

Pression des pairs et incitation à l'effort : Fondements théoriques et évidence expérimentale

Thèse de Doctorat (NR) en Sciences Economiques

Présentée et soutenue publiquement par

David MASCLET

Le 26 septembre 2002

JURY Marie-Claire VILLEVAL Directrice de Recherche CNRS, Directrice de thèse Alan P. KIRMAN
Professeur à l'Université d'Aix Marseille III, Rapporteur Louis LEVY-GARBOUA Professeur à
l'Université Paris I, Panthéon-Sorbonne, Rapporteur Charles NOUSSAIR Professeur à l'Université de
Purdue, Etats-Unis Jean-Louis RULLIERE Professeur à l'Université Lumière Lyon II, Jean-Robert
TYRAN Professeur à l'Université de St Gallen, Suisse

Table des matières

Résumé : .	1
Abstract : . .	3
Introduction générale . .	5
Chapitre 1. La tentation du passage clandestin dans une communauté de pairs . .	11
1. Introduction .	11
2. Les effets de l'observation du passage clandestin .	13
2.1. Passage clandestin et contrôle mutuel : Kandel et Lazear (1992) .	14
2.2. Le contrôle mutuel dans une équipe de travail : Dong et Dow (1993) .	15
2.3. Le contrôle mutuel dans une relation d'agence : Barron et Gjerde (1997) .	17
3. La motivation du passage clandestin dans un partenariat .	22
3.1. Hypothèses et fonctions de gain . .	22
3.2. Travail en équipe et problème de passage clandestin .	22
4. L'exercice de la pression des pairs . .	23
4.1. L'environnement de la pression des pairs .	23
4.2. La pression des pairs par le contrôle mutuel .	24
4.3. La pression des pairs par les sanctions .	28
5. Conclusion .	29
Chapitre 2. Sanctions monétaires versus sanctions non monétaires ¹¹ .	31
1. Introduction .	31
2. La pression des pairs et l'effort de contribution dans les jeux de biens public .	35
2.1. Jeux de bien public sans opportunité de sanction .	36
2.2. Jeu de bien public avec mécanismes d'incitation par sanction et/ou récompense .	40
3. Inefficacité théorique des sanctions monétaires et non monétaires .	46
3.1. Le protocole expérimental .	47

¹¹ Ce chapitre est le fruit d'une collaboration avec Charles Noussair, Steve Tucker et Marie-Claire Villeval

3.2. Conditions d'information . .	48
3.3. Prédications théoriques .	49
4. Efficacité réelle des sanctions : aversion à la baisse des gains et à la désapprobation . .	51
4.1. Observations des niveaux de contributions individuels .	51
4.2. Caractéristiques des comportements individuels de contribution . .	64
4.3. Observations des niveaux de gains .	71
5. Conclusion .	73
Chapitre 3. l'autodiscipline par la menace d'ostracisme .	77
1. Introduction .	77
2. La désignation et l'exclusion des passagers clandestins par leurs pairs .	79
2.1. Jeu de désignation par les pairs: Cahuc et Kramarz (1997) . .	79
2.2. L'exclusion par les pairs: Hirshleifer et Rasmusen (1989) .	82
3. Les conditions d'un pouvoir dissuasif de la menace d'ostracisme .	84
3.1. Prédications théoriques .	85
3.2. Le protocole expérimental .	89
4. Ostracisme et coopération .	90
4.1. Niveaux de contribution et d'exclusion .	90
4.2. Pourquoi exclure les autres membres de son groupe ? . .	98
4.3. Effet global de l'ostracisme . .	103
5. Conclusion .	105
Chapitre 4. La renonciation au passage clandestin par l'aversion à l'inégalité . .	107
1. Introduction .	107
2. Le poids théorique des préférences sociales .	110
2.1. Le rôle des intentions .	111
2.2. Le poids des considérations distributives . .	113
3. La tentation du passage clandestin contrariée par les considérations distributives .	117
3.1. Modèle sans pression des pairs . .	118
3.2. Modèle avec pression des pairs . .	120
3.3. Relâchement de certaines hypothèses . .	123

4. Le pouvoir explicatif de l'aversion à l'inégalité .	124
4.1. Prédictions théoriques des jeux de bien public avec aversion à l'inégalité .	125
4.2. Résultats expérimentaux .	128
5. Conclusion .	139
Conclusion générale .	141
Annexes . .	145
ANNEXES CHAPITRE I .	145
ANNEXE I.A 1 : Le modèle d'Holmström (1982) . .	145
ANNEXE I.A 2 : Le modèle d'Esvaran et Kotwal (1984) .	146
ANNEXE I.A 3 : Le modèle de Rasmusen (1987) . .	146
ANNEXE I.B : Internalisation de la pression des pairs .	147
ANNEXE I.C Statique comparative avec système de deux équations et deux inconnues .	148
ANNEXES CHAPITRE II . .	148
ANNEXE II.A : Instructions pour l'expérience .	148
ANNEXE II.B Niveaux de contribution de groupe (sessions MP à Purdue et à Lyon GATE) .	152
ANNEXES CHAPITRE III .	154
ANNEXE III.A : Instructions de l'expérience .	154
ANNEXE III.B1 . .	159
ANNEXE III.B2 : Définition de la coopération .	160
ANNEXE III.C : Contributions de groupe . .	162
ANNEXE III.D: tests non paramétriques .	163
ANNEXES CHAPITRE IV .	164
ANNEXE IV.A Equilibre coopératif sans pression des pairs .	164
ANNEXE IV.B Equilibre coopératif sans pression des pairs dans un jeu de bien public .	165
Bibliographie . .	167

Résumé :

Ces dernières années, la littérature sur la pression des pairs appliquée aux équipes de travail a modélisé le contrôle mutuel en supposant que l'environnement de pression des pairs se résumait à l'effort de contrôle exercé par les agents sur les autres membres de l'équipe.

L'objet de cette thèse est de dépasser les modèles de contrôle mutuel existants en considérant que les décisions des agents portent également sur les décisions de sanction. Dès lors que sanctionner est coûteux, les agents, rationnels et maximisateurs de leur gains pécuniaires ne sont pas incités à punir leurs pairs. Par conséquent, sans menace crédible la coopération ne peut pas émerger. Les modèles théoriques proposés sont testés à l'aide de l'économie expérimentale. Les expériences réalisées testent une relation d'équipe de production où les agents ont la possibilité de sanctionner leurs pairs. Les résultats des expérimentations montrent que les sujets n'hésitent pas à sanctionner les passagers clandestins et que l'opportunité de sanctionner accroît considérablement le niveau de coopération. Les résultats expérimentaux mettent également en évidence que l'efficacité de la pression des pairs est fortement corrélée au coût ainsi qu'à la nature des sanctions (sanctions monétaires, non monétaires, exclusion).

Le niveau élevé de coopération lorsque les sujets ont l'opportunité de se sanctionner mutuellement s'explique par la volonté des agents d'éviter à la fois les conséquences monétaires des sanctions mais également la désapprobation de leurs pairs. Par ailleurs, deux raisons principales sont avancées pour expliquer les comportements de sanction. La première raison est que les agents sanctionnent leurs pairs afin de les inciter à coopérer davantage dans le futur. La deuxième raison repose sur les considérations distributives des agents et l'aversion à l'inégalité auto-centrée explique la sanction par la volonté de réduire les différences de gains. La coopération peut alors apparaître si les agents sont suffisamment averses à l'inégalité.

Mots clés : pression des pairs, contrôle, travail en équipe, passager clandestin, coopération, sanction, économie expérimentale, jeu de bien public

Abstract :

In recent years, models of peer pressure which have been developed are essentially models of mutual monitoring insofar as they assume that the only choice confronting agents in determining the peer pressure environment is to specify their monitoring effort.

In this thesis, we aim at considering a model of peer pressure where the agents decide on the sanctions to impose to eventual free riders. When sanctioning is costly, a rational money maximizer will never punish. As a consequence, subjects will shirk because they anticipate that they will never be punished. We conducted several experimental treatments to explore the validity of the theoretical predictions on the efficiency of peer pressure. The experiments simulate team production and allow subject to punish the other members. Our results indicate that subjects often engage in costly punishment of shirkers and that the availability of punishment increases cooperation. Moreover, the efficiency of peer pressure is strongly correlated with the cost and the nature of the sanction.

We find that the increase in cooperation is not only due to the willingness to avoid monetary consequences of sanctions but also to avoid disapproval of the peers. This analysis also provides evidence that subjects sanction their peers for two main reasons. The first explanation is that subject punish their peers because they expect that sanctions might have positive effects on future cooperation. The second explanation relies upon considerations of fairness : subjects are willing to sanction their peers in order to reduce inequity of payoffs. So full cooperation, as observed in the experiments, would be sustained as an equilibrium outcome if subjects would be sufficiently inequity averse.

Keywords : peer pressure, monitoring, work teams, free riding, cooperation, sanction, experimental economics, public good game with opportunity of sanction

Je tiens tout d'abord à remercier ma directrice de thèse, Marie-Claire Villeval, Directrice de Recherche au CNRS pour son soutien constant et ses encouragements qui m'ont permis de surmonter les nombreuses périodes de doute qui ont ponctué ce travail. Je lui suis infiniment reconnaissant des précieux conseils qu'elle m'a prodigués et de la patience dont elle a fait preuve à mon égard. Qu'elle trouve ici l'expression de ma sincère reconnaissance pour la qualité de son encadrement et la confiance qu'elle m'a accordée.

Mes remerciements vont aussi au professeur Jean-Louis Rullière, Directeur du GATE pour m'avoir accueilli au sein du laboratoire et m'avoir fourni des conditions de travail idéales pour mener à bien mon projet de thèse. Je tiens par ailleurs à lui exprimer ma sincère gratitude pour m'avoir fait découvrir et avoir su, lors de mon cursus universitaire, éveiller mon intérêt pour l'économie expérimentale.

Je tiens également à remercier le professeur Charles Noussair qui m'a accueilli au sein de son département de l'Université de Purdue lors de mes séjours aux Etats-Unis. Je souhaite lui exprimer ici toute mon admiration et ma profonde gratitude. Qu'il trouve ici l'expression de ma sincère reconnaissance.

Mes remerciements vont également aux professeurs Louis Lévy-Garboua et Alan Kirman pour avoir accepté d'être rapporteurs de cette thèse ainsi qu'au professeur Jean-Robert Tyran pour sa participation à cette soutenance qu'il honore de sa présence.

Tout au long du parcours qui s'achève avec la soutenance de cette thèse, j'ai rencontré des

personnes extraordinaires pour leurs qualités humaines et scientifiques. Merci plus particulièrement aux professeurs David Dickinson et Claude Montmarquette. Leurs conseils et leurs réflexions pertinentes m'ont été précieux

Ce travail doit aussi beaucoup aux stages de recherche que j'ai effectués au sein du département d'économie de l'Université de Purdue. Ces stages n'auraient pas été possibles sans le soutien financier de la Région Rhône Alpes. Qu'ils soit remerciés de leur accueil et de leur aide.

Mes remerciements vont aussi à tous ceux qui ont su apporter l'aide nécessaire à la mise en œuvre des sessions expérimentales qui seront présentées dans ce document. Pour cela, je remercie l'ensemble des membres du Gate qui ont toujours répondu favorablement à mes nombreuses sollicitations. Ces remerciements vont plus spécialement à Romain Zeiliger, l'auteur du Logiciel REGATE ayant permis mes expérimentations. Merci Romain pour tes nombreux conseils. Par ailleurs, ces expérimentations n'auraient pas été possibles sans l'aide et les compétences de l'équipe informatique et de l'équipe administrative du Gate .

J'adresse une pensée particulière à l'ensemble de l'équipe du Gate. Merci plus particulièrement pour leurs coups de main et leurs encouragements à Nacer Bernou, Sandra Cavaco, Patricia Criffo Tillet, Armelle Delorme, Yannick Gabuthy, Nicolas Jacquemet, Sandrine Ollier, Yohan Pelosse, Hind Sami, Mathieu Neveu, David Risser, Mustapha, Sadni. J'attends avec impatience la fin de thèse de ceux qui doivent en passer par là, en espérant leur être aussi utile qu'ils l'ont été pour moi.

Mes pensées vont enfin vers ma famille et l'ensemble de mes amis. Merci à tous pour votre amour et votre confiance. J'ai réussi grâce à vous à surmonter bien des obstacles. Je veux également associer à ce travail celle qui a su m'encourager et me soutenir lors de cette fin de thèse. Merci Nathalie.

Introduction générale

« Aussi égoïste l'homme puisse-t-il être supposé, il y a évidemment des principes dans sa nature qui l'amènent à s'intéresser au sort des autres et lui rendent nécessaire leur bonheur, même s'il n'en tire aucun avantage personnel, hormis le plaisir de le voir. Notre esprit associe alors la honte à toute violation de la confiance instaurée. » Adam Smith, *La Théorie des Sentiments Moraux* (1759)

A l'instar d'Adam Smith, force est de constater que les décisions des individus ne sont pas uniquement guidées par des aspects purement égoïstes mais également par des considérations d'ordre social. Ils s'intéressent au sort des autres et plus particulièrement à celui de leurs pairs, c'est-à-dire l'ensemble des personnes placées sur un pied d'égalité, pour qui ils ont de l'empathie et à qui ils se comparent. Les expériences d'Hawthorne menées à partir de 1924 sous la houlette de Mayo (Mayo, 1949 ; Roethlisberger et Dickson, 1939) ont ainsi montré l'importance des facteurs proprement humains tels que l'empathie ou la désapprobation sociale entre les salariés au sein des organisations. Ces expériences réalisées dans les ateliers de la West Electric Company ¹ révélèrent

¹ Cette expérience fut réalisée de 1924 à 1927 dans les ateliers Hawthorne de la West Electric Company dans la banlieue de Chicago. La direction, avec la collaboration de chercheurs universitaires décida d'améliorer les conditions de travail tout en vérifiant leur relation à la productivité. Un groupe d'ouvrières accepta d'être isolé dans l'atelier pour continuer le même travail dans une pièce à part. On changea successivement un certain nombre de facteurs : le système de salaire (individuel, par équipe, au rendement, ect.), les pauses durant le travail, les horaires. Il apparut qu'à chaque changement, dans quelque sens qu'il se fit, la productivité augmentait. Au total, on obtint, à la fin de l'expérience, une augmentation de productivité de 20%. Ces expériences mirent également en exergue l'importance des effets de groupe au sein des ateliers.

l'importance de la vie de groupe et son influence sur le comportement de chacun de ses membres. Elles montrèrent que les employés ne se contentent pas d'obéir aux règles officielles de l'entreprise mais également à des normes informelles mises en place par les employés eux-mêmes au sein des ateliers et groupes de travail. Les salariés établissent ainsi des normes de production à atteindre mais également à ne pas dépasser, à la fois pour maintenir la cohésion dans le groupe et pour éviter la révision des standards officiels de l'entreprise.

Afin de s'assurer que les normes établies sont respectées, les membres du groupe s'engagent dans des activités de contrôle de leurs pairs pour d'observer l'effort de production des autres membres du groupe. Si un individu viole les normes établies de manière évidente et excessive, ses pairs n'hésitent pas à le sanctionner d'une manière ou d'une autre (Francis, 1985 ; Miller ; 1992 ; Franck, 1994). Les sanctions imposées par les pairs sont de nature diverse. Elles peuvent être morales et reposer sur la désapprobation des pairs. L'observation par les pairs véhicule alors la sanction qui se traduit par un sentiment de honte (Kandel et Lazear, 1992). Une autre forme de sanction est l'ostracisme, qui consiste à exclure du groupe un membre déviant (Hirshleifer et Rasmusen, 1989). Les sanctions peuvent aussi être de nature monétaire et consister à ce que l'agent déviant ne reçoive pas la totalité de la part de l'output qui lui est due (Jones, 1984 ; Dong et Dow, 1993) ou qu'il fasse l'objet d'actions de sabotage de la part de ses pairs (Lazear, 1989). Les sanctions des pairs peuvent enfin aller jusqu'à des violences physiques et des châtiments corporels (Frank, 1994 ; Francis, 1985).

L'étude des mécanismes de discipline par les pairs a fait l'objet d'un certain nombre de travaux dans des domaines aussi variés que l'octroi de crédit (Varian, 1990), les assurances (Arnott et Stiglitz, 1991), la régulation des ressources publiques (Ostrom, 1990), l'université (Carmichael, 1988 ; Sacerdote, 2000) ou l'organisation du travail en équipe (Kandel et Lazear, 1992 ; Dong et Dow, 1993 ; Barron et Gjerde, 1997). Varian (1990) a étudié l'efficacité du contrôle mutuel dans le cadre de l'octroi de crédits à partir d'un exemple particulièrement intéressant de contrôle mutuel, la *Grameen Bank* -de *Gram : village*-. La *Grameen Bank*, créée en 1977 par un économiste du Bangladesh, M. Yunus, octroie des micro-crédits aux plus démunis. Cette banque utilise un schéma incitatif particulier qui lui permet de réduire ses coûts de sélection et de contrôle. Les crédits sont accordés à des groupes de quatre ou cinq individus plutôt qu'à des individus isolés afin de faciliter le remboursement. Les individus s'engagent dans le contrôle des pairs afin de s'assurer du remboursement du prêt par chacun des membres du groupe. Cinquante huit pays dont les Philippines, le Vietnam, l'Ethiopie ou le Burkina ont appliqué des programmes de crédit inspirés de la *Grameen Bank* (Farnsworth, 1988). Arnott et Stiglitz (1991) ont étudié quant à eux l'efficacité du contrôle mutuel appliqué à l'assurance. Les auteurs considèrent la possibilité pour les agents de 'souscrire' à une assurance faite d'assistance mutuelle entre les membres d'une même famille. L'assistance mutuelle est efficace dans la mesure où elle favorise le contrôle mutuel.

Un autre exemple du contrôle mutuel est celui qui s'applique au recrutement sur le marché interne du travail. Ainsi, en Europe, dans certains orchestres, comme au Philharmonique de Berlin, les membres de l'orchestre assistent à l'audition des candidats et la décision finale est prise après le vote de l'ensemble des membres de l'orchestre

(Harris et Townsend 1981). De même, lorsque l'administration universitaire ne peut pas juger du talent et des compétences des nouveaux postulants à un poste de chercheur, les titulaires sont plus à même de le faire puisqu'ils possèdent davantage d'informations sur les postulants. Ils sont d'autant plus incités à recruter les meilleurs éléments que la réputation de l'entité à laquelle ils appartiennent en dépend (Carmichael, 1988).

De nombreux travaux se sont également intéressés au contrôle mutuel dans les équipes de production. Un exemple du contrôle mutuel dans les équipes de production est celui des Kanbans japonais (Aoki, 1986)². Les kanbans reposent sur la participation du personnel et l'idée selon laquelle la discipline est généralement plus efficace lorsqu'elle est imposée par les pairs que par les décisions hiérarchiques. Dans les entreprises japonaises où le travail est organisé en équipe, les membres des équipes n'hésitent pas à exercer un contrôle sur les activités de leurs pairs. Ainsi les performances de productivité des entreprises japonaises seraient en grande partie attribuées au contrôle mutuel. Kandel et Lazear (1992) ont modélisé le contrôle mutuel dans le cadre d'un partenariat et mettent en évidence l'importance de la désapprobation des pairs véhiculée par l'observation. A la suite des travaux de Kandel et Lazear (1992), Dong et Dow (1993) étudient le contrôle mutuel dans le cadre spécifique des équipes chinoises de production agricole. Barron et Gjerde (1997) modélisent quant à eux le contrôle mutuel dans une relation d'agence et montrent sous quelles conditions le principal peut bénéficier des mécanismes de discipline par les pairs.

L'objet de cette thèse consiste à s'interroger sur les conditions d'efficacité de la pression des pairs au sein des équipes de travail. A l'instar de Barron et Gjerde (1997), on définit la pression des pairs, comme l'ensemble de ces mécanismes de contrôle mais aussi de sanction mis en place par les agents eux-mêmes afin de discipliner leurs pairs. L'analyse de la pression des pairs présentée dans cette thèse vise d'une part, à étudier dans quelle mesure les membres d'une équipe sont incités à exercer une pression sur leurs pairs et d'autre part, à montrer sous quelles conditions la pression des pairs conduit les agents à adopter un comportement moins opportuniste et davantage coopératif au sein du groupe.

Lorsque le travail est organisé en équipe de production où les profits sont partagés à égalité entre les membres de l'équipe, les agents peuvent être incités à adopter un comportement opportuniste. Les pratiques de partage des profits conduisent ainsi les membres du groupe à tirer avantage de leur appartenance au groupe sans toutefois participer systématiquement à la génération du profit (Alchian et Demsetz, 1972 ; Jensen et Meckling, 1979). Déjà, en 1740, Hume décrivait l'inclinaison des individus à profiter des activités collectives tout en contribuant le moins possible à ces activités. Ce problème est généralement qualifié de problème de passager clandestin. La notion de passager clandestin traduit l'existence d'un conflit potentiel entre les intérêts de l'individu et ceux du groupe. Pour illustrer le problème de passager clandestin, Alchian et Demsetz (1972) donnent l'exemple d'une équipe de déménageurs. Pour déplacer un piano, ces derniers doivent être quatre. S'ils ne sont que trois, la productivité de l'équipe est de zéro. S'ils

² La méthode *Kanban* a été élaborée et promue au Japon par TOYOTA MOTOR Co. C'est une technique d'approvisionnement sur consommation approchant de près les idées du Juste à Temps.

sont quatre, elle est maximale. Mais le chef de l'équipe est incapable de déterminer si chacun des membres a fourni $\frac{1}{4}$ de l'effort total ou non. Pour cette raison, chaque déménageur est incité à fournir l'effort le plus faible possible en comptant sur l'effort des autres.

Si un schéma de rémunération, fondé sur le partage des profits, fournit aux agents des incitations fortes à adopter un comportement opportuniste, il peut également les inciter à se contrôler mutuellement dans la mesure où tout comportement de passager clandestin affecte l'ensemble de l'équipe. En effet, si un agent adopte un comportement opportuniste, il réduit la rémunération des autres membres de l'équipe puisque la rémunération de chaque membre du groupe dépend non seulement de son propre effort mais également de celui des autres membres de l'équipe. Donc, les agents seraient incités sous certaines conditions à exercer une pression sur leurs pairs afin de les inciter à coopérer davantage. Cela nous conduit donc à nous interroger sur les conditions d'efficacité de la pression des pairs. En effet, il ne suffit pas que les agents soient incités à exercer une pression sur leurs pairs mais il faut également que cette pression soit efficace, c'est à dire qu'elle réduise effectivement les incitations au passage clandestin.

La validation ou l'infirmité des modèles de pression des pairs sur la base d'observation de la réalité est difficile. En effet, les données empiriques que l'on peut récolter intègrent un nombre important de facteurs qui sont négligés dans la modélisation. Il est par ailleurs difficile d'observer les sanctions imposées par les pairs et leur impact sur le comportement des membres de l'équipe en raison de leur caractère très informel. La méthodologie retenue pour la soumission à réfutation des propriétés d'équilibre des modèles de pression des pairs est l'économie expérimentale. La démarche expérimentale peut se décomposer en trois phases (Camerer, 1997). Il s'agit dans une première phase d'observer le comportement des joueurs dans un environnement expérimental contrôlé par un protocole. L'expérimentation recrée les hypothèses avancées dans les modèles et permet d'isoler l'influence de certaines variables. Les participants aux expérimentations sont ainsi confrontés à une structure d'incitation identique à celle décrite dans le modèle. Dans une deuxième étape, il s'agit alors d'identifier et d'isoler les déterminants psychologiques et sociaux qui contribuent à l'explication de l'ensemble des observations par rapport à la prédiction théorique. En effet, l'économie expérimentale permet la collecte de données empiriques pertinentes et fiables. Il est alors possible de vérifier si les comportements observés sont conformes aux prédictions théoriques. Enfin, dans une troisième phase, il convient d'intégrer des explications plausibles aux observations dans les modèles lorsque la prédiction théorique se trouve réfutée par l'évidence expérimentale.

Le jeu expérimental utilisé afin de tester les prédictions théoriques de la pression des pairs est un jeu de bien public avec opportunité de sanction. En effet, l'organisation du travail en équipe peut facilement être transposée à un jeu de contribution à un bien public où le résultat fourni par l'équipe est le bien public et l'effort de production fourni par un agent est la contribution au bien public. Dans une expérience de bien public avec opportunité de sanction, chaque sujet doit décider dans un premier temps de son niveau de contribution à un bien public. Le financement du bien est obtenu par la somme des contributions individuelles. Une fois financé, le bien public est partagé à égalité entre tous

les membres du groupe. Dans une deuxième étape, les agents observent parfaitement les niveaux individuels de contribution et ont l'opportunité de sanctionner les autres membres du groupe. Chaque membre du groupe est individuellement incité à ne pas participer au financement du bien public, c'est à dire à adopter un comportement opportuniste. Un tel comportement constitue un dilemme social car au niveau collectif il serait plus efficient que chacun contribue au financement du bien public.

Le chapitre premier propose une modélisation de la pression des pairs afin d'étudier ses conditions d'existence et d'efficacité. Dans un premier temps, on modélise l'organisation du travail en équipe afin de mettre en évidence le problème de passager clandestin. Il apparaît que, sans opportunité à discipliner leurs pairs, les membres de l'équipe ne coopèrent pas à la production de l'output. Chacun des membres adopte un comportement opportuniste. Dans un deuxième temps, on introduit l'opportunité pour les agents à discipliner leurs pairs. On montre que l'efficacité du contrôle mutuel est fortement conditionnée par le degré de substituabilité des efforts de contrôle et de production. En effet, si le contrôle mutuel incite les agents à fournir davantage d'effort afin de ne pas subir la pression des pairs, il conduit également ceux qui contrôlent leurs pairs à réduire leur propre effort de production du fait de la substitution des efforts de production et de contrôle. On montre par ailleurs que l'exercice de la pression des pairs est fortement conditionné par le choix des actions qui visent à discipliner les pairs. Ainsi, lorsque les décisions des agents portent sur le choix des sanctions et que sanctionner est coûteux, la pression des pairs ne peut pas exister.

L'objet du **second chapitre** est de tester expérimentalement l'efficacité de la pression des pairs. Le jeu utilisé est un jeu de contribution volontaire au financement d'un bien public avec opportunité pour les sujets de sanctionner les autres membres de leur groupe. Ce jeu se compose de deux étapes. Dans une première étape, chaque sujet reçoit une dotation initiale individuelle qu'il doit répartir entre le bien privé et le bien public. Dans une deuxième étape, les sujets observent parfaitement les niveaux de contribution de chaque membre du groupe et peuvent leur attribuer des points de sanction.

Différents traitements expérimentaux ont été réalisés afin de tester l'efficacité des sanctions. Dans un premier traitement, les sujets participent à un jeu classique de bien public et n'ont pas l'opportunité de sanctionner leurs pairs. Ce traitement sert de traitement de référence. Dans un deuxième traitement, les sujets peuvent attribuer des points de sanction monétaire aux autres membres de leur groupe. Les points de sanction sont coûteux non seulement pour ceux qui les reçoivent mais également pour ceux qui les imposent. Dans un troisième traitement, les points de sanction n'affectent ni les gains des sujets sanctionnés, ni de ceux qui sanctionnent, et expriment uniquement la désapprobation morale des pairs. Les résultats expérimentaux révèlent que les sujets sont sensibles aux sanctions monétaires mais également non monétaires et contribuent davantage lorsqu'ils ont l'opportunité de se sanctionner mutuellement.

Le chapitre 3 vise à étudier l'ostracisme comme une forme particulière de pression des pairs. L'ostracisme joue un rôle important dans l'exécution des comportements socialement approuvés au sein de la plupart des groupes. Ainsi, dans certaines sociétés primitives, des individus peuvent être châtiés et mis à l'écart d'un village s'ils ne se conforment pas aux règles informelles préalablement établies. Le jeu d'ostracisme avec

projets imbriqués présenté dans ce chapitre se déroule de la façon suivante : dans une première étape, les sujets doivent décider combien contribuer à un premier projet commun. Dans une deuxième étape, les sujets observent parfaitement les niveaux de contribution des autres joueurs et ont l'opportunité d'exclure un ou plusieurs autres membres du groupe. Seuls les sujets qui n'ont pas été exclus peuvent participer à un deuxième projet.

La sanction d'ostracisme répond en partie aux interrogations suscitées par les résultats expérimentaux issus des expériences du chapitre 2. Plus précisément, comment bénéficier de l'effet incitatif des sanctions sans avoir à en subir les conséquences coûteuses ? En effet, si les sanctions incitent à davantage coopérer, elles sont généralement coûteuses pour l'équipe dans sa globalité puisqu'elles affectent à la fois les gains de ceux qui les reçoivent et de ceux qui les infligent. L'intérêt des sanctions suffisamment fortes, comme celles qui reposent sur l'exclusion, est que la menace d'être sanctionné peut être à elle seule suffisante pour inciter les sujets à coopérer. La sanction est alors rarement appliquée et n'engendre donc pas de coûts pour l'équipe. Ostrom, Walker et Gardner (1992) ont ainsi montré que l'efficacité des sanctions est positivement corrélée à l'importance de la sanction reçue. Un deuxième intérêt des sanctions qui consistent à exclure d'un deuxième projet ceux qui ne coopèrent pas à un premier projet est de « déplacer » les coûts de la pression des pairs. L'idée est la suivante : lorsque les membres d'un groupe sont engagés dans deux activités communes, alors une activité peut bénéficier des effets incitatifs de la pression des pairs tandis que l'autre activité en subit les conséquences coûteuses. Par exemple, les membres d'un groupe engagés dans une activité productive commune sont souvent amenés à partager d'autres activités non productives comme par exemple jouer dans la même équipe de football ou manger ensemble le week-end. Ceux qui ne coopèrent pas dans l'activité productive peuvent alors être exclus de l'activité non productive. L'activité productive bénéficie alors des effets incitatifs de la sanction d'exclusion sans en supporter les coûts puisqu'ils sont déplacés sur l'activité non productive. Les résultats expérimentaux montrent que les sanctions d'exclusion sont rarement appliquées mais que la menace d'ostracisme permet à elle seule de soutenir un niveau élevé de coopération dans le premier projet.

L'idée défendue dans le **quatrième chapitre** est que la modélisation de la pression des pairs gagnerait en pertinence si elle enrichissait sa représentation en considérant que les agents incorporent des considérations sociales dans leur fonction d'utilité. En effet, au regard des observations obtenues, il apparaît que les résultats expérimentaux semblent infirmer les prédictions théoriques des modèles de pression des pairs. Il convient dès lors de dégager une expression théorique des motivations sociales qui président les décisions des agents. Le choix de la fonction d'utilité modifiée adaptée à la modélisation de la pression des pairs est celui d'une approche fondée sur les considérations distributives d'aversion à l'inégalité. A l'instar de Fehr et Schmidt (1999), on définit l'aversion à l'inégalité de la façon suivante : une personne ressent de l'aversion à l'inégalité lorsqu'elle subit une certaine désutilité à percevoir des gains inférieurs à ceux des autres mais également lorsqu'elle obtient des gains supérieurs à ceux des autres.

Chapitre 1. La tentation du passage clandestin dans une communauté de pairs

« La famille est souvent dans la meilleure position pour contrôler et discipliner ses membres » Kandel et Lazear, 1992, *Journal of Political Economy*, p802.

1. Introduction

L'étude de la pression des pairs s'inscrit dans un contexte de restructuration des entreprises où la dimension collective de la compétence apparaît de plus en plus cruciale et où les schémas de rémunération traditionnels donnent place à des systèmes incitatifs reposant sur l'intéressement aux profits de l'entreprise (Greenan et Guellec, 1994). En effet, la plupart des changements organisationnels décrits dans la littérature suivent deux grandes directions. La première est une décentralisation partielle des décisions vers les unités les plus proches de l'information. La seconde est la réorganisation du travail en équipes polyvalentes et autonomes qui déterminent le partage des tâches de manière à remplir les objectifs fixés, avec une moindre intervention de maîtrise en poste

Dans ce contexte de réorganisation des entreprises, le mode de rémunération des

agents s'est également modifié. Dès lors que le travail est organisé en équipe et que les performances sont interdépendantes, les incitations salariales individuelles s'en trouvent affectées. Il est en effet plus difficile de contrôler les performances individuelles, c'est donc le travail collectif qui est récompensé. Lier les rémunérations aux performances individuelles apparaît dès lors problématique. Les schémas d'incitations basés sur l'intéressement deviennent une alternative incontournable aux schémas de rémunération traditionnels. L'intéressement consiste à relier la rémunération individuelle aux performances de l'entreprise. A ce titre, il concourt à impliquer les salariés dans le processus de production. Néanmoins, ce caractère incitatif est souvent remis en cause dans la mesure où chaque salarié doit partager le rendement de son effort avec tous les autres salariés de l'entreprise. Chaque salarié est alors tenté de se comporter comme un resquilleur (passager clandestin) en se reposant sur les efforts de ses collègues. En effet, chaque agent a intérêt à fournir l'effort de production le plus bas possible pour maximiser son utilité puisqu'il bénéficie à égalité du résultat total. Le mécanisme incitatif de partage des profits semble donc condamné du fait même du problème de passager clandestin.

Selon Alchian et Demsetz (1972) ; Jensen et Meckling (1979), la mise en place d'un mécanisme incitatif de partage des profits se justifie si et seulement si les agents mettent en œuvre des stratégies coopératives. De même, Wetzman et Kruse (1990) suggèrent que le partage des profits n'est efficace que lorsque le problème de passager clandestin est résolu. Une solution consiste à ce que l'entreprise verse un bonus d'un montant donné si et seulement si une certaine performance est réalisée. Dans le cas contraire, les salariés ne perçoivent aucun bonus. Dans de telles conditions, chaque salarié a intérêt à fournir un effort correspondant exactement à la performance moyenne requise si tous les autres font de même. Cette idée a été formalisée par Holmström (1982) qui préconise l'intervention d'un principal sanctionnant les agents en prélevant la totalité de l'output si un certain niveau d'output n'est pas atteint (voir annexe A1). Toutefois, le schéma de rémunération proposé par Holmström (1982) a fait l'objet d'un certain nombre de critiques (Esvaran et Kotwal, 1984 ; Rasmusen, 1987 ; Andolfatto et Nosal, 1997). La principale critique avance le fait que le schéma d'incitation d'Holmström ne tient pas compte du problème d'aléa moral émanant du principal. En effet, le principal peut être incité à corrompre un membre de l'équipe et lui demander de réduire son effort afin que le niveau d'output descende légèrement en dessous du niveau d'output fixé (voir annexes A2 et A3). Le principal prélève alors la globalité de l'output (Esvaran et Kotwal, 1984).

De nombreuses études empiriques ont souligné que le partage des profits ne conduit pas nécessairement à une réduction de la productivité. Ainsi, Cable et Wilson (1989, 1990), mettent en évidence, à partir d'une étude comparative entre le Royaume-Uni et l'Allemagne de l'Ouest, que l'efficacité du partage des profits est étroitement liée à l'organisation interne de l'entreprise et notamment à la possibilité pour les agents de se contrôler mutuellement. De même, selon Jones et Svejnar (1985), l'existence de groupes de pression et le contrôle mutuel devraient permettre de créer un environnement coopératif au sein des entreprises et faciliter ainsi le passage de comportements opportunistes en comportements coopératifs motivés par le partage des gains. En effet, alors que le partage des profits fournit aux agents des incitations fortes à adopter un comportement opportuniste, un tel mode de rémunération incite également au contrôle

mutuel. Le partage des profits peut inciter les agents à se contrôler mutuellement car la rémunération de chaque membre du groupe dépend non seulement de son propre effort mais également de celui des autres membres de l'équipe. Dès lors, si un agent adopte un comportement opportuniste, il réduit la rémunération des autres membres de l'équipe.

Les membres d'un groupe ne se contentent généralement pas de contrôler leurs pairs mais décident également des sanctions à imposer aux agents qui ne se conforment pas aux normes établies dans le groupe (Francis, 1985 ; Frank, 1994). L'ensemble des mécanismes de contrôle et de sanction mutuels visant à discipliner les membres du groupe est ainsi qualifié de pression des pairs. Les expériences de l'usine Hawthorne de la General Electric à la fin des années vingt ont mis en évidence l'existence de la pression des pairs au sein des équipes de travail. Ces expériences ont montré que les ateliers et les bureaux définissent des règles et des normes informelles auxquelles se réfèrent les salariés. Les contrevenants à ces normes sont alors sanctionnés par l'hostilité du groupe. Les entreprises de construction automobile japonaises implantées au nord des Etats-Unis offrent un autre exemple de pression des pairs. Dans ces usines, la production s'organise autour d'équipes de travail de petite taille facilitant la pression exercée par les pairs. Un absentéisme injustifié peut alors conduire à l'application de sanctions formelles de la part de la hiérarchie mais également à des sanctions informelles de la part des autres membres de l'équipe puisque l'équipe dans sa globalité souffre de l'absentéisme (Rehder, 1990).

L'objet de ce chapitre est d'étudier les conditions d'efficacité et d'existence de la pression des pairs. La modélisation de la pression des pairs présentée dans ce chapitre s'inspire des travaux de Kandel et Lazear (1992) et de Barron et Gjerde (1997) qui seront présentés dans la section suivante. Plus précisément, elle reprend l'analyse séquentielle de Barron et Gjerde mais s'inscrit dans le cadre d'un partenariat comme chez Kandel et Lazear (1992). Cette analyse se distingue de ces deux approches dans la mesure où elle enrichit l'étude de la pression des pairs en considérant que les décisions des agents ne portent pas uniquement sur les activités de contrôle mutuel mais également sur le choix des sanctions à imposer à d'éventuels passagers clandestins. De ce fait elle offre une vision plus complète de la pression des pairs puisqu'elle étudie à la fois les conditions d'efficacité de la pression des pairs en terme d'incitation à l'effort et ses conditions d'existence.

La deuxième section de ce chapitre présente les principaux travaux théoriques sur la pression des pairs appliquée à une équipe de production. Après cette revue de la littérature, la troisième section de ce chapitre vise à mettre en exergue le problème de passager clandestin. Ce modèle sert de référence dans l'étude de l'efficacité de la pression des pairs. La quatrième section étudie les effets incitatifs de la pression des pairs et s'interroge sur ses conditions d'existence. Enfin la cinquième section conclut le chapitre.

2. Les effets de l'observation du passage clandestin

On expose tout d'abord dans cette section le modèle fondateur de Kandell et Lazear (1992), qui occupe un rôle central dans la littérature sur la pression des pairs dans le cadre d'une équipe de production. On présente ensuite les modèles de Dong et Dow (1993) et de Barron et Gerde (1997).

2.1. Passage clandestin et contrôle mutuel : Kandell et Lazear (1992)

Kandell et Lazear (1992) sont les premiers à modéliser la pression des pairs dans le cadre d'une équipe de travail. Ils modélisent la pression des pairs appliquée à un partenariat, comme une alternative au schéma d'incitation d'Holmström (1982). Les partenariats diffèrent des organisations hiérarchiques dans la mesure où les membres de l'équipe sont tous créanciers résiduels : ils partagent les fortunes et infortunes de l'entreprise sans pouvoir déléguer le risque à une tierce personne. Les agents sont supposés identiques et neutres vis-à-vis du risque. Chaque agent $j, i=(1,...,N)$ fournit un effort de production noté $e_j, e_j \in [0,1]$ afin de réaliser un output $f(e_1, ..., e_N)$ tel que $f_1 > 0$ ³ et $f_{11} < 0$, non séparable en effort de production. Une fois réalisé, l'output est partagé entre les agents et chaque agent reçoit $f(e)/N$. Supposons qu'en plus de son activité de production, chaque agent fournit un effort de contrôle de ses pairs. On note a_j , l'effort de contrôle de l'agent i . Chaque agent i subit alors une désutilité $C(e_j, a_j)$ à fournir un effort de production mais également de contrôle de ses pairs. Cette fonction de coût est croissante et convexe ($C_1 > 0, C_{11} > 0, C_2 > 0$ et $C_{22} > 0$).

