium basically requires rational players and mutual beliefs about actions (AUMANN AND BRANDENBURGER 1995). The rationality of the players is characterized by the anticipation that the system will be affected by other factors than their own control instruments. These factors are either completely or partially observed and may be exogenous variables or unobserved random shocks. Even if the players act in an individually rational way, their strategic interactions can quite often cause collective irrationality.

Multiplayer models are generally complicated, requiring some specified form of interaction or some premise as to the outcome of this interaction. In such of situations, the problem to be solved is, on the one hand, the nature of equilibrium and its definition, and on the other hand, the existence and uniqueness of equilibruim and its stability. All these aspects must be considered in a world which changes.

It is possible that the system shifts from an equilibrium to another due to permanent random shocks. The final equilibrium will not be hence determined by the initial conditions alone. There is a path dependency of equilibria, such that past equilibria impact the trajectory of future equilibria.

The analysis is placed in the context of the closed-loop control theory, the information being utilized by the players in real time. Closed-loop games are appropriate modelling approaches to economic environmental problems, being often employed in empirical modelling and policy analyses. The closed-loop decisions are specific to some short term interactions. In this case, the players set their control vectors as functions of the system history. They thus attempt to reduce the uncertainties related to the choice of their actions by acquiring new information in the game.

The dynamic learning process is endogenous, that is, the players learn from observing and base their actions on their state of knowledge at the point where the actions are taken. By their learning process, the players affect the strategic variables they are learning about. Their decision rules are reviewed and revised in response to new signals from the system. The players thus refine the distance between the current target position and their fixed objectives at each stage of the game. The control inputs implemented by the players are purported to contribute towards equilibrium and stability of the dynamic system.

The process of control has the implementation stage of an individual plan built into it. It is in the players' interest to move in their desired direction at each stage of the game.

Although there is a large amount of literature on game theory, there is little work focused on Nash-type equilibria incorporating time-varying endogenous risk inherent to the system. Most of prior research has only considered the case of exogenous risk, that is, external to the system, and hence beyond the control of the players. There is generally an intimate relationship between endogenous risk and adaptive effort behavior. In real world situations, the fluctuating level of the risk is an input in the effort estimation process. The players' effort behavior is thus endogenous by nature. This important aspect is almost ignored in the literature, resulting in a biased analysis of equilibrium dynamics. When players are risk-and effort-averse, it is optimal for them to share risk and effort.

The purpose of the present study is to re-evaluate the traditional approach, by taking into account endogenous risk-and effort-averse behavior in dynamic strategic interactions stated as a Nash game.

The paper is organized as follows. Section 2 describes the model. Section 3 deals with the players' risk behavior. Section 4 focuses on the players' effort behavior. Section 5 defines the information structure of the game. Section 6 introduces the concept of optimal Nash equilibrium strategy sensitive to endogenous risk and effort and provides a complete analytical characterization of the closed-loop solution path. Section 7 concludes and makes suggestions for further research.

2. The Model

An environment consists of the specification of three distinct objects. First, it is the set of players. Second, the preferences of each player. And finally, the information structure, that is, the set of parameters which can be observed by each player.

The players are characterized by an individual rationality, in the sense that they will not refuse to act in accordance with the efficient outcome. They are supposed to share the same environment which generally have a different impact on their individual behavior.

In the present study, we consider the case of a dynamic Nash game composed of two playing periods, say $[-T_1, 0]$ and $[1, T_2]$. We only analyze the second period $[1, T_2]$ by taking into account the information acquired by the players in the first period $[-T_1, 0]$. The players are indexed by an ordered pair $(i, j) \in \{(1, 2), (2, 1)\}$, where the index j refers to the opponent of the player i.

Let us give the formal statement of the problem. The discrete time periods are denoted by $t = 1, ..., T_2$. The end of the planning horizon is generally dictated by the degree of uncertainty in the external environment. Higher the uncertainty, shorter the planning horizon. The variables involved in the problem are listed as follows:

Let $x_t^{(i)} \in \mathbf{R}^{q_i}$ be the player i's control instrument for the period t, and let $y_t \in \mathbf{R}^p$ be the observed target variable in t. There is generally a stochastic relationship between players' actions and observations. Let $z_t \in \mathbf{R}^r$ be an exogenous variable not subjected to the control of the players at time t, and let $e_t^{(i)}$ be the player i's effort level in t. Denote by $u_t \in \mathbf{R}^p$ an exogenous environmental "white noise" modelled by a normal random variable with zero mean-vector and finite variance-covariance matrix Ψ.

We make the following basic assumptions:

Assumption 1. The evolution of the system is modelled by a discrete-time multivariate linear stochastic process:

$$y_t = A_t y_{t-1} + C_t^{(1)} x_t^{(1)} + C_t^{(2)} x_t^{(2)} + E_t^{(1)} e_t^{(1)} + E_t^{(2)} e_t^{(2)} + B_t z_t + D_t + u_t, \ t = 1, ..., T_2$$

where $\beta_t \stackrel{not.}{=} (A_t, C_t^{(1)}, C_t^{(2)}, E_t^{(1)}, E_t^{(2)}, B_t, D_t) \in \mathbf{R}^k$ is the time-varying parameter to be estimated. This specifies the structure of the model according to the information available in t.

Assumption 2. The first step in the decision making process is to establish a set of non-conflicting objectives. The exploitation of the environment is generally asymmetric and the players have different anticipations about the system evolution. Their individual preferences are described by distinct output pathways:

$$\eta^{(i)} \stackrel{not.}{=} \{y_1^{g(i)}, y_2^{g(i)}, ..., y_{T_2}^{g(i)}\}, \ i = 1, 2$$

These represent targets or goals that the players wish to achieve during the period of the game. In the context of a partial cooperation, the players' targets are closed, while for a full cooperation, these are supposed to be identical.

In order to ensure the system stability during the entire planning process, the players are supposed to choose small values for their individual targets:

$$0 < y_t^{g(i)} \le l_t < 1, \ t = 1, ..., T_2$$

where l_t are common optimal bounds selected according to foreseeable movements in y.

For simplicity, one can suppose that $l_t = l \in (0,1)$ for all $t = 1, ..., T_2$. Smaller targets generally require a higher risk-aversion, as well as a higher effort to be invested.

Large deviations of the system from the fixed targets signify that $\| y_t - y_t^{g(i)} \| > 1$ for $t = 1, ..., T_2$. These are probably due to the measurement errors, and may be always negotiated ex-post. For this study, it is assumed that the players' targets remain unchanged for the sample observations.

Remark 1. In the case of a common optimal path $\eta = \{y_1^g, y_2^g, ..., y_{T_2}^g\}$, with $y_t^g \stackrel{def.}{=} \alpha y_t^{g(1)} + (1-\alpha)y_t^{g(2)}$ and $\alpha \in (0,1)$, the difficulty to overcome is to choice the optimal value of the parameter α that satisfies the individual preferences of the players.

