Comportement réciproque et information incomplète dans un jeu du dilemme du prisonnier

Résumé

Dans ce papier, nous introduisons la notion de déformation des paiements dans un jeu du dilemme du prisonnier. Cette notion nous permet de prendre en compte l'appréciation de chaque joueur du comportement de son adversaire. Nous déterminons les conditions sous lesquelles la coopération mutuelle constitue un équilibre. Premièrement, nos résultats montrent que contrairement à l'altruisme, la réciprocité permet de faire face au comportement opportuniste de l'égoïsme. Deuxièment, la coopération émerge avec la réciprocité, et elle peut être un équilibre de Nash dans un domaine d'information complète ou un équilibre bayésien sous certaines conditions dans un environnement d'information incomplète.

Mots-clefs : *dilemme du prisonnier, déformation des paiements, coopération.*

JEL Classification : *C7, A13.*

Ahmed Doghmi

Strategic Interaction Group, Max Planck Institute of Economics, Germany (ahmeddoghmi@hotmail.com)

Miloudi Kobiyh

Center for Research in Economics and Management, University of Caen, France (miloudi.kobiyh@gmail.com)

1. Introduction

L'objectif de ce papier est de déterminer dans quels cas la coopération peut devenir un équilibre du jeu du dilemme du prisonnier lorsque chaque joueur intègre dans son choix le comportement de l'autre joueur ainsi que son propre comportement. La notion de "comportement" renvoie ici à la dimension psychologique des choix.

Dans l'approche non coopérative de la théorie des jeux, l'expérimentation a révélé que, souvent, les joueurs n'atteignent pas la solution préconisée par la théorie (voir par exemple Dawes and Thaler 1988). Dans le cadre particulier du jeu du dilemme du prisonnier, deux explications ont été avancées pour justifier la coopération comme solution du jeu. La théorie des jeux répétés est la première d'entre elles (Kreps *and al.* (1982)). Dans ces modèles, l'hypothèse que chaque agent croit même très faiblement en la possibilité que les autres agents jouent la coopération est capitale. Cette hypothèse d'information incomplète associée à la répétition des coups entraîne en effet des incitations à coopérer. Dans un tel cadre, les acteurs

sont supposés rester égoïstes. La deuxième explication stipule que les agents ne sont pas tous égoïstes et qu'au moins une partie d'entre eux s'intéresse au bien-être des autres joueurs dans le cadre d'une interaction sociale. A partir de cette idée, deux approches différentes ont été envisagées.

La première stipule qu'une partie des agents a des "préférences sociales" ou, plus précisément, une fonction d'utilité qui dépend à la fois de leur propre utilité matérielle mais aussi des gains obtenus par les autres joueurs. Cette approche, comme le font remarquer Falk and Fischbacher (2002), ne prend pas en compte seulement les objectifs des joueurs mais aussi leurs intentions. Une deuxième approche dite de la *réciprocité* basée sur les intentions a été proposée. Dans cette approche, la décision d'un joueur est influencée par les intentions des autres joueurs. Si ce joueur estime que les autres joueurs ne sont pas coopératifs, il ne fera aucun effort non plus. Cette idée est à la base du concept d'équilibre de "fairness" développé par Rabin (1993). Il s'agit de la première contribution qui modélise la notion de réciprocité et examine les conséquences de comportements réciproques. De nombreux articles ont ainsi mis à jour le comportement réciproque des agents (voir Fehr et Schmidt (2003), Fehr et Gächter (2000) pour un *survey* sur le sujet), les comportements altruistes et égoïstes étant parfois considérés comme des cas limites de comportements réciproques. La réciprocité apparaît alors comme étant une propriété psychologique des individus qui va indubitablement altérer les caractéristiques de leur bien-être.

L'émergence de la coopération dans les jeux expérimentaux renvoie irrévocablement à la question de la psychologie des joueurs et à son influence sur la formation des utilités des paiements reçus : un joueur n'apprécie pas forcément de la même façon que son adversaire un gain obtenu. Le choix d'une action dépend de la déformation du paiement reçu que chaque joueur élabore pour former son utilité. La déformation des paiements n'est pas forcément une connaissance commune, elle dépend de la psychologie de chaque joueur. La distinction entre un jeu à information complète et un jeu à information incomplète s'avère donc essentielle. Dans le deuxième type de jeu, les déformations des paiements des autres joueurs sont inconnues ou mal connues pour au moins un joueur.

Dans un jeu du dilemme du prisonnier standard, les joueurs basent leurs choix uniquement sur le gain qu'ils retireront de ces choix. Dans notre article, le joueur détermine son choix sur la base des trois éléments suivants : états du monde, actions et conséquences pour lesquelles il procède à une évaluation personnelle de la probabilité de réalisation de chaque état du monde. Dans un contexte incertain, il est logique de penser qu'un agent en situation d'interaction explique les choix des autres agents par ses propres croyances. Dans une interaction sociale, un agent exprime ses préférences sur un ensemble d'alternatives par le choix d'une action qui peut résulter, au moins en partie, d'éléments qui relèvent de la croyance concernant ses motivations et celles d'autrui.