La fonction d'utilité de chaque agent i s'écrit de la façon suivante :

Le premier terme de la fonction d'utilité représente la part d'output reçue par l'agent i . Le deuxième terme traduit la désutilité pour l'agent à exercer un effort de production et de contrôle. Enfin le dernier terme représente la pression que subit l'agent i de la part de ses pairs.

Le jeu se déroule de la façon suivante. Dans un premier temps, les agents décident de leur effort de contrôle. Ils choisissent dans un deuxième temps leur effort de production. Résolvons le jeu à rebours. En dernière étape, les agents décident de leur effort de production afin de maximiser son utilité tel que :

Soit la condition de premier ordre :

s.c.

L'effort de contrôle optimal de l'agent i est alors déterminé à partir de la condition de premier ordre suivante⁴ :

Encadré I.1 :statique comparative : cas avec une seule variable

³ f_1 représente la dérivée première vis à vis de la première variable de la fonction $f(.)$.

⁴ Il y a une erreur typographique dans Kandell et Lazear (1992) sur le signe de $\partial e_j / \partial a_i$ (équation (16)).

La démarche à suivre est explicitée ci-dessous :
On sait que la fonction de choix optimal $x(a)$ doit satisfaire la condition :
En différenciant les deux côtés de cette identité, on obtient
On cherche $dx(a)/da$:
Au regard de la condition du second ordre, il apparaît que le dénominateur de cette expression est négatif. D'où, en tenant compte du signe « moins » précédant la fraction :
En conséquence, le signe de la dérivée du choix optimal par rapport au paramètre ne dépend que du signe de la dérivée partielle croisée de la fonction-objectif par rapport à x et a .

Kandel et Lazear (1992) soulignent l'importance de la désapprobation des pairs. Ils distinguent à ce titre la pression externe (la honte) de la pression interne (la culpabilité). La distinction entre honte et culpabilité dépend de l'observation des efforts. Alors que la honte est le fruit de l'observation de l'effort de production d'un agent par ses pairs (c'est à dire que $a_1, \dots, a_N > 0$), la culpabilité ne nécessite pas d'effort de contrôle par les pairs ($a_1, \dots, a_N = 0$) mais émerge d'un processus d'internalisation du regard. Les mécanismes psychologiques qui fondent l'internalisation d'une norme sont complexes. L'internalisation suppose en effet que les agents soient engagés dans des relations de longue durée et nécessite des investissements particuliers au préalable⁵. L'internalisation émergerait en fait de trois sources principales : les parents (transmission verticale), les influences institutionnelles (transmission oblique) et enfin les pairs (transmission horizontale) (Cavalli-Sforza et Feldman, 1981 ; Boyd et Richerson, 1985) (Voir annexe B).

Le modèle de Kandel et Lazear a le mérite de montrer d'une part l'importance des mécanismes de contrôle par les pairs comme une alternative possible aux schémas d'incitations classiques ; d'autre part, il met en exergue le rôle fondamental joué par les sentiments de désapprobation engendrés par la seule observation des pairs.

2.2. Le contrôle mutuel dans une équipe de travail : Dong et Dow (1993)

Dans le prolongement de l'analyse de Kandel et Lazear (1992), Dong et Dow (1993) étudient l'efficacité du contrôle mutuel et son implication dans le système agricole chinois

⁵ Parson (1954) montre l'importance des sentiments de culpabilité dans l'armée américaine. Il observe que les militaires américains dépensent beaucoup de temps et d'argent à instaurer un « esprit d'équipe » dans leurs rangs. L'objectif souhaité est d'inculquer un certain niveau de culpabilité aux soldats qui se trouvent seuls en mission. Parson prend l'exemple d'un pilote de chasse. Lorsque le soldat est seul dans son avion, le courage du pilote est alors difficile à appréhender. Le sentiment de culpabilité donc être suffisamment fort pour ne pas l'inciter à dévier de ses objectifs. Selon Parson, l'internalisation des standards à travers l'esprit d'équipe et autres agents de socialisation devrait conduire au développement d'un système interne de sanction équivalent au système externe et, par implication, à l'adoption de standards de comportement.

entre 1970 et 1976. L'organisation du travail en équipe dans les fermes collectives rendant difficile le contrôle des efforts individuels, le seul moyen d'accroître le niveau d'effort des membres de l'équipe consiste à mettre en place un système de contrôle mutuel (Lin, 1990). L'analyse de Dong et Dow (1993) se distingue du modèle de Kandell et Lazear (1992) par la nature de la sanction imposée à ceux qui ne coopèrent pas. Contrairement à Kandell et Lazear, la sanction n'est pas de nature morale mais de nature monétaire. En cas de détection d'une déviation, la sanction est appliquée automatiquement et se traduit par le prélèvement d'une partie de l'output.

Dong et Dow modélisent le contrôle mutuel dans une équipe de production. Ils considèrent une équipe de production composée de n travailleurs identiques et neutres vis à vis du risque. Les agents doivent prendre deux décisions successives. Ils décident tout d'abord de répartir leur temps entre activité de production et activité de contrôle. Ainsi chaque travailleur utilise une partie $\alpha \in [0,1]$ de son temps quotidien à des activités de production et $(1-\alpha)$ à des activités de contrôle. Les agents décident ensuite de leur effort de production. Soit $e_i \in [0,1]$, l'effort de production de l'agent i . Chaque agent ne peut être contrôlé que par un seul autre membre de l'équipe. On suppose qu'un travailleur qui fournit un effort positif ($e_i = 1$) n'est pas puni. Au contraire, si un travailleur i tire au flanc ($e_i = 0$), alors la probabilité de détection de son effort dépend de l'effort du travailleur j qui l'évalue. Si le travailleur j tire également au flanc ($e_j = 0$) alors j ne peut alors pas apporter de preuve contre i . Le travailleur i ne sera donc pas puni. Au contraire, si j travaille ($e_j = 1$), la probabilité $\alpha(e_j)$ que i soit puni dépend du temps passé au contrôle et donc inversement du temps passé à la production ($\alpha'(e_j) < 0$). Par ailleurs, personne n'est puni si le temps de contrôle est nul ($\alpha(1) = 0$). La fonction d'utilité de chaque travailleur i s'écrit de la façon suivante :

$v_i > 0$ traduit la désutilité de l'agent i à l'effort de production. y_i est la part du revenu de l'équipe que reçoit chaque agent i . Elle s'écrit sous la forme suivante :

Le premier terme de l'identité (2) représente la part $\alpha \in [0,1]$ du revenu net, distribuée à l'agent i selon ses performances de travail tandis que le deuxième terme représente la part $1-\alpha$ du revenu distribuée à l'agent i , indépendamment de ses performances et qui satisfait des considérations sociales. $D \geq 0$ représente les coûts associés à l'utilisation d'inputs autres que le travail ainsi que les coûts fixes. y et d sont définis respectivement par : $y = Y/n$ et $d = D/n$. Enfin Y représente le revenu global brut de l'équipe défini de la façon suivante :

Y est fonction des inputs (x), autres que le facteur travail, et des efforts de production des n agents

Le jeu se déroule de la façon suivante : dans un premier temps, chaque travailleur décide de la répartition de son temps entre activité de contrôle et activité de production ; dans un deuxième temps, les travailleurs décident de leur effort de production. Résolvons le jeu à rebours. Pour cela, considérons en dernière étape du jeu, les décisions des agents concernant leur effort de production. On montre qu'à l'équilibre, personne n'est incité à dévier de la stratégie qui consiste à fournir l'effort de production le plus élevé ($e_i = 1$). Pour cela on vérifie qu'aucun travailleur n'est incité à dévier de la stratégie d'équilibre.

Encadré I.2 Définition de l'équilibre de Nash

Dans un jeu $G = (S_1, \dots, S_n; u_1, \dots, u_n)$, de forme normale avec n joueurs, les stratégies (s_1^*, \dots, s_n^*) forment un équilibre de Nash, si pour chaque joueur i , s_i^* est la meilleure réponse du joueur i aux stratégies $(s_1^*, \dots, s_{i-1}^*, s_{i+1}^*, \dots, s_n^*)$ spécifiées des $n-1$ autres joueurs:
pour chaque stratégie possible s_i dans S_i ; c'est à dire que s_i^* soit solution de :

y^n est le revenu par agent quand les n membres coopèrent. Si i dévie de l'équilibre, et choisit $e_i = 0$, alors le fait que $e_j = 1$ pour tout $j \neq i$, implique que i est contrôlé par un agent qui ne tire pas au flanc et sera donc sanctionné. D'où :

y^{n-1} est le revenu par agent quand $(n-1)$ membres travaillent. L'agent i ne subit pas la désutilité de l'effort de production mais il est pénalisé par ses pairs et ne reçoit pas la totalité de son revenu. Coopérer à la production est un équilibre de Nash si :

Ce qui revient après calcul à la condition suivante :

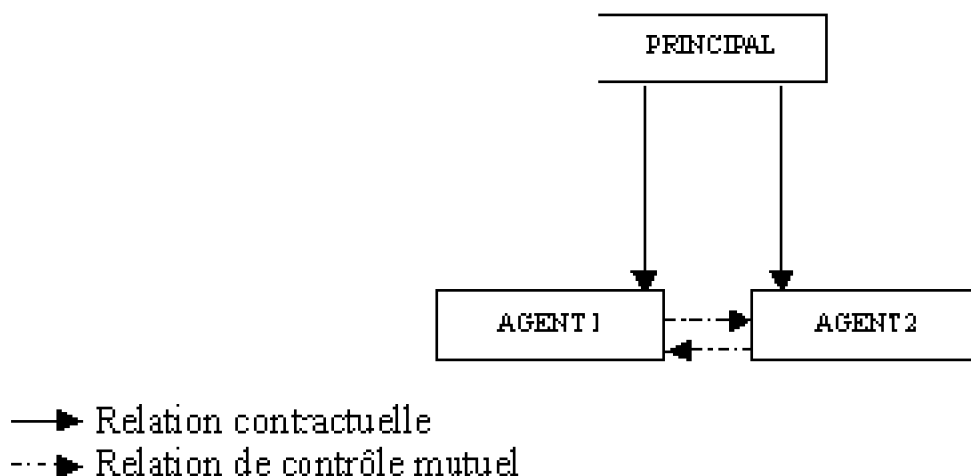
$B(i)$ représente le gain total pour l'agent i à fournir un effort positif. Le terme $(y^n - y^{n-1})$, traduit l'augmentation du revenu par agent résultant de l'effort de l'agent i . Le deuxième terme représente l'incitation pour l'agent i à fournir plus d'effort afin d'éviter la pénalité en cas de déviation. Donc si la condition (7) est vérifiée, alors chaque agent est incité à fournir l'effort le plus élevé possible.

En première étape du jeu, l'équipe décide de l'allocation du temps entre contrôle et production sous la contrainte d'effort de production issue de la condition (7). L'allocation du temps entre effort de contrôle et de production se résume au choix de α qui maximise l'utilité des agents sous la contrainte du choix d'effort de production optimal. Soit :

L'analyse de Dong et Dow (1993) apporte un éclairage intéressant sur l'étude de la pression des pairs car elle met en exergue l'émergence de la pression des pairs comme le résultat d'un arbitrage entre le temps alloué aux activités de production et aux activités de contrôle mutuel. Dong et Dow estiment ainsi qu'environ 10 à 20% du temps total de travail d'un membre de l'équipe est utilisé à des fins de contrôle mutuel et non à des fins directement productives.

2.3. Le contrôle mutuel dans une relation d'agence : Barron et Gjerde (1997)

Barron et Gjerde (1997) s'inspirent du modèle de Kandel et Lazear (1992) mais privilégient le cadre d'une relation d'agence. Les auteurs considèrent la structure composée d'un principal et de deux agents neutres vis à vis du risque⁶ suivante :



Graphique 1.1 : Contrôle mutuel dans une relation d'agence

Les agents fournissent un effort de production (respectivement e_1 et e_2) afin de produire un output noté $f(e_1, e_2)$ tel que $f_1 > 0$, $f_2 > 0$, $f_{11} < 0$ et $f_{22} < 0$. Le principal spécifie la rémunération des agents de la façon suivante:

où α_i représente une fraction de l'output total distribuée à chaque agent et α_0 une somme forfaitaire globale versée au principal par les agents. Le profit du principal est alors déterminé par :

En plus de leur effort de production, les agents exercent un effort de contrôle de leurs pairs. La désutilité pour chaque agent i ($i=1,2$) à fournir un effort de production et un effort de contrôle est notée $c(e_i, a_i)$ telle que $c_1 > 0$, $c_{11} > 0$, $c_2 > 0$, $c_{22} > 0$ et enfin $c_{12} > 0$.

Barron et Gjerde définissent l'environnement de la pression des pairs comme l'ensemble des éléments suivants : la norme de production t établie au sein du groupe par les agents eux-mêmes ; les efforts de contrôle mutuel des agents (respectivement a_1 et a_2) afin de s'assurer de l'effort de production de ses pairs ; enfin les sanctions S imposées à ceux dont l'effort de production observé est inférieur à la norme de production. Afin de simplifier l'analyse de la pression des pairs, les auteurs considèrent comme données la norme d'effort t ainsi que les sanctions S qui sont imposées automatiquement en cas d'observation d'une déviation, c'est à dire lorsque l'effort de production descend en dessous de t . La pression des pairs se résume donc au choix d'effort de contrôle des agents. Chaque agent j ($j=1,2$) doit décider de son effort de contrôle afin d'obtenir un signal sur l'effort de production de l'agent i . On note z_j le signal reçu par l'agent j concernant l'effort de production de l'agent i tel que $z_j = \alpha_i(e_i) + \alpha_{ji}$. α_{ji} est une variable aléatoire de moyenne nulle et de variance σ_{ji}^2 . $\alpha_i(e_i)$ est fonction croissante de l'effort de production de l'agent i , $\alpha_i' > 0$ et $\alpha_i'' < 0$. La précision du signal reçu par l'agent j , $1/\sigma_{ji}^2$, s'accroît lorsque a_{ji} augmente. Une sanction est alors automatiquement imposée si $z_j < t$.

⁶ Les auteurs relâchent par la suite l'hypothèse de neutralité vis-à-vis du risque et montrent que les implications en terme d'efficacité de la pression des pairs ne sont pas fondamentalement changées.

et l'agent i peut alors s'attendre à subir la pression des pairs P donnée par la fonction suivante :

Il existe un gain pour l'agent i à accroître son effort de production afin de réduire la pression des pairs qui dépend positivement de l'effort de contrôle de l'agent j ($P_{12} < 0$). Soit :

Soit la fonction d'utilité de chaque agent, séparable en output, en coût d'effort et en pression des pairs :

Le jeu séquentiel proposé par Barron et Gjerde se déroule en 3 étapes. Dans une première étape, le principal choisit le schéma de rémunération, ce qui se résume par le choix de la part α du profit attribuée aux agents ; dans une deuxième étape, les agents choisissent simultanément leur effort de contrôle (respectivement a_1 et a_2). Dans une troisième étape, les agents choisissent leur niveau d'effort de production (e_1 et e_2) en considérant comme donnés le schéma de rémunération et l'environnement de pression des pairs. A la fin de cette étape, les agents sont rémunérés en fonction du paramètre α choisi, subissent les coûts inhérents aux choix de leur effort de production et de contrôle et ressentent la pression des pairs. Les auteurs procèdent par induction à rebours afin de résoudre ce jeu :

En troisième étape, les agents décident de leur effort de production qui maximise leur utilité. Soit le programme de maximisation :

Soit la condition de premier ordre définissant le choix d'effort optimal de chaque agent i :

A partir de la condition de premier ordre (6), on détermine les efforts de production optimaux de chaque agent, respectivement $e_1 = g(a_1, a_2, \alpha)$ et $e_2 = h(a_1, a_2, \alpha)$, équilibre de Nash tels que :

g_1, h_2, g_2 et h_1 sont déterminés par la façon présentée dans les encadrés 3 et 4.

Encadré I. 3 : statique comparative : cas avec plusieurs variables

Supposons qu'on cherche à déterminer de quelle manière les fonctions de choix optimaux réagissent à des variations du paramètre a . On sait que les choix optimaux doivent satisfaire les conditions du premier ordre :
En différenciant ces deux expressions par rapport à a on obtient :
Il est plus commode d'écrire ces conditions sous la forme matricielle
Si la matrice de gauche est inversible, on peut résoudre ce système d'équations pour obtenir :
Plutôt que d'inverser la matrice, on peut également résoudre ce système par la règle de Cramer afin de trouver $\partial x_1 / \partial a$ et $\partial x_2 / \partial a$. Ainsi pour trouver $\partial x_1 / \partial a$ par exemple en appliquant la règle de Cramer, on exprime cette dérivée sous la forme du ratio de deux déterminants :
En vertu de la condition du second ordre d'un problème de maximisation, la matrice au dénominateur est semi-finie négative. Son déterminant est négatif. En conséquence, le signe de $\partial x_1 / \partial a$ est simplement le négatif du signe du déterminant du numérateur.

Encadré I.4 :statique comparative : cas avec plusieurs variables et plusieurs agents

Supposons qu'on cherche à déterminer de quelle manière les fonctions de choix optimaux réagissent à des variations du paramètre a . On sait que les choix optimaux doivent satisfaire les conditions du premier ordre :
En différenciant ces deux expressions par rapport à a on obtient :
En appliquant la règle de Cramer, on obtient le système suivant:
le signe du dénominateur est déterminé par la condition de stabilité donnée par :

Il existe une ambiguïté sur le signe de g_1 et de g_2 . Cette ambiguïté dépend notamment du signe de f''_{12} . Lorsque $f''_{12} < 0$, alors l'ambiguïté est levée : $g_1 < 0$ et $g_2 > 0$. Lorsque les efforts de production sont complémentaires ($f''_{12} > 0$), alors $g_1 < 0$ et $g_2 > 0$ pour des valeurs de c_{12} suffisamment faibles.

Dans la deuxième étape du jeu, les agents décident de leur effort de contrôle. Les auteurs considèrent deux possibilités selon que le principal spécifie ou non le niveau d'effort de contrôle mutuel. Etudions tout d'abord le cas où la pression des pairs fait l'objet d'un contrat entre le principal et les agents. Le principal spécifie alors les efforts de contrôle de chaque agent (a_1 et a_2) ainsi que le schéma de répartition de l'output entre le

en vertu de la loi du droit d'auteur.

principal et les agents (le choix de α). Le problème du principal s'écrit de la façon suivante :

La première contrainte traduit les conditions d'équilibre relatives au choix des efforts de production des agents. La deuxième contrainte est la contrainte de participation. Enfin la dernière contrainte impose un niveau d'effort de contrôle positif ou nul. Soient les conditions de premier ordre respectivement par rapport à α , a_1 et a_2 :

Lorsque la pression des pairs ne fait pas l'objet d'un contrat, chaque agent décide de son effort de contrôle qui maximise son utilité tel que :

La première contrainte traduit les conditions d'équilibre de Nash relatives au choix des efforts de production des agents. La deuxième contrainte impose un niveau d'effort de contrôle positif ou nul. Les niveaux de contrôle optimaux de chaque agent i , $i=1,2$ sont donnés par les conditions de premier ordre suivantes :

En première étape du jeu, le principal choisit le paramètre α de partage des profits. Le programme de maximisation sous contraintes du principal s'écrit de la façon suivante:

sous les contraintes :

Les deux contraintes définissent les conditions d'équilibre de Nash relatives aux choix d'effort de contrôle des deux agents, préalables à leur choix d'effort de production. Soit la condition de premier ordre par rapport à α :

L'apport majeur du modèle de Barron et Gjede est de montrer que le principal peut bénéficier des effets incitatifs de la pression des pairs lorsque celle-ci ne fait pas l'objet d'un contrat mais également qu'il peut l'influencer à travers le choix d'un schéma de rémunération optimal.

La présentation des modèles de pression des pairs a mis en évidence les résultats suivants. Il apparaît que la pression des pairs incite les agents à fournir davantage d'effort de production. L'exercice de la pression des pairs repose alors sur un arbitrage entre effort de production et effort de contrôle mutuel. Enfin, il apparaît que le principal peut influencer les décisions de contrôle mutuel des agents par l'intermédiaire d'un schéma de rémunération approprié. Ces modèles de pression des pairs suscitent toutefois un certain nombre d'interrogations. Ils conduisent notamment à s'interroger sur l'existence de limites à l'efficacité de la pression des pairs. En effet, alors que le contrôle mutuel incite les agents à fournir plus d'effort de production afin de ne pas subir la pression de leurs pairs, ne conduit-il pas ceux qui l'exercent à réduire leur propre activité productive ? On peut également s'interroger sur les conditions de l'exercice de la pression des pairs. En effet, les modèles de pression des pairs présentés dans cette section sont avant tout des modèles de contrôle mutuel dans la mesure où les décisions des agents portent uniquement sur l'effort de contrôle et que les sanctions sont appliquées automatiquement en cas de déviation. Or dans la réalité, les agents ne se contentent pas d'observer leurs pairs mais décident également des sanctions à imposer à ceux qui ne respectent pas les normes du groupe (Miller, 1992). Dans le cas où les décisions des agents portent sur les sanctions à imposer à d'éventuels passagers clandestins, les agents sont-ils toujours incités à discipliner leurs pairs? Les sections suivantes visent à répondre à ces interrogations.

3. La motivation du passage clandestin dans un partenariat

Cette section vise à étudier une relation d'équipe sans l'opportunité pour les agents d'exercer une pression sur leurs pairs. Ce modèle sert de modèle de référence afin d'étudier, par la suite, l'efficacité de la pression des pairs. On montre dans cette section que lorsque les agents n'ont pas l'opportunité de se discipliner mutuellement, l'organisation du travail en équipe avec partage des profits conduit inéluctablement à une dilution des incitations. En effet, lorsque le travail est organisé en équipe de production et que les profits sont partagés à égalité entre les membres de l'équipe, ceux-ci peuvent être incités à tirer avantage de leur appartenance au groupe sans toutefois participer systématiquement à la génération du profit (Alchian et Demsetz 1972 ; Holmström, 1982). Chaque agent est tenté de se comporter comme un resquilleur (passager clandestin) en comptant sur les efforts de ses collègues puisqu'il bénéficie à égalité du résultat de l'équipe.

3.1. Hypothèses et fonctions de gain

3.2. Travail en équipe et problème de passage clandestin

Les agents doivent décider simultanément de leur effort de production. Ils choisissent l'effort de production qui maximise son utilité tel que :

Un problème de maximisation doit satisfaire les conditions de premier ordre et de second ordre. Soit les conditions de premier ordre par rapport à l'effort de production des agents 1 et 2 (respectivement e_1 et e_2):

Les premiers termes respectifs des conditions de premier ordre (3a) et (3b) représentent l'output marginal résultant d'un accroissement de l'effort de production s'accroît. Le deuxième terme représente la désutilité marginale d'accroître l'effort de production. Soit e_1^* et e_2^* les niveaux d'effort de production solutions des conditions de premier ordre (3a) et (3b). Les conditions du second ordre sont respectivement :

PROPOSITION 1 : *sans opportunité pour les agents de discipliner leurs pairs, chaque agent adopte un comportement de passager clandestin alors que le gain total de l'équipe serait maximal si tous les agents coopéraient.*

Preuve de la proposition 1:

L'efficacité au sens de Pareto suppose que le surplus total soit maximisé tel que :

Avec la condition de premier ordre :

Notons e_i^0 le niveau d'effort solution de l'équation (6). Sous l'hypothèse de convexité de la fonction de coût, on remarque que $e_i^1 < e_i^0$. La production totale du

groupe est inférieure au niveau Pareto optimal, en raison de l'existence du problème de passager clandestin qui apparaît dans le terme $1/N$ des conditions (3a) et (3b).

Donc, sans opportunité d'exercer un contrôle mutuel, personne ne coopère à la production de l'output puisque tous les agents adoptent un comportement de passager clandestin. Dans les sections suivantes, on s'interroge sur les effets de l'introduction de la pression des pairs sur le comportement des membres de l'équipe.

4. L'exercice de la pression des pairs

Cette section vise à étudier les conditions d'efficacité et d'existence de la pression des pairs. Dans un premier temps, on considère que la pression des pairs se résume aux choix des efforts de contrôle mutuel. On suppose alors que les sanctions sont appliquées automatiquement en cas de détection d'un comportement opportuniste. On s'interroge alors sur l'efficacité du contrôle mutuel à inciter les agents à coopérer. Si l'opportunité de pouvoir se contrôler mutuellement incite les agents à coopérer davantage afin d'éviter de subir la pression des pairs, l'exercice du contrôle mutuel conduit également les agents à réduire leur propre effort de production lorsque contrôle et production sont substituables. On relâche dans un deuxième temps l'hypothèse selon laquelle les sanctions sont automatiquement appliquées. On s'interroge alors sur les incitations des agents à exercer une pression sur leurs pairs lorsque sanctionner est coûteux.

4.1. L'environnement de la pression des pairs

- Les normes de groupe

Les normes sont des règles établies par le groupe afin de pouvoir évaluer le comportement individuel de chacun de ces membres (Elster, 1989). Elles définissent les règles de comportement que les individus croient nécessaires d'adopter. Ainsi, les groupes de travail déterminent la production individuelle des agents en référence à une norme qui représente la conception que se fait le groupe de la quantité de travail qu'il est juste de produire. La norme contribue à apporter au groupe des avantages en réduisant les incitations des membres à adopter un comportement de passager clandestin.

- Les efforts de contrôle mutuel

Les efforts de contrôle mutuel sont exercés par les membres de l'équipe afin d'obtenir un signal concernant l'effort de production de leurs pairs. L'observabilité des efforts des membres du groupe est au cœur de l'environnement de la pression des pairs (Varian, 1990). Du fait de leur proximité, les agents ont une position privilégiée pour obtenir des informations concernant les activités de leurs pairs (Bowles, Carpenter et Gintis, 2001). L'organisation du travail en équipe favorise cette proximité entre les agents et facilite donc le contrôle mutuel.

· Les sanctions des pairs

Les sanctions sont imposées par les pairs lorsqu'ils observent une déviation vis à vis de la norme. Les sanctions sont généralement imposées à ceux qui produisent moins que la norme. Toutefois, on peut également observer des comportements de sanction à priori " aberrants " du point de vue de l'économiste qui consistent à sanctionner ceux qui dépassent la norme de production, à la fois pour maintenir la cohésion dans le groupe et pour éviter la révision des normes officielles si elles étaient systématiquement dépassées par les travailleurs (Miller, 1992). Dans ce cas, les sanctions assurent la survie du groupe (Roethlisberger et Dickson, 1947 ; Whyte, 1955 ; Marsden, 1986). Les sanctions imposées par les pairs sont de natures diverses. Posner et Rasmusen (1999) distinguent ainsi plusieurs sortes de sanctions. Elles peuvent être de nature morale liées à la désapprobation des pairs (honte), reposer sur l'ostracisme qui consiste à exclure du groupe un membre déviant ou être de nature monétaire et consister à ce que l'agent déviant ne reçoive pas la totalité de la part de l'output qui lui est due. C'est cette dernière forme de sanction qui sera retenue par la suite dans la modélisation de la pression des pairs.

4.2. La pression des pairs par le contrôle mutuel

Lorsqu'un agent i exerce un effort de contrôle, il subit une désutilité qui s'ajoute à la désutilité de fournir un effort de production, la désutilité totale étant notée $C(e_i, a_{ij})$. L'utilité de chaque agent i ($i=1,2$) est linéaire en output, séparable en coût d'effort, en output et en coût de pression. La fonction d'utilité de chaque agent i s'écrit de la façon suivante :

Le premier terme de la fonction d'utilité représente la part de l'output que reçoit chaque agent. Le deuxième terme est la désutilité que subit l'agent i lorsqu'il fournit un effort de production e_i et un effort de contrôle a_{ij} . La fonction de coût est une fonction croissante ($C'_{e_i} > 0, C'_{a_{ij}} > 0$) et convexe ($C''_{e_i e_i} > 0, C''_{a_{ij} a_{ij}} > 0$). On suppose par ailleurs que les efforts de production et de contrôle sont substituables de sorte que l'augmentation de l'effort de contrôle d'un agent accroît le coût marginal de son effort de production ($C''_{e_i a_{ij}} > 0$). Enfin, le troisième terme représente la pression exercée sur l'agent i par l'agent j . La pression des pairs $P(e_i, a_{ij})$ dépend de l'effort de production de l'agent i et du choix d'effort de contrôle de l'agent j . Elle diminue avec l'effort de production e_i ($P'_{e_i} < 0, P''_{e_i e_i} > 0$) et augmente avec l'effort de contrôle de l'agent j , a_{ij} ($P'_{a_{ij}} > 0$). Il existe par ailleurs un gain pour l'agent i à augmenter son effort de production afin de réduire la pression des pairs ($P''_{e_i a_{ij}} < 0$).

A l'instar de Barron et Gjerde (1997), on considère l'approche séquentielle suivante de la pression des pairs⁷ : dans une première étape, les agents décident de leur effort de

⁷ Comme le soulignent à juste titre Barron et Gjerde, cette interprétation est très réaliste puisqu'il est légitime de supposer que les agents décident de leur effort de production en considérant comme donné l'environnement de pression des pairs. Cette interprétation séquentielle de la pression des pairs conduit toutefois à supposer que les agents vont effectivement tenir leur engagement en deuxième étape sur les efforts de contrôle annoncés en première étape (Barron et Gjerde, 1997).

contrôle (respectivement a_{12} et a_{21}) ; dans une deuxième étape, ils choisissent simultanément leur effort de production (e_1 et e_2) en connaissant l'effort de contrôle de l'autre agent. A la fin de cette seconde étape, les agents subissent les coûts de leurs efforts de production et de contrôle, l'output est partagé de façon égale entre eux et chaque agent subit une pression des pairs qui dépend de son propre effort de production et de l'effort de contrôle choisi en première étape par ses pairs. Le graphique 2 décrit le déroulement du jeu.

t	t+1
<ul style="list-style-type: none"> Les agents choisissent leur effort de contrôle 	<ul style="list-style-type: none"> Les agents décident de leur effort de production Le résultat est partagé Les sanctions sont appliquées

Graphique 1.2. Les étapes du modèle de contrôle mutuel séquentiel

Le jeu est résolu par induction à rebours . En dernière étape, les agents décident simultanément de leur niveau d'effort de production.

Encadré 1.5 :Equilibre de Cournot

Soient deux agents 1 et 2 qui décident en simultané de leur effort de production respectivement e_1 et e_2 . Un équilibre de Nash (de stratégie pure) est alors un ensemble d'efforts (e_1^*, e_2^*) tel que chaque agent choisit l'effort qui maximise son utilité compte tenu de ses conjectures sur le choix de l'autre agent, et tel que les suppositions de chaque agent sur le choix de l'autre sont effectivement correctes.
En supposant un optimum intérieur pour chaque agent, cela signifie qu'un équilibre de Nash Cournot doit satisfaire les deux conditions du premier ordre suivantes :
La condition du premier ordre pour l'agent 1 exprime l'effort de production optimal de celui ci en fonction de ses anticipations sur le choix de production de l'agent 2 ; c'est la fonction de réaction de l'agent 1 ; elle montre les réactions de l'agent 1 compte tenu de ses différentes conjectures sur le choix possible de l'agent 2. En supposant une régularité suffisante, la fonction de réaction de l'agent 1 $f_1(e_2)$ est implicitement définie par l'identité
Pour déterminer dans quelle mesure l'agent 1 modifie son choix optimal de production lorsque ses conjectures sur la production de l'agent 2 changent, on différencie cette identité et on résoud pour $f'_1(e_2)$:

Encadré 1.6 :Equilibre de Stackelberg

Soit un modèle à deux temps dans lequel l'agent 1 agit en premier. L'agent 2 peut alors observer le choix de l'autre agent (e_1) et choisir son propre niveau optimal d'effort (e_2). Le jeu est donc joué en séquentiel. Résolvons ce jeu à rebours.

En deuxième étape, le problème de l'agent 2 est alors de déterminer son effort afin de maximiser son utilité, compte tenu du choix de l'agent 1. La condition de premier ordre de ce problème est simplement :

En première étape, l'agent 1 choisit son effort, anticipant et sachant comment réagira l'agent 2. Ainsi l'agent 1 souhaite maximiser son utilité. Soit la condition de premier ordre :

Chaque agent i choisit l'effort de production qui maximise son utilité :

Les conditions de premier ordre suivantes définissent les niveaux d'effort de production optimal choisis respectivement par les agents 1 et 2, étant donnée la détermination préalable des efforts de contrôle:

La condition de premier ordre pour chaque agent i exprime l'effort de production optimal de cet agent en fonction des choix d'effort de contrôle des deux agents en première étape : a_{ij} et a_{ji} . Les conditions de second ordre s'écrivent :

Puisque les agents sont identiques, un équilibre symétrique suppose que $l'_{a12} = Q'_{a21}$ et $l'_{a12} = Q'_{a21}$. Par ailleurs, $f''_{eiej} = f''_{eiej}$, $C''_{eiej} \square C''_{eiej}$ (e_j, a_j) = C''_{eiej} (e_j, a_j) et $P''_{eiej} \square P''_{eiej}$ (e_j, a_j) = P''_{eiej} (e_j, a_j). Le premier terme du numérateur dans les expressions (13a) et (14a), traduit le gain pour l'agent i à accroître son effort de contrôle. En effet, l'agent i est incité à accroître son effort de contrôle afin que l'agent j augmente son effort de production. Les agents sont alors conduits à accroître leur effort de production afin de réduire la pression des pairs ($P''_{eiaji} < 0$). Le premier terme du numérateur dans les expressions (15a) et (16a) traduit le coût pour l'agent i à augmenter son effort de contrôle. En effet, lorsque l'agent i augmente son effort de contrôle, il est incité à réduire son propre effort de production puisque l'augmentation de l'effort de contrôle accroît le coût marginal de l'effort de production ($C''_{eiaji} > 0$). Les signes de l'_{a12} , l'_{a21} , Q'_{a12} et Q'_{a21} dépendent également du deuxième terme du numérateur. Le signe du deuxième terme du numérateur dépend, quant à lui, du signe de f''_{eiej} . On distingue trois cas selon que les efforts de production des agents sont substituables, indépendants ou complémentaires : $f''_{eiej} = 0$, $f''_{e1e2} > 0$ ou $f''_{e1e2} < 0$.

Si $f''_{eiej} = 0$ on peut alors réécrire les équations précédentes de la façon suivante ⁸ :

Nous savons que le dénominateur des expressions (13b) à (16b) est négatif en vertu de la condition de second ordre. En conséquence, le signe de ces expressions ne dépend que du signe du numérateur. Puisque d'après les hypothèses du modèle, $P''_{eiaji} < 0$,

⁸ On retrouve ici la solution théorique du modèle de Kandel et Lazear (1992) (13b) et (15b). Barron et Gjerde (1997) p251-253 font remarquer à juste titre que l'hypothèse implicite dans le modèle de Kandel et Lazear (1992) selon laquelle $f''_{eiej} = 0$ est incompatible avec l'hypothèse de non séparabilité de l'output en effort.

l'ambiguïté quant au signe de l'_{a21} et de Q'_{a12} est levée. De même pour l'_{a12} et de Q'_{a21} puisque $C''_{e1sij} > 0$. $l'_{a21} \square 0$ et $Q'_{a12} \square 0$ traduisent le gain pour l'agent i à accroître son effort de contrôle afin d'inciter l'agent j à augmenter son effort de production. $l'_{a12} \square 0$ et $Q'_{a21} \square 0$ représentent le coût pour l'agent i d'augmenter son effort de contrôle dans la mesure où cela l'incite à réduire son propre effort de production. Ce coût disparaît si l'on relâche l'hypothèse de substituabilité des efforts de production et de contrôle. En effet, si l'on suppose que $C''_{e1aij} = 0$, alors $l'_{a12} = 0$ et $Q'_{a21} = 0$.

On peut toutefois s'interroger sur la pertinence de l'hypothèse selon laquelle $f''_{e1ej} = 0$. En effet, considérer que $f''_{e1ej} = 0$, revient à supposer que l'output est séparable en effort de production. Or si l'output $f(e_1, e_2)$ est séparable en effort de production, alors les problèmes d'incitation et de contrôle disparaissent dans la mesure où l'on peut rémunérer les agents en fonction de leur effort de production. L'organisation du travail en équipe de production avec partage à égalité de l'output n'est en effet justifiée que sous l'hypothèse de non séparabilité de l'output en effort.

Enfin, le dernier cas envisagé est celui où $f''_{e1e2} < 0$, c'est à dire le cas où les efforts de production des deux agents sont substituables. Dans ce cas, il n'existe alors plus d'ambiguïté sur les signes de l'_{a21} , Q'_{a12} , l'_{a12} et de Q'_{a21} : $l'_{a21} \square 0$, $Q'_{a12} \square 0$, $l'_{a12} \square 0$ et $Q'_{a21} \square 0$.

Le degré d'efficacité du contrôle mutuel, c'est à dire son effet sur l'effort de production des agents dépend de l'écart relatif entre les gains et les coûts de l'exercice du contrôle mutuel⁹. Le gain à contrôler un pair se traduit par l'augmentation de l'effort de production de l'agent contrôlé. En effet, un agent est incité à accroître son effort afin de ne pas faire l'objet de la pression de ses pairs. Le coût à contrôler un pair résulte quant à lui de la réduction du propre effort de production de l'agent qui contrôle. Etudier le degré d'efficacité du contrôle mutuel revient donc à comparer les gains et les coûts, ce qui se résume à l'étude du signe de $l'_{aji} + l'_{aij}$:

Tableau:1.1 Degré d'efficacité du contrôle mutuel

Efficacité du Contrôle mutuel	Contrôle mutuel efficace	Contrôle mutuel inefficace	Contrôle mutuel coûteux
Signe de $l'_{aji} - l'_{aij}$			

Preuve de la proposition 2 :

Afin de comparer le niveau d'effort de production choisi par les agents avec et sans contrôle mutuel, on compare les conditions de premier ordre par rapport à l'effort de production dans les modèles avec et sans contrôle mutuel :

En notant e_j^1 le niveau d'effort de production dans le modèle sans contrôle mutuel, et e_j le niveau d'effort de production dans le modèle avec opportunité de contrôle mutuel, alors :

Afin de vérifier que $e^* > e_j^1$, supposons l'inverse, c'est à dire $e^* \square e_j^1$. Comme $C(e_j)'' > 0$, on peut écrire :

⁹ On suppose que les conditions nécessaires sont réunies de sorte que $l'_{a21} \square 0$, $Q'_{a12} \square 0$, $l'_{a12} \square 0$ et $Q'_{a21} \square 0$.

Or, comme $P'_{ei^*} < 0$, alors $f'(e^*) \square f'(e^1)$, ce qui viole l'hypothèse de concavité de la fonction $f(\cdot)$. Donc $e_{ei^*} > e^1$. Le niveau d'effort de production avec pression des pairs, solution de l'équation (19) est supérieur au niveau d'effort de production sans contrôle mutuel, solution de l'équation (18). Le contrôle mutuel incite donc les agents à fournir davantage d'effort de production afin d'éviter les conséquences néfastes de la pression des pairs.

En première étape du jeu, les agents décident de leur effort de contrôle. Chaque agent i choisit l'effort de contrôle qui maximise son utilité tel que :

sous les contraintes :

La première contrainte traduit les conditions d'équilibre de Nash relatives au choix des efforts de production des agents. La deuxième contrainte impose un niveau d'effort de contrôle positif ou nul. On substitue dans l'expression (22) les niveaux d'effort de production optimaux $e_1^* = I(a_{21}, a_{12})$ et $e_2^* = Q(a_{12}, a_{21})$, solutions de la première contrainte du programme de maximisation. En supposant une solution symétrique, on obtient les conditions de premier ordre par rapport aux efforts de contrôle respectivement des agents 1 et 2 :

\square_i est le multiplicateur associé à la deuxième contrainte du programme de maximisation (effort de contrôle non négatif).