Remark 2. Arrive to an agreement generally implies a tâtonnement process from the part of the players. The stage which will put the bases of the agreement will be crucial. Even if the game starts in a non-cooperative manner, a potential cooperation can occur in every stage of the game, or the players can have the intention to play cooperatively in the future. Promises about future behavior can influence the players' current behavior. Note here that the agreements to play a Nash equilibrium are fragile when players have a strict preference over their opponent's strategy choice (AUMANN 1990).

Assumption 3. The timing of the game is as follows: At each period t, the players implement their actions $x_t^{(i)}$ ($i = 1, 2$), which are a stimulus for the system. A shock u_t is realized and they observe the output (or impulse response) y_t. A signal about the future trend of the system is thus extracted. The uncertainty is reduced only ex-post, that is, only after the informative output message has been received. This output together with the corresponding actions provide information on the data generating process. The players employ this output-signal for a strategic learning (specific to a closed-loop monitoring) in order to drive the system as close as possible to their desired path $\eta^{(i)}$.

Assumption 4. The optimality of the instrument $x_t^{(i)}$ is considered with respect to a local criterion $W_t^{(i)}$ which measures the system output deviation $\triangle^{(i)} y_t \stackrel{not.}{=} y_t - y_t^{g(i)}$ and, respectively, the instrument deviation from the current position to the desired goal state, denoted by $\triangle^{(i)} x_t \stackrel{not.}{=} x_t - x_t^{g(i)}$.

The player i's instrument target $x_t^{g(i)}$ desired to be achieved at time t is not necessarily equal with the equilibrium target realized in t. It may be possible that differences between ex-ante decisions and ex-post results exist.

Following PINDYCK (1976), we choose a quadratic asymmetric performance criterion dependent on the instruments:

$$W_t^{(i)}(y_t) \stackrel{def.}{=} \overline{W}_t^{(i)}(y_t) + \widetilde{W}_t^{(i)}(x_t^{(i)}) + \widehat{W}_t^{(i)}(x_t^{(j)}), \ i, j = 1, 2; \ i \ne j$$

where $\overline{W}_t^{(i)}, \widetilde{W}_t^{(i)}, \widehat{W}_t^{(i)}$ are quadratic functions supposed twice continuously differentiable, strictly increasing and convex:

$$\overline{W}_t^{(i)}(y_t) \stackrel{def.}{=} (y_t - y_t^{g(i)})'K_{t,i}(y_t - y_t^{g(i)}) + 2(y_t - y_t^{g(i)})'d_{t,i}, \ i = 1,2$$
$$\widetilde{W}_t^{(i)}(x_t^{(i)}) \stackrel{def.}{=} (x_t^{(i)} - x_t^{g(i)})'Q_{iit}(x_t^{(i)} - x_t^{g(i)}) + 2(x_t^{(i)} - x_t^{g(i)})'q_{iit}, \ i = 1,2$$
$$\widehat{W}_t^{(i)}(x_t^{(j)}) \stackrel{def.}{=} (x_t^{(j)} - x_t^{g(j)})'Q_{ijt}(x_t^{(j)} - x_t^{g(j)}) + 2(x_t^{(j)} - x_t^{g(j)})'q_{ijt}, \ i,j = 1,2; \ i \neq j$$

with a prime denoting transpose.

The parameters Q_{iit}, q_{iit}, and respectively Q_{ijt}, q_{ijt}, show the importance attributed to the realization of the targets $x_t^{g(i)}$, and respectively $x_t^{g(j)}$ ($t = 1,...,T_2$). In other words, it is taken into account the importance given to the reduction of the difference between the instruments and their corresponding target values. The desired values of the instruments are included in the quadratic loss function to prevent them from going too far away from realistic values. The corresponding weighting parameters are related to the relative loss attributed to the non-achievement of the target values for the instruments.

The asymmetry of the criterion derives, on the one hand, from the difference in penalty costs that the players may attach to errors (depending on whether these are errors of over-shooting or errors of shortfall about the targets) and, on the other hand, due to the additional terms, dependent on the instruments, introduced in the expression of the quadratic performance function.

The treatment of the errors to either side of the targets $y_t^{g(i)}$ and $x_t^{g(i)}$ ($t = 1,...,T_2$; $i = 1,2$) is not symmetric. In other words, the players are not indifferent as regards the sign of the deviations $\triangle^{(i)} Y \stackrel{not.}{=} (\triangle^{(i)} y_1, ..., \triangle^{(i)} y_{T_2})$ and, respectively, $\triangle^{(i)} X \stackrel{not.}{=} (\triangle^{(i)} x_1, ..., \triangle^{(i)} x_{T_2})$.

The decision for choosing certain parameters $K_{t,i}$ and $d_{t,i}$ reflects the priorities of the player i and also depends on the available amount of information concerning the future development of the system parameters. It allows to weight differently the various loss components. At each period t, the parameters $K_{t,i}$ and $d_{t,i}$ are updated and new optimal values are chosen in order to satisfy the player i's requirements. The players choose the parameters which yield less fluctuating controls.

In this formulation of the model, we deal with a reference-dependent behavior of the players. Experimental evidence in a variety of economic environments show that the behavior of decision-makers depends on strategic reference points.

Assumption 5. The player i computes his optimal policy $\widehat{x}^{(i)} \stackrel{not.}{=} (\widehat{x}_1^{(i)}, \widehat{x}_2^{(i)}, ..., \widehat{x}_{T_2}^{(i)})$ conditional to the policy rule of his opponent. In both cooperative and non-cooperative world, the players adjust their instrument settings period by period to accommodate past information errors and revised projections for the remaining planning periods.

This is the classical context, often employed in the literature, where the hypothesis of risk-neutrality for both players is adopted for the entire period of the game.

3. Endogenous Risk-Aversion

Attitudes to risk may differ across players. There is by now accumulating evidence that players differ substantially in their risk-preferences. They generally have different degrees of uncertainty to achieve the desired outputs.

We define the player i's risk-aversion indice level at time t by:

$$\varphi_t^{(i)} \stackrel{def.}{=} \frac{\| y_{t-1} - y_{t-1}^{g(i)} \|^2 \overline{L}_{t-1,i} + ... + \| y_{t-\overline{k_i}} - y_{t-\overline{k_i}}^{g(i)} \|^2 \overline{L}_{t-\overline{k_i},i}}{\sqrt{(\| y_{t-1} - y_{t-1}^{g(i)} \|^2 + ... + \| y_{t-\overline{k_i}} - y_{t-\overline{k_i}}^{g(i)} \|^2)^2 + \overline{l}_i}}, \quad t = 1,...,T_2$$

where:

- $-1 < \overline{L}_{t-1,i} < \overline{L}_{t-2,i} < ... < \overline{L}_{t-\overline{k}_i,i} < 0$ are strategic weights attached to the system deviations from the player i's optimal path $\{y_{t-1}^{g(i)}, ..., y_{t-\overline{k}_i}^{g(i)}\}$.
- $\overline{l}_i \geq 1$ is a fixed integer which characterizes the player i's type (more or less risk-averse).
- $1 \leq \overline{k}_i < T_2$ is an optimal number of periods (the most informative) for the player i.