Ainsi, les croyances des agents vont être modifiées par le comportement psychologique de chaque joueur. Nous proposons une modélisation alternative de la réciprocité proposée par Rabin, tout en maintenant l'idée que la réciprocité est basée sur les intentions des joueurs. Alors que les équilibres des jeux du dilemme du prisonnier dans un cadre standard sont basés sur une évaluation de l'utilité retirée par chaque joueur, nous allons introduire dans cet article la notion de "jeu de comportement" qui se différencie des jeux de "gains stricts" où les agents, supposés rationnels, ont pour objectif de maximiser leur utilité matérielle. Ainsi le choix de chaque joueur est-il assimilé à son choix en termes de comportement, la probabilité de coopérer ou non étant totalement liée au comportement adopté. Le comportement d'un agent est donc assimilé à son degré de "réciprocité" vis-à-vis de l'autre agent. Dans ce cadre, nous montrons que la coopération peut être le résultat du jeu sous certaines conditions relatives au comportement plus ou moins réciproque des joueurs. Si ce résultat ne paraît pas surprenant en information complète, nous creusons notre analyse en montrant la relation entre le degré de coopération et les croyances sur la coopération des joueurs dans le cas d'une information incomplète.

La deuxième section est consacrée à l'étude de la déformation des paiements selon la croyance sur le choix de l'autre joueur dans un jeu du dilemme du prisonnier. La troisième section présente la notion de réciprocité et de ses conséquences dans les cas d'information complète et d'information incomplète.

2. Croyances des joueurs et déformation des paiements

2.1. Déformation des paiements dans le jeu du dilemme du prisonnier

Le jeu du dilemme du prisonnier constitue le principal modèle utilisé pour étudier le problème de la coopération dans une population d'égoïstes (Axelrod et Hamilton, 1981). C'est l'exemple le plus expressif nous permettant de démontrer l'influence de la structure sociale et de la psychologie sur la nature des interactions et, par conséquent, sur les résultats du jeu au moment où les éléments sociaux et psychologiques peuvent être appréhendés par la nature des préférences des joueurs. Pour introduire la dimension psychologique et montrer la relation existant entre le comportement des joueurs et le résultat du jeu – qui peut différer de celui de la théorie standard – on va assimiler dans un premier temps le dilemme du prisonnier à une loterie incertaine, puis en déduire les paiements déformés qui en résultent.

Soit le dilemme du prisonnier suivant :

1/2	C^2	D^2
C^1	a,a	b,c
D^1	a,a	b,c

avec c > a > d > b et 2a > c+b.

Dans ce jeu standard du dilemme du prisonnier, nous introduisons les états du monde qui vont être considérés par les joueurs. Ainsi, la coopération (C^i) et la défection (D^i) d'un joueur représentent les états du monde pour l'autre joueur. Nous considérons donc deux joueurs 1 et 2, dont nous étudions les stratégies pures. Les états du monde auquel fait face le joueur i sont : C^i et D^i avec $i = 1, 2$; $j = 1, 2$ $j \neq i$. Nous pouvons décrire le dilemme du prisonnier comme une loterie incertaine avec l'espérance de gains suivante :

$$\mathbf{E}(L) = a.p(C^i \cap C^j) + b.p(C^i \cap D^j) + c.p(D^i \cap C^j) + d.p(D^i \cap D^j).$$

La somme des probabilités est égale à1, puisque chaque joueur choisit soit la coopération soit la défection. Il existe $E \in \{C^i, D^i\}$ et $F \in \{C^j, D^j\}$ tels que $p(E \cap F) = 1$, les autres probabilités étant nulles. En utilisant les probabilités de Bayes $p(E \cap F) = p(E/F).p(F)$, nous obtenons un paiement espéré :

$$\mathbf{E}(L) = p(C^j)[a.p(C^i/C^j) + c.p(D^i/C^j)] + p(D^j)[b.p(C^i/D^j) + d.p(D^i/D^j)].$$

Le joueur évalue seulement la probabilité de réalisation de l'état du monde $p(C^j)$ (puisque $p(D^j) = 1 - p(C^j)$) et choisit les probabilités conditionnelles qu'il attache aux conséquences liées aux différents états du monde pour maximiser **E**(L). D'après cette formule de paiement espéré, les paiements déformés qui en résultent pour le joueur i sont décrits dans la matrice suivante :

Tableau 1

Paiements déformés

Etats du monde pour joueur $i \Rightarrow$ paiement joueur i $\Downarrow$	$p(C^j)$	$p(D^j)$
i coopère	$p(C^i/C^j).a$	$p(C^i/D^j).b$
i fait défection	$p(D^i/C^j).c$	$p(D^i/D^j).d$

avec $(C^i/C^j) = 1 - p(D^i/C^j)$ et $p(C^i/D^j) = 1 - p(D^i/D^j)$.

Cette matrice des paiements déformés intègre la subjectivité du joueur puisque nous supposons ici que la psychologie des joueurs intervient dans l'évaluation des probabilités conditionnelles. Cette évaluation se base sur la croyance probabiliste $p(C^j)$ qui représente l'information dont le joueur i dispose sur la coopération du joueur j. Pour Savage (1972) « la probabilité mesure la confiance qu'un individu particulier a dans la véracité d'une position particulière ». Dans notre cadre, cela signifie que cette croyance mesure la confiance que le joueur a dans l'effectivité de la coopération du joueur j. On suppose que la croyance $p(C^j)$ est cohérente, ce qui revient à dire que la distribution de probabilité $(p(C^j), (1 - p(C^j)))$ sur laquelle sont fondées ces croyances correspond à la vraie distribution.