PROPOSITION 3 : *il existe un gain stratégique à exercer un effort de contrôle positif sur ses pairs. En effet, davantage d'effort de contrôle incite les pairs à accroître leur effort de production, ce qui se traduit par davantage d'output.*

Preuve de la proposition 3 :

L'analyse précédente a mis en évidence que sous certaines conditions, la pression des pairs incitait les agents à fournir davantage d'effort de production. Dans la sous section suivante on s'interroge sur les implications du relâchement de l'hypothèse sur les sanctions.

4.3. La pression des pairs par les sanctions

On a supposé jusqu'à présent que les sanctions étaient imposées automatiquement en cas de détection d'une déviation négative de la norme. On relâche ici cette hypothèse en considérant que les agents décident des sanctions à imposer à d'éventuels passagers clandestins. En effet, les agents ne se contentent généralement pas d'observer les efforts de leurs pairs mais décident également des sanctions à infliger aux comportements de *passagers clandestins* observés (Miller, 1992). Dès lors que les sanctions font l'objet d'une décision particulière de la part des agents, alors il convient d'associer un coût à la décision de sanctionner un pairs (Posner et Rasmusen, 1999). En effet, toute décision nécessite du temps et un effort généralement coûteux pour celui qui l'exerce¹⁰.

Considérons le jeu suivant : dans un premier temps, les agents décident simultanément de leur effort de production. Après avoir observé parfaitement et sans coût l'effort de production de l'autre agent, chaque agent décide de l'effort afin de sanctionner un éventuel passager clandestin. Le graphique 3 décrit le déroulement du jeu.

i	$i+1$
• Les agents choisissent leur effort de production	• Les agents décident des sanctions à appliquer

Graphique 3. Les étapes du modèle de pression des pairs avec sanctions coûteuses

PROPOSITION 4: *La pression des pairs est inexistante lorsque les décisions des agents portent sur les sanctions à imposer. Par conséquent, les agents adoptent un comportement opportuniste.*

Afin de vérifier que la combinaison des stratégies qui consiste à ne pas coopérer à la production de l'output et à ne jamais sanctionner forme un équilibre parfait en sous jeu, procédons par induction à rebours. En $t+1$, les agents décident de la sanction qu'ils vont imposer à un agent déviant. Chaque agent i choisit le niveau de sanction qui maximise son utilité tel que :

Les conditions de premier ordre définissent les niveaux de sanction choisis par les agents 1 et 2. Notons p_{12}^{**} et p_{21}^{**} , les niveaux de sanction optimaux nuls. Puisqu'il n'existe pas de gain à sanctionner et que cela est coûteux, les agents ne sont pas incités à sanctionner leurs pairs.

En première étape, les agents décident de l'effort de production. On note e_i^{**} l'effort de production choisi par l'agent i . Puisque aucun agent n'est incité à sanctionner en deuxième étape du jeu, personne ne produit en première étape. En effet, les agents anticipent qu'ils ne seront jamais sanctionnés quelque soit leur effort de production. De ce fait, les deux joueurs adoptent un comportement opportuniste de passager clandestin. L'effort de production choisi par les agents en première étape est donc identique à celui choisi dans le modèle sans pression des pairs.

Donc lorsque l'on relâche l'hypothèse selon laquelle les sanctions sont imposées automatiquement, la pression des pairs ne peut pas apparaître. Dès lors que la menace de sanction n'est pas crédible, personne n'est incité à coopérer à la production de l'output.

5. Conclusion

¹⁰ Un autre argument pour justifier du coût inhérent au fait de sanctionner consiste à mettre en évidence les conséquences directes ou indirectes de sanctionner ses pairs. Ainsi par exemple, Hirshleifer et Rasmusen (1989) ont montré que lorsque la sanction consiste à exclure un agent non coopératif du groupe, l'ostracisme pénalise l'ensemble du groupe et pas seulement ceux qui subissent les sanctions. En effet, l'exclusion implique la réduction de la taille du groupe ce qui est coûteux pour le groupe qui perd un ou plusieurs de ses membres. Il existe également un coût psychologique à sanctionner un pair. En effet, il peut être désagréable de sanctionner un individu avec qui on a des relations étroites. Il convient dès lors de distinguer les sanctions par les pairs des sanctions motivées par un esprit de vengeance. En effet, alors qu'il peut être désagréable de sanctionner les autres individus, certains individus peuvent ressentir au contraire une satisfaction personnelle à punir lorsque leur acte est motivé par un « esprit de vengeance ». On peut ainsi comparer cette réaction à la satisfaction que ressent une victime lorsque son agresseur est puni (Lindbeck, 1997).

La comparaison des modèles avec et sans pression des pairs (sous l'hypothèse que les sanctions sont automatiquement appliquées et sans coût pour ceux qui sanctionnent) a montré que le contrôle mutuel incite les agents à fournir davantage d'effort de production. Toutefois, lorsque l'on relâche l'hypothèse selon laquelle les sanctions sont appliquées automatiquement et sans coûts, alors la pression des pairs devient inexistante. Les agents anticipent alors qu'ils ne seront jamais sanctionnés par leurs pairs et adoptent donc un comportement opportuniste en choisissant l'effort de production le plus bas possible qui maximise leur utilité. Comment alors expliquer que, dans la réalité, les individus n'hésitent pas à punir ceux qui dévient des normes établies au sein du groupe même si cela est coûteux pour l'équipe dans son ensemble (Posner et Rasmusen, 1999 ; Falk, Fehr et Fischbacher, 2000) ?

Il convient de vérifier les prédictions théoriques des modèles de pression des pairs. Comment tester l'efficacité de la pression des pairs ? La pression des pairs est difficilement mesurable en raison notamment du caractère souvent informel des sanctions imposées par les pairs. C'est notamment le cas lorsque les sanctions reposent sur la désapprobation morale des pairs et relèvent alors des sentiments de honte. La méthodologie retenue pour la soumission à réfutation des propriétés d'équilibre des modèles proposés est l'économie expérimentale. L'objectif des expérimentations en économie est d'analyser et de comprendre l'éventuelle différence qui existe entre les prédictions théoriques à l'équilibre et les résultats tant expérimentaux qu'observés dans la vie quotidienne (Ledyard, 1995). En effet, l'expérimentation permet de recréer, dans un environnement contrôlé, les hypothèses exactes postulées dans les modèles et d'isoler l'influence de certaines variables

Chapitre 2. Sanctions monétaires versus sanctions non monétaires¹¹

“Les sanctions matérielles sont mieux comprises comme véhiculant l’émotion du mépris, qui est le déclencheur direct de la honte” John Elster (1998)

1. Introduction

Les problèmes d'action collective s'apparentent souvent à la structure des biens publics. L'organisation du travail en équipe et le problème posé par celle-ci peuvent facilement être transposés en économie expérimentale à un jeu de contribution à un bien public où l'output fourni par l'équipe est le bien public et l'effort de production fourni par un agent est la contribution volontaire au bien public. Afin de tester l'effet de la pression des pairs sur le comportement des agents et le niveau de coopération au sein d'une équipe de travail, un jeu de contribution à un bien public a été développé puis testé à l'aide d'un protocole expérimental.

Dans une expérience de bien public, chaque sujet doit réaliser un niveau de

11

Ce chapitre est le fruit d'une collaboration avec Charles Noussair, Steve Tucker et Marie-Claire Villeval

contribution individuel en vue d'atteindre un résultat global qui dépend de l'effort agrégé de l'ensemble des membres du groupe. Les contributions au bien public doivent satisfaire trois conditions : elles sont volontaires ; elles imposent un coût au contribuable ; enfin, elles confèrent un bénéfice à chacun des membres du groupe. Un jeu de contribution volontaire repose sur un arbitrage individuel des choix d'investissement entre deux biens, l'un privé, l'autre public. Au début du jeu, chaque sujet i , $i=1,...,N$ reçoit une dotation initiale individuelle notée $D_i = D$ qu'il doit répartir entre le bien privé et le bien public. On note c_i , le montant investi par le sujet i dans le bien public. La dotation que le sujet i n'investit pas dans le bien public ($D - c_i$) est investie dans le bien privé. La fonction de paiement du sujet i s'écrit sous la forme suivante¹² :

(1)

où a représente le rendement marginal du bien public. Cette fonction a deux composantes : une composante en bien privé et une composante en bien public, $(D - c_i)$, est le terme associé au bien privé et $g(.)$ est la fonction de paiement associée au bien public. Chaque unité allouée au bien privé a un rendement marginal constant. Le bien public a un rendement unitaire sur investissement plus faible que le bien privé. L'investissement dans le bien public rapporte à tous les membres du groupe a/N par unité investie. Ainsi toute unité investie dans le bien public rapporte le même montant à chacun des membres du groupe. La fonction de paiement du sujet i dépend non seulement de son investissement personnel dans chacun des deux biens mais aussi de l'investissement des autres membres du groupe dans le bien public. Les sujets doivent décider du montant de la dotation investie dans le bien public. Soit le programme de maximisation suivant :

En supposant que les sujets sont rationnels et cherchent à maximiser leur utilité, la prédiction théorique du jeu indique qu'aucun sujet ne contribue au bien public. Ils adoptent tous la stratégie de passager clandestin qui est la stratégie dominante. En effet, chaque sujet est individuellement incité à ne pas participer au financement d'un bien public alors que le profit agrégé serait maximisé si toutes les dotations étaient allouées au bien public. Ce problème constitue donc un dilemme social car au niveau collectif il serait plus efficient que chacun contribue au financement du bien public.

Contrairement aux prédictions théoriques, la plupart des expériences réalisées sur des jeux de contribution volontaire au financement d'un bien public montrent que les individus contribuent généralement entre 40 et 60% de leur dotation initiale au bien public dans les premières périodes du jeu. Toutefois, lorsque le jeu est répété un certain nombre de fois, le niveau des contributions converge vers la prédiction théorique (Marwell et

¹² Cette fonction de gain s'inspire de celle d'Isaac, Walker et Thomas, (1984) qui est devenue un standard dans les jeux de biens publics. Cette fonction a l'avantage d'être simple et de faire apparaître clairement l'opposition d'intérêt individuel et collectif. Les études qui ont précédé celle d'Isaac et al. (1984) proposent un cadre méthodologique différent avec une fonction de paiement caractérisée par des seuils où le retour marginal du bien public est plus élevé à chaque palier supérieur. Ainsi, Bohm (1972) s'est intéressé à une expérience où les sujets devaient financer collectivement la diffusion d'un programme de télévision. Le programme était diffusé si le montant récolté dépassait le coût de la diffusion. Les résultats obtenus suggèrent que les individus peuvent facilement être amenés à contribuer au financement d'un bien public. En 1979 et 1980, Marwell et Ames réalisent une série d'expérimentations en période unique. Les auteurs observent que le taux de contribution moyen est relativement élevé (41% environ de la dotation).

Ames, 1979 ; Isaac, McCue et Plott, 1985 ; Kim et Walker, 1985 ; Fehr et Schmidt, 1997). A partir de cette constatation, les études portant sur les biens publics ont suivi deux grandes directions. La première vise à étudier les comportements de contribution volontaire au financement du bien public (Isaac, Walker et Thomas, 1984 ; Isaac et Walker 1983, 1984, 1987, 1988). Les auteurs cherchent alors à identifier les facteurs qui influencent les comportements de contribution. Plus récemment, une deuxième vague de travaux s'est intéressée aux conséquences de l'introduction d'opportunité de sanction sur le niveau des contributions (Fehr et Gächter, 2000 ; Falk, Fehr et Fischbacher, 2000, 2001). Fehr et Gächter (2000) ont développé un jeu de bien public avec possibilité de sanctions. Ce jeu se déroule en deux étapes. La première étape est similaire à un jeu de bien public standard mais, dans une deuxième étape, les sujets sont informés des contributions individuelles des autres membres de leur groupe et ont l'opportunité de réduire les gains des autres sujets en leur attribuant des points de sanction. Toutefois, l'introduction d'une opportunité de sanction ne modifie pas l'équilibre parfait en sous jeu. Puisqu'il est coûteux de punir, aucun sujet n'est incité à sanctionner les autres agents. La menace n'étant pas crédible, les sujets devraient adopter une stratégie de *passagers clandestins* en première étape du jeu puisqu'ils anticipent qu'aucune sanction ne sera prise en deuxième étape. Contrairement aux prédictions théoriques, Fehr et Gächter observent que les agents sanctionnent les autres sujets même si cela est coûteux pour eux. La possibilité de sanctionner les autres sujets accroît considérablement la coopération au sein des groupes et permet d'approcher l'optimum de Pareto.

L'objet de ce chapitre est de tester l'effet de la pression des pairs sur le comportement des agents et le niveau de coopération au sein d'une équipe de travail à partir d'un jeu de contribution à un bien public avec opportunité de sanction. Plus précisément, on cherche à étudier les raisons pour lesquelles les individus contribuent davantage lorsqu'ils ont l'opportunité de se sanctionner mutuellement alors que, comme on l'a vu, dans le chapitre précédent, la menace de sanction n'est pas crédible lorsque sanctionner est coûteux. Afin d'expliquer pourquoi les sujets contribuent au bien public, Fehr et Gächter (2000) avancent l'idée suivante : les sujets contribuent au bien public afin d'éviter l'effet direct des sanctions monétaires qui se traduit par une réduction de leurs gains. Cependant, Fehr et Gächter ne considèrent pas l'aspect non monétaire des sanctions. En effet, les sanctions monétaires se traduisent par une réduction des gains de ceux qui les reçoivent mais véhiculent également la désapprobation des pairs.

Un certain nombre de travaux ont mis en évidence que la désapprobation sociale joue un rôle considérable dans les décisions des agents. Ainsi, Elster (1998) met en exergue le fait que les sanctions matérielles doivent être comprises comme des manifestations du mépris, qui engendrent directement le sentiment de honte. D'après l'auteur, la honte peut alors être perçue comme un coût psychologique. Ainsi lorsque la désapprobation sociale est trop forte, la seule issue possible peut être le suicide¹³. Elster

¹³ Une personne qui a été l'objet du mépris public a le choix entre trois options : le suicide, un déménagement vers un autre lieu ou simplement rester sur place en attendant que s'estompe le mépris des autres et ainsi son propre sentiment de honte. Si dans l'instant, ce sentiment peut être très élevé, le suicide sera peut être préféré à la valeur actuelle escomptée des autres options. Ce choix est d'autant plus plausible que la honte comme d'autres sentiments profonds sape notre capacité à prédire nos états mentaux futurs. Lorsqu'on éprouve de la honte, on peut difficilement envisager que cet état ne durera pas toujours (Loewenstein, 1996).

relate ainsi l'exemple d'un amiral de la Navy qui se suicida après avoir fait l'objet d'un ostracisme social. La pression morale joue également un rôle de garant de la paix sociale d'un pays. En effet, Tyler (1990) met en évidence que les individus se conforment aux lois de leurs pays non seulement pour éviter les sanctions légales mais également pour éviter une désapprobation de la société. Knack (1992) observe quant à lui que les décisions de vote sont conditionnées en grande partie par la pression morale exercée par les pairs. A partir d'une enquête post électorale (Social Sanction Survey ¹⁴), réalisée sur des ménages, l'auteur observe que ceux qui ne vont pas voter sont désapprouvés par leurs pairs.

On cherche dans ce chapitre à dissocier les raisons pour lesquelles les agents contribuent au bien public avec opportunité de sanction : une première explication repose sur l'idée selon laquelle les sujets contribuent afin d'éviter les conséquences monétaires des pénalités. Par la suite, on appellera cet effet, *effet direct des sanctions*. Une deuxième explication repose sur l'idée que le seul fait de recevoir des points de sanction, en dehors de leurs conséquences monétaires, permet également d'accroître le niveau de contribution. En effet, les points de sanction expriment la désapprobation de ceux qui les octroient. Si les sujets sont sensibles à la désapprobation de leurs pairs, ils ressentent alors de la désutilité à recevoir des points de sanction en dehors de leurs conséquences monétaires. Donc, les sujets contribueraient davantage au bien public pour éviter la désapprobation des autres membres du groupe. On appellera par la suite cet effet, *effet indirect des sanctions*.

Afin de tester l'importance respective des effets directs et indirects des sanctions, plusieurs traitements ont été réalisés. Dans un premier traitement intitulé traitement SP, les sujets participent à un jeu classique de bien public et n'ont pas l'opportunité de sanctionner leurs pairs. On observe que les niveaux de contributions des sujets convergent vers une contribution nulle au bien public avec répétition du jeu. Ce résultat est similaire aux résultats observés dans la plupart des expériences de bien public répété (Isaac, Walker et Thomas, 1984 ; Andreoni, 1988 ; Weimann, 1994 ; Croson 1996 ; Fehr et Schmidt, 1997). Dans un deuxième traitement, intitulé traitement MP, les sujets peuvent attribuer des points de sanction monétaire aux autres membres de leur groupe. Les points de sanction sont coûteux non seulement pour ceux qui les reçoivent mais également pour ceux qui les imposent. Ce traitement est similaire à celui de Fehr et Gächter (2000).

Nous obtenons des résultats similaires à ceux de Fehr et Gächter. Nous observons que l'opportunité de sanctions monétaires incite les sujets à davantage contribuer. Les effets directs et indirects des sanctions permettent d'expliquer ce niveau élevé de contribution dans le traitement MP. En effet, dans ce traitement, les sujets peuvent contribuer davantage au bien public parce qu'ils veulent éviter la désapprobation de leurs pairs mais également parce qu'ils veulent se prémunir d'une réduction de leurs gains monétaires. Afin de dissocier ces deux effets, nous avons réalisé un troisième traitement appelé traitement NP. Ce traitement est identique au précédent à une exception près : les

¹⁴ L'enquête pose entre autres les questions suivantes: « *Do you have any friends, neighbors, or relatives who would be disappointed or angry with you if they knew that you had not voted?* » ou encore « *If a friend told you that he was not going to vote, would you disapprove and let him know?* »

points de sanction n'affectent les gains ni des sujets sanctionnés ni de ceux qui sanctionnent. Dans ce traitement, seul l'effet indirect des sanctions peut influencer le niveau des contributions puisque les gains des agents sanctionnés ne sont pas modifiés. Nous observons également dans ce traitement NP, des niveaux élevés de contribution. Néanmoins, le niveau moyen des contributions est plus faible que dans le traitement MP.

Afin de tester l'effet d'une relation durable (par analogie à une relation dans une équipe de travail) sur l'efficacité de la pression des pairs, nous avons comparé les effets des sanctions non monétaires selon que les sujets sont en interaction avec les mêmes autres sujets pour l'ensemble des périodes (condition *partenaire*) ou qu'ils jouent avec des sujets différents à chaque période (condition *étranger*). L'idée est de tester l'hypothèse de Kandell et Lazear (1992) selon laquelle l'efficacité de la pression des pairs est fortement conditionnée par l'importance des liens sociaux qui nécessitent des interactions répétées entre les agents. En effet, selon Kandell et Lazear (1992), lorsque les membres du groupe ont des interactions répétées, l'empathie est plus forte que si le groupe était constitué d'individus ne se connaissant pas et toute déviation suscite des sentiments de honte d'autant plus forts. La notion de « pair » prend alors ici toute sa signification. Cette idée est déjà présente chez Adam Smith (1759) qui écrit :

« On s'attend à moins de sympathie à l'égard d'une connaissance qu'à l'égard d'un ami....On s'attend encore à moins de sympathie d'une assemblée constituée d'étrangers. » Adam Smith, *The Theory of Moral sentiments* (1759)

Dans la première section de ce chapitre, une revue de la littérature existante sur les jeux de contribution volontaire au financement d'un bien public est présentée. L'accent est mis sur les jeux de bien public avec opportunité de sanction. Dans la deuxième section le protocole expérimental et les prédictions théoriques du jeu sont exposés. Dans une troisième section les résultats expérimentaux sont étudiés. Enfin la quatrième et dernière section présente les conclusions du chapitre.

2. La pression des pairs et l'effort de contribution dans les jeux de biens public

Le comportement rationnel de *l'homo oeconomicus* de la théorie néoclassique est individualiste et instrumental : il décide de son comportement de façon à maximiser son utilité personnelle, sous les contraintes que son environnement lui impose. Autrement dit, il est opportuniste et n'hésite pas à mentir ou à tricher si les gains espérés de la tricherie sont supérieurs à ceux attendus de l'honnêteté. Les agents au sein d'un groupe peuvent donc adopter des comportements de passager clandestin afin de maximiser leur utilité.

Un certain nombre d'expériences contredit les prédictions théoriques relatives à l'existence et l'importance du problème de passager clandestin. En effet, plusieurs études en économie expérimentale ont montré que la plupart des individus contribue largement au financement des biens publics dans les premières périodes d'un jeu répété (Ledyard, 1995). Ainsi les sujets contribuent généralement entre 40 et 60% de leur dotation initiale

au bien public dans les premières périodes d'un jeu de bien public répété (Marwell et Ames, 1979 ; Isaac, McCue et Plott, 1985 ; Kim et Walker, 1985 ; Fehr et Schmidt, 1997). De nombreux travaux se sont alors attachés à déterminer les variables susceptibles d'influencer les comportements de contribution afin d'expliquer les divergences entre la prédiction théorique et les observations (Isaac, Walker et Thomas, 1984 ; Isaac et Walker, 1983, 1984, 1987, 1988). Toutefois, lorsque les jeux de biens publics sont répétés un nombre fini de fois, le niveau de contribution moyen diminue considérablement pour atteindre des niveaux proches d'une contribution nulle. Plus récemment, d'autres travaux se sont intéressés à l'opportunité d'introduire un mécanisme d'incitation (de sanction ou de récompense) afin d'inciter les sujets à contribuer davantage au bien public (Fehr et Gächter, 2000 ; Sefton, Shupp et Walker, 2000).

2.1. Jeux de bien public sans opportunité de sanction

Plusieurs facteurs influencent les comportements de contribution au financement d'un bien public. Le tableau 1 résume les effets observés de ces différentes variables.

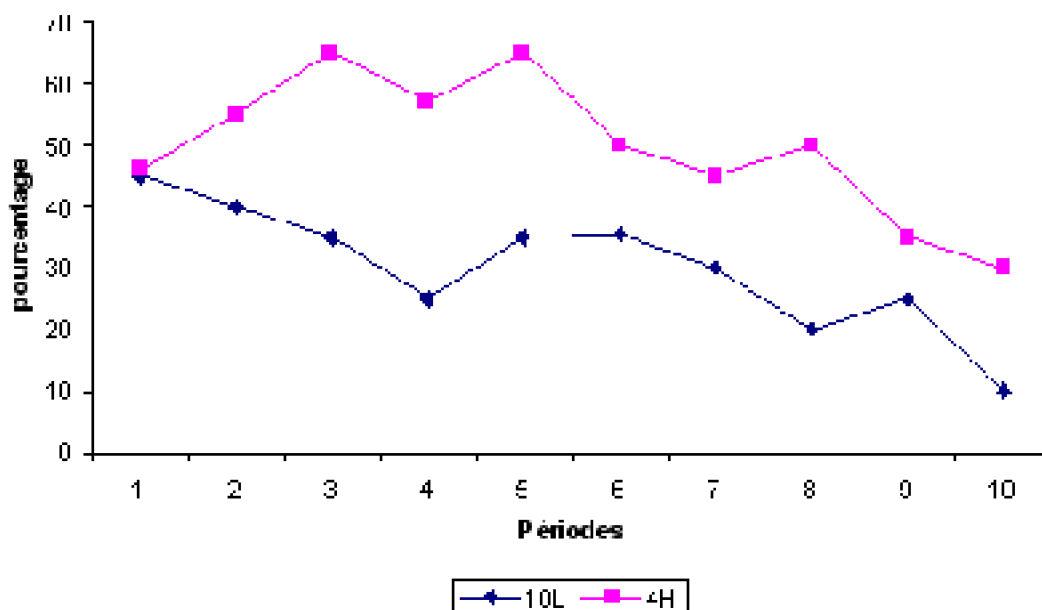
	Effet sur le niveau de contribution (en pourcentage)
Variables liées à l'environnement (facile à contrôler)	
Rendement marginal du bien public (MPCR)	++
Nombre	00
Régénération	
Connaissance commune	-
Sexe	0
Homogénéité	+
Scouts	-
Variables systémiques (difficile à contrôler)	
Différences	
Formation en économie	--
Expérience	--
Identification de groupe	+
Apprentissage	0
Autres variables	
Communication	—

Note: ++ signifie un effet positif, 0 : aucun effet, -- un effet négatif
D'après le *Handbook of Experimental Economics*

Tableau II.1 : faits stylisés sur les variables affectant les comportements de contributions volontaires au financement d'un bien public

2.1.1. Le rendement marginal et la taille du groupe

Isaac, Walker et Thomas (1984) et Isaac et Walker (1988) ont étudié l'impact du niveau du rendement marginal du bien public (*Marginal Per Capita Return*) sur le montant des contributions. Ils observent qu'une augmentation du rendement marginal accroît considérablement les niveaux de contributions. Au contraire, une diminution du rendement marginal augmente la proportion des comportements de *passagers clandestins* dans le groupe. La taille du groupe est également une variable facilement contrôlable. L'effet de la taille du groupe a suscité un débat parmi les expérimentalistes. Marwell et Ames (1979) ne trouvent aucun effet de la taille du groupe sur les niveaux de contribution. Isaac, Walker et Williams (1990) observent au contraire que le niveau de contribution est inversement proportionnel à la taille du groupe. Ainsi, plus le groupe est de petite taille et plus les contributions sont importantes.



Graphique II.1 : contributions moyennes au financement du bien public

Source : Isaac, Walker et Williams, 1990.

Le graphique 1 montre l'impact de « l'effet taille du groupe » ainsi que de « l'effet gain » sur les contributions au financement d'un bien public. Ainsi, les individus d'un groupe de petite taille qui espèrent un rendement élevé de leur contribution (4H pour un groupe de quatre personnes avec un rendement marginal élevé), participent davantage que les individus d'un groupe plus grand, espérant un retour plus faible de leur contribution (10L pour un groupe de dix personnes avec un rendement marginal faible).

2.1.2. Répétition et apprentissage

Une caractéristique des jeux de bien public répétés réside dans le fait que le niveau des contributions diminue avec la répétition (Isaac, Walker et Thomas, 1984 ; Isaac McCue et Plott, 1985 ; Isaac Walker et Williams, 1990). Isaac, Walker et Thomas (1984) ont réalisé une série d'expériences de bien public répété sur 10 périodes. Ils observent que le niveau de contribution moyen est environ de 40% de la dotation. Ils observent par ailleurs une grande hétérogénéité des résultats selon les groupes et la période du jeu. Alors qu'au début du jeu, les sujets contribuent assez largement au bien public, au fil des périodes, le niveau de contribution diminue et converge vers le passager clandestin en fin du jeu. Ainsi, le niveau des contributions décline dans le temps pour atteindre un niveau de 10 à 30% de la dotation initiale en fin de jeu. Andreoni (1988) ; Keser et Van Winden (1998) observent une décroissance similaire du niveau des contributions individuelles en jeu répété. Andreoni rapporte des contributions en début de jeu à environ 48% de la dotation initiale et à 12% de la dotation en dernière période. Keser et Van Winden, (1998) observent quant à eux que les sujets contribuent initialement à hauteur de 56% de leur dotation puis à environ 15% de leur dotation en dernière période. Le tableau 2 synthétise les résultats des expériences sur jeu de bien public répété.

Tableau II.2. Pourcentage de sujets qui adoptent un comportement de passager clandestin en dernière période d'un jeu répété de bien public.

Etudes	Pays	Taille du groupe (n)	Retour d'invest. dans le bien public	Nombre total de sujets	% de passagers clandestins
Isaac et Walker 1988 a	USA	4 et 10	0.3	42	83
Isaac et Walker 1988 b	USA	4 et 10	0.75	42	57
Andreoni 1988	USA	5	0.5	70	54
Andreoni 1995a	USA	5	0.5	80	55
Andreoni 1995b	USA	5	0.5	80	66
Crozon 1995	USA	4	0.5	48	71
Croson 1996	USA	4	0.5	96	65
Keser et van Winden 1996	Holland	4	0.5	160	84
Ockenfels et Weiman 1996	Germany	5	0.33	200	89
Burlando et Hey 1997	UK, Italy	6	0.33	120	66
Falkinger, Fehr, Gächter et Winter-Ebner 1997	Switzerland	8	0.2	72	75
Falkinger, Fehr, Gächter et Winter-Ebner 1997	Switzerland	16	0.1	32	84
Nombre total de sujets et pourcentage de <i>passagers clandestins</i>				1042	73
<i>Fehr et Schmidt, 1997</i>					

Sur l'ensemble des expériences de jeu de bien public répété, la contribution moyenne varie entre 40% et 60% de la dotation initiale lors des premières périodes du jeu. En dernière période, 73% des individus choisissent de ne pas contribuer au bien public et une grande majorité des sujets restants choisit un niveau de contribution très faible et proche de zéro. Le déclin considérable et particulièrement prononcé du niveau des contributions dans les dernières périodes du jeu caractérise donc les jeux répétés de bien public. Selon les termes de Stoecker (1983) et Selten et Stoecker (1986), ces observations peuvent être interprétées comme un comportement de fin de jeu.

Comment expliquer la diminution des niveaux de contribution dans un jeu répété ? Les expérimentalistes ont avancé deux explications possibles : les effets stratégiques et les effets de l'expérience. Isaac, Walker et Thomas (1984) ont testé l'effet de l'expérience sur les niveaux de contribution des sujets en comparant les comportements de contribution selon que les joueurs ont déjà participé ou pas au jeu. Ils observent que les

sujets expérimentés contribuent moins que les sujets non expérimentés. Cependant, les sujets expérimentés continuent à contribuer un montant non nul au bien public. L'expérience n'est donc pas la seule explication. Andreoni (1988) a également testé les effets de l'expérience. Pour ce faire, il a réalisé l'expérimentation suivante : les sujets jouent tout d'abord 10 périodes d'un jeu de bien public. A la fin de ces 10 périodes, les sujets rejouent 10 nouvelles périodes dans les mêmes conditions. Le taux moyen de contribution au bien public en première période des 10 premières périodes est d'environ 50% et il diminue au cours du temps. Il n'apparaît pas de différence significative dans la seconde série d'expériences. De même, Palfrey et Prisbey (1993) n'observent aucune différence significative entre les niveaux de contributions des sujets expérimentés et ceux des sujets inexpérimentés.

Une autre explication possible à la décroissance du niveau des contributions individuelles est d'ordre stratégique et repose sur l'idée que les sujets réduisent leur propre contribution en réponse aux niveaux de contributions précédents des autres sujets de leur groupe (Kreps et al. ,1982 ; Keser et van Winden, 1996). Dans un jeu de bien public répété, les sujets potentiellement coopérateurs n'ont aucun moyen direct de discipliner ceux qui adoptent un comportement de passager clandestin. La seule réponse possible vis à vis d'un tel comportement consiste à adopter également un comportement de passager clandestin. Cela suppose toutefois que les sujets soient en interaction avec les mêmes joueurs dans toutes les périodes (traitement *partenaire*).

Donc dans un jeu de bien public répété sans opportunité de sanctionner les autres membres du groupe, la seule façon de discipliner les *passagers clandestins* consiste à adopter également un comportement de passager clandestin. Ce comportement conduit à la diminution progressive des niveaux de contribution et à la convergence vers le niveau de contribution prédit par la théorie. On peut alors s'interroger sur la possibilité de discipliner les autres membres du groupe sans avoir à réduire sa propre contribution.

2.2. Jeu de bien public avec mécanismes d'incitation par sanction et/ou récompense

Plus récemment, des travaux ont cherché à savoir si l'opportunité de sanctionner et/ou de récompenser les autres membres du groupe permettrait de soutenir la coopération dans un jeu de bien public répété à horizon fini. Alors qu'un certain nombre de travaux s'est intéressé à l'effet des sanctions et/ou des récompenses lorsqu'elles sont attribuées de façon exogène (Yamagishi, 1986 ; Dickinson, 2000 ; Falkinger, Fehr, Gächter et Winter-Ebner, 2000), peu de travaux ont étudié la possibilité pour les sujets de se sanctionner et/ou se récompenser mutuellement. Fehr et Gächter (2000) ont étudié l'impact des sanctions sur les niveaux de contribution volontaire à un bien public. Sefton, Shupp et Walker (2000) ont comparé l'efficacité relative des sanctions et des récompenses sur les niveaux de contribution. Bowles, Carpenter et Gintis (2001) étudient quant à eux l'impact d'une variation du rendement marginal du bien public et de la taille du groupe sur la propension à sanctionner les autres membres du groupe.

2.2.1. Jeu de bien public avec opportunité de sanction : Fehr et Gächter

en vertu de la loi du droit d'auteur.

(2000)

Fehr et Gächter (2000) étudient la possibilité d'introduire dans un jeu de bien public l'opportunité d'attribuer des points de sanction monétaire aux autres sujets. L'idée est que lorsque les agents ont la possibilité de se sanctionner mutuellement, ils peuvent alors discipliner les *passagers clandestins* sans avoir à réduire leur propre contribution. Les auteurs ont modifié un jeu de bien public classique en ajoutant une deuxième étape au jeu standard. Ce jeu se déroule en deux étapes. La première étape est similaire à un jeu de bien public standard. Dans une deuxième étape, les sujets sont informés des contributions individuelles des autres membres de leur groupe et ont l'opportunité de réduire les gains des autres sujets en leur attribuant des points de sanction. Les points de sanction sont coûteux pour ceux qui les reçoivent mais également pour ceux qui sanctionnent les autres sujets. Pour chaque point de sanction reçu, le gain de première étape du sujet sanctionné est réduit de 10%. Le gain du sujet sanctionné est nul lorsqu'il reçoit au total 10 points de sanction ou plus. Les sujets jouent 10 périodes d'une session sans sanction puis 10 périodes d'une session avec sanction. Afin d'étudier la persistance de l'effet de la sanction dans le temps, d'autres sujets jouent 10 périodes avec sanction puis 10 périodes sans sanction. Soit π_i^1 la fonction de gain du sujet i , $i=1, \dots, N$ dans le jeu de bien public sans sanction :

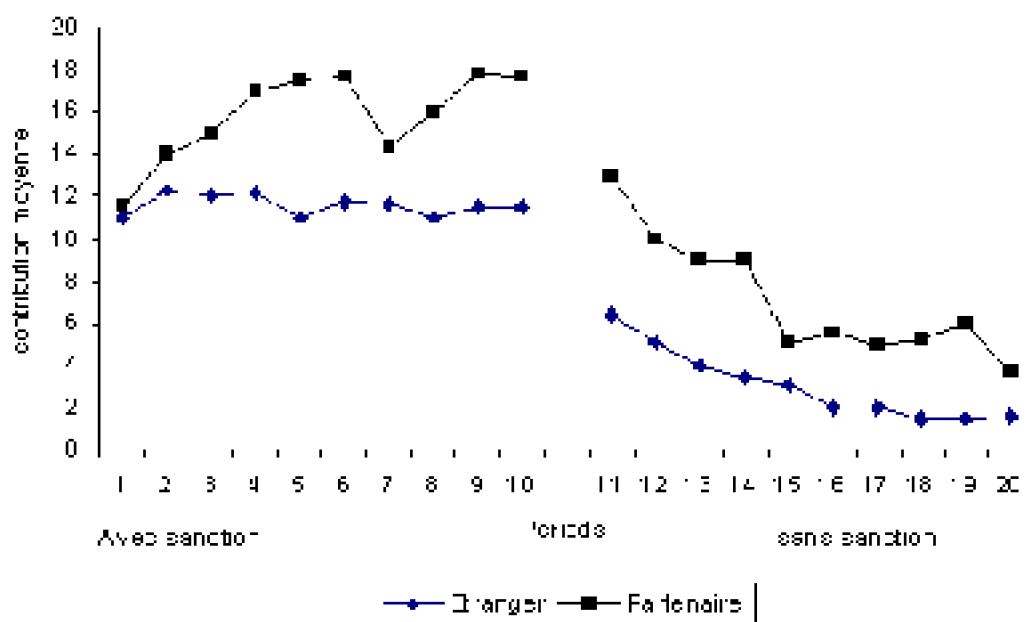
(3)

a est le rendement marginal du bien public ; y représente la dotation initiale que reçoit chaque sujet au début de chaque période. g_i est la contribution de l'agent i au bien public. Soit π_i^2 la fonction de gain de chaque sujet dans le jeu de bien public avec possibilité de sanction :

(4)

p_{ij} représente les points de punition que le sujet i affecte au sujet j . Le nombre total de points que l'agent i reçoit est noté P_i^1 . Le coût inhérent au fait de sanctionner les autres sujets est donné par $c(p_{ij})$. Dans une première étape, les sujets décident de leur effort de contribution au bien public. Dans une deuxième étape, les sujets sont informés des niveaux de contribution de chaque membre du groupe et décident de les sanctionner ou non. Puisqu'il est coûteux de punir les autres, aucun agent n'est incité à le faire. Donc les sujets devraient adopter une stratégie de passager clandestin en première étape du jeu puisqu'ils anticipent qu'aucune sanction ne sera prise envers les *passagers clandestins*.

Les auteurs étudient trois conditions différentes : une condition *étranger*, une condition *parfait étranger* et une condition *partenaire*. En condition *partenaire*, la composition du groupe est fixe sur l'ensemble des périodes. À l'inverse, en condition *étranger*, la composition du groupe est modifiée à chaque période. La différence entre la condition *parfait étranger* et la condition *étranger* est qu'en condition *parfait étranger*, la probabilité pour un joueur de se retrouver avec le même joueur au cours du jeu est nulle. La comparaison des sessions *étranger* et *partenaire* permet d'identifier l'importance de la durée des interactions dans l'augmentation des contributions avec sanctions monétaires, comme le montre le graphique 2.



Graphique 11.2 : Contribution moyenne dans les sessions partenaire et étranger

Contrairement aux prédictions théoriques, Fehr et Gächter observent que les sujets sanctionnent les autres membres de leur groupe même si cela est coûteux pour eux. La possibilité de sanctionner les autres sujets accroît considérablement la coopération au sein des groupes. Par exemple, dans la session *partenaire*, les sujets contribuent en moyenne 37.5% de leur dotation initiale durant les 10 premières périodes d'un jeu de bien public classique et 85% de leur dotation dans les 10 périodes suivantes avec possibilité de sanction. A la dernière période, la contribution moyenne représente 16% de la dotation initiale mais augmente jusqu'à 91% à la vingtième période. Les sujets seraient donc capables d'atteindre un niveau élevé de coopération lorsqu'ils ont la possibilité de sanctionner leurs pairs.

2.2.2. Sanction et récompenses : Sefton, Shupp et Walker (2000)

Sefton, Shupp et Walker (2000) ont comparé les effets des sanctions et des récompenses sur les comportements de contribution volontaire à un bien public répété. Le traitement avec sanction est similaire à celui de Fehr et Gächter (2000). Les sujets ont la possibilité de sanctionner les autres membres du groupe. Sanctionner réduit les gains de ceux qui subissent la sanction mais également de ceux qui l'imposent. Les auteurs comparent ce traitement à un deuxième traitement où les sujets ont la possibilité de récompenser les autres membres du groupe. Les récompenses se traduisent par un transfert de gain de celui qui récompense à celui qui est récompensé. Enfin, dans un troisième traitement, ils étudient un jeu avec opportunité de sanctionner et également de récompenser.