Denote by $\{y_0, ..., y_{1-\overline{k}_i}, ...\}$ and $\{y_0^g, ..., y_{1-\overline{k}_i}^g, ...\}$ the history of the process and, respectively, the targets of the player i during the first playing period $[-T_1, 0]$. The set of parameters $\{\overline{L}_{0,i}, ..., \overline{L}_{1-\overline{k}_i,i}, ...\}$ represents strategic weights attached to the system deviations with respect to the player i's optimal path for the period $[-T_1, 0]$.

One can imagine several scenarios when comparing the degree of risk-aversion exhibited by the players: i) when both players have similar coefficients of risk-aversion during the period of the game (i.e., $\mid \varphi_t^{(1)} - \varphi_t^{(2)} \mid$ is small, with $t \in \{1, ..., T_2\}$); ii) when both players have the same coefficient of risk-aversion at given periods of time (i.e., $\varphi_t^{(1)} = \varphi_t^{(2)}$, with $t \in \{1, ..., T_2\}$ taking fixed values); and iii) when players have very different degrees of risk-aversion during the period of the game (i.e., $\mid \varphi_t^{(1)} - \varphi_t^{(2)} \mid$ is high, with $t \in \{1, ..., T_2\}$).

Let $\varphi_{\min}^{(i)}$ be a risk-aversion threshold fixed by the player i before starting the control and for the entire period $[1, T_2]$. When this limit threshold is exceeded, the player i becomes excessively risk-averse for the current control period, being characterized by an extreme pessimism.

It is important to distinguish between $\varphi_t^{(i)}$ and $\varphi_{\min}^{(i)}$ $(i = 1, 2)$. In other words, it must distinguish between local risk-aversion (at time t) and global risk-aversion (over the entire period $[1, T_2]$).

It is thus possible to differentiate the types of common /distinct players as follows:

$$-1 < \varphi_{\min,\,more}^{(i)} < -\frac{l\overline{k}_i}{\sqrt{[l\overline{k}_i]^2 + \overline{l}_{i,more}}} \text{ (more risk-averse player)}$$

and

$$-\frac{l\overline{k}_i}{\sqrt{[l\overline{k}_i]^2 + \overline{l}_{i,less}}} < \varphi_{\min,\,less}^{(i)} < 0 \text{ (less risk-averse player)}$$

with $1 \leq \overline{l}_{i,more} < \overline{l}_{i,less}$ two strategic parameters which characterize the risk-averse type of the player i. For further details, see PROTOPOPESCU (2007).

The number of players in the game can modify their individual attitude towards risk but not their types.

4. Endogenous Effort Behavior

It is well-known that the uncertainty diminishes with time and effort. If the players are supposed to be risk-averse, there is no obvious reason why they should also be assumed to be effort-averse. We could measure the effort-aversion just as the risk-aversion is measured. In general, the players differ in their effort-preferences.

We define the player i's effort level at time t by:

$$e_t^{(i)} \overset{def.}{=} -\frac{\parallel y_{t-1} - y_{t-1}^{g(i)} \parallel^2 \overline{L}_{t-1,i}}{\parallel y_{t-1} - y_{t-1}^{g(i)} \parallel^2 + \overline{s}_i} \cdot \frac{\varphi_{t-1}^{(i)} - 1}{\varphi_{t-1}^{(i)} - \overline{d}_i}, \quad t = 1, ..., T_2$$

where $\overline{s}_i \in (0, 1]$ and $\overline{d}_i \in (1, 2]$ are two strategic parameters which characterize the player i's type (more or less effort-averse).

Each player gives a greater importance to the system deviation which is closer to the moment of implementation of a new optimal action. The optimal effort level at each stage of the game is dictated by the system evolution and the players' effort behavior profiles. It does not always take a minimum value, as the intuition would suggest. The system will not incite equally the players to invest effort.

There exist situations where the players are interested in increasing the effort by enforcing the active learning in order to improve their objectives. Further effort is necessary if the risk is to be reduced or avoided in the future. In a fluctuating system, increasing the effort is not always sufficient to avoid a high risk. It may be the case of a system with a high inertia. Increases in the effort cost (measured in terms of disutility) result in reduced effort levels. This is a direct consequemce of the importance the players place on the system deviations with respect to their fixed targets.

Optimal effort-sharing generally depends on the players' fixed objectives. The closer their objectives, the closer their attitudes to effort over time. The equilibrium and stability of the dynamic system is dependent on the cumulative efforts of the players.

It is easy to see that $e_t^{(i)}$ is a non-monotonous function of t. There exist periods where the players need to allocate a higher effort with respect to others. The effort invested by the players is no more seen as a pure disutility, like in the traditional approach, but rather as an efficient instrument in optimizing the equilibrium solution.

One can imagine several scenarios when comparing the effort exhibited by the players: i) when both players exert similar efforts during the period of the game (i.e., $\mid e_t^{(1)} - e_t^{(2)} \mid$ is small, with $t \in \{1, ..., T_2\}$); ii) when both players exert the same effort level at given periods of time (i.e., $e_t^{(1)} = e_t^{(2)}$, with $t \in \{1, ..., T_2\}$ taking fixed values); and iii) when players exert very different efforts during the period of the game (i.e., $\mid e_t^{(1)} - e_t^{(2)} \mid$ is high, with $t \in \{1, ..., T_2\}$).

The players' utilities for the period t are assumed to be multiplicative separable in the effort level $e_t^{(i)}$ and the loss function $W_t^{(i)}$:

$$U_t^{(i)}(W_t^{(i)}, \varphi_t^{(i)}, e_t^{(i)}) \overset{def.}{=} \frac{2D^{(i)}(e_t^{(i)})}{\varphi_t^{(i)}}[\exp(-\frac{\varphi_t^{(i)}}{2}W_t^{(i)}) - 1], \quad i = 1, 2; \ t = 1, ..., T_2$$

where $D^{(i)}$ are twice continuously differentiable functions such that:

$$D^{(i)}(e_t^{(i)}) > 0, \quad D^{(i)\prime}(e_t^{(i)}) > 0, \quad D^{(i)\prime\prime}(e_t^{(i)}) > 0, \ \forall \ e_t^{(i)} \in [0, 1), \quad \forall \ t = 1, ..., T_2$$

Dynamic games in which each player has an exponential utility function are referred to as risk-sensitive dynamic games. Due to the imperfect information about the system reaction, it is perfectly reasonable to consider a maximization on a short-time horizon for the players' utility functions. The expected utility level attained by a player may be measured exactly by his opponent only under idealized conditions.

A higher effort does not necessarily ensure a lower utility level. It is also important to note that a risk-averse player does not necessarily invest a higher effort level compared to a risk-neutral one at the same stage of the game:

$$\varphi_{t-1}^{(i)} < \varphi_{t-1}^{(i)r.n.} = 0 \nRightarrow e_t^{(i)} > e_t^{(i)r.n.}$$

Suppose that the player i will fix before starting the game an optimal effort-aversion threshold $e_{\max}^{(i)}$ which must not be exceeded during the entire control period. If it happens, he becomes excessively effort-averse for the current period of the game.