A la différence des stratégies déterministes, les stratégies mixtes ont plusieurs interprétations possibles. Leur utilité dans ce problème tient au caractère aléatoire induit par les croyances des joueurs sur les stratégies et le comportement des autres joueurs. Même si un joueur décide de faire un choix particulier, il ne souhaite pas, dans la plupart des cas, que son adversaire connaisse ce choix. De plus, chaque joueur ne connaît pas exactement le comportement de l'autre joueur et ne fait que formuler des croyances sur ce comportement. Dans ce cas, même si le premier joueur décide de jouer la stratégie qu'il a choisi, le second a l'impression qu'il joue aléatoirement, et il représente son comportement par une stratégie mixte.

2.2. Comportements et paiements déformés

Traditionnellement, dans le jeu du dilemme du prisonnier, un joueur n'a que deux options possibles : coopérer ou faire défection. Axelrod (1984) a montré que l'émergence de la coopération est étroitement liée à l'existence des comportements réciproques. L'une de ses suggestions est qu'il faut pratiquer la réciprocité dans la coopération comme dans la défection.

Des articles récents ont démontré la possibilité d'élargir ce jeu classique en permettant un degré variable de coopération (Sherratt et Robert, 1998 ; Wahl et Nowak, 1999).

Considérons donc les trois principaux comportements : **altruisme, réciprocité et égoïsme**.

2.2.1. Altruisme (Al)

En étant altruiste, le joueur choisit de coopérer quel que soit le choix de l'autre joueur. Sa probabilité de coopérer est :

$$p(C^i/C^j) = 1 \; \forall p(C^j) \Rightarrow p(D^i/C^j) = 0$$

$$p(C^i/D^j) = 1 \; \forall p(D^j) \Rightarrow p(D^i/D^j) = 0$$

Dans un jeu où les deux joueurs sont altruistes, il existe donc un seul équilibre de Nash(C,C) puisque la coopération est le seul profil de stratégie joué par les deux joueurs. Le paiement reçu par les joueurs est alors (a,a).

2.2.2. Réciprocité (R)

La réciprocité, dans sa définition la plus classique et la plus stricte, consiste pour un joueur à choisir une stratégie identique à celle choisie par son adversaire. Dans ce cas, le joueur i coopère avec une probabilité identique à la probabilité de coopération de l'autre joueur. Donc $p(C^i) = p(C^j) = p$ et $p(D^i) = p(D^j) = (1-p)$. Les paiements déformés de deux joueurs réciproques sont donnés par la matrice suivante :

$$\begin{pmatrix} p.a, p.a & p.b, (1-p)c \\ (1-p)c, p.b & (1-p).d, (1-p).d \end{pmatrix}$$

Le paiement espéré de chaque joueur est :

$$\begin{aligned} r(p) &= p[p.a + (1-p)c] + (1-p)[p.b + (1-p).d] \\ &= p^2.a + p(1-p)(c+b) + (1-p)^2.d \end{aligned} \quad (1)$$

2.2.3. Egoïsme (Eg)

En supposant le comportement égoïste, le joueur préfère faire défection inconditionnellement.

$$p(C^i/C^j) = 0 \; \forall p(C^j) \Rightarrow p(D^i/C^j) = 1$$
$$p(C^i/D^j) = 0 \; \forall p(D^j) \Rightarrow p(D^i/D^j) = 1$$

Dans un jeu où chaque joueur est égoïste, choisir de faire défection est le seul profil de stratégie possible (*D,D*) et est l'équilibre de Nash du jeu. Le joueur égoïste obtient le paiement d en jouant contre un joueur égoïste ou réciproque et le paiement c contre un altruiste.

2.2.4. Analyse des comportements

Dans ces trois cas on a pu déterminer, selon les comportements des joueurs, la matrice des paiements déformés. Selon l'information sur les probabilités de coopération de son concurrent, le joueur i réagit avec une probabilité $p(C^i/C^j)$. Cette information peut-être tirée du passé du joueur i en se basant sur d'autres interactions. Néanmoins, le joueur i peut la formuler objectivement en estimant l'attitude de son adversaire d'après des expériences semblables. Sa réaction reflète sa psychologie par rapport à cette information.

En considérant une information complète, la matrice des paiements déformés est une connaissance commune des deux joueurs. Dans le cas contraire, l'information est quasi complète et le choix difficile.

Dans un premier temps, nous restons dans le cadre d'une information complète où chaque joueur connaît la structure du jeu.

En tenant compte des trois comportements qui viennent d'être présentés et des paiements déformés présentés dans le tableau 1, l'interaction entre les deux joueurs peut être décrite par la matrice 3 x 3 suivante :

Tableau 2

1/2	Al	R	Eg
Al	a, a	a, a	b, c
R	a, a	$r(p), r(p)$	d, d
Eg	c, b	d, d	d, d

Un joueur égoïste obtient donc un paiement d si l'autre joueur est égoïste ou réciproque, et un paiement c si l'autre joueur est altruiste. En revanche, un joueur altruiste reçoit a si l'autre joueur est altruiste ou réciproque et b si l'autre joueur est égoïste.