Les auteurs considèrent tout d'abord un jeu de bien public sans opportunité de sanction ou de récompense. Chaque sujet d'un groupe de quatre reçoit une dotation de 6 jetons et doit décider combien allouer au bien public. Soit la fonction de gain de chaque

sujet :

La stratégie dominante est ici de ne contribuer aucun jeton au bien public. Les auteurs considèrent ensuite un jeu de bien public avec opportunité de sanction. Ce jeu consiste à ajouter une deuxième étape au jeu précédent. Ainsi, après avoir observé les niveaux de contribution des autres sujets, chaque sujet reçoit 6 jetons supplémentaires et peut décider de les utiliser ou non pour sanctionner les autres sujets (Le sujet garde pour lui les jetons qu'il n'utilise pas pour sanctionner). Le gain d'un sujet est réduit d'autant de points de sanction reçus. Soit la fonction de gain de chaque sujet i dans ce traitement :

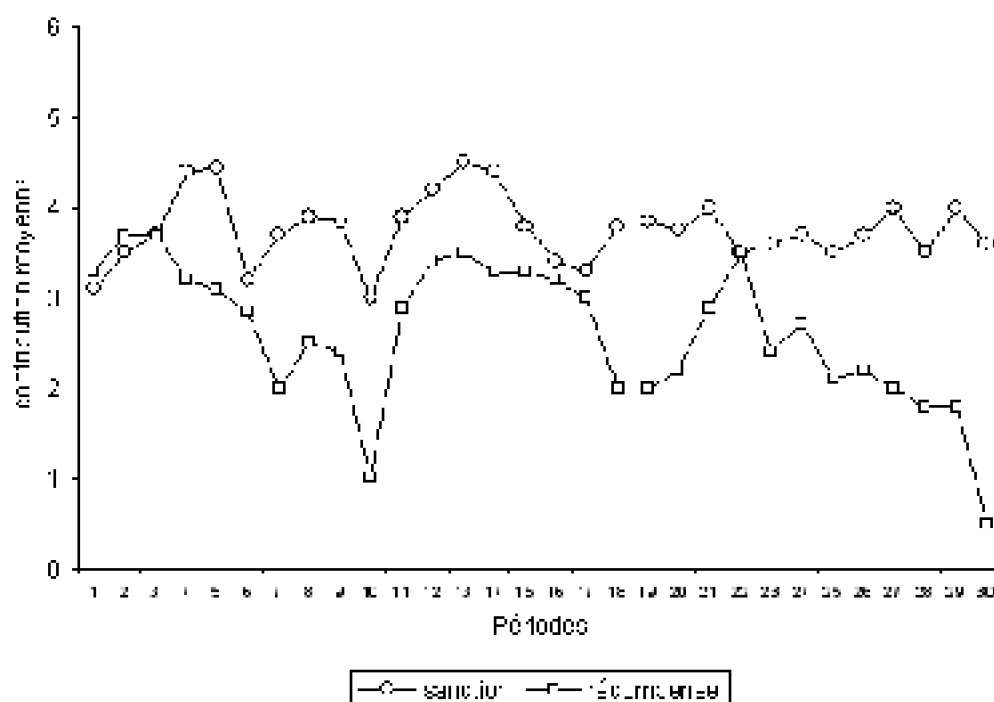
A l'équilibre, personne ne contribue au bien public et personne ne récompense.

Dans un dernier traitement, les sujets ont la possibilité de sanctionner et de récompenser les autres membres de leur groupe. Soit la fonction de gain du sujet i :

L'unique équilibre de Nash est de ne jamais sanctionner, ne jamais récompenser et ne pas contribuer au bien public. Les auteurs distinguent deux séquences de 10 périodes chacune. Dans la première séquence, les sujets de toutes les sessions participent au traitement sans opportunité de sanction ou de récompense. Dans la deuxième séquence de 10 périodes, les sujets participent à l'un des quatre traitements : jeu de bien public sans sanction ni récompense, avec sanction, avec récompense ou avec l'opportunité de sanctionner et de récompenser.

Ils observent que les niveaux de contribution diminuent progressivement lorsque les sujets continuent à jouer un jeu sans opportunité de sanction ou de récompense. L'opportunité de sanctionner ou de récompenser accroît considérablement le niveau de contribution des sujets. Ils obtiennent des résultats similaires à ceux de Fehr et Gächter (2000). Ils observent en effet que l'opportunité de sanctionner affecte positivement le niveau de contribution des sujets. Sefton, Shupp et Walker observent par ailleurs que l'opportunité de récompenser accroît également le niveau des contributions. Dans le traitement où les sujets ont la possibilité de choisir entre sanctionner ou récompenser, les auteurs observent que les sujets choisissent initialement davantage les récompenses aux sanctions. Avec répétition du jeu, les sujets choisissent autant les deux modes d'incitation.

Afin de comparer plus distinctement les deux modes d'incitation, Sefton, Shupp et Walker ont conduit des sessions supplémentaires où les sujets jouent 20 périodes des traitements avec sanction ou avec récompense dans la deuxième séquence.



Source : Sefton, Shupp et Walker (2000)

Graphique II.3 : Contribution moyenne dans les sessions avec sanction ou avec récompense

Sefton, Shupp et Walker observent des niveaux de contribution similaires entre les deux traitements lors des premières périodes du jeu. Dans les dernières périodes, au contraire, les niveaux de contribution diminuent brutalement dans le traitement avec récompense alors qu'ils restent relativement stables dans le traitement avec sanctions. Les auteurs concluent que les mécanismes de sanction seraient plus efficaces à long terme que les incitations par les récompenses.

2.2.3. Effet du rendement et de la taille sur les sanctions : Bowles, Carpenter et Gintis (2001)

Bowles, Carpenter et Gintis (2001) ont répliqué l'analyse de Fehr et Gächter (2000). Contrairement à Fehr et Gächter, toutes leurs sessions sont en condition *étranger*. Chaque session est composée de 10 périodes. Au début de chaque période du jeu, chaque sujet reçoit une dotation $w=20$ francs (convertis en dollars à la fin de l'expérience). Les sujets décident combien contribuer à un projet commun. Dans un deuxième temps, ils observent les niveaux de contribution des autres membres de leur groupe et ont l'opportunité de les sanctionner en leur attribuant des points de sanction. Ces points sont coûteux pour ceux qui les reçoivent mais également pour ceux qui les imposent.

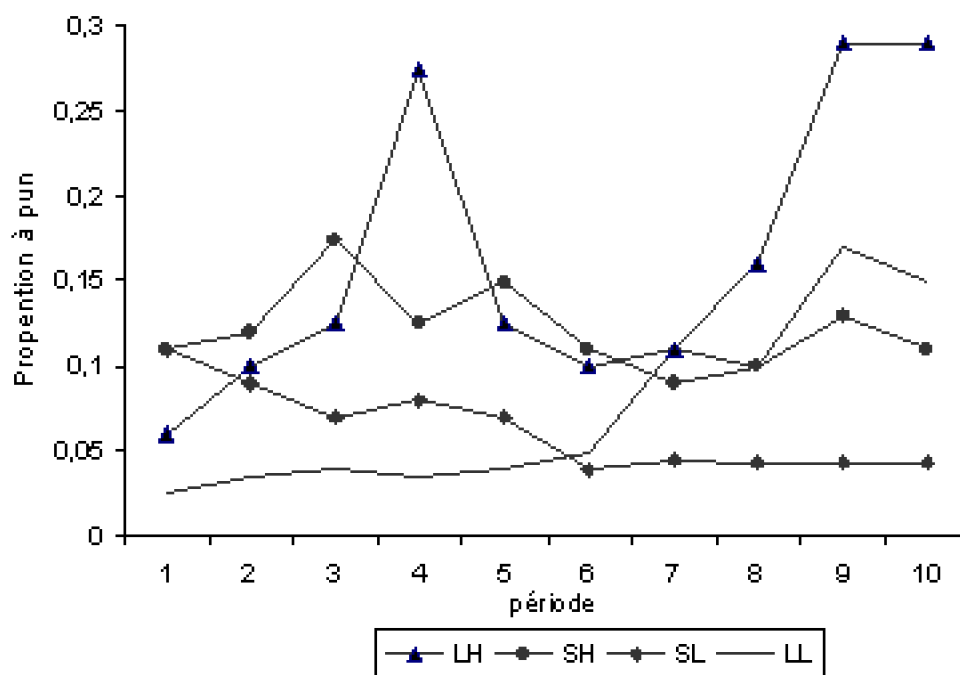
Les auteurs ont étudié l'effet de la taille du groupe et du rendement marginal du bien public sur les comportements de sanction et de contribution des sujets. Les groupes sont composés de 5 sujets dans un traitement et de 10 dans un autre. Pour étudier l'effet d'une variation du rendement marginal du bien public, ce dernier est fixé à 0.3 dans un

traitement et à 0.75 dans un autre. Le tableau ci-dessous résume la répartition des sujets selon la taille du groupe et le rendement marginal.

Tableau II.3 : Protocole expérimental (Bowles, Carpenter et Gintis, 2001)

	5 sujets par groupe	10 sujets par groupe
m=0.3	(SL) 7 groupes 35 sujets	(LL) 4 groupes 40 sujets
m=0.75	(SH) 6 groupes 30 sujets	(LH) 4 groupes 40 sujets

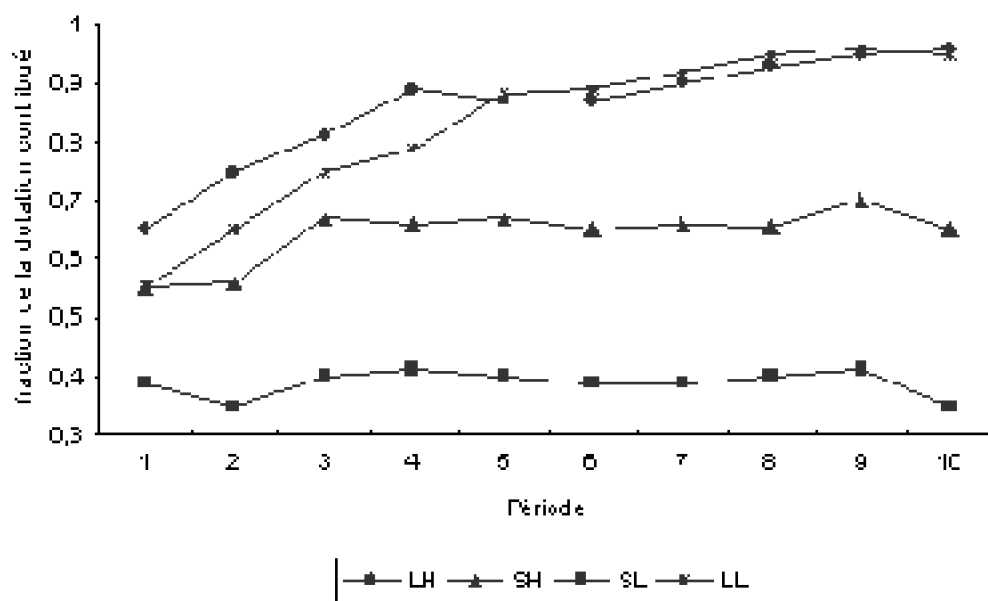
Bowles, Carpenter et Gintis observent que la propension à punir les autres membres du groupe augmente avec le rendement marginal du bien public et ne décline pas lorsque la taille du groupe s'accroît.



Source : Bowles, Carpenter et Gintis (2001)

Graphique II.4. : Propension à punir mesurée en points moyens attribués

Les auteurs observent par ailleurs que la contribution avoisine 100% dans les grandes équipes (dix joueurs).



Source : Bowles, Carpenter et Gintis (2001)

Graphique 11.5 : Contribution moyenne avec opportunité de sanction

Les modèles présentés dans cette section ont mis en évidence l'efficacité des sanctions monétaires comme support de la coopération dans le groupe. Toutefois, ils ne permettent pas de dissocier les raisons pour lesquelles les sujets contribuent davantage lorsqu'ils reçoivent des sanctions. Deux explications peuvent être avancées : la première est que les sujets contribuent afin d'éviter les conséquences néfastes des sanctions en terme de réduction de leurs gains. La deuxième explication est qu'ils contribuent davantage afin d'éviter la désapprobation de leurs pairs, véhiculée par les sanctions.

3. Inefficacité théorique des sanctions monétaires et non monétaires

L'expérience présentée ici cherche à dissocier les conséquences monétaires des sanctions imposées aux passagers clandestins des conséquences morales relatives à la désapprobation des pairs. Afin de tester l'importance respective de ces effets, plusieurs traitements ont été réalisés. Dans un premier traitement appelé SP, les sujets participent à un jeu classique de bien public et n'ont pas l'opportunité de sanctionner leurs pairs. Ce traitement sert de traitement de *référence* (benchmark) afin d'étudier l'effet des sanctions. Dans un deuxième traitement, appelé MP, les sujets peuvent attribuer des points de sanction monétaire aux autres membres de leur groupe. Les points de sanction sont coûteux non seulement pour ceux qui les reçoivent mais également pour ceux qui les imposent. Ce traitement est similaire à celui de l'expérience de Fehr et Gächter (2000).

Nous avons réalisé un troisième traitement appelé traitement avec sanctions non monétaires (NP). Ce traitement est identique au précédent à une exception près : les points de sanction n'affectent les gains ni des agents sanctionnés ni ceux des agents qui sanctionnent. Afin de tester l'importance de la durée de la relation, nous avons comparé les effets des sanctions non monétaires selon que les sujets sont en interaction avec les mêmes autres sujets pour l'ensemble des périodes (condition *partenaire*) ou qu'ils jouent avec des sujets différents à chaque période (condition *étranger*). La comparaison des conditions *étranger* et *partenaire* permet d'identifier l'effet de la répétition des interactions dans l'augmentation des contributions avec sanctions non monétaires.

3.1. Le protocole expérimental

Le protocole expérimental définit les conditions de déroulement de l'expérimentation. Il renseigne sur les informations disponibles au cours du jeu et détermine les valeurs des dotations initiales, des fonctions de paiement et de la valeur des différents paramètres. Pour l'expérimentation, des groupes composés de 4 personnes sont formés. Les participants sont répartis dans un groupe de façon aléatoire au début de l'expérience. Les sujets recrutés sont des étudiants de premier cycle dans des disciplines diverses. Les étudiants n'ont jamais participé à d'autres expériences et n'ont jamais suivi de cours de théorie des jeux ni d'économie expérimentale. Par ailleurs, avant le début de l'expérience, aucun étudiant ne connaît le but de l'expérience. Aucun sujet ne peut participer à plus d'une session expérimentale. L'identité de chacun des sujets reste anonyme pendant toute la durée de l'expérience. En moyenne l'expérience dure 90 minutes (lecture des instructions et paiement inclus). Le déroulement du protocole expérimental est le suivant : les instructions de jeu sont distribuées aux sujets sur papier et lues à haute voix (voir annexe A pour les instructions détaillées du jeu). Puis les participants disposent de quelques minutes pour relire ces instructions et poser des questions. Il est répondu publiquement aux questions¹⁵.

Afin d'obtenir un nombre suffisant d'observations indépendantes, plusieurs expérimentations ont été réalisées. L'expérience a été réalisée aux Etats-Unis (Université de Purdue) et en France au GATE (Université de Lyon II). Le logiciel expérimental utilisé est REGATE. Les informations communiquées aux sujets et toutes les décisions prises sont informatisées. Dès l'entrée dans le laboratoire, les sujets n'ont pas été autorisés à communiquer entre eux sous peine de se voir exclure du jeu et de tout paiement. Au total onze sessions ont été conduites dont sept à l'Université de Purdue et quatre autres au GATE. Chaque session comprend 3 parties de 10 périodes chacune : le traitement SP est joué sur les périodes 1-10, le traitement MP ou NP est joué sur les périodes 11-20 et enfin le traitement SP est joué des périodes 21 à 30. La comparaison des 10 premières périodes avec les périodes 11-20 permet d'étudier l'impact de la sanction monétaire ou non monétaire sur le niveau de contribution des sujets tandis que la comparaison des périodes 11-20 avec les périodes 21-30 permet d'étudier la persistance de l'effet de la sanction. Le tableau 1 retrace les caractéristiques des différentes sessions. Les quatre premières colonnes représentent respectivement le numéro de la session, le nombre de

¹⁵ Les réponses faites par l'expérimentaliste ne doivent comporter aucun élément faisant référence à une stratégie de jeu.

sujets par session ainsi que le lieu où s'est déroulé la session. La cinquième colonne informe sur le traitement joué lors des périodes 11 à 20 :

Tableau II.1: caractéristiques des sessions expérimentales

Expérience	Nombre de sujet	Localisation	Traitement	Condition		
				Périodes 1-10	Périodes 11-20	Périodes 21-30
1	12	Purdue	MP	Partenaire	Partenaire	Partenaire
2	12	Purdue	MP	Partenaire	Partenaire	Partenaire
3	12	GATE	MP	Partenaire	Partenaire	Partenaire
4	12	GATE	MP	Partenaire	Partenaire	Partenaire
5	16	Purdue	NP	Partenaire	Partenaire	Partenaire
6	8	Purdue	NP	Partenaire	Partenaire	Partenaire
7	12	GATE	NP	Partenaire	Partenaire	Partenaire
8	8	GATE	NP	Partenaire	Partenaire	Partenaire
9	16	Purdue	NS	Partenaire	Étranger	Étranger
10	16	Purdue	NS	Partenaire	Étranger Périodes 11-15	Étranger
11	16	Purdue	NS	Partenaire	Étranger	Étranger

MP pour le traitement monétaire, NP pour le traitement non monétaire en condition partenaire et NS pour le traitement non monétaire en condition étranger.

Sous la condition *étranger*, la composition du groupe de 4 sujets est modifiée de façon aléatoire à chaque période du jeu. Ainsi chaque sujet ne peut pas interagir avec les mêmes personnes à chaque période du jeu. Sous la condition *partenaire*, les groupes conservent la même structure tout au long de la session ; ainsi un sujet interagit avec les mêmes personnes sur l'ensemble de la session. Cette information est communiquée aux sujets pendant la lecture des instructions. Cette condition offre la possibilité d'envoyer des signaux aux autres membres du groupe sur leurs intentions de coopération. Dans les sessions expérimentales jouées sous la condition *partenaire*, le nombre de participants varie entre 8 participants, 12 participants et 16 participants. Dans chaque session jouée en condition *étranger*, le nombre de participants est de 16. Les sujets reçoivent les instructions avant chaque session immédiatement jouée. Ils sont informés que la session dure 10 périodes mais ne savent pas que l'expérience est composée de trois sessions successives.

3.2. Conditions d'information

Pour chaque session, la dotation initiale de chaque sujet à chaque période est $y=20$ ECU (Experimental Currency Unit), avec un taux de conversion de 30 ECU=1 dollar ou de 5 ECU=1 franc. Le taux de rendement marginal du bien public par individu est fixé à 0.4. En traitement SP, la fonction de gain ainsi que la valeur des paramètres y , n , N et le gain marginal du bien public sont de connaissance commune. A la fin de chaque période, les sujets de chaque groupe sont informés de la contribution totale de leur groupe ($\sum g_j$) au

bien public. En revanche, le niveau de contribution individuelle des autres membres du groupe n'est pas communiqué. A la fin de chaque période, chaque sujet est informé de son gain total.

En traitement MP, les sujets sont également informés de la valeur des paramètres y , n , N et a ainsi que des coûts associés aux points distribués. En fin de première étape de chaque période, les sujets ont connaissance de la contribution totale au projet commun et de leurs gains de première étape. Dans la deuxième étape de chaque période, les sujets sont informés de la contribution individuelle de chacun des autres sujets de leur groupe. Afin d'éviter les phénomènes de réputation, l'ordre des sujets qui apparaît sur l'écran est modifié de façon aléatoire à chaque période. Ainsi les sujets n'ont pas l'information leur permettant de faire le lien entre un sujet donné et l'historique de sa contribution individuelle. Après avoir été informés des contributions individuelles, les sujets décident alors des points de sanction qu'ils attribuent aux autres sujets de leur groupe. Ils sont ensuite informés de leur gain pour la période qui prend en compte les points reçus et le coût des points distribués. Les sujets ne sont jamais informés des activités individuelles de sanction des autres sujets. Ils connaissent uniquement le nombre total de points qu'ils reçoivent.

Dans le traitement NP, les sujets sont également informés de la valeur des paramètres y , n , N . En fin de première étape de chaque période, les sujets ont connaissance de la contribution totale au projet commun et de leurs gains de première étape. Dans la deuxième étape de chaque période, les sujets sont informés de la contribution individuelle de chacun des autres membres de leur groupe et décident alors des points de sanction non monétaire qu'ils attribuent aux autres sujets de leur groupe. Ils sont ensuite informés de leur gain pour la période. Ce gain est identique au gain de première période puisque les points n'affectent ni les gains des sujets sanctionnés ni les gains des sujets qui sanctionnent. Comme dans le traitement précédent, les sujets ne sont jamais informés des activités individuelles de punition des autres sujets. Ils connaissent uniquement le nombre total de points qu'ils reçoivent.

3.3. Prédictions théoriques

Dans le traitement SP, les sujets doivent décider simultanément à chaque période combien d'ECU (*Experimental currency Unit*) allouer à un bien public. Chaque ECU affecté au bien public procure un rendement de 0.4 ECU à chacun des 4 sujets. Les sujets conservent les ECU qu'ils n'ont pas affecté au bien public. Soit la fonction de gain individuel pour chaque sujet i à chaque période :

où g_i représente la contribution au bien public du sujet i . Puisque $0.4 < 1$, l'investissement dans le bien public génère une perte monétaire de $(1-0.4)$.

La stratégie dominante d'un agent supposé rationnel est donc de ne pas contribuer au bien public (soit $g_i = 0$ pour tout sujet $i \in \{1, \dots, 4\}$). Cependant, le gain agrégé monétaire serait maximisé si chaque sujet choisissait d'affecter toute sa dotation (soit $g_i = 20$). Donc, d'après les prédictions théoriques, aucun sujet ne devrait contribuer au financement du bien public.

Encadré II.1 Définition de l'équilibre parfait en sous jeu (Selten, 1965)

Un équilibre de Nash est parfait en sous jeu si les stratégies des joueurs constituent un équilibre de Nash à chaque sous-jeu.

Dans le traitement MP, les sujets jouent un jeu répété de 10 périodes. A chaque période, le jeu comprend 2 étapes. La première étape est identique à la session sans sanction. Dans la deuxième étape, chaque sujet est informé de la contribution totale au bien public mais également de la contribution individuelle de chacun des autres membres de son groupe. Les sujets ont alors l'opportunité de sanctionner les autres sujets en leur affectant des points de sanction (entre 0 et 10 points). Chaque point que reçoit un sujet réduit son gain issu de la première étape du jeu de 10%. Le sujet qui décide de sanctionner un ou plusieurs autres sujet subit un coût proportionnel au nombre de points imposés. Le tableau 2 illustre la relation entre le nombre de points donnés par un sujet et le coût de sanctionner ¹⁶.

Tableau II.2 : Coût associé aux points de sanction donnés

Points	0	1	2	3	4	5	6	7	8	9	10
Coût des points	0	1	2	4	6	9	12	16	20	25	30

La fonction de gain de chaque joueur i à chaque période dans cette session s'écrit de la façon suivante :

où P_{ik} représente le nombre de points que l'agent i assigne à chaque sujet k et $K(P_{ik})$ le coût pour l'agent i à assigner ces points.

Résolvons ce jeu par induction à rebours. En dernière étape de chaque période du jeu, les sujets décident de leur sanction envers les autres sujets. Puisque punir est coûteux, le modèle prédit qu'un sujet rationnel et égoïste ne punira jamais. En première étape de chaque période du jeu, les sujets, supposés rationnels s'attendent à ne jamais être sanctionnés en deuxième étape. Par conséquent, ils choisiront de contribuer le moins possible au bien public en première étape tel que $g_i = 0$. L'existence de la deuxième étape du jeu ne modifie donc en rien le comportement des sujets en première étape du jeu. Il ne devrait donc pas exister de différence entre les traitements sans sanction et avec sanction monétaire. La combinaison de stratégies qui consiste à ne pas contribuer au bien public et à ne pas sanctionner se répète sur l'ensemble des périodes du jeu répété un nombre fini de fois.

Encadré II.2 Jeux répétés un nombre fini de fois (Gibbons, 1992)

Si un jeu G a un seul équilibre de Nash, alors l'équilibre de Nash est joué à chaque période du jeu $G(T)$ répété un nombre fini T de fois.

Le traitement NP consiste également en un jeu répété de 10 périodes. Comme dans le traitement avec sanction monétaire, le jeu se déroule en deux étapes. La première

¹⁶ En notant P_{ij} le nombre de points que le sujet i assigne à un sujet j , le tableau 3 illustre le coût $K(P_{ij})$.

étape de ce traitement est identique à la première étape du traitement précédent. Dans la deuxième étape, les sujets observent le niveau de contribution de leurs pairs et peuvent leur affecter des points de sanction non monétaire. Contrairement aux points de sanction monétaire, les points de sanction non monétaire n'affectent ni les gains du sujet sanctionné ni ceux du sujet qui sanctionne. Ils indiquent seulement le sentiment de désapprobation que ressent celui qui sanctionne. Soit la fonction de gain de chaque sujet i à chaque période pour cette session :

La fonction de gain dans le traitement avec sanction non monétaire est strictement identique à celle du traitement sans sanction puisque les sanctions non monétaires n'ont aucune influence sur les gains des sujets. Si l'on suppose que les sujets ne cherchent que la maximisation de leurs gains monétaires, le problème de maximisation dans ce traitement est alors le même que dans la session sans sanction. Les agents fournissent l'effort de contribution le plus faible possible qui maximise leurs gains. L'équilibre parfait en sous jeu est le même quel que soit le nombre de points de sanction.

4. Efficacité réelle des sanctions : aversion à la baisse des gains et à la désapprobation

4.1. Observations des niveaux de contributions individuels

Nous avons dans un premier temps étudié l'impact des sanctions monétaires et non monétaires sur le niveau de contribution des sujets. Si les sujets étaient uniquement motivés par leurs gains pécuniaires, on ne devrait pas observer de différence entre les traitements avec et sans sanction. Contrairement aux prédictions théoriques, on observe que les sanctions monétaires ont un effet positif sur le niveau de contribution des sujets. Ce résultat confirme le résultat de Fehr et Gächter (2000). On observe également que les sanctions non monétaires augmentent le niveau de contribution des sujets.

RESULTAT 1 : *lorsque les agents ont l'opportunité d'affecter des points de sanction monétaire aux autres membres du groupe, alors le niveau des contributions augmente fortement.*

Le tableau 3 ci-dessous présente les niveaux moyens de contribution dans la session MP. Dans les colonnes 2 à 3 apparaissent les niveaux des contributions individuelles moyennes des expériences menées à Purdue pour chacune des trois sessions (SP, MP et de nouveau SP). Les colonnes 4 à 6 donnent les contributions individuelles moyennes des expériences conduites au GATE. Le tableau 4 reporte les contributions individuelles moyennes en dernière période de chaque session (voir annexe B).

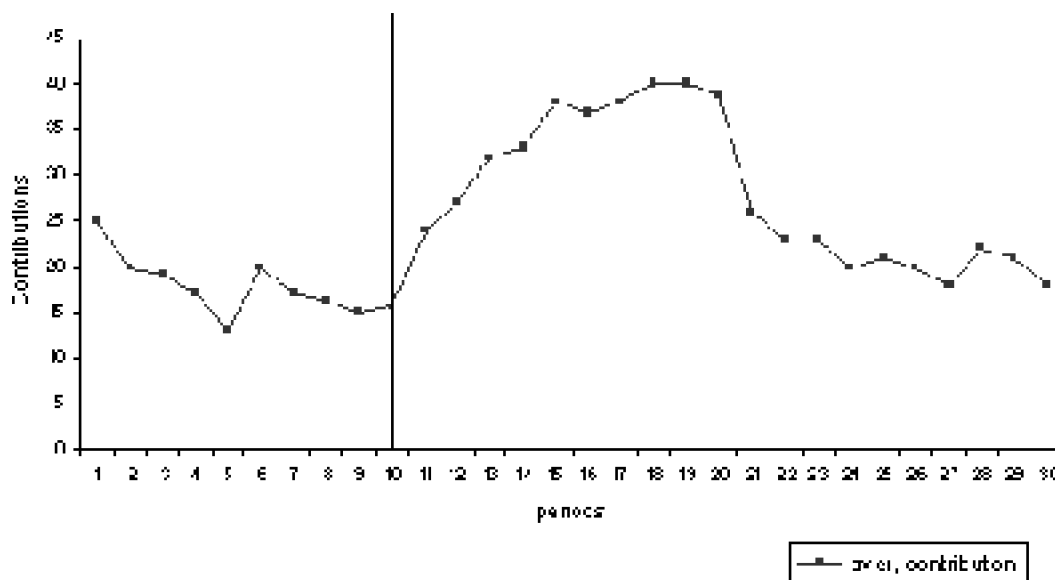
Tableau II.3 : contributions moyenne dans les sessions avec sanctions monétaires

	Sessions avec traitement MP (Purdue)			Sessions avec traitement MP (GATE)		
Groupes	Sans sanction Périodes 1-10	Sanctions monétaires Périodes 11-20	Sans sanction Périodes 21-30	Sans sanction Périodes 1-10	Sanctions monétaires Périodes 11-20	Sans sanction Périodes 21-30
1	4.20	2.73	1.50	6.63	18.83	10.63
	(3.74)	(2.69)	(2.12)	(4.72)	(2.83)	(7.16)
2	7.15	7.33	4.28	5.33	2.33	1.35
	(6.29)	(4.69)	(4.66)	(4.96)	(3.18)	(1.59)
3	9.68	15.89	7.05	5.50	12.68	5.25
	(8.01)	(7.41)	(8.78)	(4.62)	(3.45)	(3.73)
4	6.00	19.05	3.85	6.05	6.73	4.33
	(6.14)	(3.19)	(5.63)	(6.30)	(6.25)	(5.42)
5	4.20	9.28	0.75	7.20	9.48	6.63
	(5.76)	(2.16)	(1.98)	(4.97)	(4.34)	(4.05)
6	3.05	16.35	7.09	7.35	13.03	5.44
	(3.37)	(5.62)	(5.74)	(4.82)	(6.38)	(6.13)
Moyenne	5.71	11.77	4.09	6.34	10.51	5.60
	(5.55)	(4.29)	(4.81)	(5.06)	(4.40)	(4.68)
Note : Les écarts types sont en parenthèse						

Tableau II.4 : contributions moyenne en dernière période dans les sessions avec sanctions monétaires

	Sessions avec traitement MP (Purdue)			Sessions avec traitement MP (GATE)		
Groupes	Sans sanctions Période 10	Sanctions monétaires Période 20	Sans sanction Période 30	Sans sanctions Période 10	Sanctions monétaires Période 20	Sans sanction Période 30
1	2.25	0.75	0.5	3	20	10
	(1.7)	(0.5)	(1)	(2.5)	(0)	(10.4)
2	3.75	8	1.5	1.75	0.5	0.5
	(4.7)	(2.4)	(2.4)	(2)	(0.5)	(0.5)
3	10	20	5	2.5	15.75	5
	(8.1)	(0)	(10)	(2)	(3.4)	(5.7)
4	3	20	0.75	6.75	6.5	1.5
	(2.4)	(0)	(0.9)	(9)	(9.2)	(2.3)
5	5	9.5	0	3.75	9.75	5.5
	(10)	(0.5)	(0)	(2.5)	(5)	(4.4)
6	1	15	2.5	8.5	19	2.5
	(1.4)	(10)	(3.3)	(6.4)	(1.4)	(2.8)
Moyenne	4.16	12.2	1.7	4.375	11.91	4.166
	(4.71)	(2.23)	(2.93)	(4.06)	(3.25)	(4.35)
Note : les écarts types sont en parenthèse						

Le graphique 1 montre l'évolution dans le temps des contributions des groupes pour l'ensemble des sessions.



Graphique 11.1 : contribution totale des groupes pour les sessions avec sanctions monétaires (données américaines et françaises cumulées)

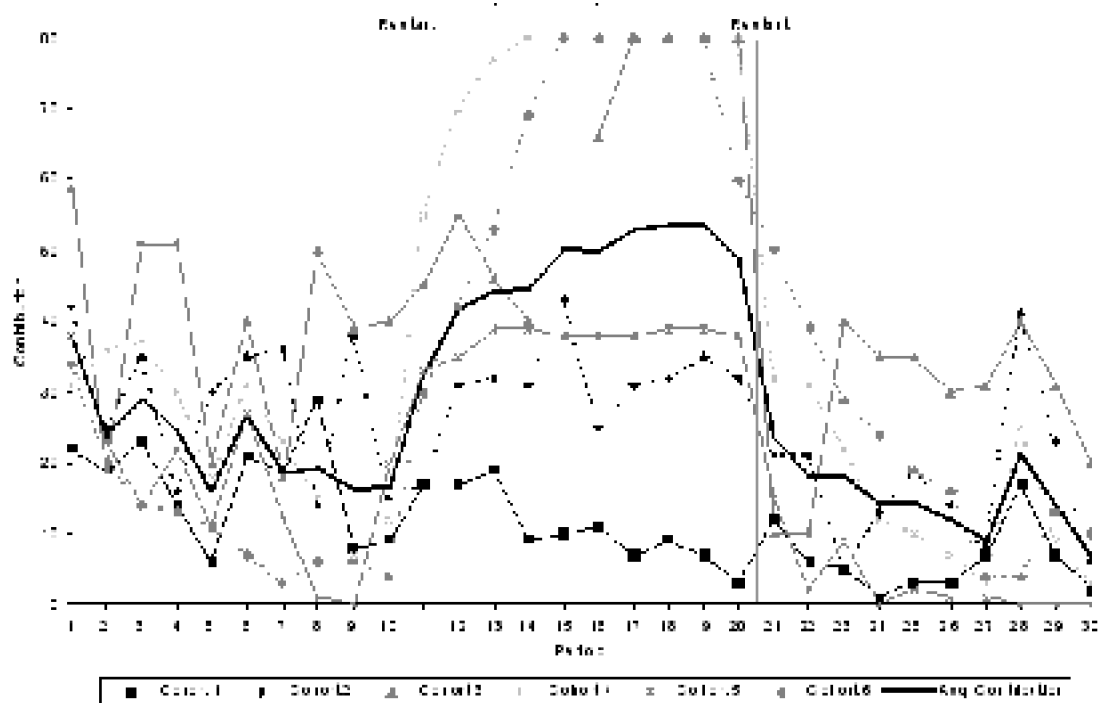
Le niveau moyen de contribution de groupe pour les données française et américaines cumulées est environ de 24 ECU sur l'ensemble des périodes (1) à (10) et de 44 ECU sur l'ensemble des périodes (11) à (20). L'opportunité d'attribuer des points de sanction monétaire accroît de façon significative le niveau moyen de contribution. Dans 9 groupes sur 12 le niveau de contribution augmente lorsque les sujets ont l'opportunité de sanctionner les autres membres de leur groupe. D'après le test non paramétrique *Wilcoxon signed rank*, les niveaux de contribution pour les périodes (11) à (20) sont significativement plus élevés que pour les périodes (1) à (10), ($p < 0.05$). Si l'on compare maintenant le traitement MP (périodes 11-20) aux deux traitements SP (périodes 1-10 et 21-30), il apparaît que les sujets contribuent deux fois plus en traitement MP qu'en SP. Cette différence est significative ($p < 0.05$).

L'effet des sanctions monétaires est immédiat. Ainsi, le niveau moyen de contribution de groupe augmente significativement entre les périodes (10) et (11), ($p < 0.01$) pour atteindre un niveau similaire à celui de la période (1), [(11)=(1), $p > 0.1$]. Il convient toutefois d'être prudent quant à l'interprétation de ce résultat. En effet, il faut tenir compte de l'effet « *redémarrage du jeu* », selon les termes employés par Andreoni (1988)¹⁷. En dernière période du traitement MP (période 20), la moitié des groupes contribue 15 ECU ou plus au bien public. La comparaison des périodes (10) et (20) montre que les niveaux de contribution en période (20) sont significativement plus élevés qu'en période (10) ($p < 0.01$).

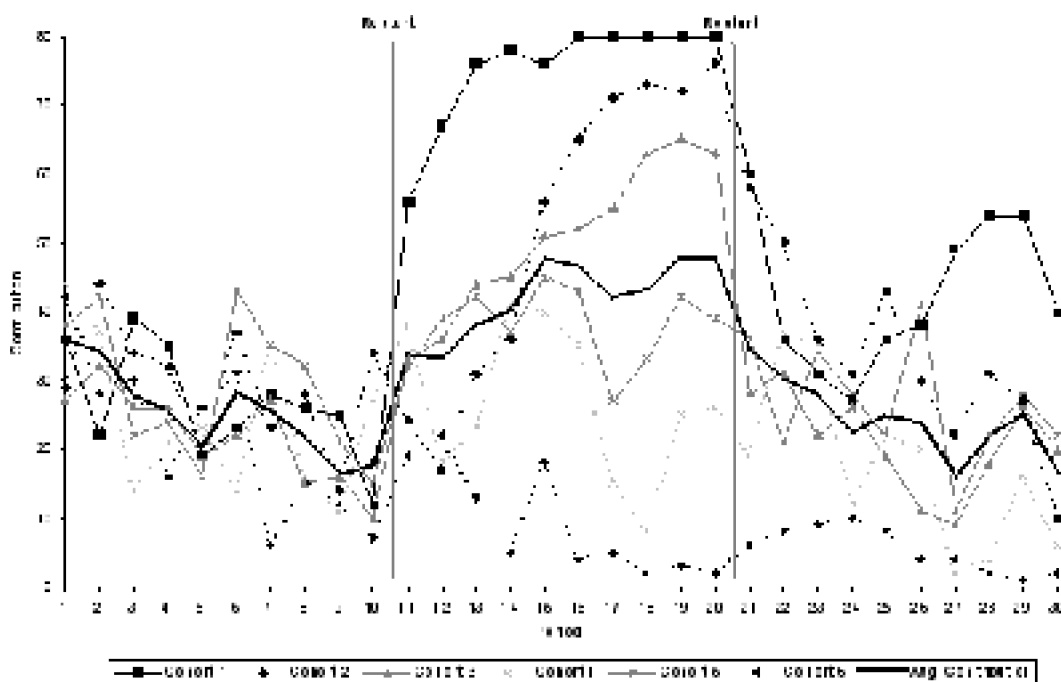
¹⁷ Andreoni (1988) s'est intéressé à l'impact du changement de session dans un jeu et met en évidence un « effet stratégique », positif sur le niveau de coopération. Croson (1996) obtient des résultats similaires.

On observe que le niveau de contribution chute de façon significative lorsque l'opportunité de sanction monétaire disparaît, $[(21-30) < (11-20), p < 0.01]$. Les niveaux de contribution pour les périodes (21) à (30) ne sont pas significativement différents des niveaux de contribution pour les périodes (1) à (10), $[(21-30) = (1-10), p > 0.05]$. Lorsque les sanctions disparaissent, les sujets adoptent de nouveau un comportement non coopératif. Donc l'effet des sanctions monétaires ne semble pas être persistant dans le temps. Ce résultat met en évidence la fragilité de la coopération. En effet, il apparaît que les agents sont potentiellement coopérateurs et que cette coopération dépend fortement de l'environnement dans lequel évoluent les sujets.

Existe-il des différences significatives entre les observations françaises (GATE) et américaines (Purdue) pour les sessions avec traitement MP ? Les graphiques 2 et 3 montrent l'évolution dans le temps des contributions des groupes ayant participé à l'expérimentation à Purdue et au GATE.



Graphique 11.2 : contribution totale des groupes pour les sessions avec sanctions monétaires (Purdue)



Graphique II.3 : contribution totale des groupes pour les sessions avec sanctions monétaires (GATE)

La comparaison des niveaux moyens de contribution de groupe sur l'ensemble des périodes (1) à (10) ne montre pas de différence significative entre la France et les Etats-Unis (d'après le test de Wilcoxon Mann-Whitney [(1-10,F)=(1-10,US), $p>0.1$]). Les comparaisons entre pays sur les périodes (11) à (20) et (21) à (30) donnent des résultats identiques [(11-20,F)=(11-20,US), $p>0.1$; (21-30, F)=(21-30, US), $p>0.1$]. Il n'apparaît pas non plus de différence significative entre la France et les Etats-Unis si l'on considère uniquement la première période du jeu [(1,F)=(1,US), $p>0.1$] ou les dernières périodes des trois traitements [(10,F)=(10,US), $p>0.1$; (20,F)=(20,US), $p>0.1$; (30,F)=(30,US), $p>0.1$]. Il n'existe donc pas de différences significatives entre les deux pays pour le traitement MP.