The effort thresholds $e_{\max}^{(i)}$ $(i = 1, 2)$ are chosen such that these characterize the best the player i's type. It is important to distinguish between $e_t^{(i)}$ and $e_{\max}^{(i)}$. In other words, it must distinguish between local effort level (at time t) and global effort-aversion (over the entire period $[1, T_2]$).

Define the effort-aversion threshold of the player i by:

$$e_{\max}^{(i)} \stackrel{def.}{=} \frac{1}{\overline{s_i} + 1} \cdot \frac{\varphi_{\min}^{(i)} - 1}{\varphi_{\min}^{(i)} - \overline{d_i}}, \quad i = 1, 2$$

Note that an exceeding of the risk-aversion threshold $\varphi_{\min}^{(i)}$ does not necessarily imply an exceeding of the effort-aversion threshold $e_{\max}^{(i)}$. If the player i is less (more) risk-averse by nature, then he will choose a higher (smaller) effort-aversion threshold $e_{\max}^{(i)}$. There is thus a trade-off between less risk-aversion and higher effort investment.

We can give a complete characterization of the players' types in terms of effort-averse behavior. We have:

$$1 > e_{\max,\,less}^{(i)} > \frac{1}{\overline{s_i} + 1} \cdot \frac{-\frac{l\overline{k_i}}{\sqrt{[l\overline{k_i}]^2 + \overline{l_{i,more}}}} - 1}{-\frac{l\overline{k_i}}{\sqrt{[l\overline{k_i}]^2 + \overline{l_{i,more}}}} - \overline{d_i}} \quad \text{(less effort-averse player)}$$

and

$$0 < e_{\max,\,more}^{(i)} < \frac{1}{\overline{s_i} + 1} \cdot \frac{-\frac{l\overline{k_i}}{\sqrt{[l\overline{k_i}]^2 + \overline{l_{i,less}}}} - 1}{-\frac{l\overline{k_i}}{\sqrt{[l\overline{k_i}]^2 + \overline{l_{i,less}}}} - \overline{d_i}} \quad \text{(more effort-averse player)}$$

In the case where the player i is characterized by a high risk-aversion in $t-1$ and the system deviation $\| y_{t-1} - y_{t-1}^{g(i)} \|$ is large, he will allocate a high effort level for the period t. In contrast, if $\varphi_{t-1}^{(i)}$ exceeds $\varphi_{\min}^{(i)}$ but the value of $\| y_{t-1} - y_{t-1}^{g(i)} \|$ is small, then the player i is not incited to invest a good effort in t.

Note that $e_{\max}^{(i)} < e_{\max}^{(j)}$ $(i, j = 1, 2; i \neq j)$ does not necessarily imply that $e_t^{(i)} < e_t^{(j)}$ at time t. Moreover, an exceeding of the threshold $e_{\max}^{(i)}$ for the player i does not necessarily induce an exceeding of the threshold $e_{\max}^{(j)}$ for his opponent. In other words, the players do not necessarily have an imitative behavior.

Denote by $U_t^{(i)}(W_t^{(i)}, \varphi_{\min}^{(i)}, e_{\max}^{(i)})$ the disutility threshold of the player i at time t. We have:

$$\varphi_{\min}^{(i)} \leq \varphi_t^{(i)} \text{ and } e_t^{(i)} \leq e_{\max}^{(i)} \Rightarrow U_t^{(i)}(W_t^{(i)}, \varphi_t^{(i)}, e_t^{(i)}) \geq U_t^{(i)}(W_t^{(i)}, \varphi_{\min}^{(i)}, e_{\max}^{(i)})$$

We can thus distinguish between two disutility thresholds according to the player i's individual type: i) a smaller disutility threshold, denoted by $U_{t\text{-}s}^{(i)}(W_t^{(i)}, \varphi_{\min,\,more}^{(i)}, e_{\max,\,less}^{(i)})$; and ii) a higher disutility threshold, denoted by $U_{t\text{-}h}^{(i)}(W_t^{(i)}, \varphi_{\min,\,less}^{(i)}, e_{\max,\,more}^{(i)})$.

During the entire game period, the player i has three strategic objectives: i) to optimally track the targets $y_t^{g(i)}$; ii) not to exceed the risk-aversion threshold $\varphi_{\min}^{(i)}$; and iii) not to exceed the effort-aversion threshold $e_{\max}^{(i)}$.

The relationship between the players' reaction to the perceived states of nature and their attitude to effort is complex. In general, effort aversion makes the reaction stronger than effort neutrality.

In the case of a cooperative game with common fixed objectives, the evolution of the system generates close attitudes to effort for the players. It is also useful to note that the length of

the game has a non-negligible effect on the players' effort behavior. The shorter the planning horizon, the higher the players' effort-aversion before starting the game. A closed-loop Nash game is therefore correlated with smaller effort-thresholds fixed by the players. For further details, see PROTOPOPESCU (2008).

Note that the equilibrium of the game is not necessarily realized with minimum-effort coordination, as other authors have assumed in related issues (ANDERSON ET AL. 1996, 2001).

5. Game's Information Structure

Most commonly studied asymmetric Nash games in the literature are games where there are not identical strategy sets for the players. However, it is possible for a Nash game to have identical strategies for all players, yet be asymmetric.

The player i's behavior at time t is characterized by a strategy $s_t^{(i)} : I^{(i)} \to \mathbf{R}^{q_i}$ which prescribes, according to the information set $I^{(i)}$, an action /distribution over the feasible strategy set $S^{(i)}$. This determines the ability of the player i to affect the set of feasible strategies of his opponent.

The player i's information set $I^{(i)}$ can contain: i) the system parameters of interest; ii) an a priori distribution on the initial state of the system; iii) the distribution of the input errors; iv) his fixed output targets; v) his desired instruments targets; vi) his local utility /loss levels (and maybe of his opponent); vii) his individual risk distribution; viii) his individual effort distribution; and ix) his individual strategy (and maybe of his opponent).

The player i's information set at time t is given by the whole observable history up to that period. This is endogenously determined by the equilibrium behavior itself. More exactly, the information set is reached along the equilibrium path of the dynamic system. In equilibrium, the information sets differ across players.

Let $C \overset{def.}{=} \underset{i=\overline{1,2}}{\cap} I^{(i)}$ be the common knowledge of the game. It transmits information about how the game has been played up to each information set $I^{(i)}$. Even if the players have access to the same information, the losses caused by information errors (although errors common to both players) can easily be larger for non-cooperative rather than for cooperative solutions because under cooperation the risks are effectively spread over the players.

The players receive imperfect information about the system and this information is not shared. They adjust their individual actions until their decisions are mutually consistent. The player i is not generally informed about all the knowledge of his opponent even if the private information is of common interest. If the players differ in their degree of information accurracy, then the equilibrium of the game can be seen as non-symmetrical.