Nous sommes alors en mesure de classer les gains issus des jeux en fonction des différents comportements étudiés.

Proposition 1 *Le classement des paiements, pour $0 < p < 1$, peut être donné de la façon suivante:*
Si $2.d \leq (b+c)$ alors, $a > r(p) > d$
Si $2.d > (b+c)$ alors, $a > r(p) > d$ si $p \in]p^*, 1[$ avec $p^* = \frac{2d-c-b}{a+d-c-b}$. Sinon $a > d \geq r(p)$.

Preuve : voir annexe 1

D'après ce classement, il s'avère que le résultat d'un jeu entre deux joueurs non égoïstes, dont au moins un est altruiste, est meilleur que celui obtenu par deux joueurs qui ont tous les deux un comportement réciproque.

De plus, la réciprocité ne représente un intérêt pour un joueur que sous certaines conditions sur la croyance (p) qui représente la probabilité de la coopération de son adversaire et les paiements (a, b, c, d). En effet, si $2.d \leq (b+c)$, le joueur a intérêt à adopter un comportement réciproque. En revanche, dans le cas où $2.d > (b+c)$, ce comportement ne représente un intérêt pour le joueur que si la croyance p sur la probabilité de coopérer de son adversaire vérifie $p > \frac{2d-c-b}{a+d-c-b}$.

Ce résultat signifie que le comportement réciproque représente un intérêt pour les joueurs si les gains des joueurs issus de la défection commune sont relativement faibles comparés aux gains issus de la défection d'un joueur et de la coopération de l'autre. Dans le cas contraire, il faut que la croyance du joueur sur la coopération de l'autre joueur soit relativement élevée pour que ce type de comportement soit profitable.

Ce résultat est un premier élément d'explication à l'émergence de la coopération au sein d'une population où existent les trois comportements (*Al, R, Eg*). Pour illustrer d'avantage cette conclusion, il convient de considérer le cas où le joueur est impliqué dans une interaction sociale représentée par une population caractérisée par l'existence de ces trois comportements *Al, R* et *Eg*.

2.3. Réciprocité et interaction sociale

Lorsqu'un individu évolue dans un cadre social, quel qu'il soit, il est en interaction avec tous les autres individus un à un. Son paiement résulte donc de toutes ces interactions ainsi que de la répartition des comportements dans la population.

L'objectif de cette section est de montrer les conditions sous lesquelles le comportement réciproque domine les autres comportements.

Soit q_1, la proportion d'altruistes ; q_2, la proportion des individus ayant un comportement réciproque et $1 - q_1 - q_2$, la proportion d'égoïstes. On appelle l'état de la population le vecteur $q = (q_1, q_2, 1 - q_1 - q_2)$.

Nous désignons par le symbole $\mathbf{E}(J, q)$ l'espérance mathématique de la variable J en considérant l'état de la population q. Dans le cas qui nous intéresse, la variable J caractérise le comportement du joueur. D'après le tableau 2 on obtient :

$$\mathbf{E}(Al, q) = q_1(a-b) + q_2(a-b) + b$$

$$\mathbf{E}(R, q) = q_1(a-d) + q_2(r(p)-d) + d$$

$$\mathbf{E}(Eg, q) = q_1(c-d) + d$$

Proposition 2 *Supposons que $p \neq 1$, il existe $\hat{p}$ et $\bar{p}$ avec $0 < \hat{p} < \bar{p} < 1$ tels que:*
i) Le comportement réciproque domine l'altruisme si et seulement si $p > \hat{p}$.
ii) Le comportement réciproque domine l'égoïsme si et seulement si $\frac{q_1}{q_2} < \frac{(a-d)}{(c-a)}$ et $p > \bar{p}$.

Preuve : voir annexe 2

Proposition 3 *Si l'égoïsme et l'altruisme sont les seuls comportements présents dans la population, alors la coopération ne peut émerger.*

Preuve : voir annexe 3

La façon dont sont répartis les différents comportements au sein de la population a un rôle primordial dans l'émergence de la coopération. Ceci a deux implications : premièrement, l'explication de l'émergence de la coopération au sein d'une population sans l'intervention d'une autorité centrale. Deuxièmement, l'importance de la réciprocité dans l'instauration de la coopération puisqu'il s'agit d'un comportement qui consiste à coopérer avec ceux qui coopèrent et punir ceux qui dévient de la coopération.

3. Modélisation de la réciprocité

Le rôle crucial du comportement réciproque des joueurs dans l'apparition de la coopération nécessite une attention particulière. C'est pour cette raison que, dans cette section, notre analyse va se concentrer sur le comportement réciproque d'un agent, dans une définition plus large que celle qui vient d'être présentée et dans les cas d'information complète et d'information incomplète.

Définissant le comportement comme étant ce qui conduit un joueur i à coopérer avec une probabilité qui est fonction de celle de l'autre joueur, il s'agit donc d'une application $\alpha_i : [0,1] \mapsto [0,1]$ telle que $\alpha_i(p(C^j)) = p(C^i)$. Le comportement réciproque, ou réciprocité, est défini par l'inégalité suivante :

$$p(C^i) = \alpha_i(p(C^j)) \geq p(C^j)$$

Un joueur, ayant un comportement α_i, choisit une probabilité de coopération $p(C^i)$ qui est en accord avec son comportement. Si son choix ne respecte pas son comportement, son utilité baisse. Dans cette écriture, $p(C^j)$ formalise les intentions de l'autre joueur, alors que $p(C^i)$ représente les intentions du joueur i assimilables à son comportement.