Considérons maintenant la régression suivante :

(4)

où t représente la période (de 1 à 30), C_j est la contribution totale du groupe j à la période t , et une variable *dummy* en cas d'opportunité de sanctionner.

α est le coefficient qui mesure l'évolution du niveau des contributions dans le temps sur l'ensemble des périodes. On devrait observer un coefficient

négatif si le niveau des contributions diminue dans le temps comme cela a déjà été observé dans un certain nombre d'études expérimentales de jeu de bien public répété sur plusieurs périodes. β mesure l'effet des sanctions sur le niveau de contribution, indépendamment de l'effet trend. Un coefficient positif traduit le fait que le niveau des contributions augmente en présence de sanction. Enfin, le coefficient mesure le trend du niveau des contributions individuelles durant les 10 périodes du traitement avec

opportunité de sanction. Un coefficient significativement positif signifie que la diminution du niveau des contributions individuelles (si $\beta_0 < 0$) est plus modérée dans le traitement avec sanction que dans les traitements sans sanction [périodes (1-10) et (21-30)]. Si β_1 , alors le trend des contributions dans le traitement avec sanction est positif. L'estimation de cette régression pour le traitement avec sanctions monétaires est donnée dans le tableau 5 (données cumulées) et dans le tableau 6 (données non cumulées).

Tableau II.5: Régression sur le niveau des contributions (données françaises et américaines cumulées)

Données cumulées	Variable expliquée : Contribution totale du groupe		
	MP	NP	NS
Constante (β_0)	26.854*** (2.058)	31.153*** (1.842)	9.294*** (0.348)
β_1 trend sur l'ensemble des périodes	-0.317*** (0.110)	-0.640*** (0.099)	-0.270*** (0.021)
β_2 effet des sanctions	11.638*** (3.742)	18.057*** (3.344)	1.688*** (0.644)
β_3 trend sur les périodes 21-30	1.998*** (0.575)	-0.620 (0.515)	-0.391*** (0.107)
R^2	0.292	0.263	0.424
F	48.759	38.699	72.127
Observations	359	330	298
***significativité au seuil de 1% , ** : significativité au seuil de 5%, * : significativité au seuil de 10% . Les écarts types sont entre parenthèses			

Tableau II.6: Régression sur le niveau des contributions (données non cumulées)

Données non cumulées Variable expliquée : Contribution totale du groupe				
	Purdue		GATE	
	MP	NP	MP	NP
Constante (α_0)	26.191*** (2.937)	28.007*** (2.761)	27.562*** (2.824)	34.927*** (2.000)
(α_1) trend sur l'ensemble des périodes	-0.466*** (0.157)	-0.431*** (0.148)	-0.172 (0.151)	-0.880*** (0.107)
(α_2) effet des sanctions	15.725*** (5.348)	20.451*** (5.012)	7.594* (5.128)	15.184*** (3.632)
(α_3) trend sur les périodes 21-30	2.261*** (0.821)	0.109 (0.772)	1.736** (0.790)	-1.500*** (0.560)
R^2	0.389	0.289	0.203	0.405
F Statistic	37.120	23.802	14.965	33.082
Observations	179	180	180	150
***significativité au seuil de 1% , ** : significativité au seuil de 5%, * : significativité au seuil de 10% . Les écarts types sont entre parenthèses				

Considérons dans un premier temps uniquement les estimations pour le traitement MP. La constante est positive et significative. Cela montre que le niveau initial des contributions individuelles est non nul. Le coefficient est significativement négatif pour les données cumulées (Purdue et GATE), ce qui illustre la décroissance des contributions dans le temps généralement observée dans les expériences menées sur des jeux de biens publics répétés. Le coefficient est significativement positif, ce qui traduit que le niveau des contributions individuelles augmente en présence des sanctions monétaires. Le coefficient est également significativement positif pour les sessions MP. Il indique que le niveau de contribution augmente entre les périodes 11-20 avec possibilité de sanction monétaire.

L'étude du traitement MP a donc mis en évidence que les résultats observés dans la session avec sanction monétaire sont semblables à ceux obtenus par Fehr et Gächter (2000). Ils montrent que les sujets s'engagent dans des activités de sanctions coûteuses et que l'opportunité de sanction amène les agents à augmenter leur effort de contribution. Une explication possible des niveaux de contribution élevés est que les sujets souhaitent éviter les conséquences monétaires des sanctions (*effet direct des sanctions*). Cette explication est-elle la seule explication du comportement de contribution des sujets ? Une façon de vérifier s'il existe ou pas d'autres explications consiste à étudier le traitement NP. En effet, si l'*effet direct des sanctions* est la seule explication possible pour expliquer le niveau élevé des contributions individuelles dans les traitements avec sanction, on ne devrait alors pas observer de différences significatives en terme de contribution dans les traitements NP et SP.

Contrairement aux prédictions théoriques, on observe que les sanctions non monétaires ont également un effet positif sur le niveau des contributions. Donc *l'effet direct des sanctions* ne permet pas à lui seul d'expliquer les niveaux de contribution observés. La possibilité pour les agents d'exprimer leur désapprobation semble jouer également un rôle majeur dans le niveau élevé des contributions (*effet indirect des sanctions*). Toutefois, les niveaux de contributions observés dans les sessions NP (périodes 10-20) sont significativement plus faibles que dans les sessions MP (périodes 10-20). Donc l'effet indirect des sanctions n'est pas non plus la seule explication possible.

RESULTAT 2 : *Les sanctions non monétaires ont également un effet positif sur les niveaux de contribution. Contrairement aux sanctions monétaires, l'effet des sanctions non monétaires s'atténue au cours du temps.*

Les tableaux 7 et 8 rapportent les niveaux de contribution individuelle moyenne pour la session avec sanction non monétaire.

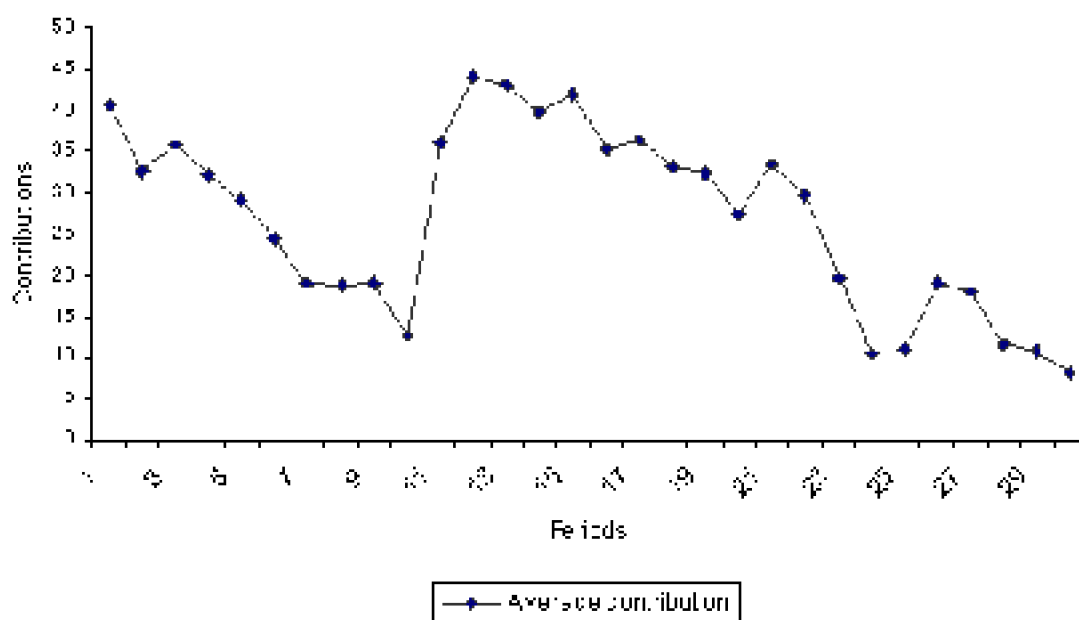
Tableau II.7 : Niveaux de contribution moyenne dans les sessions avec traitement NP

	Sessions avec traitement NP (Purdue) (condition partenaire)			Sessions avec traitement NP (GATE) (condition partenaire)		
Groupes	Sans sanction Périodes 1-10	Sanctions non monétaires Périodes 11-20	Sans sanction Périodes 21-30	Sans sanction Périodes 1-10	Sanctions non monétaires Périodes 11-20	Sans sanction Périodes 21-30
1	3.98 (4.37)	10.90 (4.72)	3.30 (3.35)	7.08 (6.45)	7.70 (7.14)	5.33 (6.20)
2	5.25 (5.91)	3.50 (3.75)	1.38 (2.19)	8.60 (5.09)	6.60 (5.48)	4.80 (5.07)
3	6.05 (4.68)	7.25 (4.07)	5.68 (6.05)	7.08 (6.01)	5.63 (5.52)	3.30 (3.61)
4	3.85 (4.12)	17.68 (5.15)	3.55 (6.26)	6.20 (4.78)	5.83 (3.67)	2.10 (2.48)
5	5.95 (5.53)	10.15 (8.51)	2.30 (3.81)	6.38 (4.68)	9.55 (4.22)	2.35 (2.59)
6	11.60 (6.49)	13.90 (6.98)	4.79 (7.84)			
Moyenne	6.11 (5.81)	10.56 (7.30)	3.50 (6.06)	7.07 (5.46)	7.06 (4.17)	3.58 (4.39)
Ecart type						
Les écarts types sont en parenthèse						

Tableau II.8 : Niveaux de contribution moyenne de dernière période dans les sessions avec traitement NP

Groupes	Sessions avec traitement NP (Purdue) (condition partenaire)			Sessions avec traitement NP (GATE) (condition partenaire)		
	Sans sanction Période 10	Sanctions non monétaires Période 20	Sans sanction Période 30	Sans sanction Période 10	Sanctions non monétaires Périodes 20	Sans sanction Période 30
1	0.5 (1)	11 (8.7)	0.25 (0.5)	1.75 (2.8)	0.25 (0.5)	6.25 (9.8)
2	0.75 (0.9)	5.5 (5.2)	0.25 (0.5)	3 (6)	4.74 (4.4)	0.75 (1.5)
3	1.75 (0.9)	6 (6.4)	7.75 (8.5)	3.25 (3.9)	0 (0)	0 (0)
4	0 (0)	15 (10)	0 (0)	3 (2.9)	3.25 (3.4)	0.25 (0.5)
5	8.25 (9.1)	8 (9.8)	0.5 (1)	1.5 (1.9)	6.25 (6.6)	0.5 (1)
6	9.25 (8.7)	13 (8.8)	7.5 (9.2)			
Moyenne	3.41	9.75	2.708	2.5	2.9	1.35
Ecart type	(3.43)	(8.15)	(3.28)	(3.5)	(2.98)	(2.56)
Les écarts types sont en parenthèse						

Le graphique 4 décrit l'évolution des contributions de groupe dans les sessions cumulées (France et Etats-Unis) avec traitement NP en condition partenaire.



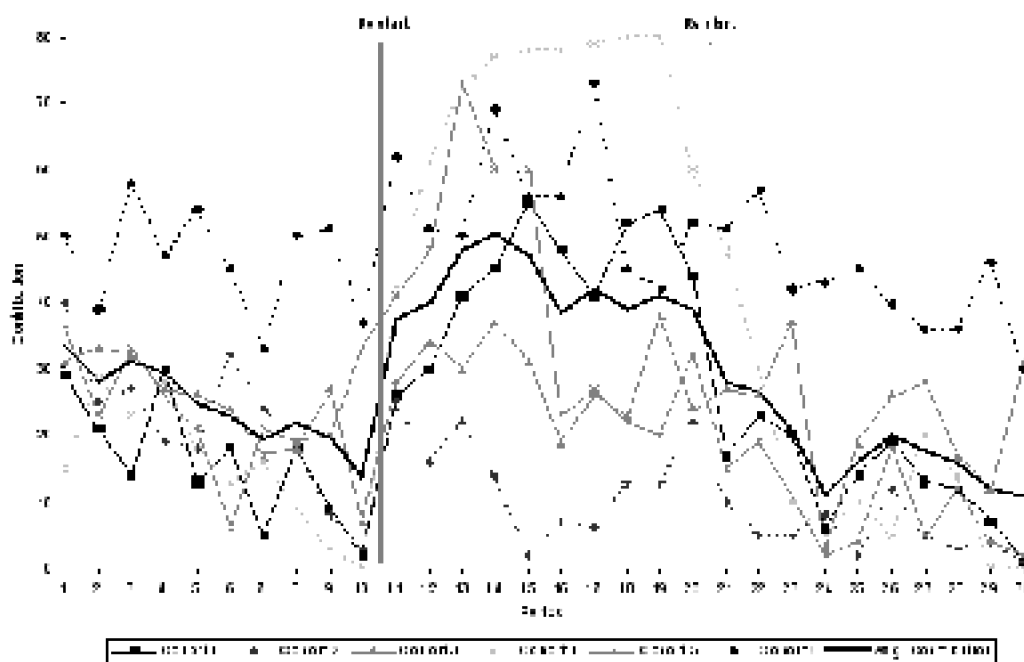
Graphique II.4 : Niveaux de contribution totale du groupe dans les sessions avec sanctions non monétaires en condition partenaire (données américaines et françaises)

cumulées)

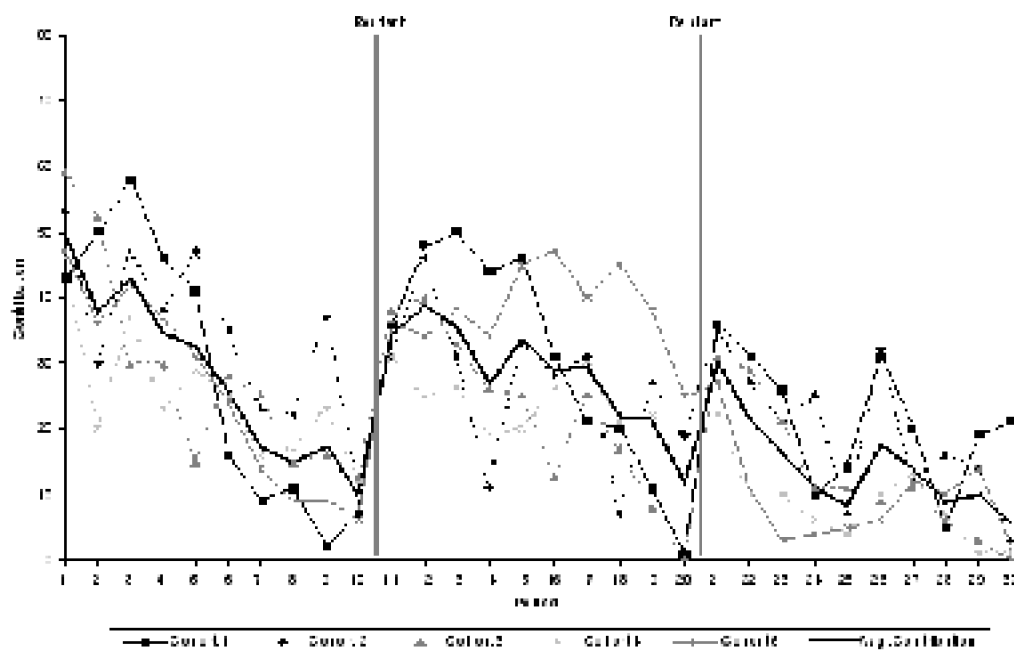
La contribution moyenne du groupe est environ de 26 ECU sur l'ensemble des périodes (1) à (10). Lorsque les sujets ont l'opportunité de sanctionner leurs pairs par l'attribution de points de sanction non monétaire, le niveau moyen de contribution du groupe est de 35.24 ECU. La contribution moyenne du groupe est significativement plus élevée lorsque les agents ont l'opportunité d'assigner à leurs pairs des points de sanction non monétaires, d'après le test *Wilcoxon signed rank* [(1-10, 21-30) < (11-20), $p < 0.01$]. L'effet des sanctions non monétaires est immédiat et accroît significativement le niveau des contributions entre les périodes (10) et (11) ($p < 0.01$). On observe que le niveau de contribution à la période (11) n'est pas significativement différent de celui observé en première période du jeu [(11)=(1), $p > 0.1$].

Comme dans les sessions avec sanctions monétaires, l'effet des sanctions non monétaires n'est pas persistant dans le temps. Ainsi, lorsque les sujets n'ont plus la possibilité de sanctionner les autres membres de leur groupe, le niveau moyen de contribution du groupe diminue. Le niveau de contribution est significativement plus élevé dans les périodes (11) à (20) que dans les périodes (21) à (30) [(21-30) < (11-20), $p < 0.001$]. En moyenne, les sujets contribuent 14.16 ECU sur l'ensemble des périodes (21) à (30).

Les graphiques 5 et 6 décrivent l'évolution des contributions de groupe dans les sessions NP (*partenaire*) conduites respectivement à Purdue et au GATE.



Graphique 11.5 : Niveaux de contribution totale du groupe dans les sessions avec sanction non monétaires en condition partenaire (Purdue)



Graphique II.6 : Niveaux de contribution de groupe dans les sessions avec sanction non monétaires en condition partenaire (GATE)

Existe-t-il des différences significatives pour les sessions en traitement NP entre les observations américaines (GATE) et françaises (Purdue)? Contrairement au traitement MP, il existe des différences significatives sur les données américaines et françaises dans le traitement NP. Les niveaux de contribution dans les périodes (1) à (10) sont significativement plus élevés en France qu'aux Etats-Unis, d'après le test de *Wilcoxon Mann-Whitney* $[(1-10, F) > (1-10, US), p < 0.05]$. Ces différences sont également significatives si l'on considère uniquement la première période $[(1, F) > (1, US), p < 0.05]$. Si l'on compare maintenant les niveaux de contribution dans les périodes (11) à (20), il apparaît qu'ils sont significativement plus élevés aux Etats-Unis qu'en France $[(11-20, F) < (11-20, US), p < 0.05]$. La comparaison entre les Etats-Unis et la France montre que, pour la période (20), le niveau de contribution est significativement plus élevé aux Etats-Unis $[(20, F) < (20, US), p < 0.01]$. La comparaison des deux pays pour les périodes (21) à (30) ne fait pas apparaître de différences significatives entre les pays $[(21-30, F) = (21-30, US), p > 0.05]$.

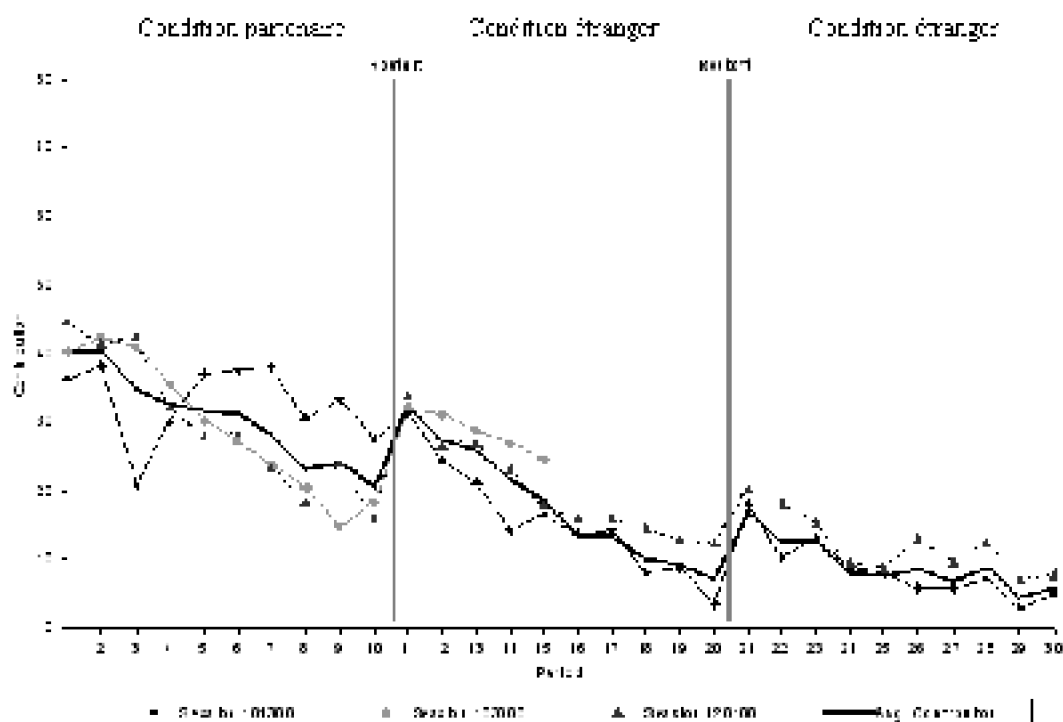
$p > 0.05$]. Il n'y a pas non plus de différence significative en comparant les niveaux de contribution en période 30 [(30,F)=(30,US), $p > 0.05$].

Les niveaux de contribution sont-ils différents entre les sessions avec sanction monétaires et sanctions non monétaires ? Les niveaux moyens de contribution de groupe sur l'ensemble des périodes (1) à (10) avec sanctions monétaires et avec sanctions non monétaires sont respectivement 6.025 ECU et 6.54 ECU. D'après le test de Wilcoxon Man Whitney, il n'apparaît pas de différences significatives (quel que soit le seuil de significativité) entre ces niveaux moyens de contribution ($z = 0.246$). Les niveaux moyens de contribution pour les périodes (11) à (20) avec sanctions monétaires et avec sanctions non monétaires sont respectivement 11.11 et 8.97 ECU. Il n'y a pas non plus de différence significative (quel que soit le seuil de significativité) entre le niveau de contribution moyen dans les périodes (11) à (20) des sessions avec traitement MP et traitement NP ($z = -0.89$). On n'observe pas de différences significatives entre les sessions avec traitement MP et avec traitement NP sur les niveaux de contribution des périodes (21) à (30), ($z = -0.739$).

Les estimations de la régression (4) pour les sessions NP et NS sont données par les tableaux 5 et 6. Comme dans la session avec sanction monétaire, la constante est positive et significative. Cela traduit le fait que le niveau initial des contributions est non nul. Le coefficient est également significativement négatif pour les données cumulées, ce qui illustre la décroissance des contributions dans le temps. Le coefficient est significativement positif pour l'ensemble des sessions avec sanctions non monétaires, ce qui signifie que les sanctions non monétaires ont un effet positif sur le niveau des contributions moyennes. Contrairement aux sessions MP, dans les sessions NP, le coefficient est négatif pour l'ensemble des sessions. Il indique que l'effet des sanctions non monétaires sur le niveau de coopération n'est pas aussi durable que celui de sanctions monétaires. Ainsi, l'effet des sanctions non monétaires s'atténue dans le temps.

RESULTAT 3 : *le niveau de contribution est plus important dans les sessions NP que dans les sessions NS.*

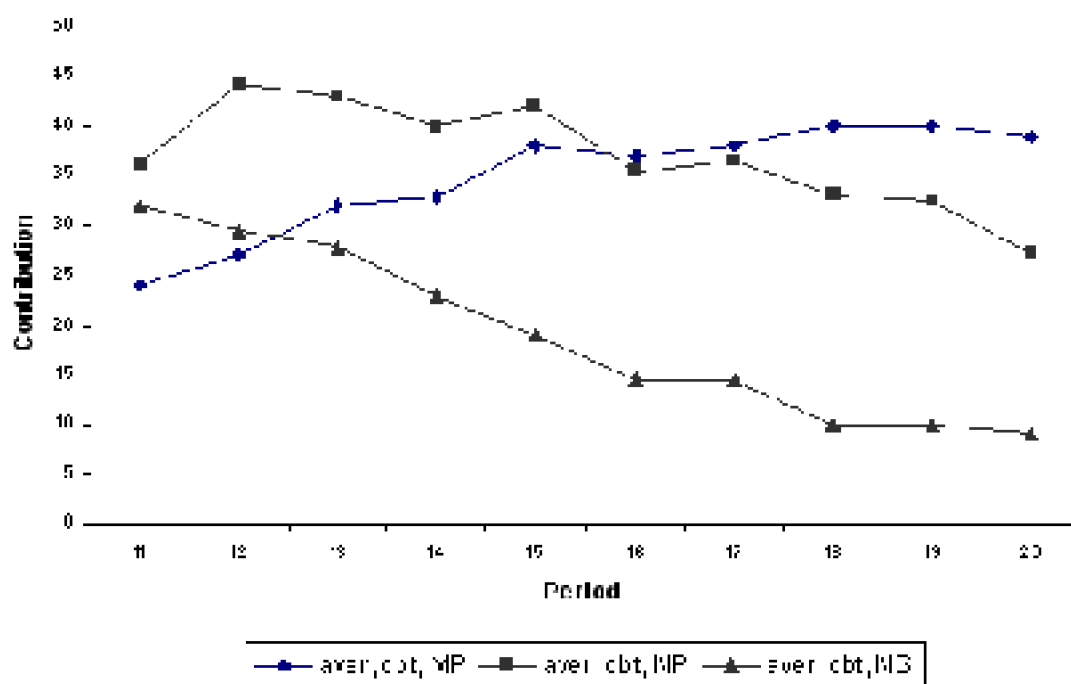
Le graphique 7 montre l'évolution des contributions totales des groupes pour les sessions NS.



Graphique 11.7 : Niveaux de contribution totale du groupe dans les sessions avec sanctions non monétaires en condition étranger (Purdue)

L'intérêt de faire jouer les sujets en condition *partenaire* lors des 10 premières périodes est de pouvoir facilement comparer les comportements de contribution aux sessions NP et de vérifier qu'ils sont identiques. L'inconvénient de ce choix est qu'il rend impossible une comparaison entre les niveaux de contribution lors des périodes 1 à 10 avec les autres périodes du jeu.

Le graphique 8 compare l'évolution des contributions moyennes sur les périodes 11 à 20 pour les trois traitements : MP, NP et NS.



Graphique 11.8 : Niveaux moyens de contribution dans les traitements MP, NP et NS

Le niveau moyen des contributions individuelles sur l'ensemble des périodes 1 à 20 est de 8.97 ECU en traitement NP et de 4.97 ECU en traitement NS. D'après le test non paramétrique Wilcoxon Man Whitney, les niveaux de contribution moyenne sur les périodes 11 à 20 sont significativement plus élevés dans le traitement NP que dans le traitement NS ($z=1.776$, $p<.05$). La comparaison des sessions NP et NS révèle donc davantage de coopération sous la condition *partenaire* que sous la condition *étranger*.

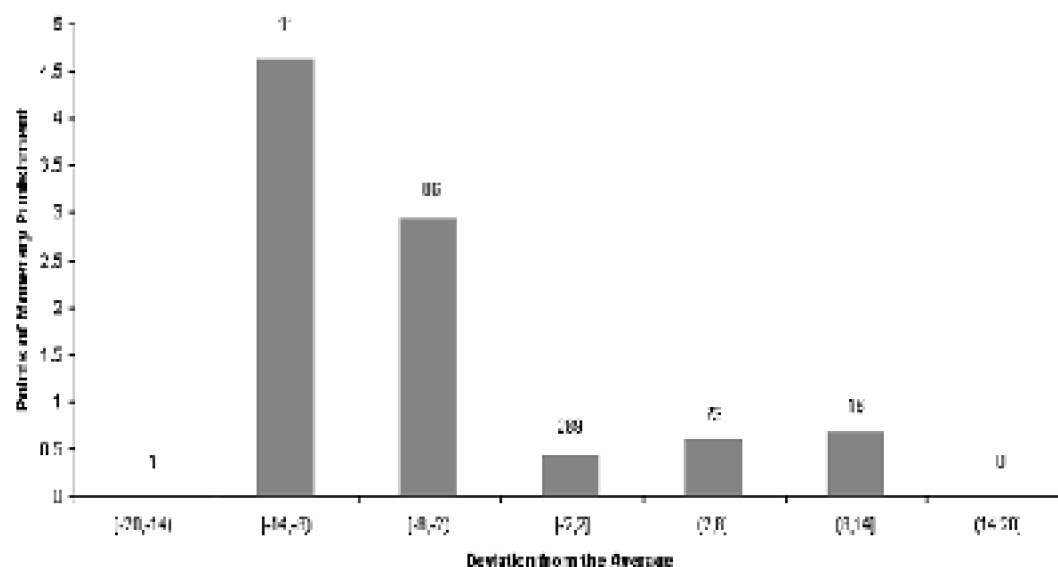
En conclusion de cette sous-section, il apparaît que les sanctions monétaires et non monétaires ont un impact positif sur le niveau de contribution des sujets. Les effets *direct* et *indirect* des sanctions expliquent les niveaux de contribution observés. Ainsi les sujets contribuent davantage afin de ne pas subir les conséquences monétaires des sanctions mais également afin d'éviter la désapprobation de leurs pairs. On observe par ailleurs que l'effet des sanctions non monétaires est toutefois moins persistant dans le temps.

4.2. Caractéristiques des comportements individuels de contribution

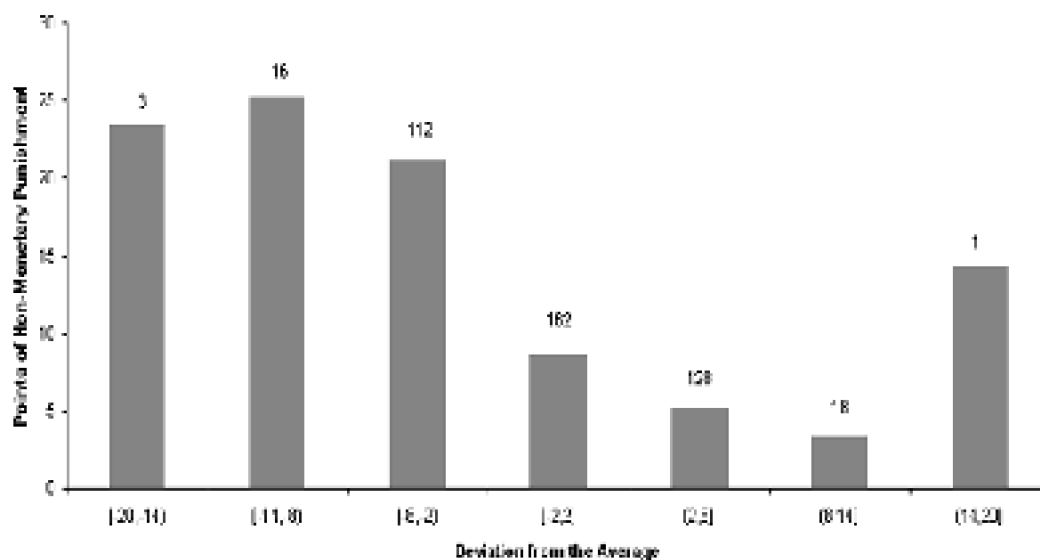
On s'intéresse dans cette sous-section à la relation existant entre les sanctions et les niveaux de contribution individuelle. Fehr et Gächter (2000) ont mis en évidence une relation positive entre le nombre de points reçus par un sujet et la déviation négative de sa contribution vis-à-vis de la contribution moyenne du groupe. Ainsi, les sujets sont d'autant plus sanctionnés que leur niveau de contribution est inférieur à la contribution moyenne des autres membres du groupe.

RESULTAT 4 : *les sujets reçoivent des points de sanctions (monétaires ou non monétaires) si leur contribution est plus faible que la contribution moyenne du groupe.*

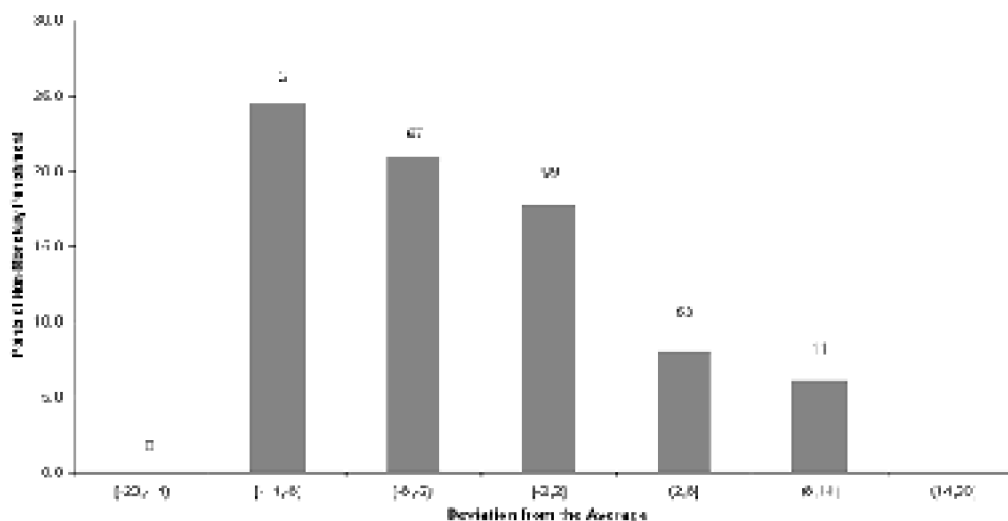
Les graphiques 9, 10 et 11 montrent la relation entre le nombre de points de sanction reçus et la déviation en termes de contribution vis-à-vis de la contribution moyenne des autres membres du groupe pour les sessions MP, NP et NS. Les déviations vis-à-vis de la contribution moyenne apparaissent sur l'axe horizontal du graphique. L'axe vertical donne le nombre de points de sanction reçus. Le nombre au-dessus des barres correspond au nombre d'observations pour chaque intervalle.



Graphique II.9 : Points de sanction monétaire reçus pour une déviation vis à vis de la moyenne Traitement MP



Graphique II.10 : Points de sanction non monétaire reçus pour une déviation vis à vis de la moyenne Traitement NP (condition partenaire)



Graphique II.11 : Points de sanction non monétaires reçus pour une déviation vis-à-vis de la moyenne traitement NS (condition étranger)

Les graphiques 9, 10 et 11 montrent que pour les trois sessions (MP, NP et NS), toute déviation négative vis-à-vis de la moyenne est fortement sanctionnée par les autres joueurs. Ainsi, un joueur qui contribue entre 14 et 8 ECU de moins que la moyenne reçoit en moyenne 4.5 points de sanction monétaire et environ 25 points de sanction non monétaire (partenaire et étranger). A l'inverse, les contributions qui se rapprochent de la moyenne du groupe, c'est à dire dans l'intervalle $[-2, 2]$ sont peu sanctionnées (moins de 0.5 points de sanction monétaires). Les joueurs reçoivent environ 10 points de sanction non monétaire dans les sessions NP et environ 20 points de sanction non monétaires en session NS lorsque leur contribution se trouve dans l'intervalle $[-2, 2]$. Lorsque les joueurs contribuent davantage que la moyenne, ils sont également sanctionnés par les autres membres du groupe, mais dans une moindre mesure qu'en cas de déviation négative. Ainsi un joueur qui contribue plus de 14 ECU que la moyenne du groupe ne reçoit aucun point de sanction dans les traitements MP et NS et moins de 15 ECU dans le traitement NP.

Nous avons étudié la régression suivante afin de tester la relation entre les points de sanction reçus par un sujet i et les déviations vis à vis de la contribution moyenne des autres membres du groupe:

La variable expliquée est le nombre de points que le joueur i reçoit du joueur j . Le coefficient traduit la relation entre le fait que le joueur j contribue moins que la moyenne et le nombre de points que i impose à j . β_2 traduit à l'inverse le fait que j contribue plus que la moyenne. β_3 traduit la relation entre le nombre de points reçus et la moyenne des autres membres du groupe. Le tableau 10 donne les estimations de la régression ci-dessus. Les variables explicatives sont respectivement : « Valeur absolue d'une déviation négative par rapport à la contribution moyenne », « Déviation positive par rapport à la contribution moyenne » et « Contribution moyenne du groupe ». La variable « valeur absolue d'une déviation négative par rapport à la contribution moyenne » est construite de

la façon suivante : cette variable prend la valeur absolue de la déviation négative de la contribution d'un joueur i vis-à-vis de la contribution moyenne lorsque la contribution de i est inférieure à la contribution moyenne. A l'inverse, cette variable prend la valeur zéro lorsque la contribution de i est supérieure à la contribution moyenne.

Tableau II.9: relation entre le nombre de points de sanctions reçus et les niveaux de contribution

	MP [#] Périodes 11-20	NP Périodes 11-20	NS Périodes 11-20	MP [#] Période 20	NP Période 20	NS Période 20
<i>Constante</i>	-2.018*** (.228)	4.76*** (0.219)	6.823*** (0.151)	-3.746*** (1.314)	2.638*** (0.411)	7.527*** (0.6942)
Déviation négative (en valeur absolue)	0.488*** (0.0352)	0.765*** (0.0299)	0.6475*** (0.0458)	0.6501*** (0.228)	0.863*** (0.0698)	0.6529** (0.1534)
Déviation positive	0.0449 (0.0398)	-0.0008* (0.002)	-0.3017*** (0.2977)	-0.1841 (0.344)	0.143** (0.070)	-0.3968*** (0.0843)
Contribution moyenne du groupe	-0.05656*** (0.0155)	-0.2479*** (0.0199)	-0.43946*** (0.03801)	-0.033 (0.06715)	-0.00553 (0.0045)	-0.4153** (0.1172)
	N=1308 Log lik. F. -873.89	N=1200 F[3,1196]= 271.62*** R2 0.405	N=1176 F[3,1172]= 168.25*** R2 0.301	N=132 Log lik. F. -75.75	N=122 F[3,118]= 16208.41*** R2 0.99	N=95 F[3,91]= 21.48*** R2 0.414
Note : ***significativité au seuil de 1% , ** : significativité au seuil de 5%, * : significativité au seuil de 10% . Les écarts-types sont en parenthèses : des estimations TOBIT ont été réalisées pour le traitement MP. Les quatre autres régressions sont des estimations OLS						

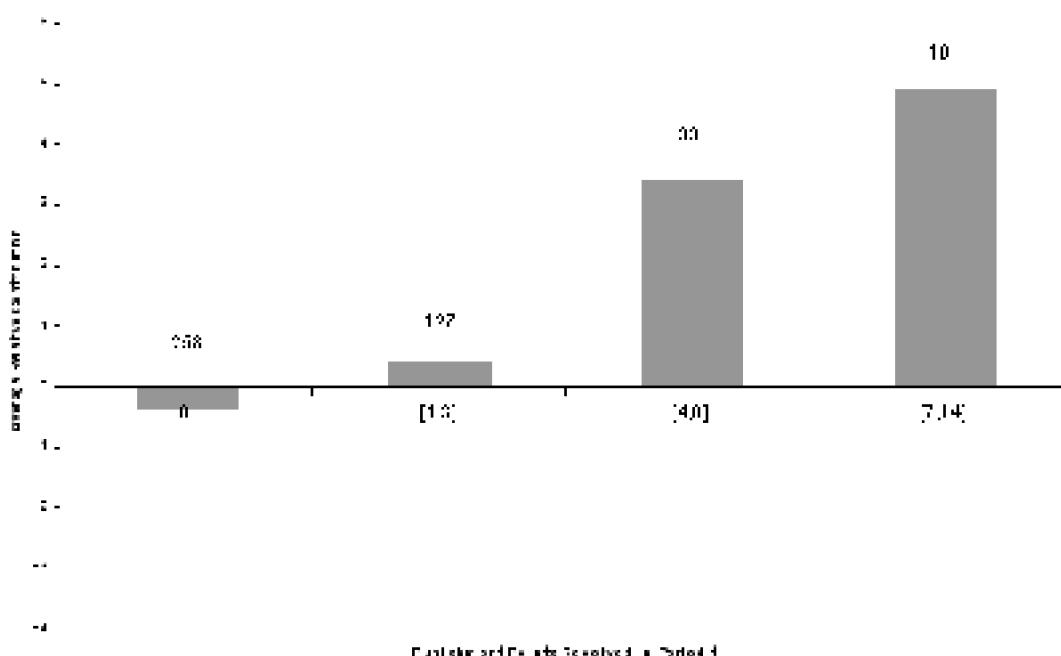
Le coefficient β_1 est positif et très significatif dans les trois traitements (MP, NP et NS). Ainsi les sujets sont fortement sanctionnés par l'attribution de points de sanction (monétaire ou non monétaire) lorsque leur contribution est plus faible que la contribution moyenne du groupe. Le coefficient β_2 n'est pas significatif pour le traitement MP. Les « déviations positives » n'ont donc pas de pouvoir explicatif fort quant à l'attribution des points de sanction monétaire. Pour les traitements avec sanction non monétaire, β_2 est négatif, ce qui signifie qu'un joueur reçoit moins de points de sanction non monétaire lorsqu'il contribue plus que la moyenne. Cette relation est davantage prononcée dans le traitement NS que dans le traitement NP. Le coefficient β_3 est de signe négatif et significatif pour l'ensemble des traitements (sur l'ensemble des périodes). Cela signifie que plus la contribution moyenne du groupe est élevée, moins un sujet reçoit de points de sanction. Ces résultats pour le traitement MP sont similaires à ceux obtenus par Fehr et Gächter (2000).