Note that in most games, common knowledge of the structure of the game and of players' rationality are not sufficient to predict that an equilibrium will be played, even if it is the unique equilibrium of the game. Nash equilibrium concept requires not only common knowledge of rationality, and of the game, but also that each player has the right expectations about what his opponent is going to play (especially when there are multiple equilibria).

When both players are rational and their rationality is common knowledge, then the resulting outcome will be rationalizable (BERNHEIM 1984; PEARCE 1984).

6. Strategies for a Nash Equilibrium

In this section, closed-loop Nash equilibrium solutions sensitive to endogenous risk and

effort are derived for a discrete time linear-quadratic game under partial information sharing environment. One deals with a strategic equilibrium in which each player takes the reactions of his opponent at least partially into account.

From the game theory point of view, the equilibrium is interpreted as a state from which no player, given the rules, has any incentive to deviate. In other words, the players cannot benefit from deviating from their announced strategy. Therefore, we are looking for combinations of strategies which are such that if one player deviates from his strategy, this one will only lose.

If a player deviates from the equilibrium control rule, the state vector will be driven to a different position. The deviators will be observed only ex-post, as it is most often the case in practice.

Each player learns about his opponent's strategies by observing the actions that have been played. This information is employed to anticipate the behavior of his opponent. New information resolves uncertainty step by step.

The Nash equilibrium concept reduces the set of all possible choices of the players to a much smaller set of those choices that are stable in the sense that no player can improve his payoff by deviating unilaterally (i.e., changing his strategy while the opponent player holds his strategy fixed). Each player is assumed to know the equilibrium strategy of his opponent, not being affected by his own strategy. The strategic interactions are only possible in the space which determines the set of feasible equilibria.

Let $E_{t-1}(\cdot) \stackrel{def.}{=} E(\cdot \mid I_{t-1})$ be the operator of conditional expectation, where I_{t-1} is the information available up to time $t-1$. In what follows, denote by $s_t^{(-i)}$ the strategy of the opponent player at time t. For the present study, we consider the case of a closed-loop (state-dependent) control, when $I_t \in C_t$ for $t = 1, ..., T_2$. A feedback revision process is thus required.

Definition: A set of strategies $(s_t^{o(i)}, s_t^{o(-i)})$ and the corresponding payoffs constitute the Nash equilibrium of the game at period t if each player's strategy yields the player at least as high a payoff as any other strategy, given the strategy of his opponent:

$$E_{t-1}U_t^{(i)}(W_t^{(i)}(s_t^{o(i)}, s_t^{o(-i)})) \geq E_{t-1}U_t^{(i)}(W_t^{(i)}(s_t^{(i)}, s_t^{o(-i)})) \ \ \forall \ i = 1, 2 \ ; \forall \ s_t^{(i)} \in S^{(i)}$$

The player i chooses the strategy $s_t^{o(i)}$ in order to force the opponent player to play $s_t^{o(-i)}$, so as to secure himself an expected quantity. He could gain more by not playing $s_t^{o(i)}$ if the expectation that his opponent cooperates were fulfilled, but he could lose more if it were not. It is precisely in this sense that one can say that a Nash solution is generally risk-averse.

The concept of Nash equilibrium embodies a notion of individual rationality, since each player's equilibrium strategy is the best reply to the opponent's strategy. Unfortunately, this does not specify how each player can arrive to form beliefs about his opponent's strategies that support equilibrium play.

The closed-loop Nash equilibrium is the result of learning from past experience. The first-period actions will alter the next-period decisions. In other words, dynamically strategic players anticipate the outcome of the continuation of the game in making their period-one actions. The strategy of each player is thus justifiable, in the sense that it is optimal, given the anticipation of each player over the possible strategy of his opponent. This dynamic-strategic effect will be integrated in the computation of the closed-loop equilibrium.

In most practical economic applications, besides the model equation constraint for each policy period, additional dynamic constraints for smoothing and bounding the controls may also be present (SANDBLOM & BANASIK 1985).

Define by:

$$\mathbf{R}^{q_i} \supset \Lambda_t^{(i)} \overset{def.}{=} \{ s_t^{(i)} \mid \underbrace{0 < \nu_t^{(i)} \leq s_t^{(i)} \leq \xi_t^{(i)}}_{\text{(amplitude bounds)}} \text{ and } \underbrace{\lambda_t^{(i)} \leq s_t^{(i)} - s_{t-1}^{(i)} \leq \mu_t^{(i)}}_{\text{(change bounds)}} \}$$

the feasible strategies space for the player i at time t.

The first set of constraints is imposed in order to keep the instruments within specific bounds. The wider the bound on the instruments, the higher the importance given by the players to the variation of the instruments in that direction, so that they fit to the active learning process. It is assumed that each player chooses these bounds at each iteration of their control algorithm. In this way, they can exploit the information on the previous instruments when fixing the bounds for the next instruments by allowing a greater variability for an efficient instrument rather than for an inefficient one.

The bounds on the instruments are simply the limits up to which the players decide to extend the research of the optimal solution at each iteration. As regards the last set of constraints, this indicates that the variation of control variables between two consecutive periods lies within prespecified bounded intervals. The values of this variation can be either positive or negative. The two sets of constraints are called boundary conditions. These restrict the set of potential Nash equilibria. In a full cooperative game, the players are supposed to choose the same amplitude /change bounds for their individual controls.

We are now in a position to derive the closed-loop Nash equilibrium sensitive to endogenous risk-and effort-aversion of the players.

Proposition 1: Suppose that the matrices $M_{t,i}(\varphi_t^{(i)}) \overset{not.}{=} \Psi^{-1} + \varphi_t^{(i)} H_{t,i}$ and $N_{t,i}(\varphi_t^{(i)}) \overset{not.}{=} C_t^{(i)\prime} \widetilde{H}_{t,i} C_t^{(i)} + Q_{iit}$, with $\widetilde{H}_{t,i} \overset{not.}{=} K_{t,i} - \varphi_t^{(i)} H_{t,i} M_{t,i}^{-1}(\varphi_t^{(i)}) H_{t,i}$, are inversible for each $t = 1, ..., T_2$ and $i = 1, 2$. Under the hypotheses stated in Section 2, the Nash equilibrium solution of the game is described by the linear equations:

$$s_t^{o(i)} = G_t^{(i)} y_{t-1} + g_t^{(i)}, \ t = 1, ..., T_2; \ i = 1, 2$$

where the matrices $(G_t^{(i)}, g_t^{(i)})$ are given by:

$$G_t^{(i)} = -N_{t,i}^{-1}(\varphi_t^{(i)})(C_t^{(i)\prime} \widetilde{H}_{t,i} A_t)$$

$$g_t^{(i)} = -N_{t,i}^{-1}(\varphi_t^{(i)})\{(Q_{iit} x_t^{g(i)} - q_{iit}) + C_t^{(i)\prime}[\widetilde{H}_{t,i}(C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - R_{t,i} h_{t,i}]\}$$

with

$$R_{t,i} = I_p - \varphi_t^{(i)} K_{t,i} M_{t,i}^{-1}(\varphi_t^{(i)})$$

$$H_{t,i} \overset{not.}{=} K_{t,i}, \ h_{t,i} \overset{not.}{=} K_{t,i} y_t^{g(i)} - d_{t,i}$$

and -1 power denoting inverse.