Dans l'analyse qui suit, nous allons développer un jeu dit de "comportement" dans lequel l'équilibre du jeu sera uniquement déterminé par le comportement des joueurs.

3.1. Information complète

L'information complète suppose que chaque joueur i est informé de son comportement α_i, de celui de l'autre joueur α_j, de l'ensemble des stratégies [0, 1] et des fonctions des paiements $u_i(s_i, s_j)$ et $u_j(s_j, s_i)$.

Avec $s_i = \alpha_i(p(C^j)) = p(C^i)$, $s_j = \alpha_j(p(C^i)) = p(C^j)$, où i et j sont les stratégies des joueurs i et j.

3.1.1. Comportements homogènes $\alpha_i = \alpha_j$

– Cas 1 : $p(C^i) = \alpha_i(p(C^j)) = p(C^j)$

Pour chaque joueur i, la fonction de paiement atteint son maximum si sa meilleure réponse $p(C^i)$ est égale à $p(C^j)$. On peut démontrer que chaque point $(p(C^i), p(C^j))$ est un équilibre de Nash. Dans ce cas, aucun joueur n'a intérêt à dévier du comportement de son adversaire.

– Cas 2 : $p(C^i) = \alpha_i(p(C^j)) < p(C^j)$

Ce type de comportement implique que chaque joueur a une volonté de coopérer inférieure à celle exprimée par l'autre joueur. Ainsi, si $p(C^j) \neq 0$, on a : $\alpha_i(p(C^j)) = p(C^i) < p(C^j)$ et $\alpha_j(p(C^i)) = p(C^j) < p(C^i)$. Ceci donne $p(C^i) < p(C^j)$ et $p(C^j) < p(C^i)$. Le seul point d'équilibre est donc la défection mutuelle.

– Cas 3 : $p(C^i) = \alpha_i(p(C^j)) > p(C^j)$

Cette situation est l'inverse de la précédente. Chacun a une volonté de coopérer supérieure à l'autre. $\alpha_i(p(C^j)) = p(C^i) > p(C^j)$ et $\alpha_j(p(C^i)) = p(C^j) > p(C^i)$. par conséquent le seul point d'équilibre est la coopération mutuelle.

3.1.2. Des comportements hétérogènes $\alpha_i \neq \alpha_j$

Dans le cas de comportements hétérogènes, on a :

$$p(C^i) = \alpha_i(p(C^j)) \geq p(C^j) \text{ et } \alpha_j(p(C^i)) = p(C^j) < p(C^i)$$

Le paiement du joueur j diminue si $p(C^j) \geq p(C^i)$, ce qui implique qu'il n'a pas intérêt à accorder à la coopération une probabilité qui soit au moins égale à celle du joueur i. Le seul point d'équilibre est donc la défection mutuelle.

3.1.3. Des comportements changeants

Ces comportements sont compliqués par rapport aux précédents. On se limite donc au cas où un joueur change de comportement pendant que l'autre coopère avec la même probabilité de coopération que son adversaire.

On a donc :

$$\begin{aligned} p(C_i) &= \alpha_i(p(C^j)) = p(C^j) \\ \text{et } \alpha_j(p(C^i)) &< p(C^i) \text{ si } p(C^i) < p(C^i)^* \\ \text{et } \alpha_j(p(C^i)) &> p(C^i) \text{ si } p(C^i) > p(C^i)^* \end{aligned}$$

où $p(C^i)^*$ est une probabilité de coopération seuil.

Pour des probabilités $p(C^i) < p(C^i)^*$, le seul équilibre de Nash est $(p(C^i), p(C^j)) = (0,0)$. Par contre, pour des probabilités $p(C^i) > p(C^i)^*$, on a $\alpha_j(p(C^i)) > p(C^i)$ et $\alpha_i(p(C^j)) = p(C^j)$. Alors $(p(C^i), p(C^j)) = (1,1)$ est le seul équilibre de Nash.

D'après cette discussion sur les différentes possibilités des comportements des deux joueurs, on peut énoncer le résultat qui caractérise l'équilibre de la coopération mutuelle.

Proposition 4 *Pour que $(p(C^i), p(C^j)) = (1,1)$ soit un équilibre de Nash il faut qu'il existe, au moins, un voisinage de $(1,1)$ où les deux joueurs sont réciproques.*

Ce qui détermine la coopération des deux joueurs est leurs comportements au voisinage de (1,1). Si un joueur est tenté de faire défection et s'il constate que la probabilité de coopération de l'autre joueur est proche de 1, (1,1) ne sera pas un équilibre de Nash.