Les résultats de l'analyse précédente mettent en évidence les deux faits suivants : d'une part, les sujets n'hésitent pas à sanctionner ceux qui contribuent moins que la

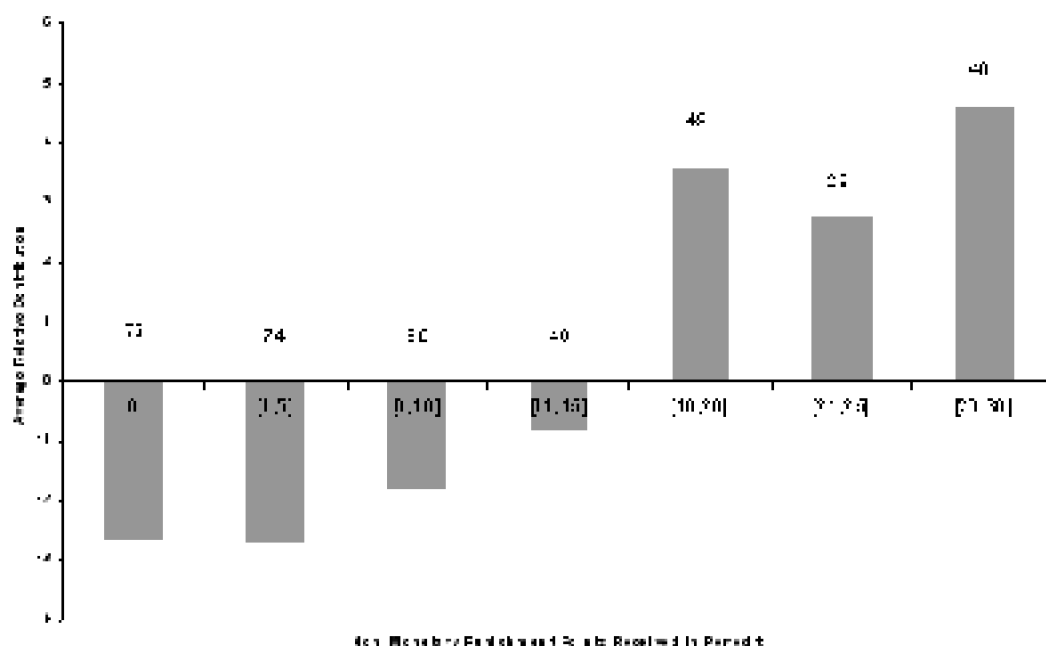
contribution moyenne. D'autre part, les sujets peuvent échapper aux sanctions (monétaires et non monétaires) en contribuant largement au bien public. Ils évitent ainsi les conséquences monétaires et la désapprobation de leurs pairs véhiculés par les points de sanction.

RESULTAT 5 : *Les sujets augmentent leur niveau de contribution en période $t+1$ en réaction aux points de sanction monétaire ou non monétaire reçus en période t .*

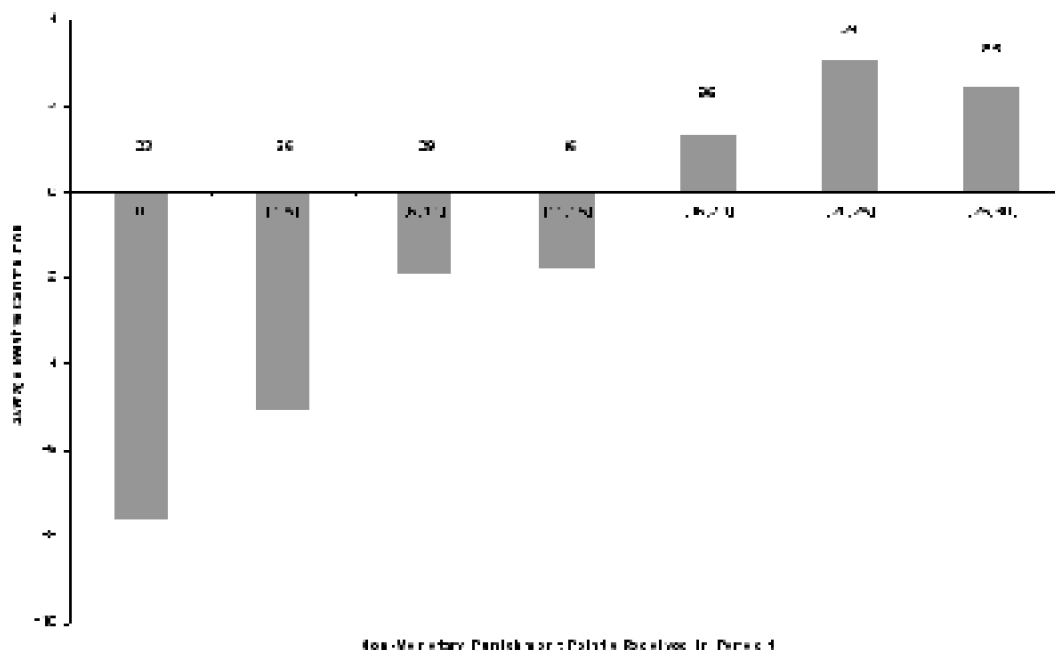
Les graphiques 12, 13 et 14 illustrent la relation entre le nombre de points reçus et la déviation en termes de contribution par rapport à la contribution de celui qui sanctionne. Il apparaît une relation positive entre le nombre de points de sanction monétaire ou non monétaire reçus à une période et l'augmentation de la contribution de celui qui les reçoit à la période suivante. Il apparaît que les sanctions monétaires mais également non monétaires reçues en période t ont un impact positif sur le niveau de contribution des sujets à la période $(t+1)$.



Graphique II.12 : *Contribution relative entre les périodes $(t+1)$ et (t) selon le nombre de points de sanction monétaire reçus en période (t)*



Graphique II.13 : Contribution relative entre les périodes $(t+1)$ et t selon le nombre de points de sanction non monétaire reçus en période (t) (Condition Partenaire)



Graphique II.14 : Contribution relative entre les périodes $(t+1)$ et t selon le nombre de points de sanction non monétaire reçus en période (t) (condition Etranger)

Le graphique 12 montre que les sujets qui reçoivent des points de sanction monétaire en période (t) tendent à augmenter leur niveau de contribution en période $(t+1)$. A l'inverse, ceux qui ne reçoivent pas de points de sanction à une période donnée ont tendance à réduire leur contribution à la période suivante. Dans les traitements NP et NS,

les joueurs qui reçoivent plus de 15 points de sanction, ce qui représente la moitié des points de sanction que peut recevoir un joueur, augmentent leur contribution en période suivante. A l'inverse, ceux qui reçoivent moins de 15 points de sanction en période (t) réduisent leur niveau de contribution en (t+1). Les coefficients de corrélation entre le nombre de points de sanction reçus et la différence de contribution entre les périodes (t+1) et (t) sont respectivement de 0.28 et 0.42 pour les périodes 11-20 des sessions MP et NP.

Soit la régression suivante :

Le coefficient traduit l'effet de la somme totale des points reçus par le joueur i en période t sur sa contribution relative entre les périodes t et $t+1$. Le coefficient β_2 traduit l'effet de l'observation de la contribution moyenne en période t . Le tableau 10 donne les estimations de la régression ci-dessus.

Tableau II.10: Effets des points reçus en t sur la contribution relative entre $t+1$ et t

variable expliquée : $c_i^{t+1} - c_i^t$	MP	NP	NS
Constante	-0.783*	-4.698***	-4.896***
	(0.46619)	(0.866)	(0.9843)
Points (β_1)	0.8359***	0.2957***	0.3380*** ((
	(0.1167)	(0.0322)	(0.03586)
Moyenne (β_2)	0.02966	0.1257*	-0.19816*
	(0.0345)	(0.0663)	(0.1165)
R^2	0.1077	0.181	0.2466
Observations	428	397	344
***significativité au seuil de 1% , ** : significativité au seuil de 5%, * : significativité au seuil de 10% Les écarts types sont en parenthèses			

L'étude du tableau 10 fait apparaître un certain nombre de caractéristiques sur les comportements de contribution des sujets. Le coefficient β_1 est positif et significatif pour l'ensemble des traitements. Il existe donc une relation positive et significative entre le nombre de points de sanction monétaire et non monétaire reçus en période (t) et l'évolution de la contribution de celui qui les reçoit entre les périodes (t) et (t+1). Ainsi les sanctions reçues à une période conduisent les sujets à augmenter leur contribution à la période suivante. Les coefficients β_2 ne sont pas très significatifs pour les trois traitements. Cela signifie que l'observation de la contribution moyenne du groupe n'influence pas leur comportement de contribution.

Comment expliquer ces résultats ? Une explication possible consiste à interpréter les coefficients β_1 et β_2 à partir de la terminologie utilisée par Kandell et Lazear (1992) afin de dissocier la pression des pairs externe de la pression des pairs interne. Alors que la pression externe requiert le regard de l'autre qui montre sa désapprobation suscitant ainsi

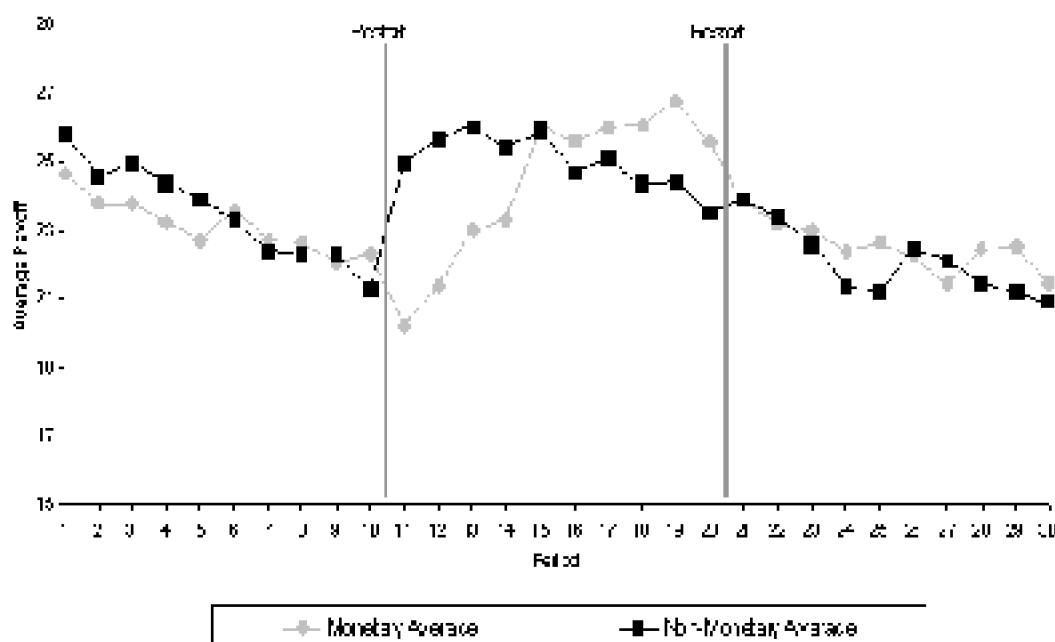
des sentiments de honte, la pression interne est l'internalisation de ce regard. C'est le regard que l'on porte sur soi qui engendre de la culpabilité et l'observation par les pairs n'est pas nécessaire. Les points de sanctions attribués par les pairs sont une pression externe puisqu'ils reflètent la désapprobation. A l'inverse, la comparaison de sa contribution à la contribution moyenne en période (t) traduit le regard que les sujets portent sur leurs propres actions. L'analyse du tableau 10 montre que seule la pression externe joue ici. En effet, les points de sanction incitent les sujets qui les reçoivent à contribuer davantage dans le futur. A l'inverse, la pression interne n'influence pas le comportement des joueurs puisqu'un joueur qui observe la moyenne du groupe n'est pas incité à contribuer davantage à la période suivante. Il ne ressent pas de culpabilité. Donc, seule la pression des pairs externe semble efficace et permet de soutenir la coopération au sein d'un groupe de pairs.

L'analyse de la relation entre les comportements de sanction et de contribution met donc en évidence le fait que les points de sanction monétaire et non monétaire sont attribués de la même façon. Ainsi, les sujets sont d'autant plus sanctionnés qu'ils contribuent moins que la moyenne. En outre, ils augmentent leur niveau de contribution en réaction aux points de sanction qu'ils ont reçus à la période précédente. La sous-section suivante considère l'impact des sanctions monétaires et non monétaires sur les gains des sujets.

4.3. Observations des niveaux de gains

RESULTAT 6 : *Le gain moyen des sujets est plus élevé lorsqu'ils ont l'opportunité de sanctionner les autres sujets, et ceci quelle que soit la nature des sanctions.*

Le graphique 12 illustre l'évolution des gains moyens dans le temps pour les sessions avec sanction monétaire et non monétaire.



Graphique II.12 : Gain moyen par période dans les sessions MP et NP.

Les gains moyens avec opportunité de sanction (périodes 11 à 20) sont plus élevés que les gains sans opportunité de sanction (périodes 11 à 20 et 21 à 30). D'après le test non paramétrique Wilcoxon signed rank, ces différences sont significatives pour les deux traitements MP ($p < 0.05$) et NP ($p < 0.05$). Donc, l'opportunité de sanctionner par l'attribution de points monétaires ou non monétaires accroît considérablement les gains moyens des sujets. Ces résultats s'expliquent par le fait que les sanctions monétaires ou non monétaires conduisent les sujets à accroître leur niveau de contribution, ce qui conduit à augmenter les gains des sujets.

Existe-t-il des différences significatives entre les sessions avec traitement MP et traitement NP ? Les gains moyens dans les périodes (1) à (10) des sessions MP ne sont pas significativement différents de ceux observés dans les sessions NP d'après le test de Wilcoxon-Mann-Whitney ($z = 1.6$; $p > 0.05$). Le gain moyen sur l'ensemble des périodes (1) à (10) est de 23 ECU pour les deux sessions.

Si l'on compare maintenant les gains moyens entre les sessions MP et NP pour les périodes (11) à (20) ils sont respectivement de 24.8 et 25 ECU. Les gains moyens pour ces périodes ne sont pas significativement différents selon les sessions ($z = 0.52$; $p < 0.05$). Toutefois, il apparaît que les gains avec sanction non monétaire sont supérieurs à ceux avec sanction monétaire dans les périodes (11) à (15) et l'ordre s'inverse pour les cinq périodes suivantes. Ainsi, les gains moyens en session NP pour les périodes (11) à (15) sont significativement plus élevés que ceux obtenus avec en session MP ($z = 1.416$; $p < 0.1$). Ils sont respectivement de 22.79 ECU avec sanction monétaire et 25.62 ECU avec sanction non monétaire. Au contraire, les gains moyens sur les périodes (16) à (20), sont plus élevés en session MP qu'en session NP. Ils sont respectivement de 26 ECU dans les sessions MP et de 24 ECU dans les sessions NP. Les gains moyens diminuent de façon similaire dans les dix dernières périodes du jeu. Ainsi, dans les périodes (21) à (30), les gains moyens sont respectivement de 22.5 ECU en sessions MP et 22 ECU en sessions NP. Ces gains ne sont pas significativement différents entre les deux traitements ($z = 1.44$, $p > 0.05$).

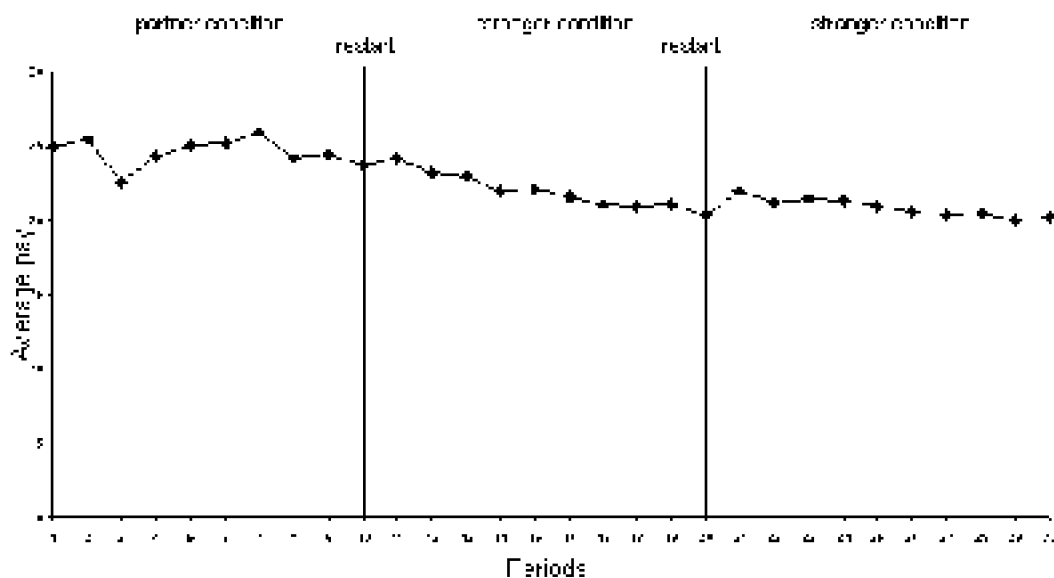
Comment expliquer les évolutions opposées des gains moyens avec sanction monétaire et avec sanction non monétaire ? L'évolution décroissante des gains en session NP entre les périodes (11) à (20) peut s'expliquer de la façon suivante : l'effet incitatif des sanctions non monétaires semble s'affaiblir au cours du temps. L'effet des sanctions non monétaires sur les contributions s'atténue progressivement.

Afin d'expliquer pourquoi les gains moyens en traitement MP augmentent progressivement lors des périodes (11) à (20), il convient d'étudier de façon plus détaillée l'impact des sanctions monétaires sur les gains des sujets. Les sanctions monétaires ont deux effets opposés sur les gains : le premier effet est positif et se traduit par l'augmentation des niveaux de contribution des joueurs qui cherchent à éviter d'être sanctionnés. Le deuxième effet est négatif et se traduit par la réduction des gains de ceux qui sont sanctionnés mais également de ceux qui sanctionnent. Dans un premier temps (périodes 11 à 15), l'effet positif des sanctions monétaires est en partie annulé par l'effet négatif puisque les sujets subissent le coût de discipliner les autres membres du groupe. Dans un deuxième temps (périodes 16 à 20), l'effet négatif des sanctions monétaires

disparaît au fur et à mesure que les sanctions deviennent suffisamment crédibles et n'ont plus besoin d'être appliquées pour que les agents augmentent leurs contributions.

Donc, pour les deux sessions MP et NP, les gains moyens augmentent avec l'opportunité de sanction. Les gains moyens dans les sessions MP et NP ne sont pas significativement différents. Cependant les gains dans les sessions avec sanction (MP et NP) sont significativement plus importants que sans sanction (SP). Du fait de l'existence d'un coût monétaire associé aux points reçus et distribués dans la session avec sanction monétaire, le gain moyen dans les 5 premières périodes de la session avec sanction non monétaire est plus important que dans la session avec sanction monétaire. Cependant, dans les 5 dernières périodes, le gain moyen est plus élevé en session MP qu'en session NP. Alors que dans les sessions NP, l'accroissement des gains moyens est immédiat mais diminue dans le temps ; à l'inverse, dans les sessions MP, l'augmentation des gains moyens est plus lente mais constante dans le temps.

Le graphique 13 illustre l'évolution des gains moyens dans le temps pour les sessions avec sanction non monétaire en traitement « étranger ».



Graphique 11.13 : Gain moyen par période dans les sessions NS.

On observe une décroissance des gains moyens au fil des périodes. Le gain moyen (traitement sans sanction) est de 25 ECU dans les périodes 1-10 et de 22 ECU pour les 11-20 (traitement avec sanction non monétaire) et enfin de 21 ECU pour les dix dernières périodes (traitement sans sanction).

5. Conclusion

Cette étude est une tentative d'identification de l'effet des sanctions sur le niveau de coopération dans une situation de jeu expérimental de contribution volontaire à un bien public. L'expérience réplique celle de Fehr et Gächter (2000) et montre que l'opportunité

de sanctionner monétairement les autres membres du groupe a un effet positif sur le niveau des contributions. Nous obtenons des résultats similaires à ceux obtenus par Fehr et. Gächter (2000). Lorsque les sujets ont la possibilité de sanctionner les autres membres de leur groupe en leur attribuant des points de sanction monétaire, le niveau moyen de contribution augmente considérablement. Deux raisons sont avancées pour expliquer les niveaux élevés de contribution. Les sujets sont incités à contribuer davantage afin d'éviter les conséquences directes monétaires des sanctions (effet direct des sanctions) mais également non monétaires en termes de désapprobation des pairs (effet indirect des sanctions).

Afin d'étudier l'importance relative de ces deux explications, nous avons conduit une session avec des sanctions non monétaires. Les sanctions non monétaires ne modifient pas les gains des agents sanctionnés ni de ceux qui sanctionnent. Nous observons que les sanctions non monétaires ont également un effet positif significatif sur le niveau de contribution des sujets. Donc l'effet direct des sanctions ne permet pas à lui seul d'expliquer la totalité du niveau des contributions puisque la coopération peut également être obtenue avec des sanctions non monétaires. Si les sanctions non monétaires ont un impact positif sur les niveaux de contribution, leur effet incitatif s'atténue toutefois progressivement au cours du temps. Leur efficacité est par ailleurs fortement conditionnée par la composition du groupe. Ainsi, lorsque la composition du groupe est modifiée à chaque période (traitement étranger), les niveaux de contribution observés sont moins importants qu'en traitement partenaire où la composition du groupe est inchangée à chaque période. L'effet des sanctions monétaires sur le niveau de contribution est plus fort et plus persistant dans le temps que celui des sanctions non monétaires.

L'influence des points de sanction non monétaire sur le niveau de coopération des joueurs exige une explication d'ordre non-stratégique. L'importance attachée aux émotions (honte) apparaît comme une explication possible pour comprendre pourquoi les joueurs augmentent leur niveau de contribution en période $t+1$ lorsqu'ils reçoivent des points de sanctions non monétaires en période t . Selon Kandell et Lazear (1992), les individus seraient sensibles à la désapprobation de leurs pairs qui induirait des sentiments de honte. Ainsi la honte serait un stimulus conduisant les agents qui en font l'expérience à repérer les situations qui l'engendrent afin de les éviter dans le futur. De ce fait, la honte conférerait deux types d'avantage à ceux qui en font l'objet. D'une part, elle alerterait l'individu sur les conséquences négatives de ses actions et l'inciterait à adopter un comportement différent dans le futur. D'autre part, elle permettrait d'atteindre la coopération à moindre coût dans les groupes où elle est susceptible d'apparaître (Bowles et Gintis, 20001). Cette hypothèse ne peut pas être rejetée au regard des résultats obtenus qui montrent que les sujets qui ont contribué moins que la moyenne à une période, sont incités à accroître leur contribution à la période suivante, lorsqu'ils reçoivent des points de sanction non monétaires. La pression des pairs externe (désapprobation des pairs) apparaît donc comme un facteur susceptible de soutenir la coopération au sein d'une équipe de travail.

Une extension possible de ces travaux sur les sanctions non monétaires pourrait consister à dissocier les effets de l'observation de ceux de l'attribution des points de sanction non monétaire. En effet, si les résultats expérimentaux mettent en évidence

l'efficacité de la pression des pairs d'ordre non monétaire, ils ne permettent toutefois pas de dissocier les effets respectifs de l'attribution des points de sanction non monétaires de ceux inhérents à la seule observation des niveaux individuels de contribution des autres membres du groupe. La seule observation des contributions individuelles des autres membres du groupe est-elle un facteur suffisant pour inciter les sujets à contribuer davantage au bien public. L'étude des effets de l'observation des niveaux individuels de contribution sur le niveau de coopération dans un jeu de bien public a fait l'objet d'un certain nombre de travaux récents. Rege et Telle (2001) observent que l'introduction de sanctions sociales consistant à révéler le niveau de contribution des autres sujets accroît significativement le niveau de contribution au bien public. Barr (2001) obtient des résultats similaires. Moir (1998) observe au contraire que la seule observation des comportements de contribution conduit à réduire le niveau de coopération au sein du groupe. Il existe donc une incertitude sur l'effet de la seule observation sur le niveau de contribution des sujets.

Une autre extension serait d'étudier l'efficacité des sanctions non monétaires lorsqu'elles sont coûteuses pour ceux qui les attribuent. Brennan et Pettit (1993) ; Loewenstein (2000) ou Rege et Telle (2001) suggèrent que les sanctions sociales n'ont pas besoin d'être coûteuses pour ceux qui les affectent pour être efficaces. Elster (1998), Coleman (1990) suggèrent quant à eux que les sujets sont d'autant plus sensibles aux sanctions sociales qu'elles sont coûteuses pour ceux qui les attribuent. L'idée est la suivante : lorsque la sanction est coûteuse pour le sanctionneur, la désapprobation est alors d'autant plus forte dans la mesure où la sanction véhicule le sacrifice réalisé par celui qui sanctionne.

Chapitre 3. l'autodiscipline par la menace d'ostracisme

« Dérober un objet sans être vu ne nous délivre pas du regard le plus important, celui que l'on porte sur soi-même » De La Tour

« Pendant le grève des mineurs de 1984...Pour isoler ceux qui refusaient de faire grève, les cinémas et les boutiques furent boycottées et ils furent expulsés des équipes de football. Les jaunes furent ainsi témoins de leur propre mort dans une communauté qui ne les acceptait désormais plus. » Francis, 1985, p269.

1. Introduction

Le mot «ostracisme » provient du grec ancien. Il désignait les morceaux d'argile sur lesquels les citoyens de l'ancienne Athènes indiquaient les individus qu'ils souhaitaient exclure de la cité parce qu'ils constituaient une menace pour la cité. Plus généralement, l'ostracisme est une pratique d'exclusion des individus désapprouvés par leurs pairs. Cette pratique joue un rôle important dans le respect des comportements socialement approuvés au sein de la plupart des groupes. Ainsi, dans certaines sociétés primitives, des individus peuvent être châtiés et mis à l'écart d'un village s'ils ne se conforment pas aux règles informelles préalablement établies. Lin (1990) donne un autre exemple de

l'ostracisme appliqué aux équipes de production dans l'agriculture chinoise. Jusqu'à la fin des années soixante-dix, l'agriculture chinoise était organisée en système de fermes collectives, ce qui rendait difficile le contrôle formel par supervision. Une façon privilégiée d'inciter les membres des équipes de production à fournir un effort suffisant consistait pour les membres de l'équipe eux-mêmes à menacer leurs pairs qui « tiraient au flanc » de représailles informelles sous la forme d'une exclusion du groupe.

L'objet de ce chapitre est d'étudier cette forme particulière de sanction par les pairs qui consiste à exclure un ou plusieurs membres du groupe. Les sanctions d'exclusion sont des sanctions très fortes pour ceux qui en font l'objet puisqu'elle privent les agents des avantages d'appartenir à un groupe. Quelle est l'efficacité de telles sanctions ? D'après Ostrom, Walker et Gardner (1992) ; Tyran et Feld (2002), l'efficacité des sanctions serait positivement corrélée à l'importance de la sanction reçue. La sanction d'exclusion présentée dans ce chapitre repose sur l'utilisation des liens sociaux des agents. Lorsque les agents sont engagés ensemble dans des activités de production, ils sont souvent amenés à partager également des activités sociales comme par exemple jouer dans la même équipe de basket ou manger ensemble le week-end (Spagnolo, 1999). Les activités productives et sociales sont généralement étroitement liées. En effet, tout comportement non coopératif dans une de ces activités peut avoir des répercussions sur l'autre. Ainsi, lorsqu'un agent adopte un comportement non coopératif dans l'activité productive, il peut être, en représailles exclu de l'activité sociale. L'activité productive bénéficie alors des effets incitatifs des sanctions sans en subir directement les coûts puisque ceux-ci sont « déplacés » vers les activités sociales. Un exemple de l'utilisation des activités sociales comme sanction des pairs est celui de la grève des mineurs anglais en 1984. Afin d'isoler les mineurs qui refusaient de suivre le mouvement de grève, leurs pairs les expulsèrent des équipes de football, des cinémas et d'autres associations. Les « briseurs de grève » furent ainsi les témoins de leur propre « mort sociale » dans une communauté qui ne les acceptait désormais plus (Francis, 1985).

L'ostracisme est modélisé dans ce chapitre sous la forme d'un jeu de contribution au financement de bien public avec possibilité d'exclusion¹⁸. Dans un premier temps, les sujets participent à un jeu de bien public classique. Dans un deuxième temps, ils observent les efforts de contribution des autres membres du groupe et ont la possibilité d'exclure un ou plusieurs autres sujets. Seuls les sujets qui n'ont pas été exclus participent à un deuxième jeu de bien public similaire au premier. Pour que les agents soient suffisamment incités à contribuer au premier projet, la sanction d'exclusion doit être une menace crédible, dissuadant les agents d'adopter un éventuel comportement de passager clandestin. Il faut donc que les agents soient suffisamment incités à sanctionner les comportements de « passager clandestin ». Or, exclure un membre de son groupe est coûteux puisque le deuxième projet se trouve amputé des efforts de contribution des sujets exclus. Exclure un membre du groupe réduit donc l'intérêt même de se regrouper, c'est à dire les économies d'agrégation. Cependant, sous certaines conditions, l'ostracisme constitue une menace crédible et la coopération au sein du groupe peut alors être atteinte. Les résultats de l'expérience montrent que les agents contribuent davantage

¹⁸ Buchanan (1965) définit un bien club comme un bien public exclusif. La terminologie de bien public avec possibilité d'exclusion est utilisée ici par extension en référence au jeu expérimental utilisé.

au bien public lorsqu'ils ont la possibilité d'exclure leurs pairs. Par ailleurs, il apparaît que les sujets excluent pour deux raisons essentielles. D'une part, ils décident d'exclure certains membres de leur groupe pour une raison stratégique, afin de les encourager à contribuer davantage à l'avenir. D'autre part, ils décident d'exclure certains de leurs pairs pour des raisons non stratégiques de *fairness*.

La deuxième section présente la littérature existante sur les sanctions d'exclusion. La section troisi me est consacr e   l' tude du mod le et de ses pr dictions th oriques. La section suivante pr sente le protocole exp rimental. La cinqui me section analyse les r sultats exp rimentaux. Enfin, la derni re section pr sente les conclusions et les extensions du mod le.

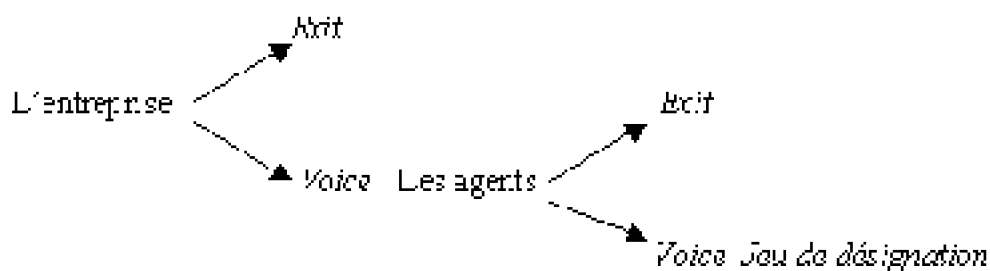
2. La d signation et l'exclusion des passagers clandestins par leurs pairs

Peu de travaux ont  tudi  l'efficacit  des sanctions d'exclusion par les pairs. Deux mod les d'exclusion sont pr sent s dans cette section. Le premier mod le est celui de Cahuc et Kramarz (1997). Ce mod le n'est pas   proprement parl , un mod le d'exclusion mais un jeu de d signation. En effet, les agents n'excluent pas eux-m mes leurs pairs mais se contentent de les d signer au principal qui applique les sanctions d'exclusion. Le deuxi me mod le pr sent  dans cette section est celui d'Hirshleifer et Rasmusen (1989). Les auteurs s'interrogent sur l'efficacit  de l'ostracisme dans une  quipe de travail.

2.1. Jeu de d signation par les pairs: Cahuc et Kramarz (1997)

Cahuc et Kramarz (1997) ont  tudi  les sanctions d'exclusion dans le cadre d'une relation d'agence. Les agents n'excluent pas eux-m mes leurs pairs mais les d signent au principal qui se charge d'appliquer les sanctions. Cahuc et Kramarz comparent deux types de contrat : un contrat de type *Exit* et un contrat de type *Voice*¹⁹. Dans le premier contrat, *Exit*, les agents ont la possibilit  de quitter l'entreprise s'ils trouvent une meilleure proposition d'embauche sur le march  du travail externe. Le deuxi me contrat, *Voice*, illustre la situation dans laquelle le principal abandonne une partie de son autorit  laissant ainsi les agents fixer collectivement le salaire. En contrepartie, les agents s'engagent    tre loyaux, c'est   dire   la fois   ne pas rechercher une autre offre d'emploi sur le march  du travail et   d signer leurs pairs d loyaux. La loyaut  est auto-entret nue puisqu'elle est soutenue par l'exclusion des agents d loyaux.

¹⁹ D'apr s la terminologie d'Hirschman A.(1970).



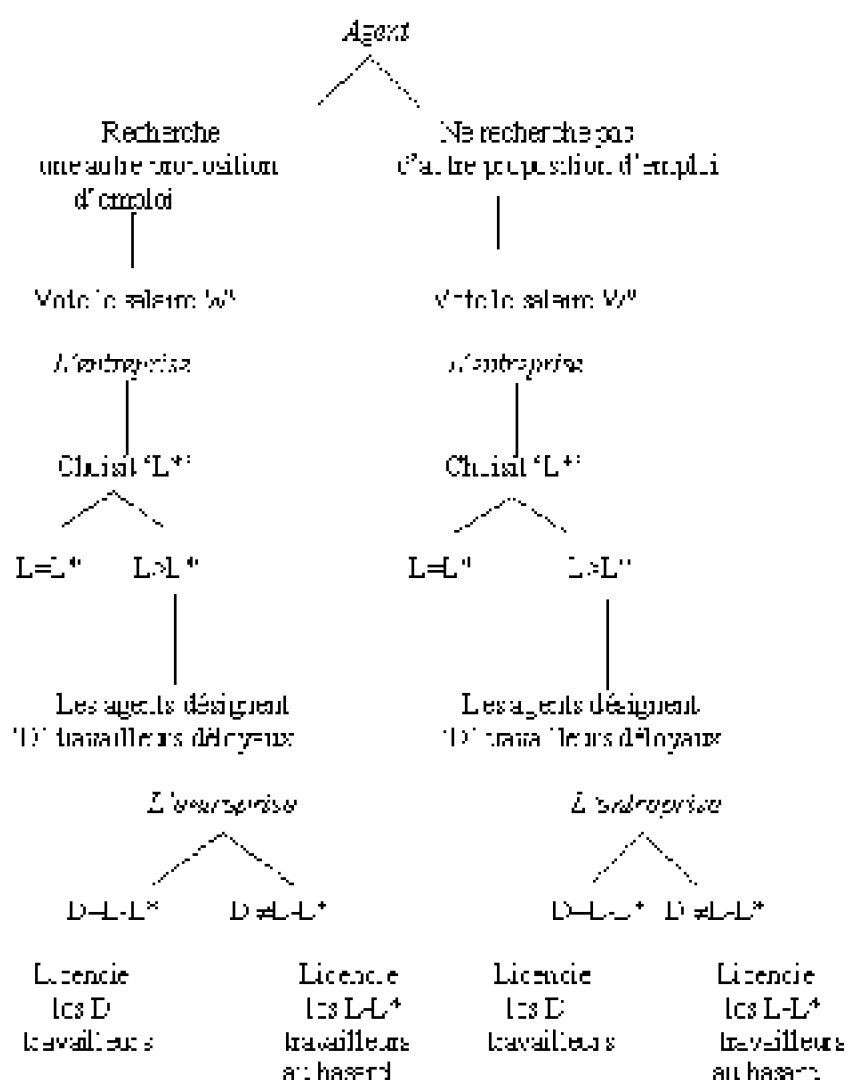
Graphique III.1. L'arbre de décision

(1)

En $t=1$, les agents sont informés de la proposition de salaire du principal. Ils ont alors la possibilité de chercher une autre offre sur le marché du travail. Soit R le coût de cette recherche pour l'agent. Les agents préalablement sélectionnés par le principal qui n'ont pas trouvé d'autre emploi sur le marché, signent alors le contrat avec le principal. Si le principal désire licencier un agent, une fois le contrat signé, il doit lui verser des indemnités Z . La fonction d'utilité des agents employés est une fonction décroissante des coûts de recherche R et fonction croissante du coût d'embauche. Le salaire est corrélé positivement au coût d'embauche. Plus le coût d'embauche augmente, plus l'entreprise est incitée à conserver ses agents. Pour les inciter à rester dans l'entreprise, elle doit augmenter les salaires. Le salaire est donc corrélé positivement aux coûts d'embauche. Soit la fonction d'utilité des agents :

(2)

Enfin, en $t=2$, les agents s'engagent dans la production.



Graphique 11.2: le jeu de désignation

(3)

La fonction de gain de chaque agent est une fonction croissante de Z et de \underline{U} :

La loyauté supporte un équilibre parfait en sous jeu du jeu de désignation. A l'équilibre, tous les agents sont loyaux, c'est à dire qu'ils ne recherchent pas d'autre offre d'emploi sur le marché. En effet, en considérant I agents, si un agent est déloyal, les autres agents vont choisir un salaire collectif tel que le niveau d'emploi annoncé par le principal soit égal à $I-1$. Les $I-1$ agents vont alors désigner unanimement l'agent déloyal qui sera licencié par l'entreprise.

Les deux parties contractantes préféreront un contrat *Voice* si et seulement si les gains associés à ce contrat sont supérieurs à ceux résultant d'un contrat *Exit*, soit :

(5)

avec :

D'après Cahuc et Kramarz, les entreprises qui ont des coûts d'embauche T élevés (du fait par exemple d'investissements en formation importants) sont davantage incitées à proposer un contrat *Voice*. A l'inverse les entreprises ayant une main d'œuvre peu qualifiée et des coûts d'embauche faibles devraient davantage être incitées à signer un contrat *Exit*.

Le modèle de Cahuc et Kramarz (1997) a le mérite de montrer que la loyauté, définie ici comme le fait de ne pas rechercher une autre offre d'emploi sur le marché du travail, peut émerger d'un mécanisme de désignation par les pairs.