Proof. The dynamic programming problem is made in discrete time and uncertain future. The optimal policy is computed step by step starting from $s_1^{o(i)}$ to $s_{T_2}^{o(i)}$ (forward through time). We maximize period by period, working every time conditionally to the information acquired.

At each period t, the player i will solve the following stochastic optimization programme:

$$s_t^{o(i)} = \underset{s_t^{(i)} \in \Lambda_t^{(i)}}{\arg \max} E_{t-1} U_t^{(i)}(W_t^{(i)}(s_t^{(i)}, s_t^{o(j)}) \mid s_t^{o(j)})$$

$$s.t.: \begin{cases} y_t = A_t y_{t-1} + C_t^{(i)} s_t^{(i)} + C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t + u_t \\ y_0, y_t > 0, \ t = 1, ..., T_2 \quad \text{(economic constraints)} \end{cases}$$

One can write:

$$s_t^{o(i)} = \underset{s_t^{(i)} \in \Lambda_t^{(i)}}{\arg\max} E_{t-1} U_{t,i}(W_{t,i}(y_t), \varphi_t^{(i)}) = \underset{s_t^{(i)} \in \Lambda_t^{(i)}}{\arg\max} E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2} W_t^{(i)}(y_t))]$$

For the computation of $E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2} W_t^{(i)}(y_t))] \overset{not.}{=} V_t^{(i)}$ (which is supposed to exist), we proceed as follows:

$$E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2} W_t^{(i)}(y_t))] = E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2}(\overline{W}_t^{(i)}(y_t) + \overline{W}_t^{(i)}(s_t^{(i)}) + \overline{W}_t^{(j)}(s_t^{(j)})))]$$

$$= E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2}(\Delta^{(i)} y_t' K_{t,i} \Delta^{(i)} y_t + 2 \Delta^{(i)} y_t' d_{t,i} + \overline{W}_t^{(i)}(s_t^{(i)}) + \overline{W}_t^{(j)}(s_t^{(j)})))]$$

$$= E_{t-1}[\exp(-\frac{\varphi_t^{(i)}}{2}(y_t' H_{t,i} y_t - 2 y_t' h_{t,i} + f_{t,i}) + \overline{W}_t^{(i)}(s_t^{(i)}) + \overline{W}_t^{(j)}(s_t^{(j)}))]$$

where:

$$\Delta^{(i)} y_t \overset{not.}{=} y_t - y_t^{g(i)}, \quad H_{t,i} \overset{def.}{=} K_{t,i}, \quad h_{t,i} \overset{not.}{=} K_{t,i} y_t^{g(i)} - d_{t,i}, \quad f_{t,i} \overset{not.}{=} y_t^{g(i)'}(h_{t,i} - d_{t,i})$$

$$\overline{W}_t^{(i)}(s_t^{(i)}) \overset{def.}{=} (s_t^{(i)} - x_t^{g(i)})' Q_{iit}(s_t^{(i)} - x_t^{g(i)}) + 2(s_t^{(i)} - x_t^{g(i)}) q_{iit}$$

$$\overline{W}_t^{(j)}(s_t^{(j)}) \overset{def.}{=} (s_t^{(j)} - x_t^{g(j)})' Q_{ijt}(s_t^{(j)} - x_t^{g(j)}) + 2(s_t^{(j)} - x_t^{g(j)}) q_{ijt}$$

Substituting $A_t y_{t-1} + C_t^{(i)} s_t^{(i)} + C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t + u_t$ for y_t, one obtains:

$$V_t^{(i)} = E_{t-1}[\exp(\omega_2(u_t))] \exp \omega_1(i_t)$$

$$= \exp \omega_1(i_t) \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det \Psi \mid^{-\frac{1}{2}} \exp(-\frac{1}{2} \widetilde{u}_t' \Psi^{-1} \widetilde{u}_t) \exp \omega_2(\widetilde{u}_t) d\widetilde{u}_t$$

with

$$i_t \overset{not.}{=} (I_{t-1}, s_t^{(i)}, s_t^{(j)}, s_t^{o(j)}, x_t^{g(i)}, x_t^{g(j)}, z_t, \beta_t, Q_{iit}, q_{iit}, Q_{ijt}, q_{ijt}, K_{t,i}, d_{t,i}, y_t^{g(i)})$$

and $\omega_2(\widetilde{u}_t)$ a quadratic function in $\widetilde{u}_t$.

One can write:

$$\widetilde{\mathbf{I}}_t^{(i)} \overset{not.}{=} \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det \Psi \mid^{-\frac{1}{2}} \exp(-\frac{1}{2} \widetilde{u}_t' \Psi^{-1} \widetilde{u}_t) \exp \omega_2(\widetilde{u}_t) d\widetilde{u}_t$$

$$= \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det \Psi \mid^{-\frac{1}{2}} \exp(-\frac{1}{2} \widetilde{u}_t'(\Psi^{-1} + \varphi_t^{(i)} H_{t,i}) \widetilde{u}_t + (\text{linear in } \widetilde{u}_t') + (\text{independent of } \widetilde{u}_t)_1) d\widetilde{u}_t$$

$$= \mid \det(\Psi^{-1} + \varphi_t^{(i)} H_{t,i}) \mid^{-\frac{1}{2}} \mid \det \Psi \mid^{-\frac{1}{2}} \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det(\Psi^{-1} + \varphi_t^{(i)} H_{t,i}) \mid^{\frac{1}{2}} \exp \omega_3(\widetilde{u}_t) d\widetilde{u}_t$$

with $\omega_3(\widetilde{u}_t)$ a quadratic function in $\widetilde{u}_t$. Now, we find $\overline{u}_t \in R^p$ such that:

$$\omega_3(\widetilde{u}_t) = -\frac{1}{2}(\widetilde{u}_t - \overline{u}_t)'(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})(\widetilde{u}_t - \overline{u}_t) + (\text{independent of } \widetilde{u}_t)_2$$

It follows that:

$$(\text{independent of } \widetilde{u}_t)_2 = \frac{1}{2}\overline{u}_t'(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})\overline{u}_t - \frac{\varphi_t^{(i)}}{2} f_{t,i} - \frac{\varphi_t^{(i)}}{2}[\overline{W}_t^{(i)}(s_t^{(i)}) + \overline{W}_t^{(j)}(s_t^{(j)})] \overset{not.}{=} \omega_4\,(\overline{u}_t)$$

and

$$-\varphi_t^{(i)}\widetilde{u}_t'[K_{t,i}(A_t y_{t-1} + C_t^{(i)} s_t^{(i)} + C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - h_{t,i}] = \widetilde{u}_t'(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})\overline{u}_t$$

that is,

$$\overline{u}_t = -\varphi_t^{(i)}(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})^{-1}[K_{t,i}(A_t y_{t-1} + C_t^{(i)} s_t^{(i)} + C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - h_{t,i}]$$

Thus, the integral becomes:

$$\widetilde{\mathbf{I}}_t^{(i)} = \quad |\det(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})|^{-\frac{1}{2}} |\det \Psi|^{-\frac{1}{2}} \exp \omega_4\,(\overline{u}_t) \cdot$$

$$\int_{R^p} (2\pi)^{-\frac{p}{2}} |\det(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})|^{\frac{1}{2}} \exp(-\frac{1}{2}(\widetilde{u}_t - \overline{u}_t)'(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})(\widetilde{u}_t - \overline{u}_t))d\widetilde{u}_t$$

The last integral is equal to 1 because the integrand is the probability density function of a p-dimensional normal random variable:

$$\widetilde{u}_t \sim \mathcal{N}(\overline{u}_t, (\Psi^{-1} + \varphi_t^{(i)} H_{t,i})^{-1})$$

with -1 power denoting inverse.