3.2. Information incomplète

Dans le dilemme du prisonnier, un joueur i peut avoir un doute sur la réciprocité de son partenaire j, donc une incertitude sur α_j. Ceci dit, pour une probabilité donnée p, le joueur i connaît $\alpha_i(p)$ puisqu'il sait son propre type α_i. En revanche, lorsqu'il prend p comme probabilité de coopération, il ne connaît pas $\alpha_j(p)$. Il s'agit donc d'un jeu en information incomplète, ce qui nous remène a un jeu dit Bayésien. L'approche bayésienne (Harsanyi) consiste à attacher une distribution de probabilité à tout élément incertain. Après réception d'une nouvelle information, un agent révise ses croyances à l'aide de la règle de Bayes pour les probabilités conditionnelles. Dans ce qui va suivre, nous considérons un jeu du dilemme du prisonnier, dans lequel aucun des deux joueurs ne connaît avec exactitude le comportement de l'autre. Nous allons également distinguer deux types de réciprocité, ce qui nous amène à considérer que chaque joueur peut être de trois types : réciproque stricte (RS) $(\alpha(p) = p, \forall p)$, réciproque large (RL) $(\alpha(p) > p, \forall p)$ ou non-réciproque (NR) $(\alpha(p) < p, \forall p)$.

Les croyances du joueur i sur le comportement du joueur j sont définies de la façon suivante : i pense que j va être (RS) avec probabilité p_1, (RL) avec probabilité p_2 et (NR) avec probabilité p_3. De même, le joueur j a des croyances sur le comportement du joueur i, avec des probabilités et q_1, q_2 et q_3. Nous pouvons écrire ce jeu sous une forme extensive présentée dans la figure 1.

Dans cette représentation, les nœuds terminaux représentent les paiements, par exemple (*RS*, *RS*) donne (p, p)(les joueurs jouent la même probabilité de coopération, le paiement dans ce cas n'est autre que les paiements espérés $(r(p), r\ (p))$ donnés dans la sous-section 2.1. Après (*RL*, *RS*) on a le profil (1, 1), les joueurs coopèrent avec une probabilité égale à 1, soit (a,a) comme paiement. Après (*NR*, *RS*) on a le profil (0,0), aucun joueur ne coopère, ce qui donne le paiement (d, d), etc. On obtient le résultat suivant :

Proposition 5 *Si chacun des joueurs croit que l'autre joueur peut être de comportement réciproque large, i.e. si $p_1 > 0$ et $q_1 > 0$, alors le seul équilibre de Nash bayésien est le couple (RL, RL).*

Preuve : voir annexe 5

4. Conclusion

La déformation des paiements mise en œuvre dans ce travail s'avère constituer une explication à l'apparition de la coopération dans un jeu du dilemme du prisonnier. Elle reflète *l'optimalisme* du joueur et sa tendance à augmenter son paiement sans diminuer celui des autres. Cet article a mis en évidence le rôle de la réciprocité dans l'émergence de la coopération. Cette dernière est étroitement liée à la façon dont sont répartis les différents comportements au sein de la population.

Notre analyse se focalise sur les incitations qui mènent les agents à coopérer. L'action de chaque joueur dépend du comportement adopté par l'autre joueur. Si les joueurs agissent selon leurs comportements rationnels, leurs actions constituent un équilibre.

La modélisation présentée dans ce papier est basée sur les intentions des joueurs et démontrent le rôle de la réciprocité dans l'émergence de la coopération. L'utilité de la déformation des paiements proposée réside dans l'explication des motivations qui dictent le choix des agents. Notre approche de déformation des paiements est basée sur la psychologie d'un joueur impliqué dans le jeu du dilemme du prisonnier. Ses croyances sur le comportement de l'autre joueur est la meilleure méthode pour capturer cette psychologie.

Dans l'information complète ou incomplète, les croyances mutuelles sur la réciprocité sont cruciales dans l'émergence de la coopération. La complexité des comportements individuels nous oblige à étudier la psychologie des agents et l'origine de leurs motivations. Un agent peut construire ses motivations en donnant de l'importance à l'intérêt général, l'équité peut remplacer l'individualisme et la confiance peut dominer l'opportunisme.

L'extension de cette approche des paiements déformés aux autres jeux à forme normale ou extensive constitue une voie de recherche intéressante à explorer.

5. Annexes

5.1. Annexe 1 : preuve de la proposition 1

D'après la matrice des paiements, on a d'une part, si $p \neq 1$, $a > r(p)$. En effet, puisque $2.a > (b+c)$ et $a > d$, donc, $p^2.a + p(1-p)(c+b) + (1-p)^2.d < p^2.a + p(1-p)(2.a) + (1-p)^2.a = a$. Alors $a > p^2.a + p(1-p)(c+b) + (1-p)^2.d = r(p)$ pour tout $p \in [0,1[$. D'autre part, si $d \leq \frac{1}{2}(b+c)$ et $p > 0$ on a: $p^2.a + 2p(1-p)\frac{1}{2}(b+c) + (1-p)^2 d > d[p^2 + 2p(1-p) + (1-p)^2] = d$. Donc $r(p) > d$. Par contre, si $d > \frac{1}{2}(b+c)$, alors $p^2.a + p(1-p)(b+c) + (1-p)^2.d > d \Leftrightarrow p^2(a+d) + p(b+c) - p^2(c+b) - 2.p.d > 0$ et donc $p > \frac{2.d-c-b}{a+d-c-b}$. Or $2.d > (b+c)$ et comme $2a > b+c$, donc on a aussi $(a+d-c-b) > 0$ et par conséquent $1 > \frac{2.d-c-b}{a+d-c-b} > 0$. Il s'ensuit qu'il existe $p^* = \frac{2.d-c-b}{a+d-c-b}$ telle que si $p > p^*$ alors $r(p) > d$. Q.E.D.