2.2. L'exclusion par les pairs: Hirshleifer et Rasmusen (1989)

Hirshleifer et Rasmusen (1989) modélisent l'ostracisme comme une forme particulière de pression des pairs dans le cadre d'une équipe de travail. A la différence du modèle de Cahuc et Kramarz (1997), les agents ne se contentent pas de désigner leurs pairs mais les excluent eux-mêmes du groupe. L'ostracisme est coûteux pour le groupe puisque l'expulsion d'un membre du groupe réduit la taille de l'équipe et donc les gains potentiels pour chacun des agents. Cependant, sous certaines conditions, l'ostracisme peut être une menace crédible et la coopération peut apparaître. Les auteurs considèrent le jeu séquentiel suivant :

Graphique II.3. Les étapes du jeu

Membres	n	n_1	n_2	n_3	n_3	n_T	n_T
Phase	os	pd	os	pd	os	os	os	pd
Etape	1	1	2	2	3	3	T	T

Chaque étape du jeu est composée de deux phases : une phase *os* et une phase *pd*. La phase *os* se déroule de la façon suivante : les agents décident d'exclure ou pas les autres membres de leur groupe. Un agent exclu en phase *os* d'une étape donnée du jeu ne peut participer ni à la phase *pd* de cette même étape ni à la phase *os* de l'étape suivante. S'il est exclu, il reçoit une pénalité Y , telle que $Y > X$, qui traduit le coût de ne plus être un membre du groupe. A chaque étape du jeu, les sujets qui n'ont pas été exclus en phase *os*, participent à la phase *pd*.

En phase *pd* (dilemme du prisonnier), les sujets doivent décider de coopérer ou pas à la production d'un output qui sera partagé à égalité entre les membres du groupe. L'output est d'autant plus important que le nombre de coopérateurs dans le groupe est élevé. Toutefois, le gain individuel d'un passager clandestin est supérieur à celui d'un coopérateur tel que :

(1)

$f(n^C, n)$ est l'output moyen reçu par chaque membre du groupe. n est le nombre d'agent dans le groupe à cette étape du jeu, n^C est le nombre de coopérateurs. La désutilité de coopérer est notée X .

A l'équilibre de Nash parfait en sous jeu, tous les joueurs adoptent une combinaison de stratégie appelée *bannissement* jusqu'à la dernière étape du jeu. Cette stratégie

consiste à coopérer à la production de l'output et à exclure tous ceux qui dévient de l'équilibre. Ceux qui dévient sont ceux qui ne coopèrent pas, ceux qui excluent un coopérateur, ceux qui n'excluent pas un non-coopérateur mais aussi ceux qui n'excluent pas quelqu'un qui n'a pas exclu un non-coopérateur et ainsi de suite. On montre que la combinaison des stratégies de *bannissement* est un équilibre de Nash parfait en sous jeu. Pour cela, on vérifie qu'aucun sujet n'est incité à dévier de la stratégie d'équilibre à chaque sous-jeu. Résolvons le jeu à rebours.

En phase *pd* de la dernière étape du jeu (étape T), personne ne coopère puisque le jeu s'arrête. En phase *os* de l'étape T, les agents décident d'exclure ou pas les autres membres de leur groupe. Puisque l'output n'est jamais réalisé en T^{pd} , quel que soit le nombre d'agents dans le groupe, alors il n'existe plus de coût à exclure. Les sujets sont donc incités à exclure ceux qui n'ont pas coopéré à l'étape précédente. Hirshleifer et Rasmusen (1986) montrent que, du fait de l'indifférence entre exclure ou ne pas exclure, il existe un autre équilibre qui consiste à « ne jamais coopérer », « ne jamais exclure ». Cet équilibre étant pareto inférieur à l'équilibre coopératif, les joueurs vont se positionner sur le second. Les auteurs se réfèrent à l'idée de point focal tout en admettant l'aspect arbitraire de cette démarche. L'idée est la suivante : si un individu ne subit aucun coût à punir un voleur ou à ne pas le punir, il choisira de le punir.

Encadré III.1. Définition du point focal (Thomas Schelling (1960) ; "The strategy of conflict")

Un point focal est un équilibre de Nash qui sera choisi pour des raisons d'ordre psychologiques. Ce qui rend fait d'une stratégie ou d'une combinaison de stratégie un point focal n'est pas chose aisée et dépend du contexte.

Les auteurs considèrent ensuite les étapes $t < T$ et étudient tout d'abord les choix d'un joueur A en t^{pd} selon qu'il ait ou non dévié en t^{os} . En t^{pd} , le joueur A décide de coopérer ou non à la production de l'output. Deux cas doivent être considérés selon que A a dévié ou pas en t^{os} . Si le joueur A a dévié en t^{os} , par exemple s'il n'a pas exclu un non-coopérateur, il est également incité à dévier en t^{pd} . En effet, il sait qu'il sera exclu en $(t+1)^{os}$ et ne subit donc pas de pénalité supplémentaire à dévier en t^{pd} . Si A n'a pas dévié en t^{os} , il est incité à coopérer si la condition suivante est vérifiée :

En t^{os} , le joueur A doit décider de dévier ou non de la stratégie d'équilibre sachant que tous les autres sont à l'équilibre. Dévier à cette étape peut consister par exemple à ne pas exclure un sujet déviant²⁰. Deux cas se présentent selon que le joueur A pense qu'il sera ou pas exclu à cette étape. Si le joueur A pense être exclu à cette étape, son gain en t^{pd} est le même, quelle que soit sa décision en t^{os} . Par contre, s'il dévie en t^{os} , il sera exclu en $(t+1)^{os}$ et ne pourra donc pas participer au jeu en $(t+1)^{pd}$. Donc A choisit de ne pas dévier en t^{os} , s'il pense être exclu à cette étape. Si A pense au contraire ne pas être exclu à cette étape et dévie en t^{os} , il sera exclu en $(t+1)^{os}$. Par ailleurs, s'il dévie en t^{os} , il déviara également en t^{pd} . Donc A ne dévie pas en t^{os} si la condition suivante est vérifiée :

²⁰ Les auteurs montrent qu'une déviation en t^{os} , consistant à exclure un joueur coopérateur, n'est pas non plus profitable puisqu'elle conduit à réduire la taille du groupe.

Les travaux présentés ici mettent en évidence l'efficacité des menaces d'ostracisme comme support de la coopération au sein d'un groupe, que l'application des sanctions se fasse par les agents eux-mêmes (Hirshleifer et Rasmusen, 1989) ou par le principal (Cahuc et Kramarz, 1997). L'étude de l'ostracisme présentée dans ce chapitre s'inspire des travaux d'Hirshleifer et Rasmusen (1989) dans la mesure où la structure du jeu est plus appropriée à l'étude de la pression des pairs dans une équipe de travail. Plus précisément, les menaces de sanctions sont étudiées comme une forme particulière de sanction des pairs appliquée à un jeu de bien public.

3. Les conditions d'un pouvoir dissuasif de la menace d'ostracisme

Le jeu expérimental utilisé afin d'étudier l'efficacité des menaces d'exclusion comme support de la coopération au sein d'une équipe de travail est un jeu de contribution volontaire au financement d'un bien public. Ce jeu apparaît le plus approprié à l'étude des comportements individuels au sein d'une équipe de travail engagée dans des projets communs (voir instructions en annexe A).

Le jeu comprend trois traitements : un traitement sans possibilité d'exclusion (ES) ; un traitement avec possibilité d'exclusion sans coût direct à exclure (ESC) et enfin, un traitement avec possibilité d'exclusion avec coût direct (EC). Dans les deux derniers traitements, les sujets participent à deux projets successifs. Alors que tous les sujets peuvent participer au premier projet, seuls les sujets qui ne sont pas exclus peuvent participer au deuxième projet. L'exclusion sans coût direct signifie que le fait d'exclure un ou plusieurs autres membres du groupe ne réduit pas le gain issu du premier projet de celui qui exclut. Toutefois, les sujets qui excluent leurs pairs subissent un coût indirect à exclure dans la mesure où ils réduisent leur gain potentiel au deuxième projet. En effet, la taille du groupe au deuxième projet est amputée du nombre des sujets exclus. En revanche, l'exclusion avec coût direct signifie que les sujets qui excluent un ou plusieurs autres membres du groupe subissent non seulement le coût indirect inhérent à la réduction de la taille du groupe au deuxième projet mais également un coût direct qui réduit leur gain issu du premier projet.

L'intérêt d'étudier ces trois traitements est multiple. Le traitement (ES) sert de traitement de référence afin de tester l'efficacité des sanctions d'exclusion sur les comportements de contribution des sujets. Le traitement (ESC) permet de tester l'hypothèse d'Hirshleifer et Rasmusen (1989) selon laquelle les menaces d'ostracisme sont des menaces crédibles qui permettent de supporter la coopération au sein d'un groupe d'agents. Les auteurs se réfèrent à la notion de point focal (Schelling, 1960) pour montrer que, lorsque les sujets sont indifférents entre exclure ou non un passager clandestin, ils choisissent de l'exclure. Hirshleifer et Rasmusen admettent toutefois l'aspect arbitraire de cette démarche et mettent en évidence l'existence d'un autre équilibre qui consiste à ne jamais coopérer et à ne jamais exclure. Le traitement (ESC)

permet de vérifier si les sujets vont effectivement se positionner sur l'équilibre coopératif ou non coopératif. Enfin le traitement (EC) vise à étudier la sensibilité des comportements de sanction et de contribution au coût d'exclusion.

3.1. Prédictions théoriques

Cette sous-section décrit en détail les prédictions théoriques des trois traitements expérimentaux.

3.1.1. Traitement sans exclusion (ES)

Dans ce premier traitement, les sujets sont répartis par groupe de 4. Les sujets n'ont pas la possibilité d'exclure les autres membres de leur groupe. Ce jeu est similaire au jeu de bien public sans opportunité de sanction présenté dans le chapitre précédent. Au début de chaque période, les sujets disposent d'une dotation initiale $y = 20$ ECU (*Experimental Currency Unit*). Les sujets doivent décider simultanément de leur contribution au bien public notée g_i ($0 \leq g_i \leq y$). Ils choisissent leur niveau de contribution à l'aide d'un curseur sur l'écran d'ordinateur. Les sujets conservent les ECU qu'ils n'ont pas contribué au projet. Le bien public réalisé est la somme des contributions individuelles de chacun des membres du groupe tel que :

(1)

Une fois réalisé, le bien public est partagé entre les différents membres du groupe. La fonction de gain de chaque sujet i pour chaque période est donnée par :

Le premier terme de la fonction de gain représente la dotation initiale du sujet i . Le deuxième terme représente la contribution de l'agent i au projet. Enfin, le dernier terme représente la part du projet que chaque sujet i reçoit. Le profit total sur l'ensemble des dix périodes est donné par :

Chaque sujet i choisit simultanément le niveau de contribution qui maximise son utilité. La stratégie consistant à ne pas contribuer au projet ($g_i=0$) est une stratégie dominante. L'efficacité au sens de Pareto suppose au contraire que chaque sujet contribue l'ensemble de sa dotation au projet commun ($g_i=y$) afin de maximiser le surplus total $\sum_i g_i$ soit maximisé.

3.1.2. Traitement avec exclusion sans coût direct(ESC)

Le fait d'exclure un ou plusieurs membres du groupe n'est pas directement coûteux pour celui qui exclut dans la mesure où cela ne modifie pas ses gains du premier projet. Il existe cependant un coût indirect à exclure. En effet, exclure un ou plusieurs membres du groupe réduit la taille du groupe contribuant au deuxième projet. En conséquence, les gains potentiels attendus de ce dernier sont réduits puisque le deuxième projet se trouve amputé des contributions individuelles des sujets exclus. Ce coût résulte de la réduction des économies d'agrégation pour le deuxième projet. Les économies d'agrégation traduisent le fait que les gains individuels des sujets sont une fonction croissante du nombre de membres présents dans le groupe, telle que :

Toutefois, la présence d'un sujet qui adopte une stratégie de passager clandestin n'accroît pas l'output (d'où le signe non strictement inférieur de l'inégalité ci-dessus). En effet, les économies d'agrégation sont nulles en l'absence de coopération, c'est à dire que les gains des sujets sont indépendants de la taille du groupe lorsque personne ne contribue au bien public. Le gain de chaque sujet i pour les deux projets dans le traitement ESC s'écrit sous la forme suivante :

$\pi_{i,1}^{esc}$ représente le gain de l'agent i résultant du premier projet et $\pi_{i,2}^{esc}$ son gain résultant du deuxième projet. Le rendement marginal du bien public est identique dans les deux projets. n^{os} est le nombre de sujets qui ont été exclus et ne peuvent donc pas participer au deuxième projet. Le gain total d'un sujet i pour l'ensemble des dix périodes du jeu est donné par :

(6)

Quelle est la prédiction théorique de ce jeu ? On montre que la combinaison de stratégies appelée *bannissement*, qui consiste à contribuer au premier projet un montant non nul²¹, exclure tous ceux qui ne coopèrent pas au premier projet et enfin ne pas contribuer au deuxième projet, forme un équilibre parfait en sous jeu. A l'équilibre, chacun coopère au premier projet. Si un sujet dévie de l'équilibre, alors les autres continuent à coopérer mais excluent le sujet déviant. Dans le deuxième projet, la stratégie de passager clandestin est une stratégie dominante puisque les décisions de contribution à ce deuxième projet ne peuvent faire l'objet de représailles. Il apparaît que le comportement non coopératif dans le deuxième projet permet d'atteindre la coopération dans le premier projet. En effet, si les agents qui n'ont pas été exclus ne contribuent pas au deuxième projet, il n'existe plus de coût indirect à exclure les autres membres du groupe puisque le deuxième projet ne sera jamais financé quelle que soit la taille du groupe. Les économies d'agrégation sont alors nulles pour ce deuxième projet indépendamment du nombre de membres. En l'absence de coût direct à exclure quelqu'un, la menace de l'exclusion devient donc une menace crédible, qui permet de soutenir la coopération dans le premier projet.

Proposition 1: *En l'absence de coût direct à exclure quelqu'un, la coopération peut émerger dans le premier projet.*

Preuve de la proposition 1 : Procédons par induction à rebours afin de montrer que la combinaison de stratégies où tous les sujets suivent les stratégies de *bannissement* est une stratégie d'équilibre. Pour cela, montrons qu'aucun sujet n'est incité à dévier de cet équilibre à chacune des étapes du jeu.

En dernière étape du jeu, les sujets qui n'ont pas été exclus décident de leur niveau de contribution au deuxième projet. La stratégie de passager clandestin est une stratégie dominante dans cette étape. En effet, aucun sujet n'est incité à contribuer au projet dans la mesure où la décision à cette étape du jeu ne peut pas faire l'objet de représailles.

²¹ Il existe en fait une multiplicité de combinaison de stratégies d'équilibre selon la définition de la coopération dans le jeu. Ainsi, lorsque la coopération consiste à contribuer toute sa dotation au premier projet, la combinaison de stratégies d'équilibre consiste à contribuer toute sa dotation au premier projet, exclure tous ceux qui contribuent moins que leur dotation et enfin, ne pas contribuer au deuxième projet. On montre, en annexe B2, que la condition (7) est vérifiée quelle que soit la définition de la coopération.

En deuxième étape, chaque sujet observe la contribution individuelle de chacun des autres membres du groupe et décide de les exclure ou non. Puisque les sujets s'attendent à ce que le deuxième projet ne soit pas produit quelle que soit la taille du groupe, le coût indirect d'exclusion d'un membre est nul. Chaque sujet i est donc incité à exclure tous les sujets qui ne coopèrent pas au premier projet ²².

En première étape, les sujets choisissent de contribuer ou pas au premier projet. On étudie l'incitation pour un sujet i à dévier de la stratégie d'équilibre, c'est à dire à ne pas coopérer au premier projet alors que tous les autres suivent la stratégie d'équilibre. Si un sujet adopte un tel comportement, il reçoit un gain immédiat supérieur à ce qu'il aurait obtenu en coopérant puisqu'il bénéficie de tout le résultat sans coopérer au projet commun. Cependant, ce sujet sera exclu par les autres membres du groupe et ne pourra donc pas participer au deuxième projet. Son profit résultant du deuxième projet est alors nul. A l'inverse, si un sujet coopère au premier projet, son profit en première étape est inférieur à celui obtenu en adoptant un comportement de passager clandestin. Toutefois, il n'est pas exclu et peut donc participer au deuxième projet. Puisqu'à l'équilibre le deuxième projet n'est jamais produit, le profit issu du deuxième projet se résume à y_2 , c'est à dire à la dotation initiale du deuxième projet.

Par conséquent chaque sujet i préfère contribuer au premier projet si la condition suivante est vérifiée:

Le premier terme de l'inégalité représente le profit du sujet i s'il coopère au premier projet et adopte un comportement de passager clandestin au deuxième projet. Le deuxième terme de l'inégalité ci-dessus traduit le gain du sujet i s'il adopte un comportement de passager clandestin au premier projet et est exclu du deuxième projet. Ses gains pour la période se résument donc aux gains issus du premier projet. La condition ci-dessus est-elle toujours vérifiée? Selon l'hypothèse sur les économies d'agrégation, on sait que :

De plus, on sait que $y \geq g_i$. Par conséquent, la condition ci-dessus est toujours vérifiée ²³. Ainsi, les sujets choisissent de coopérer au premier projet. La menace d'exclusion est donc une menace crédible permettant d'atteindre la coopération dans le premier projet. A

²² Il existe toutefois une deuxième stratégie qui consiste à « *ne jamais exclure* ». Quelle stratégie sera choisie par les joueurs ? La sanction d'exclusion peut être utilisée comme menace en cas de non-coopération. Cette menace est crédible puisque la sanction n'est pas coûteuse pour ceux qui l'appliquent du fait de l'existence du deuxième projet non réalisé (cf l'annexe B1). Une preuve de ce résultat, généralement appliquée aux jeux répétés mais qui s'applique également aux jeux séquentiels est donnée dans van Damme, (1991, pp.198-206), Gibbons, (1992, pp82-88), Fudenberg et Tirole, (1986, p167) et Eichberger (1993, pp.224-229). Considérons un jeu séquentiel Γ en deux étapes avec deux équilibres parfaits en sous jeu en deuxième étape, notés E1 et E2, tels que l'équilibre E1 procure un gain plus élevé que E2 (au sens de Pareto). S'il existe en première étape, une issue S qui procure un gain plus élevé à chaque sujet, si tous la choisissent, alors une stratégie d'équilibre du jeu Γ consiste à retenir E1 dans la deuxième étape du jeu si S est choisie dans la première étape, et de choisir E2 en représailles si S n'est pas choisie. La menace est crédible ici puisqu'elle retient un équilibre de Nash dans la deuxième étape du jeu (van Damme, 1991). Ce résultat illustre le fait que lorsqu'il y a deux ou plusieurs équilibres de Nash à la dernière étape d'un jeu séquentiel, la coopération peut être atteinte du fait de la multiplicité d'équilibre.

²³ On suppose ici qu'il n'y a pas de préférence pour le présent ; le facteur d'escompte est tel que : $\delta = 1$.

l'équilibre, chaque sujet i coopère au premier projet et personne n'est exclu. En troisième étape du jeu, personne ne contribue au deuxième projet. La coopération²⁴ peut donc ainsi être atteinte dans le premier projet du fait de l'existence d'une menace crédible de sanction d'exclusion en deuxième étape du jeu. La coopération au premier projet est alors obtenue au prix d'une non-coopération au deuxième projet permettant d'exclure sans coût en deuxième étape du jeu.

Multiplicité d'équilibre

Un problème relatif au traitement ESC, et plus généralement à la théorie des jeux, est que certaines catégories de jeu possèdent de nombreux équilibres. Ainsi l'équilibre où tous les sujets coopèrent au premier projet n'est pas un équilibre unique. En effet, il existe un autre équilibre où personne ne coopère au premier projet, personne n'exclut et personne ne coopère au deuxième projet. L'existence de ce deuxième équilibre provient du problème d'indifférence entre les stratégies d'exclusion ou de non exclusion en deuxième étape du jeu. En effet, dès lors qu'il n'existe plus de coût à exclure un sujet, les agents sont indifférents entre exclure ou non un passager clandestin. Si les agents avaient la possibilité de se coordonner, ils choisiraient donc de se positionner sur l'équilibre pareto dominant. Une preuve formelle est donnée dans Fudenberg et Tirole(1991), (p18-22 ; p175-179) et dans Rasmusen (1994), (p106-107). L'équilibre avec coopération au premier projet est Pareto supérieur. Cet équilibre devrait donc être choisi par les joueurs comme point focal (Hirshleifer et Rasmusen, 1989, p100) (voir annexe B1). Il convient toutefois de souligner l'aspect plus ou moins arbitraire d'une telle démarche comme pour tout raffinement d'équilibre. L'économie expérimentale devrait nous permettre de vérifier sur quel équilibre les agents vont effectivement se positionner

25 .

3.1.3. Traitement avec coût direct d'exclusion (EC)

On a supposé jusqu'à présent que, lorsque les agents décidaient de sanctionner leurs pairs, ils subissaient uniquement un coût indirect en termes de réduction de la taille du groupe participant au deuxième projet. Or, comme on l'a vu dans le chapitre 1, l'effort de sanction peut être directement coûteux, au même titre qu'un effort de contribution. Un traitement avec coût direct d'exclusion a donc été considéré, dans lequel ce coût direct C est proportionnel au nombre de membres qu'un sujet décide d'exclure. Il réduit le gain issu du premier projet du sujet qui exclut un ou plusieurs autres membres du groupe. La fonction de gain du sujet i dans ce traitement s'écrit de la façon suivante :

(9)

Proposition 2 : *lorsque exclure un ou plusieurs autres membres du groupe est*

²⁴ On a supposé qu'à l'équilibre un sujet exclu tous ceux qui ne contribuent pas au premier projet. Cette définition de la coopération est cependant arbitraire. On vérifie en annexe B2 que la condition (7) est vérifiée pour toute définition de la coopération.

²⁵ On a supposé ici que les agents sont engagés dans deux projets successifs. Il est évident que lorsque les agents sont engagés dans N projets successifs avec N suffisamment grand, alors le coût indirect pour le groupe de non coopération dans le projet final devienne marginal comparé aux bénéfices attendus des $N-1$ projets précédents.

directement coûteux, alors la menace d'ostracisme n'est plus une menace crédible.

Vérifions que la combinaison de stratégies qui consiste à ne jamais contribuer au premier projet, ne jamais exclure, ne jamais contribuer au deuxième projet forme un équilibre parfait en sous jeu :

Preuve de la proposition 2:

Procédons de nouveau par induction à rebours :

En dernière étape, les agents qui n'ont pas été exclus choisissent leur niveau de contribution au deuxième projet. Comme dans le traitement ESC, la stratégie de passager clandestin est une stratégie dominante.

En deuxième étape, chaque sujet choisit d'exclure ou non un ou plusieurs autres membres de son groupe. Puisque les sujets s'attendent à ce que le deuxième projet ne soit pas réalisé, le coût indirect de l'exclusion est nul. Cependant, il existe désormais un coût direct à exclure les autres membres du groupe. Si les sujets cherchent à maximiser leur gain pécuniaire issu du premier projet alors, la stratégie dominante consiste à ne jamais exclure les autres membres du groupe en raison du coût direct à exclure. Donc, la menace d'exclusion n'est pas une menace crédible en présence de coût direct à exclure.

En première étape du jeu, les agents décident de leur contribution au premier projet. Ils anticipent qu'ils ne seront jamais exclus s'ils adoptent une stratégie de passager clandestin. Par conséquent, la stratégie dominante consiste à ne pas contribuer au premier projet.

3.2. Le protocole expérimental

L'expérience se compose de 11 sessions. Toutes les sessions expérimentales ont été conduites au GATE. Les sujets sont des étudiants de premier cycle de disciplines diverses. Ces étudiants n'ont jamais participé à d'autres expériences et n'ont jamais suivi de cours de théorie des jeux ni d'économie expérimentale. Chaque sujet participe seulement à une session expérimentale. Les sessions expérimentales ont duré en moyenne une heure et demi et les sujets ont gagné en moyenne 77 francs comprenant un forfait de 15 francs. Les gains et dotations initiales libellés en ECU sont convertis en francs à la fin de l'expérience au taux de conversion de 100 ECU = 4.5 francs. Chaque session est divisée en trois traitements (voir le tableau 1).

Tableau III.1-Traitements et composition des groupes

Traitements Composition des groupes	
Sans exclusion (10 périodes)	11 groupes de taille n=4
Avec exclusion (sans coût direct) (10 périodes)	11 groupes de taille n=4
Avec exclusion (avec coût direct) (10 périodes)	11 groupes de taille n=4

Tous les traitements ont été réalisés en protocole *partenaire*, c'est à dire que la

composition du groupe ne change pas d'une période à l'autre. La dotation que reçoit chaque sujet à chaque période est $y_1=20$ ECU pour le premier projet et $y_2=20$ ECU pour le deuxième projet. Le gain marginal du bien public est fixé à 0,4 dans les deux projets. Dans tous les traitements, les sujets sont publiquement informés qu'un traitement dure 10 périodes mais ne savent pas qu'une session se compose de trois traitements. Après la dernière période du premier traitement de la session, les sujets sont informés qu'ils vont jouer une nouvelle expérience qui va durer également 10 périodes. De la même façon, en dernière période du deuxième traitement, les sujets sont informés qu'ils vont jouer une nouvelle expérience composée de 10 périodes.

Dans les trois traitements les valeurs des paramètres y , n , α ainsi que la contribution totale au premier projet sont de connaissance commune. Dans les traitements ESC et EC, les sujets sont également informés des contributions individuelles au premier projet des autres membres de leur groupe. Pour empêcher la possibilité de formation de réputation individuelle à travers les périodes, la contribution individuelle du sujet apparaît toujours dans la première colonne de son écran d'ordinateur et les contributions des autres sujets apparaissent de façon aléatoire sur les trois autres colonnes. Ainsi les sujets ne disposent pas d'information leur permettant de déterminer l'évolution des niveaux de contribution des autres sujets au cours des périodes du jeu. Par ailleurs, les sujets ne sont jamais informés des décisions individuelles d'exclusion des autres sujets. Ils sont seulement informés s'ils sont exclus ou pas du deuxième projet.

Dans six sessions, les sujets jouent dix périodes du traitement sans exclusion, puis dix autres périodes d'un jeu avec exclusion sans coût direct, et enfin dix périodes d'un jeu d'exclusion avec coût direct. Afin de contrôler un éventuel « effet d'ordre » entre les différents traitements, l'ordre des deux derniers traitements est inversé dans 5 autres sessions. Le coût associé à la décision d'exclure s'élève à 5 ECU par individu exclu dans le traitement avec coût direct d'exclusion.

4. Ostracisme et coopération

Si les sujets suivent les prédictions théoriques du jeu, on devrait observer dans le traitement ES des niveaux de contribution individuels nuls. Dans le traitement ESC, les sujets devraient coopérer à la réalisation du premier projet et ne pas coopérer à la réalisation du deuxième projet. Enfin, concernant le traitement EC, les contributions individuelles observées devraient être nulles dans les deux projets. On ne devrait donc observer aucune différence entre le premier projet du traitement EC et le traitement ES.

4.1. Niveaux de contribution et d'exclusion

4.1.1. Niveaux de contribution sans possibilité d'exclusion

Le niveau de contribution diminue au fil du jeu et tend vers un niveau nul de contribution.

RESULTAT 1 : *Le niveau moyen des contributions converge vers un niveau de contribution nul lorsque les sujets n'ont pas la possibilité d'exclure leurs pairs.*

Les contributions moyennes par groupe sont données dans le tableau 2. Les colonnes 2 et 5 du tableau 2 décrivent respectivement la contribution moyenne sur l'ensemble des périodes et en dernière période du traitement ES.

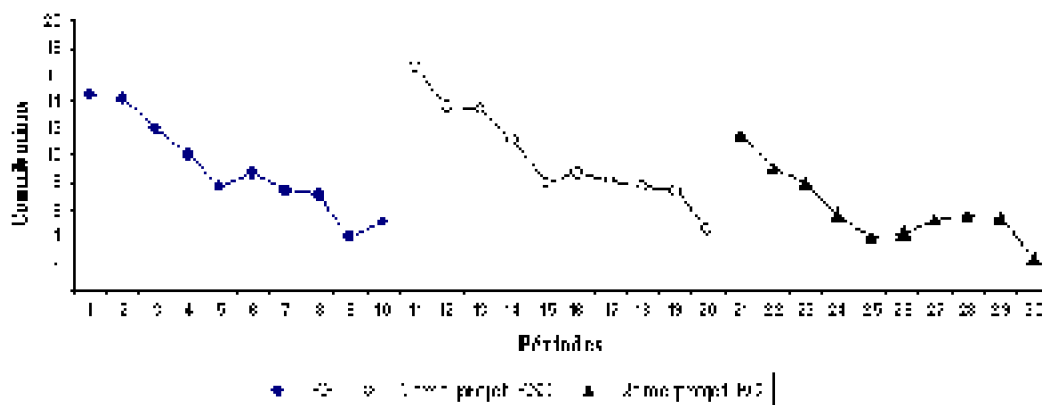
Tableau III.2 : Niveaux moyens de contribution par groupe

Groupes	Contribution moyenne sur toutes les périodes			Contribution moyenne en dernière période		
	ES	Second projet ESC	Second projet EC	ES	Second projet ESC	Second projet EC
1 2 3 4 5	9.725	15 (5.123)	9.85	7.5 (9.57)	5.33	7 (9.45)
6 7 8 9 10	(4.94)	9.84	(4.88) 8.1	0.75 (0.9)	(7.34)	0.5 (1)
11	5.175	(5.98)	(4.08)	3 (2.94)	8.75	0.33
Moyenne	(2.78)	7.18	2.66	9.25 (8.3)	(7.63) 0	(0.5) 0
	7.15	(6.69)	(4.67)	6.25 (7.5)	(0) 0	(0) 5.75
	(3.78)	8.08	4.125	3.75 (4.5)	(0) 8	(9.6) 0
	10.8	(4.26)	(3.37)	6.75	(9.75) 5	(0) 1.75
	(3.76)	12.35	9.75	(6.99) 4.5	(0) E*	(2.87)
	11.52	(3.14)	(2.04)	(4.2) 8.75	E* E*	3.25
	(3.99)	7.06	0.61	(8.057) 5	0 (0)	(3.77) 0
	9.95	(1.89)	(0.84)	(10) 5	0.5 (1)	(0) 5
	(6.19)	1.02	4.075	(10) 5.5	3.44	(10)
	9.47	(0.62) 3.1	(3.33)	(6.41)	(3.22)	6.25
	(3.67)	(2.15) 5	5.55			(9.46)
	9.975	(8.66)	(2.83)			2.71
	(2.71)	6.08	0.79			(4.29)
	6.55	(4.47)	(1.96)			
	(2.08)	3.69	12.28			
	14.35	(3.54)	(6.22)			
	(5.5)	6.845	12.0.75			
	12.325.	(4.22)	(5.51)			
	(4.16) 9.79		6.39			
	(3.96)		(3.61)			

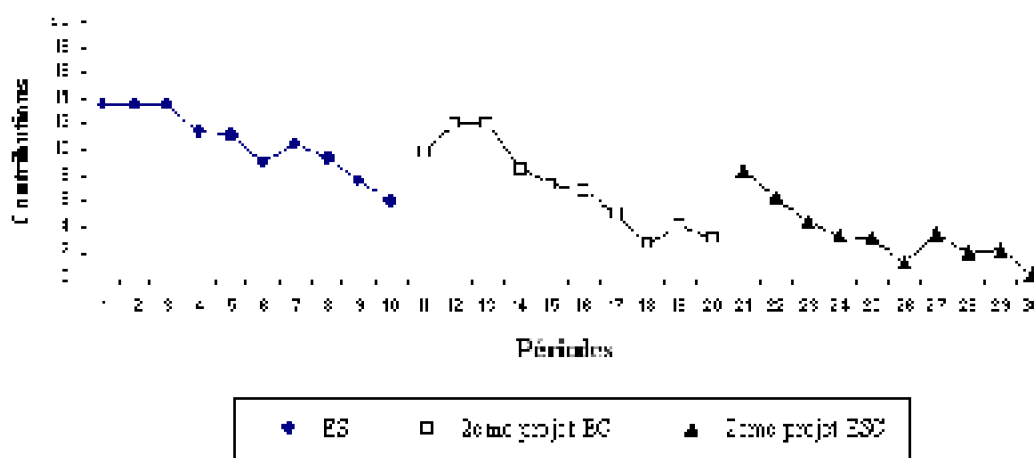
* E signifie que tous les membres du groupes ont été exclus. Les écarts types sont en parenthèse. Dans les sessions 7 à 11 l'ordre des traitement ESC et EC est inversé.

La comparaison des colonnes 2 et 5 indique que la contribution moyenne dans la période finale (5.5 ECU) est inférieure à la contribution moyenne sur l'ensemble des périodes (9.79 ECU) pour le traitement ES. En dernière période, presque 40% des sujets adoptent un comportement de passager clandestin et pour la moitié des sujets, la contribution individuelle est inférieure ou égale à 2 ECU. Les graphiques 1 et 2 décrivent l'évolution du niveau des contributions moyennes des traitements ES, ESC et EC

(uniquement le deuxième projet pour les deux derniers traitements) pour les sessions (1 à 6) et (7 à 11) (voir annexe C pour les contributions de groupe). Ces deux schémas mettent en évidence que la contribution moyenne dans le traitement ES diminue dans le temps et converge vers un niveau de contribution proche de zéro dans les dernières périodes.



Graphique III.1 : Evolution des contributions moyennes sans possibilité d'exclusion (sessions 1 à 6)



Graphique III.2 : Evolution des contributions moyennes sans possibilité d'exclusion (sessions 7 à 11)

RESULTAT 2 : Le niveau de contribution moyen dans le deuxième projet des traitements avec exclusion (EC et ESC) converge vers un niveau de contribution nul.

Les colonnes 3 et 4 du tableau 2 décrivent respectivement le niveau des contributions moyennes du deuxième projet des traitements ESC (6.84 ECU) et EC (6.39 ECU) sur l'ensemble des périodes. En considérant uniquement la dernière période, le niveau de contribution moyen du deuxième projet chute et passe respectivement à 3,44 ECU et 2,71 ECU pour les traitements ESC et EC. D'après le test non paramétrique *Wilcoxon signed*

rank sur les niveaux moyens de contribution, l'hypothèse selon laquelle les niveaux de contribution pour les périodes (1) à (10) sont significativement différents des niveaux de contribution dans le deuxième projet pour les périodes (11) à (20) et (21) à (30), ne peut pas être rejetée pour les sessions 1 à 6, ($p > 0.05$). Si l'on considère maintenant les sessions 7 à 11, le test non paramétrique *Wilcoxon signed rank* montre que les niveaux de contribution moyenne observés dans les périodes (1) à (10) sont significativement plus élevés que ceux observés dans le deuxième projet du traitement EC (périodes 11-20) et ESC (périodes 21 à 30)²⁶.

4.1.2. Niveaux de contributions avec possibilité d'exclusion

On considère ici les niveaux de contribution individuelle dans le premier projet des traitements EC et ESC.

RESULTAT 3 : *Le niveau moyen de contribution du premier projet du traitement ESC est plus élevé que dans le traitement ES.*

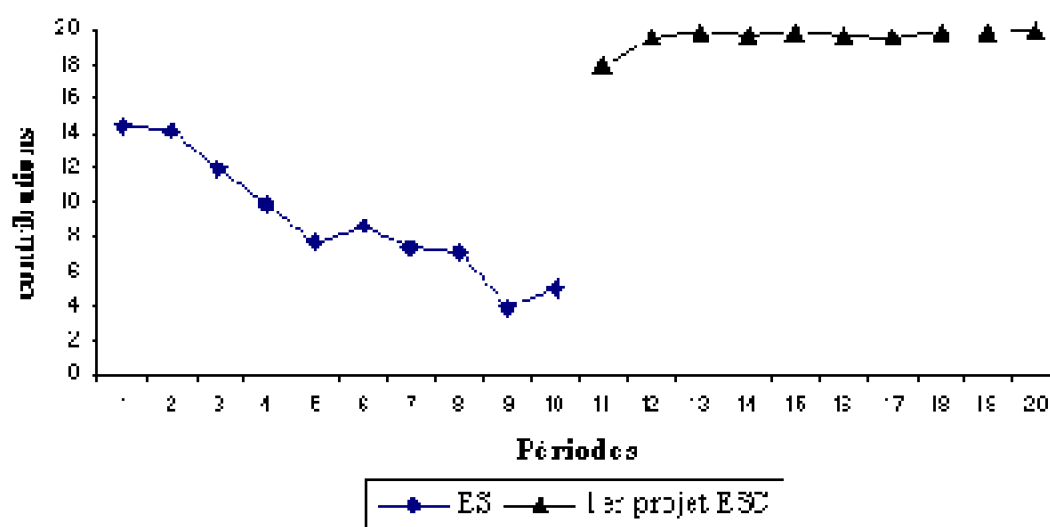
Le résultat 3 montre que la coopération peut apparaître dans le premier projet du traitement ESC. Il confirme les prédictions théoriques selon lesquelles, sans coût direct à exclure, la menace d'exclusion est crédible et supporte la coopération. Le tableau 3 donne les niveaux de contribution moyenne dans le premier projet des traitements ESC et EC.

Table III.3 : Niveaux de contribution moyens

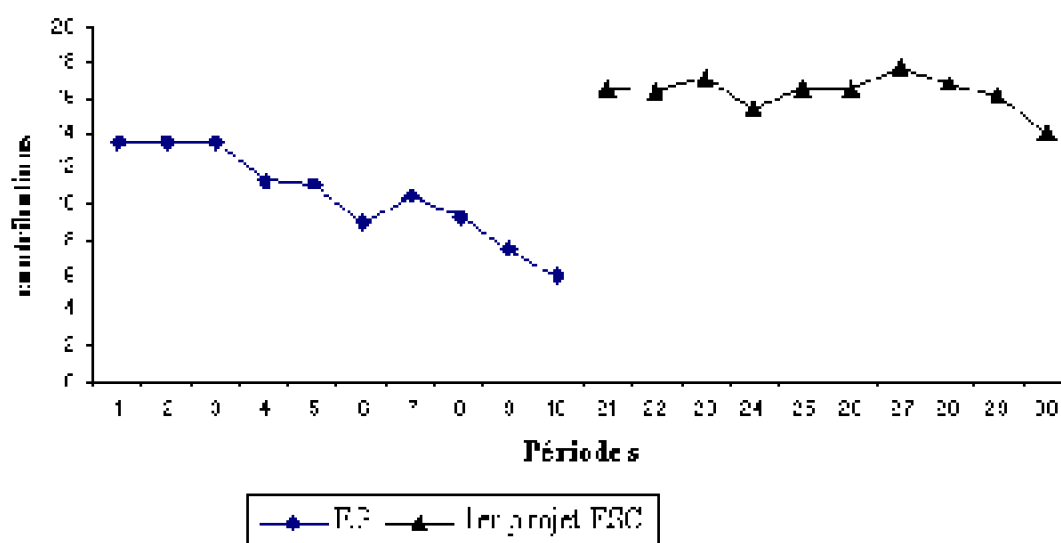
²⁶ Il convient toutefois d'être prudent quant à une possible interprétation de ce résultat. En effet, une comparaison des niveaux de contribution dans le premier et le deuxième projet est problématique puisque le nombre de joueurs participant au deuxième projet n'est pas forcément identique à celui du premier.