If we replace $\overline{u}_t$ by its value, we find without difficulty:

$$\widetilde{\mathbf{I}}_t^{(i)} = \quad |\det(I_p + \varphi_t^{(i)} \Psi H_{t,i})|^{-\frac{1}{2}} \exp \omega_4(i_t)$$

where I_p is a $p \times p$ indentity matrix.

By consequence, we have:

$$V_t^{(i)} = \exp \omega_1(i_t) \cdot \widetilde{\mathbf{I}}_t^{(i)} = |\det(I_p + \varphi_t^{(i)} \Psi H_{t,i})|^{-\frac{1}{2}} \exp \omega_5(i_t)$$

with

$$\omega_5 \overset{not.}{=} \omega_1 + \omega_4$$

After several algebraic manipulations, one obtains:

$$\omega_5(I_{t-1}, s_t^{(i)}, s_t^{o(j)}, x_t^{g(i)}, x_t^{g(j)}, z_t, \beta_t, K_{t,i}, d_{t,i}, y_t^{g(i)}) = -\frac{\varphi_t^{(i)}}{2}[y_{t-1}' A_t' \widetilde{H}_{t,i}(C_t^{(j)} s_t^{o(j)} + C_t^{(i)} s_t^{(i)})$$

$$+(s_t^{o(j)'} C_t^{(j)'} + s_t^{(i)'} C_t^{(i)'})\widetilde{H}_{t,i}(A_t y_{t-1} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t)$$

$$+(s_t^{o(j)'} C_t^{(j)'} + s_t^{(i)'} C_t^{(i)'})\widetilde{H}_{t,i}(C_t^{(j)} s_t^{o(j)} + C_t^{(i)} s_t^{(i)})$$

$$+(e_t^{(i)'} E_t^{(i)'} + e_t^{(j)'} E_t^{(j)'} + z_t' B_t' + D_t')\widetilde{H}_{t,i}(C_t^{(j)} s_t^{o(j)} + C_t^{(i)} s_t^{(i)}) + \overline{W}_t^{(i)}(s_t^{(i)})]$$

$$+\varphi_t^{(i)}(s_t^{o(j)'} C_t^{(j)'} + s_t^{(i)'} C_t^{(i)'})[I_p - \varphi_t^{(i)} K_{t,i}(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})^{-1}]h_{t,i} + \text{independent of } s_t^{(i)}$$

where:

$$\widetilde{H}_{t,i} \overset{not.}{=} K_{t,i} - \varphi_t^{(i)} H_{t,i} M_{t,i}^{-1}(\varphi_t^{(i)}) H_{t,i}$$
$$M_{t,i}(\varphi_t^{(i)}) \overset{not.}{=} \Psi^{-1} + \varphi_t^{(i)} H_{t,i}$$

The first order condition in $s_t^{(i)}$ writes:

$$-\frac{\varphi_t^{(i)}}{2} C_t^{(i)\prime} \widetilde{H}_{t,i} A_t y_{t-1} - \frac{\varphi_t^{(i)}}{2} C_t^{(i)\prime} \widetilde{H}_{t,i}(A_t y_{t-1} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - \varphi_t^{(i)} C_t^{(i)\prime} \widetilde{H}_{t,i} C_t^{(i)} s_t^{(i)}$$

$$-\frac{\varphi_t^{(i)}}{2} C_t^{(i)\prime} \widetilde{H}_{t,i}' C_t^{(j)} s_t^{o(j)} - \frac{\varphi_t^{(i)}}{2} C_t^{(i)\prime} \widetilde{H}_{t,i} C_t^{(j)} s_t^{o(j)} - \frac{\varphi_t^{(i)}}{2} C_t^{(i)\prime} \widetilde{H}_{t,i}(E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t)$$

$$-\varphi_t^{(i)} Q_{iit}(s_t^{(i)} - x_t^{g(i)}) - \varphi_t^{(i)} q_{iit} + \varphi_t^{(i)} C_t^{(i)\prime}[I_p - \varphi_t^{(i)} K_{t,i}(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})^{-1}] h_{t,i} = 0$$

$$\overset{\left(\varphi_t^{(i)} \neq 0\right)}{\Leftrightarrow} -C_t^{(i)\prime} \widetilde{H}_{t,i} A_t y_{t-1} - C_t^{(i)\prime} \widetilde{H}_{t,i}(E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - C_t^{(i)\prime} \widetilde{H}_{t,i} C_t^{(j)} s_t^{o(j)}$$

$$+Q_{iit} x_t^{g(i)} - q_{iit} + C_t^{(i)\prime}[I_p - \varphi_t^{(i)} K_{t,i}(\Psi^{-1} + \varphi_t^{(i)} H_{t,i})^{-1}] h_{t,i} = (C_t^{(i)\prime} \widetilde{H}_{t,i} C_t^{(i)} + Q_{iit}) s_t^{(i)}$$

It follows that:

$$s_t^{o(i)} = G_t^{(i)} y_{t-1} + g_t^{(i)}$$

with

$$G_t^{(i)} = -N_{t,i}^{-1}(\varphi_t^{(i)})(C_t^{(i)\prime} \widetilde{H}_{t,i} A_t)$$

and

$$g_t^{(i)} = -N_{t,i}^{-1}(\varphi_t^{(i)})\{(Q_{iit} x_t^{g(i)} - q_{iit}) + C_t^{(i)\prime}[\widetilde{H}_{t,i}(C_t^{(j)} s_t^{o(j)} + E_t^{(i)} e_t^{(i)} + E_t^{(j)} e_t^{(j)} + B_t z_t + D_t) - R_{t,i} h_{t,i}]\}$$

It is important to note that the strategies adopted depend on the players' types or characteristics. Accurate estimates are necessary for an efficient implementation of the players' strategies. These represent a basic information for the optimal decision making process. Policy changes (or other exogenous changes) will cause changes in the players' decision rules, and thus in their behavioral relationships. This reveals the role played by the information structure in the measurements of the Nash equilibrium. In the case where the policy instruments employed by the players require information on all states of the system, the policy rule becomes complicated from the viewpoint of implementation and computing cost.