5.2. Annexe 2 : preuve de la proposition 2

• D'un côté on a: $\mathbf{E}(R,q) > \mathbf{E}(Al,q) \Leftrightarrow \frac{1-q_1-q_2}{q_2} > \frac{a-r(p)}{d-b}$, qui est équivalent à $r(p) > a - \frac{1-q_1-q_2}{q_2}(d-b)$. En utilisant l'expression de $r(p)$, nous obtenons $p^2(a+d-c-b) + p(c+b-2d) + d > a - \frac{1-q_1-q_2}{q_2}(d-b)$. Soit $g_1(p) = p^2(a+d-c-b) + p(c+b-2d) + d - a + \frac{1-q_1-q_2}{q_2}(d-b)$. Étudions le signe de $g_1(p)$ pour savoir sur quel domaine elle est positive: $g_1'(p) = 2p(a+d-c-b) - (2d-c-b) > 0 \Leftrightarrow p > (<) \frac{2d-c-b}{2(a-b-c+d)} = \bar{p}$ si $(a-b-c+d) > (<) 0$. Dans le cas où le terme $(a-b-c+d)$ est nul, $g_1(p)$ devient une fonction que l'on peut étudier aisément. On a $\frac{2d-c-b}{2(a-b-c+d)} < 1$, en effet, $\frac{2d-c-b}{2(a-b-c+d)} < 1 \Leftrightarrow (-c-b) < (2a-2b-2c) \Leftrightarrow 2a > (b+c)$ et $2a > (b+c)$ est une condition liée aux paiements du dilemme du prisonnier. Nous avons trois cas:

Cas 1. Si $2d > c+b$ donc $(a-b-c+d) > 0$ puisque $a+d > 2d > c+b$, donc $g_1'(p) > 0$ pour $p > \bar{p} > 0$.

Cas 2. Si $2d < c+b$ et $(a-b-c+d) > 0$ donc $g_1'(p) > 0$ $\forall p$ puisque $\bar{p} < 0$.

Cas 3. Si $2d < c+b$ et $(a-b-c+d) < 0$ ainsi $g_1'(p) > 0$ $\forall p \in]0, \bar{p}[$.

Nous avons aussi $g_1(1) = \frac{1-q_1-q_2}{q_2}(d-b) > 0$ et $g_1(0) = \frac{(1-q_1)(d-b)-q_2(a-b)}{q_2} > 0 \Leftrightarrow q_2 < \frac{(1-q_1)(d-b)}{a-b}$. Donc pour les trois cas, puisque $g_1(1) > 0$ et g_1 est continue, il existe $]\tilde{p}, 1[$ sur lequel $g_1(p)$ est positive. Par conséquent $\mathbf{E}(R,q) > \mathbf{E}(Al,q)$ sur $]\tilde{p}, 1[$.

• D'un autre côté, $\mathbf{E}(R,q) > \mathbf{E}(Eg,q) \Leftrightarrow \frac{q_1}{q_2}(a-c) > d - r(p) \Leftrightarrow r(p) > d + (c-a)\frac{q_1}{q_2}$. Ceci implique $p^2(a+d-c-b) + p(c+b-2d) + d > d + (c-a)\frac{q_1}{q_2}$. Soit $g_2(p) = p^2(a+d-c-b) + p(c+b-2d) + (a-c)\frac{q_1}{q_2}$.

$g_2'(p) = 2p(a+d-c-b) + (c+b-2d) > 0 \Leftrightarrow p > (<) \frac{2d-c-b}{2(a-b-c+d)} = \bar{p}$ si $(a-b-c+d) > (<) 0$. Trois cas se présentent comme dans le cas de $g_1(p)$. En outre $g_2(1) = (a-d) + (a-c)\frac{q_1}{q_2} > 0 \Leftrightarrow \frac{q_1}{q_2} < \frac{a-d}{c-a}$ et $g_2(0) = (a-c)\frac{q_1}{q_2} < 0$. Pour les trois cas, si $\frac{q_1}{q_2} < \frac{a-d}{c-a}$, alors $g_2(1) > 0$. Donc, il existe $1 > \hat{p} > 0$ tel que pour chaque $p > \hat{p}$, nous avons $g_2(p) > 0$. Donc $r(p) > d + (c-a)\frac{q_1}{q_2}$ et $\mathbf{E}(R,q) > \mathbf{E}(Eg,q)$. Q.E.D.

5.3. Annexe 3 : preuve de la proposition 3

En l'absence du comportement réciproque ($q_2 = 0$), l'état de la population est $(q_1, 1-q_1)$. Aussi, $\mathbf{E}(Al,q) = q_1(a-b) + b$ et $\mathbf{E}(Eg,q) = q_1(c-d) + d$. On remarque que $\mathbf{E}(Eg,q) > \mathbf{E}(Al,q)$ quelque soit l'état de la population. Dans le cas où le comportement réciproque n'existent pas dans la population, le joueur fait le même choix contre tous les autres joueurs de la population. En tenant compte de l'information q sur la distribution du comportement dans la population, la probabilité de coopération du joueur i est

$$p(C^i/q) = p(\mathbf{E}(Al,q) > \mathbf{E}(Eg,q)) = 0$$

Ceci démontre l'impossibilité de l'émergence de la coopération. Q.E.D.

5.4. Annexe 4 : preuve de la proposition 4

Raisonnant par rapport au comportement dans la détermination de l'équilibre. Puisque ces comportements peuvent changer sur l'intervalle [0,1], la réciprocité au voisinage de (1,1) est déterminante pour la coopération.
Si les deux joueurs sont réciproques, il existe donc un p^*, $0 < p^* < 1$ mais suffisamment proche de 1 tel que pour un joueur i et quel que soit $p(C^j) > p^*$, $\alpha_i(p(C^j)) \geq p(C^j)$. Donc $\alpha_j(p(C^i)) \geq p(C^i)$ et $\alpha_i(p(C^j)) \geq p(C^j)$. Il est évident que, pour chaque joueur i, sa meilleure réponse à $p(C^j) = 1$ est $p(C^i) = \alpha_i(1) \geq 1$ donc $p(C^i) = 1$, par conséquent la coopération mutuelle est un équilibre de Nash.
Dans le cas contraire, si au moins un joueur n'est pas réciproque au voisinage de (1,1), ce point ne peut être un équilibre de Nash. Q.E.D.

5.5. Annexe 5 : preuve de la proposition 5

Preuve. Le type de chaque joueur est donné par t_k où $t_k \in \{RS, RL, NR\}$. Le paiement espéré du joueur $k = i, j$ est alors donné par

$$U_k = \sum_{(t_{-k} \in T_{-k})} u_k(s_k, s_{-k}(t_{-k}))p_k(t_{-k}/t_k)$$

Pour déterminer l'équilibre du jeu, on compare les paiements espérés pour tous les types de comportements possibles :

- Pour le joueur i,
si $t_i = RS$ alors $U_i = p_1 r(p) + p_2 a + p_3 d$,
si $t_i = RL$ alors $U_i = p_1 a + p_2 a + p_3 d$,
si $t_i = NR$ alors $U_i = p_1 d + p_2 d + p_3 d$.

- Pour le joueur j,
si $t_j = RS$ alors $U_j = q_1 r(p) + q_2 a + q_3 d$,
si $t_j = RL$ alors $U_j = q_1 a + q_2 a + q_3 d$,
si $t_j = NR$ alors $U_j = q_1 d + q_2 d + q_3$.

Puisque $a > d$ et $a > r(p)$, alors si $p_1 > 0$ et $q_1 > 0$, les deux paiements $p_1 a + p_2 a + p_3 d$ et $q_1 a + q_2 a + q_3 d$ sont les plus élevés pour les joueurs i et j respectivement. Le seul équilibre de Nash bayésien est donné par (RL, RL). Q.E.D.

Références bibliographiques

Allais M. (1952), « La psychologie de l'homme rationnel devant le risque : la théorie de l'expérience », *Journal de la société statistique de Paris*, 47-73.

Allais M. (1953), « Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine », *Econometrica*, (21) : 503-546.

Axelrod R. (1984), « The Evolution of Cooperation ». NY : *Basic Books*.

Axelrod R. and D. Hamilton (1981), « The Evolution of Cooperation », *Sciences* 211, 1390-1396.

Dawes R. and Thaler R. (1988), « Anomalies : Cooperation », *Journal of Economic Perspective*, 3, 187-197.

Falk A. and Fischbacher U. (2002), « The Economics of Reciprocity – Theory and Evidence », *in Inequality Around the World*, ed. by Richard Freeman, New York : Palgrave, in association with the IEA, 207-233.

Fehr E. and Schmidt K.M. (2003), « Theories of Fairness and Reciprocity : Evidence and Economic Applications » , *in* Dewatripont, M. *et al.* (eds), *Advances in Economic Theory* , Eigth World Congress of the Econometric Society, vol. I, p. 208-257, Cambridge, Cambridge University Press.

Fehr E. and Gächter S. (2000), « Fairness and Retaliation : The Economics of Reciprocity », *Journal of Economic Perspectives* 14, 159-181.

Harsanyi J. (1977), « Morality and the theory of Rational Behavior », *Social Research* , 44, 623-656.

Kreps D., Milgrom P., Roberts J. and Wilson R., (1982), « Rational Cooperation in the Finitely Repeated Prisoner's Dilemma », *Journal of Economic Theory*, 27, p. 245-252.

Mauroy H. (2002), « Mixed-Strategy Nash Equilibrium, Criteria of Decision Making under Uncertainty and Payoff Distorsion », *Economie Appliquée* , 3, 91-104.

Rabin M. (1993), « Incorporing Fairness into Game Theory and Economics », *American Economic Review* , 83, 1281-1302.

Savage L. (1972), *Foundations of Statistics* , NY : Dover Books.

Sherratt T. and Roberts G. (1998), « The Evolution of Generosity and Choosiness in Cooperative Exchanges », *Journal of Theorical Biology* 193, 167-177.

Wahl M. and Novak A. (1999), « The Continuous Prisoner's Dilemma : Linear Reactive Strategies », *Journal of Theorical Biology* 200, 307-321.