As opposed to the Stackelberg equilibrium, the Nash equilibrium is not necessarily defined uniquely (NASH 1950, 1951). Several equilibria may co-exist. In this case, the players are expected to choose the equilibrium which leads to the lowest output level ensuring the system stability. Dynamical systems with several equilibria occur in various fields of economics. It may be the case of a non-cooperative Nash game. The outcome of the game may thus be difficult to predict. It may also exist situations when computing a Nash equilibrium is an intractable problem. A possibility to overcome this limitation is to compute approximate Nash equilibria using randomized algorithms. It is useful to note that Nash equilibrium and Stackelberg equilibrium may give very different outcomes.

Different contexts of decision making generally call for different strategies. When the number of players in the game is large, the equilibrium is generally of non-cooperative type, whereas when the number of players is small, the outcome may be cooperative.

Because the source of randomness may differ from an application to another, the players' strategies may vary. It is of great interest to know whether small changes in the problem

statement will cause significative changes in the solution equilibrium. It comes to analyze the sensitivity of the results to the choice of policy instruments and rules employed by the players.

The present approach improves the traditional solution concept of Nash strategy with closed-loop information structure for discrete-time linear-quadratic games, by considering the case when the players exhibit endogenous risk-and effort (PAPAVASSILOPOULOS 1981; CARAVANI AND PAPAVASSILOPOULOS 1990; KRAJEWSKI 1991; BASAR 1999; KLOMPSTRA 2000; KRAWCZYK 2000; KRAWCZYK AND TIDBALL 2006; JIMENEZ-LIZARRAGA ET AL. 2009).

7. Concluding Remarks

The present approach extends the solution concept of Nash equilibrium to strategic dynamic games with asymmetric players characterized by endogenous preferences over risk and effort. A complete analytical description of discrete-time closed-loop Nash strategies sensitive to endogenous risk and effort is provided. The borderline case of excessive risk /effort-averse players is considered in the analysis. It reveals important aspects about the close relationship that exists between excessive risk /effort-averse behavior and deviating equilibrium strategy. A cooperative game is generally correlated with an optimal risk-and effort-sharing between the players.

The analysis can be easily extended to a dynamic coalition game model with cooperative and non-cooperative elements. It is the case where a group of players decide to act together, as one unit, relative to the rest of the players, each group being characterized by a joint strategy. One can suppose that players organize themselves in coalitions which form a partition, that is, each player belongs to one and only one coalition. When players form conjectures, they expect to obtain better results than in the individual case. Forming a coalition, this does not eliminate the individual players as decision-makers. In all interactions with the other players, each coalition is represented by an exponent of its members. This arrangement will continue only as long as each player finds it desirable to act in this way. All the coalitions are then examinated by each player, the rationality of the game being to choose or not one particular coalition, given the obtained results. Each coalition will play a game of common interests, that is, one with a strongly Pareto dominant payoff vector. However, each player may control his level of cooperation. Therefore, it is possible to exist another contract between the players within each coalition. Most of the existing models in game theory assume that the coalition structure is given exogenously. More realistic, it can be given as an endogenous outcome, being possible to predict which coalitions will form in each given situation. The players' effort behavior profiles will determine the structure of strategic coalitions. New information in the game affects the players' effort preferences in the coalition formation process.

References

[1] ANDERSON, S. P., GOERRE, J. K. AND HOLT, C. A., 1996. "Minimum-Effort Coordination Games: An Equilibrium Analysis of Bounded Rationality", Manuscript, University of Virginia.

[2] ANDERSON, S. P., GOERRE, J. K. AND HOLT, C. A., 2001. "Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium", *Games and Economic Behavior*, **34**, 177-199.

[3] AUMANN, R. 1990. "Nash Equilibria Are Not Self-Enforcing", in Economic Decision Making: Games, Econometrics and Optimisation, edited by J.J. Gabszewicz, J.-F. Richard and L.A. Wolsey, Elsevier.

[4] AUMANN, R. AND A. BRANDENBURGER, 1995. "Epistemic Conditions for Nash Equilibrium", *Econometrica*, **63** (5), 1161-1180.

[5] BASAR, T., 1999. "Nash Equilibria of Risk-Sensitive Nonlinear Stochastic Differential Games", *Journal of Optimization Theory and Applications*, **3**, 479-498.

[6] BERNHEIM, B. D., 1984. "Rationalizable Strategic Behavior", *Econometrica*, **52**, 1007-1028.

[7] CARAVANI, P. AND G. PAPAVASSILOPOULOS, 1990. "A class of risk-sensitive non-cooperative games", *Journal of Economic Dynamics and Control*, **14**, 117-149.

[8] KLOMPSTRA, M. B., 2000. "Nash equilibria in risk-sensitive dynamic games", *IEEE Transactions on Automatic Control*, **45**, 1397-1401.

[9] KRAKEWSKI, W., 1991. "A note on the dynamic risk sensitive Nash games", Springer Berlin/Heidelberg, Vol. **157**, 260-268.

[10] KRAWCZYK, J. B., 2000. "An Open-Loop Nash Equilibrium in an Environmental Game with Coupled Constraints", *Symposium of the International Society of Dynamic Games Adelaide*, South Australia Symposium Proceedings, pp. 325-339.

[11] KRAWCZYK, J. B. AND M. TIDBALL, 2006. "A Discrete-Time Dynamic Game of Seasonal Water Allocation", *Journal of Optimization Theory and Applications*, **2**, 2006.

[12] NASH, J., 1950. "Equilibrium points in n-person games", *Proceedings of the National Academy of Sciences*, **36**, 48-49.

[13] NASH, J., 1951. "Non-cooperative games", *Annals of Mathematics*, **54**, 286-295.

[14] PAPAVASSILOPOULOS, G. P.,1981. "Solution of some stochastic quadratic Nash and leader-follower games", *SIAM Journal on Control and Optimization*, **19**, 651-666.

[15] PEARCE, D. G., 1984. "Rationalizable Strategic Behavior and the Problem of Perfection", *Econometrica*, **47**, 1029-1050.

[16] PROTOPOPESCU, D., 2007. "Improving the Risk Concept: A Revision of Arrow-Pratt Theory in the Context of Controlled Dynamic Stochastic Environments", *Working Paper* **No. 727.08**, CODE, Universitat Autònoma de Barcelona, Spain.

[17] PROTOPOPESCU, D., 2008. "Improving the Effort Concept: A Revision of the Traditional Approach in the Context of Controlled Dynamic Stochastic Environments", *Working Paper* No. **739.08**, CODE, Universitat Autònoma de Barcelona, Spain.

[18] PINDYCK, R. S., 1976. "The Cost of Conflicting Objectives in Policy Formulation", *A. E. S. M.*, 239-247.

[19] SANDBLOM, C. L. AND J. L. BANASIK, 1985. "Economic policy with bounded controls", *Economic Modelling,* **2**, 135-148.

[20] JIMENEZ-LIZARRAGA, M., BASIN, M., ALCORTA-GARCIA, A., 2009. "Equilibrium in Linear Quadratic Stochastic Games with Unknown Parameters", *ICIC Express Letters*, **2**, 107-113.