Groupes	Contribution moyenne sur toutes les périodes			Contribution moyenne en dernière période		
	ES	1er projet ESC	1er projet EC	ES	1er projet ESC	1er projet EC
1 2 3	9.725	19.725	20 (0)	7.5 (9.57)	20 (0)	20 (0) 19
4 5 6	(4.94)	(0.78)	19.35	0.75	20 (0)	(0.81)
7 8 9	5.175	19.6	(0.39)	(0.9) 3	20 (0)	14.25
10 11	(2.78)	(0.69)	19.3	(2.94)	20 (0)	(9.6) 20
moyenne	7.15	19.4	(1.80) 20	9.25	20 (0)	(0) 20
	(3.78)	(1.65) 20	(3.37) 20	(8.3)	20 (0)	(0) 5
	10.8	(0) 20	(0) 14.5	6.25	10.5	(10)
	(3.76)	(0)	(5.98)	(7.5)	(7.04)	10.25
	11.52	18.675	12.525	3.75	20 (0)	(1.25)
	(3.99)	(1.52)	(1.266)	(4.5)	0 (0)	18.75
	9.95	13.175	16.3	6.75	20 (0)	(0.5)
	(6.19)	(1.09)	(2.10)	(6.99)	20 (0)	13.25
	9.47	19.4	11.3	4.5 (4.2)	17.31	(4.57)
	(3.67)	(0.84)	(1.67)	8.75	(0.64)	18.75
	9.975	9.92	18.55	(8.057) 5		(2.5) 20
	(2.71)	(4.36)	(2.27)	(10) 5		(0) 16.29
	6.55	19.175	19.125	(10) 5.5		(2.65)
	(2.08)	(1.76)	(1.43)	(6.41)		
	14.35	19.85	17.21			
	(5.5)	(0.24)	(1.84)			
	12.325.	17.93				
	(4.16)	(1.175)				
	9.79					
	(3.96)					
Les écarts types sont en parenthèse. Dans les sessions 7 à 11, l'ordre des traitement ESC et EC est inversé.						

Le tableau 3 indique que les sujets contribuent presque la totalité de leur dotation initiale (20 ECU) au premier projet dans le traitement ESC. En dernière période, 9 groupes sur 11 contribuent toute leur dotation initiale au premier projet (colonne 6). La comparaison des colonnes 2 et 3 indique que les sujets contribuent deux à trois fois plus dans le traitement ESC que dans le traitement ES. D'après le test non paramétrique *Wilcoxon signed rank* cette différence est significative ($p < 0.05$) pour l'ensemble des sessions (sessions 1-6 et 7-11). Les graphiques 3 et 4 décrivent l'évolution des contributions moyennes au cours du temps pour les traitements ES et ESC pour les sessions 1 à 6 et 7 à 11. Ils montrent que le niveau de contribution moyen s'accroît brusquement lorsque les sujets ont l'opportunité d'exclure leurs pairs et atteint un niveau élevé (presque 20 ECU) jusqu'à la fin du traitement.



Graphique III.3 : Evolution des contributions moyennes dans le traitement ESC (sessions 1 à 6)



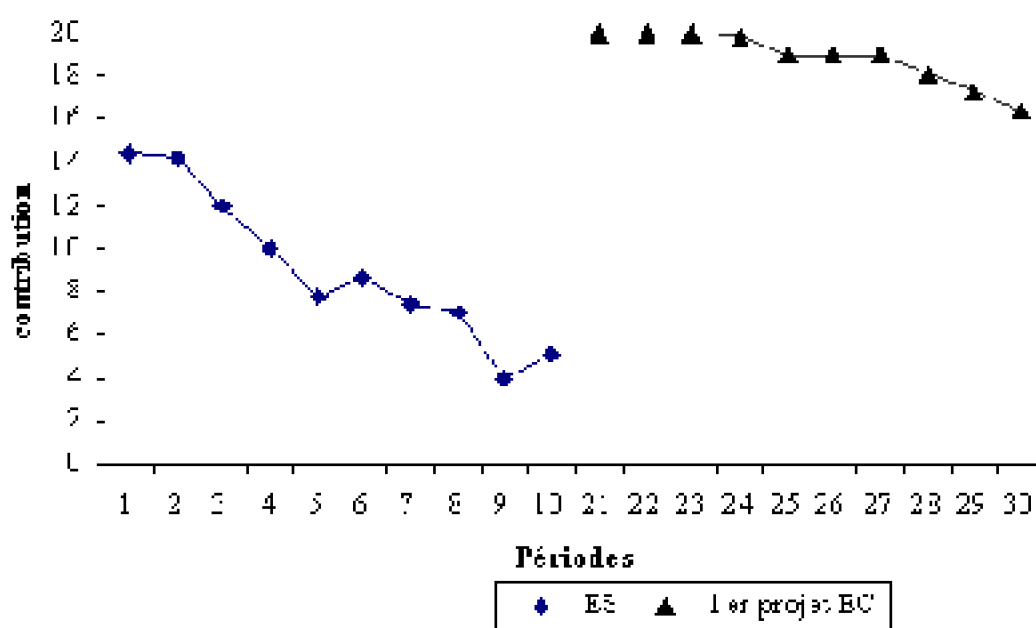
Graphique III.4 : Evolution des contributions moyennes dans le traitement ESC (sessions 7 à 11)

RESULTAT 4 : la possibilité d'exclure les autres membres du groupe incite les sujets à davantage coopérer même lorsque exclure est directement coûteux (traitement EC).

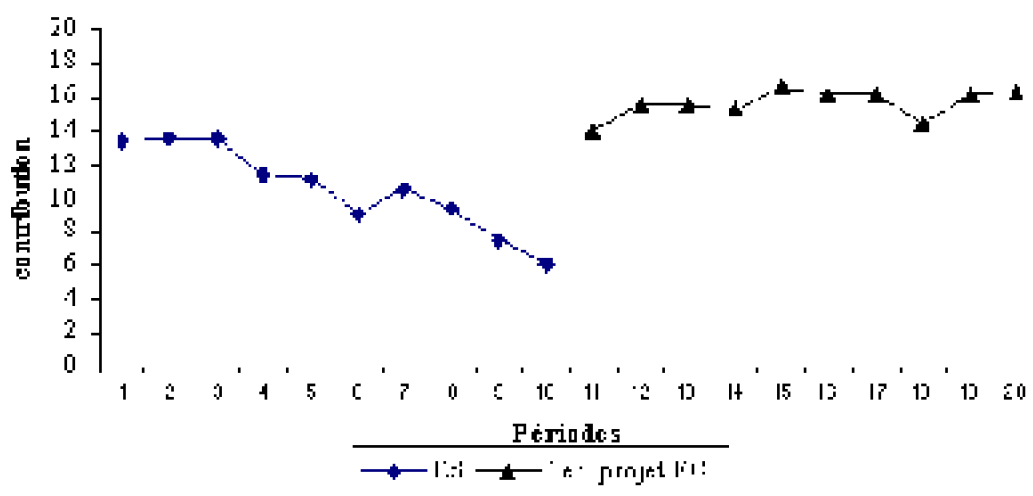
Le tableau 3 et les graphiques 5 et 6 montrent que l'opportunité d'exclusion avec coût direct provoque une augmentation de la contribution moyenne pour tous les groupes. La contribution moyenne pour tous les groupes dans le premier projet du traitement EC est de 17,21 ECU contre 9,79 ECU pour le traitement ES. D'après le test non paramétrique *Wilcoxon signed rank*, cette différence de contribution est significative ($p < 0.05$) et ceci

pour l'ensemble des sessions (sessions 1 à 6 et 7 à 11). Ce résultat infirme la prédiction théorique pour ce traitement selon laquelle les menaces d'exclusion avec coût direct ne sont pas crédibles et la coopération ne peut pas être soutenue. Si l'on considère la dernière période du traitement EC, il apparaît que 4 groupes contribuent la totalité de leur dotation au premier projet et 3 autres groupes contribuent plus de 18 ECU.

La comparaison des colonnes 3 et 5 du tableau 5 met en évidence que la contribution moyenne dans le traitement ESC est plus élevée que la contribution moyenne dans le traitement EC (excepté pour les groupes 1 et 9 et est égale pour les groupes 4 et 5). D'après le test *Wilcoxon Mann-Whitney* sur la contribution moyenne, cette différence de contribution est significative ($p < 0.05$).

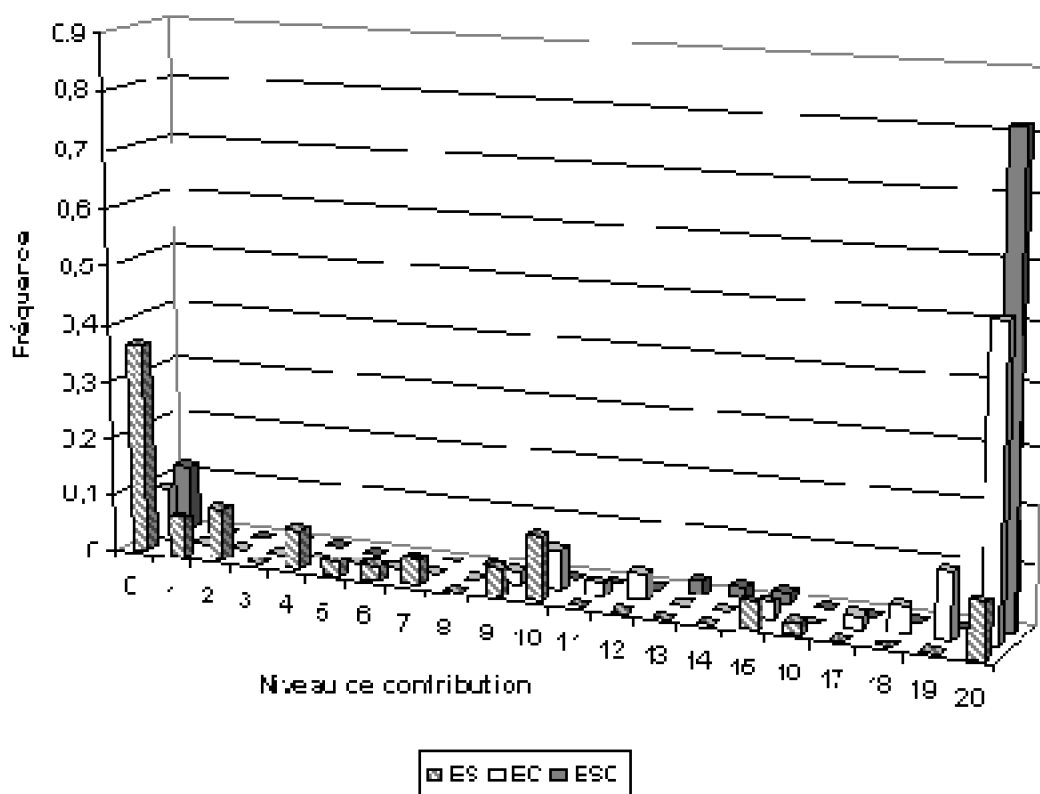


Graphique III.5 : Evolution des contributions moyennes traitement EC (sessions 1 à 6)



Graphique III.6 : Evolution des contributions moyennes traitement EC (sessions 7 à 11)

Le graphique 7 donne des informations sur la distribution des niveaux de contribution au cours du temps pour le traitement ES, et le premier projet des traitements ESC et EC.



Graphique III.7 : Distribution des niveaux de contribution

Dans les traitements ESC et EC, plus de 60% des sujets contribuent toute leur dotation au projet commun ($y_1=20$). A l'inverse, dans le traitement ES, environ 20% des sujets ne contribuent rien au projet.

Les résultats 3 et 4 mettent en évidence le fait que les sujets coopèrent quand l'opportunité d'exclure les autres sujets existe, même si l'exclusion est coûteuse pour ceux qui l'appliquent.

4.2. Pourquoi exclure les autres membres de son groupe ?

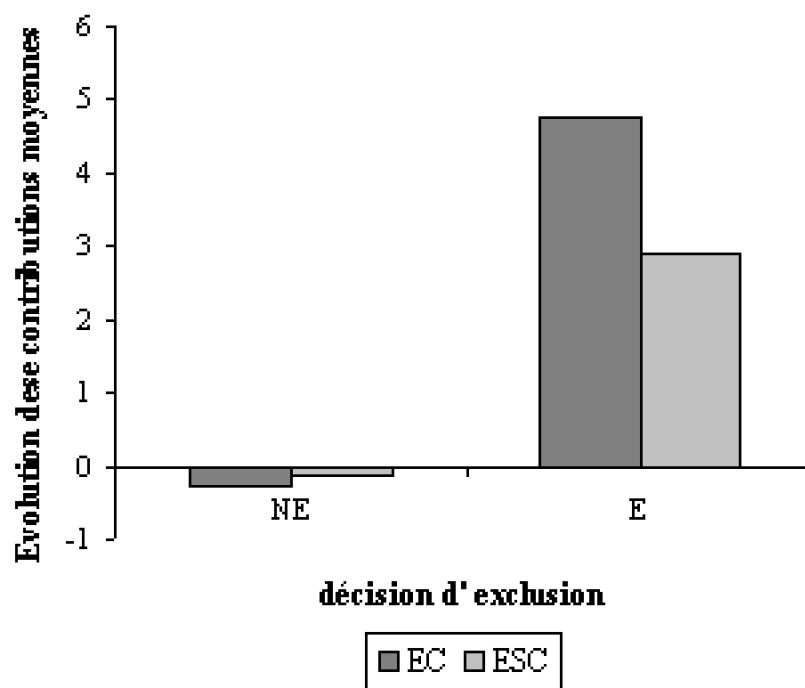
Comme dans le chapitre précédent, on s'intéresse à la relation entre contribution et sanction. Plus précisément, cette sous-section analyse les raisons pour lesquelles les sujets décident ou non d'exclure les autres membres de leur groupe. Deux raisons sont avancées pour expliquer pourquoi les sujets excluent leurs pairs même si cela est directement coûteux pour eux comme dans le traitement EC. La première explication est d'ordre stratégique : les sujets excluent leurs pairs afin de les inciter à coopérer davantage à l'avenir, ce qui nécessite que les agents soient partenaires, c'est à dire que la composition des groupes reste inchangée au cours du jeu. Cette explication repose sur le postulat selon lequel les agents ne peuvent pas anticiper tout le jeu et ne peuvent en particulier pas anticiper la dernière période. La deuxième explication, d'ordre non-stratégique, reflète les préoccupations sociales des sujets. Plus précisément, les sujets excluent leurs pairs lorsque ces derniers adoptent un comportement jugé déloyal.

4.2.1. Raison stratégique

Le résultat 5 traduit l'impact de l'exclusion sur le niveau des contributions futures.

RESULTAT 5 : *il existe un gain à exclure un passager clandestin afin de l'inciter à augmenter sa contribution dans le futur.*

Le graphique 8 décrit la relation existante entre l'évolution de la contribution moyenne entre la période $t+1$ et la période t et le fait d'être exclu (ou pas). Il montre que les sujets augmentent leur contribution future en réponse à une exclusion passée. Ainsi en moyenne, un sujet qui a été exclu à une période donnée, augmente sa contribution de 5 ECU à la période suivante dans le traitement EC et de 3 ECU dans le traitement ESC. A l'inverse, les sujets qui n'ont pas été exclus en période t ont tendance à réduire leur contribution de 0,28 ECU en $t+1$ dans le traitement EC et de 0,13 ECU dans le traitement ESC.



E : exclu ; NE : non exclu

Graphique III.8 : Evolution des contributions entre t et $t+1$ conditionnée par l'exclusion en t .

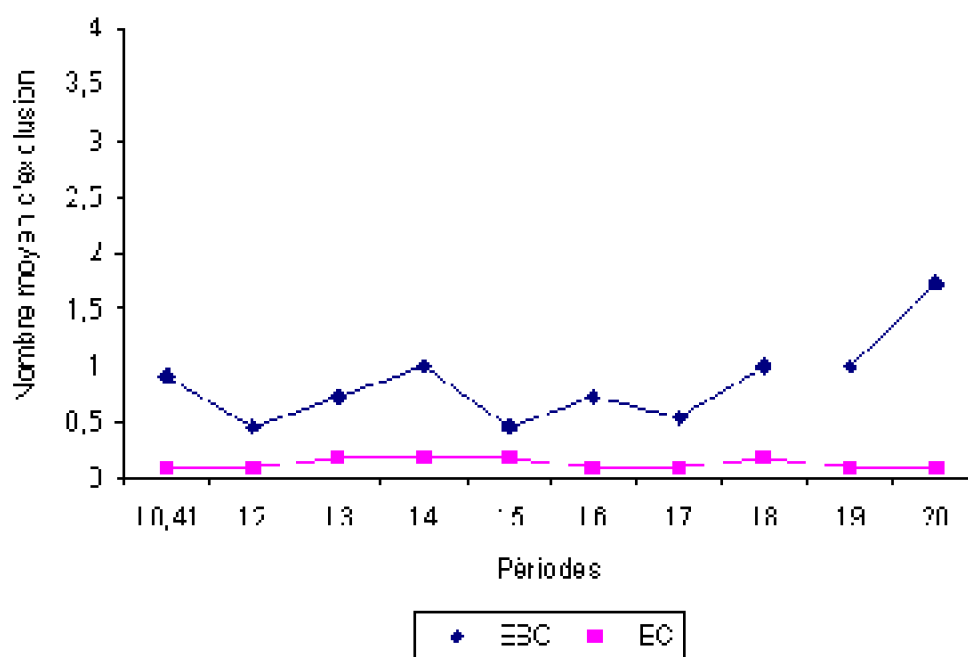
L'analyse de régression linéaire ci-dessous offre une preuve formelle du résultat 5. Le tableau 4 donne les estimations du modèle suivant pour les traitements ESC et EC.

La variable dépendante $C_i^{t+1} - C_i^t$ est "la différence de contribution entre les périodes t et $t+1$ ". Les variables indépendantes sont respectivement "les décisions d'exclusion de la part des autres sujets", "la contribution moyenne des autres membres du groupe à la période t ". Le coefficient β_1 traduit l'effet d'avoir été exclu en période t sur l'évolution de la contribution entre t et $t+1$. Le coefficient β_2 mesure l'effet de la contribution moyenne des autres membres du groupe en période t sur l'évolution de la contribution entre t et $t+1$.

Tableau III.4: Effets des exclusion reçus en t sur la contribution relative entre $t+1$ et t

variable expliquée : $c_i^{t+1} - c_i^t$	ESC	EC
Constante	-6.109*** (1.048)	-3.013*** (1.075)
Exclusion (α_1)	1.7664*** (0.2962)	1.398*** (0.428)
Moyenne(α_3)	0.3112*** (0.0550)	0.1512** (0.0585)
R^2	0.1057	0.314
Observations	396	396
***significativité au seuil de 1% , ** : significativité au seuil de 5%, * : significativité au seuil de 10% Les écarts types sont en parenthèses.		

Le coefficient α_1 est positif et significatif pour les deux traitements. Ainsi lorsqu'un sujet est exclu, il est incité à accroître sa contribution à la période suivante. Le coefficient α_2 est également positif et significatif pour les deux traitements. Cela signifie qu'un sujet est d'autant plus incité à accroître sa contribution que la contribution moyenne est élevée. Comment interpréter les résultats du tableau 5 et notamment le fait d'observer la contribution moyenne incite les agents à contribuer davantage à la période suivante ? Une interprétation possible consiste à distinguer, à l'instar de Kandell et Lazear (1992) la pression des pairs externe (honte) de la pression des pairs interne (culpabilité). Les décisions d'exclusion traduisent ici la désapprobation des pairs et véhiculent des sentiments de honte. A l'opposé, l'observation de la contribution moyenne peut susciter des sentiments de culpabilité, même en l'absence de sanction des pairs. Contrairement aux résultats obtenus dans le chapitre précédent, il apparaît ici que la pression des pairs n'est pas uniquement externe mais qu'elle est également internalisée. En effet, le seul fait pour un sujet d'observer la contribution moyenne l'incite à contribuer davantage dans le futur. Le fait que la pression des pairs soit internalisée signifie que la sanction d'exclusion n'a plus besoin d'être appliquée pour que les agents coopèrent. L'idée est que la sanction est suffisamment forte de sorte que la menace d'exclusion soit à elle seule suffisante pour inciter les agents à coopérer. On vérifie cette hypothèse en étudiant le nombre d'exclusion qui sont effectivement appliquées. Le graphique 9 décrit le nombre moyen d'exclusion par période sachant qu'au total les quatre membres du groupe peuvent être exclus à chaque période.



Graphique III.9 : Nombre moyen d'individus exclus par période

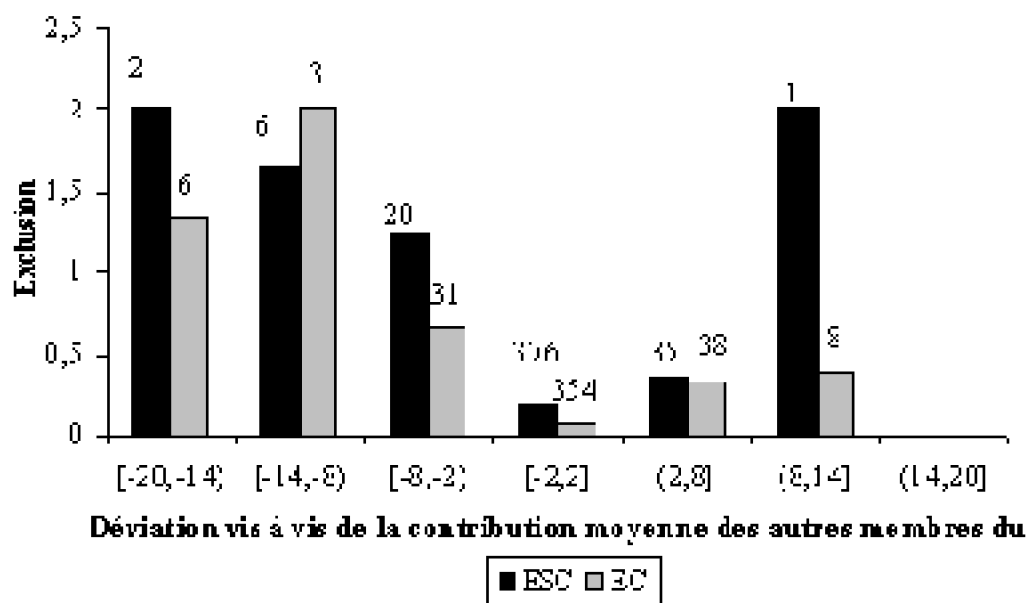
On remarque que l'exclusion est rarement appliquée dans le traitement EC. Dans le traitement ESC, moins d'un individu en moyenne est exclu du groupe à chaque période. Le nombre d'exclusion dans le traitement ESC reste très faible malgré l'inexistence d'un coût direct à exclure. Le nombre d'exclusion augmente en dernière période du traitement ESC sans toutefois dépasser deux exclusions par groupe. Le nombre d'exclusion est donc relativement faible dans les deux traitements. Ce résultat confirme l'hypothèse selon laquelle la menace d'exclusion est tellement crédible qu'elle n'a pas besoin d'être appliquée.

4.2.2. Raison non stratégique

La deuxième raison avancée pour expliquer pourquoi les sujets excluent les autres membres de leur groupe est d'ordre non-stratégique. Les sujets seraient incités à exclure leurs pairs lorsqu'ils jugent leur comportement déloyal, c'est à dire ceux qui ne se conforment pas à la contribution moyenne du groupe.

RESULTAT 6 : *les sujets sont exclus lorsqu'ils ne se conforment pas au niveau de contribution moyen du groupe*

Le graphique 10 décrit la relation qui existe entre la déviation de la contribution d'un sujet par rapport à la contribution moyenne du groupe et les sanctions en termes d'exclusion qu'il reçoit.



Graphique III.10 : Décision d'exclusion fonction de la déviation de la contribution d'un sujet vis-à-vis de la contribution moyenne

Dans les traitements avec exclusion (ESC et EC), les déviations négatives sont fortement sanctionnées par l'exclusion. Il apparaît que les sujets excluent également ceux qui contribuent davantage que la moyenne. Néanmoins, les déviations positives sont toujours moins fréquemment et moins fortement punies que les déviations négatives. Le tableau 5 contient les estimations de la régression sur l'analyse des déterminants des exclusions.

Table III.5 : relation entre les décisions d'exclusion et les déviations de contribution vis à vis de la contribution moyenne(modèle PROBIT)

Variable expliquée: décision d'exclusion		
	Trait. ESC	Trait. EC
Constante	2.5331***	-0.5976 (0.5830)
Contribution moyenne des autres	(0.629) -0.1978	-0.0609* (0.3121)
Valeur absolue D'une déviation négative	(0.032) 0.3597***	0.3431*** (0.070)
Déviations positives	(0.0751) -0.1076 (0.066)	0.10649*** (0.0448)

²⁷ Selon la méthode de Mac Fadden, $\text{pseudo } R^2 = 1 - \log L / \log L_R$

N= 440 N=440 Log L -152.51 Log L -114.20 LogL_R-228.24 LogL_R-178.90 Pseudo R² 0.331²⁷ Pseudo R² 0.361 % de var. correc. % de var. correc. prédites 87.04% prédites 91.36%

Notes Ecart type en parenthèse. *** : significatif au seuil de 1% , ** significatif au seuil de 5%, * significatif au seuil de 10% .

Encadré III.2 Rappel sur le PROBIT classique

Dans un modèle qualitatif à choix discrets, la variable expliquée ne peut prendre que deux valeurs, décision d'exclure un autre joueur ou absence d'exclusion.
On définit une variable latente , inobservable définie par la relation:
(1)
Avec x, vecteur des caractéristiques individuelles exogènes
La variable dichotomique observée est liée à la variable latente par la relation:
avec y =1, si le joueur décide d'exclure un autre joueur; y =0 sinon.
Alors Prob(=1)=Prob(>0)=Prob()
Pour un PROBIT classique, le logarithme de la fonction de vraisemblance s'écrit:
(2)
avec,
, distribution de la loi normale centrée réduite.
La recherche des paramètres maximisant la fonction de vraisemblance conduit à définir le vecteur des dérivées premières de la fonction:

Le tableau 5 montre que le coefficient " *déviatio*n* négative en valeur absdue* " est positif et significatif dans les deux traitements. Ainsi un agent est d'autant plus souvent exclu qu'il contribue moins que la moyenne. En revanche, les déviations positives semblent avoir un impact plus faible et ne sont pas significatives dans le traitement ESC. La variable « *contribution moyenne des autres joueurs* » est négative et peu significative dans les deux traitements. Le graphique 10 et le tableau 5 montrent donc que les membres du groupe peuvent éviter d'être exclus s'ils contribuent une part élevée de leur dotation au premier projet. La contribution moyenne est de 18 ECU et de 17 ECU respectivement pour les traitements ESC et EC.

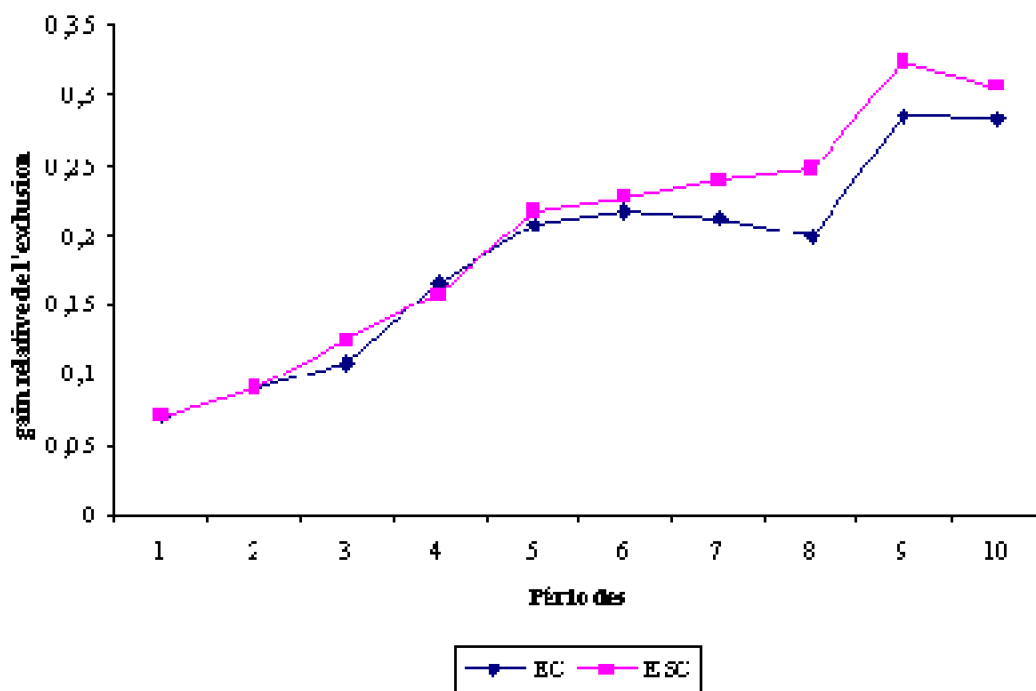
4.3. Effet global de l'ostracisme

Quel est l'effet global de l'ostracisme sur les gains des sujets ? Il existe deux effets opposés de l'ostracisme sur les gains des joueurs. Le premier effet est positif et repose sur le fait que les menaces d'exclusion incitent les sujets à augmenter leur niveau futur de contribution. On devrait donc observer des niveaux de gains supérieurs dans les traitements avec possibilité d'exclusion. Le deuxième effet est négatif. L'idée est que lorsque l'exclusion n'est plus seulement une menace mais qu'elle est appliquée, elle réduit les gains de l'ensemble du groupe. D'une part, le gain potentiel issu du deuxième

²⁷ Selon la méthode de Mac Fadden, pseudo R² = 1 - logL / logL_R

projet est réduit puisque le deuxième projet est amputé des contributions de ceux qui ont été exclus. D'autre part, l'ostracisme réduit les gains de ceux qui sont exclus et de ceux qui excluent. Quel effet l'emporte sur l'autre ? Pour répondre à cette question, il est nécessaire de comparer le gain moyen dans les traitements avec et sans exclusion.

RESULTAT 7 : *l'opportunité d'exclusion accroît le gain des sujets.*



Graphique III.11 : *gain relatif de l'exclusion (possibilité d'exclusion/sans possibilité d'exclusion)*

Le graphique 11 décrit la différence du profit moyen entre le traitement ES et les traitements ESC et EC normalisée au profit moyen dans le traitement ES. Il montre que l'opportunité d'exclusion accroît le gain moyen des sujets. En première période, le gain moyen des traitements avec exclusion (ESC et EC) est proche de celui du traitement ES. En revanche, le gain relatif est de 30 % plus élevé dans le traitement ESC et de 28 % dans le traitement EC à la dernière période. D'après le test non paramétrique *Wilcoxon signed rank*, ces différences de gain entre le traitement sans exclusion (EC) et avec exclusion (ESC ou EC) sont significatives pour l'ensemble des sessions ($p < 0.05$). Existe-il une différence de gain significative entre les traitements ECS et EC ? La comparaison des deux traitements ne montre pas de différence significative selon le test de Wilcoxon mann Withney, $p > 0.05$. L'existence du coût direct à exclure ne semble donc pas affecter considérablement les gains des sujets. Cette observation conforte le fait que les exclusions sont rarement appliquées et que la menace est suffisante pour faire émerger la coopération. L'effet incitatif des sanctions semble donc prévaloir.

5. Conclusion

Sous certaines conditions, la menace d'exclusion apparaît donc comme une menace crédible qui supporte la coopération au sein d'une équipe. En effet, dès lors qu'il n'existe plus de coût direct à exclure un sujet en dernière période, la coopération peut être soutenue durant toutes les périodes. Les sujets exploitent alors le problème de passager clandestin en dernière étape du jeu afin de supporter la coopération au premier projet (Hirshleifer et Rasmusen, 1989). Quand les sujets subissent un coût direct à exclure leurs pairs, l'ostracisme devient théoriquement inefficace. Les résultats expérimentaux montrent au contraire que l'opportunité d'exclure incite les agents à augmenter leur niveau de contribution même lorsque exclure est directement coûteux. Les résultats expérimentaux mettent également en évidence le fait que l'opportunité d'exclure les autres membres du groupe (traitement EC et ESC) augmente le gain moyen des sujets. L'ostracisme est donc un mécanisme d'incitation efficace puisqu'il conduit les sujets à adopter un comportement coopératif et augmente le gain moyen des sujets.

Deux explications sont avancées dans ce chapitre pour expliquer pourquoi les agents sanctionnent leurs pairs. La première explication est d'ordre non stratégique : les sujets décident d'exclure ceux qui contribuent moins que la moyenne. La deuxième explication au comportement de sanction est d'ordre stratégique et repose sur l'idée que les sujets excluent leurs pairs car ils attendent des gains futurs des sanctions d'exclusion. L'idée est que les sujets n'anticipent pas parfaitement l'ensemble du jeu mais seulement quelques périodes. Keser et van Winden (1996) qualifient un tel comportement de *myope*. Selon cette explication, les sujets exclus en période t seraient incités à modifier leur comportement et à davantage contribuer dans les périodes futures. Si les exclusions incitent les agents à davantage contribuer (pression des pairs externe), l'observation de la contribution moyenne des autres membres du groupe, incite également les joueurs à accroître leur propre niveau de contribution (pression des pairs interne). L'idée est que la menace de sanction est internalisée par les agents. (Kandel et Lazear, 1992).

Dans le jeu présenté ici, il suffit qu'un seul membre du groupe choisisse d'exclure un autre membre pour que ce dernier le soit. Une extension possible au modèle serait de considérer que la décision d'exclure un membre exige une décision à la majorité ou même à l'unanimité. Une autre prolongation serait de considérer que les membres du groupe procèdent à un vote afin de décider qui doit être exclu du groupe.

Chapitre 4. La renonciation au passage clandestin par l'aversion à l'inégalité

« Un homme seul est en très mauvaise compagnie » Paul Valéry

« L'iniquité existe pour une personne chaque fois que celle-ci perçoit comme inégal le rapport entre sa contribution et ses résultats par comparaison avec les contributions et les résultats d'une autre personne » Adams 1965, Inequity in Social exchange, Advances in Experimental Psychology, Academic Press

1. Introduction

Les chapitres précédents ont établi que la pression des pairs est théoriquement inefficace lorsque les agents sont uniquement motivés par la maximisation de leurs gains monétaires et que sanctionner ses pairs est coûteux. Les résultats expérimentaux montrent au contraire que les sujets n'hésitent pas à sanctionner ceux qui adoptent un comportement opportuniste et que les sujets coopèrent davantage au bien public lorsqu'ils ont la possibilité de discipliner leurs pairs.

Comment rendre compte que, dans la réalité, les individus ne se comportent pas toujours comme le prédit la théorie ? Ainsi, comment justifier les comportements de sanction mutuels au sein des équipes de production alors que, sous l'hypothèse de

maximisation des gains pécuniaires, personne ne devrait sanctionner. Une première interprétation de ces comportements consiste à mettre en avant les motivations stratégiques des agents lorsque les interactions sont répétées (Fudenberg et Maskin, 1986). Un certain nombre d'auteurs se sont attachés à modéliser ces motivations stratégiques en supposant que les agents ont une information incomplète sur les options ou motivations des autres agents (Kreps, Wilson, Milgrom et Robert, 1982 ; Fudenberg et Maskin, 1986).

Une autre interprétation consiste à supposer que les agents ne peuvent pas anticiper l'ensemble des interactions futures et notamment la fin du jeu. Keser et van Winden (1996) qualifient un tel comportement de *myope*. De même, selon Güth (1995), le comportement des agents résulterait d'un processus dynamique de prise de décision dans un contexte de rationalité limitée. Ces différentes interprétations constituent des pistes intéressantes pour expliquer les comportements individuels de sanction d'ordre stratégique mais conduisent à relâcher l'hypothèse fondamentale de rationalité des agents. Par ailleurs, elles ne permettent pas d'expliquer le comportement des agents lorsque les interactions ne sont pas répétées.

Une autre voie d'analyse complémentaire à l'interprétation précédente consiste, tout en conservant l'hypothèse de rationalité, à considérer que les agents ne maximisent pas uniquement leurs gains pécuniaires. Les sociologues ont depuis longtemps établi que les motivations individuelles des salariés ne sont pas uniquement d'ordre pécuniaire. La plupart des économistes ont souvent considéré ces phénomènes comme irrationnels, relevant de la « logique des sentiments » (Dickson et Roethlisberger, 1945). A l'opposé, un certain nombre d'économistes ont mis en évidence le fait que les individus ne se préoccupent pas uniquement de leur propre intérêt pécuniaire mais également du bien-être des autres (Becker, 1974 ; Arrow, 1981 et Samuelson, 1993). Déjà en 1759, Adam Smith mettait en exergue l'ambiguïté existant entre la prise en compte des seuls intérêts individuels et l'existence d'un individu doté de motivations plus riches²⁸. Une lecture de l'œuvre de Smith montre que les sentiments moraux et la poursuite de l'intérêt personnel ne sont pas aussi éloignés que cela. Ainsi, la notion de sympathie, élément central du concept de moralité, fait référence chez Smith à un processus psychologique complexe reposant sur un mode d'identification aux autres (Fontaine, 1997)²⁹. Comme le note Binmore (1994), l'*homo oeconomicus* doit souscrire dans une certaine mesure au principe d'empathie décrit ci-dessus puisque son expérience des autres doit être

²⁸ En référence au débat engagé sur la cohérence des motivations individuelles défendues dans deux ouvrages majeurs d'Adam Smith, « La théorie des sentiments moraux » et « La richesse des nations ». Après avoir exposé la richesse des motivations individuelles dans « La théorie des sentiments moraux », Adam Smith conclut que « les principaux ressorts de l'action humaine finissent par converger, dans le besoin d'un bien être matériel accru », (Hirshman, 1980, p99-100). La poursuite de l'intérêt personnel est alors justifiée par la doctrine libérale, elle-même fondée sur la métaphore de la main invisible selon laquelle « la somme des intérêts particuliers conduit à l'intérêt général ».

²⁹ Les sentimentaux moraux peuvent se définir comme le résultat d'un dialogue imaginaire entre l'individu et un spectateur impartial conduisant l'individu pour un instant à s'identifier à l'autre. Dans « La théorie des sentiments moraux », A. Smith écrit : « Par l'imagination, nous nous mettons nous mêmes dans la situation de l'autre, nous endurons les mêmes tourments, nous entrons dans son corps et nous formons alors dans une certaine mesure la même personne.. » extrait de Fontaine, 1997, p265.

suffisamment riche pour lui permettre de se mettre à leur place et de voir la situation de son propre point de vue. Les sentiments moraux se présentent alors comme un élément de réconciliation entre les prédictions théoriques et les régularités comportementales issues des expérimentations.

Ces dernières années, les travaux en économie expérimentale ont également conduit à s'interroger sur l'hypothèse de maximisation stricte des gains monétaires. Ainsi le jeu d'Ultimatum³⁰ inventé par Guth, Schmittberger et Schwarze (1982), les jeux de don et contre don (Fehr Kirchsteiger et Riedl, 1993)³¹, le jeu de confiance (Berg, Dickhaut et McCabe, 1995)³², le jeu du dictateur³³ et finalement le jeu de contribution volontaire au financement d'un bien public ont joué un rôle important dans la remise en cause partielle de l'hypothèse de stricte maximisation des gains matériels. Toutefois, l'interprétation de ces résultats fait l'objet de discussion. En effet, des incertitudes subsistent quant aux facteurs influençant le comportement individuel.

La perspective de l'introduction des sentiments moraux dans l'analyse économique pose le problème de la multiplicité des motivations. Il est donc nécessaire de déterminer quelle motivation est dominante. Deux approches principales visant à expliquer les résultats expérimentaux doivent être distinguées. La première approche met l'accent sur les intentions des joueurs (Geanakoplos, Pearce et Stacchetti, 1989 ; Rabin, 1993 ; Dufwenberg et Kirchsteiger, 1998, Rabin et Charness, 2000). Cette approche vise à introduire dans la fonction d'utilité une réaction possible du joueur par rapport aux intentions qu'il prête à son partenaire lorsqu'il prend une décision. La deuxième approche est axée sur les considérations distributives (Bolton et Ockenfels, 2000 et Fehr et Schmidt, 1999). Cette approche suppose que l'utilité d'un agent ne dépend pas

³⁰ Dans le jeu d'ultimatum, deux joueurs doivent s'accorder sur le partage d'une somme. Le premier propose au second une répartition de la somme qui accepte ou refuse cette offre. S'il accepte, la somme est partagée selon la proposition initiale. A l'inverse, s'il refuse, les deux joueurs ont des gains nuls. Selon les prédictions théoriques de ce jeu, le premier joueur devrait proposer de donner au second la somme non nulle la plus faible possible, attendu que le second acceptera toute offre positive. Contrairement aux prédictions théoriques, les centaines d'expériences de ce jeu donnent des résultats similaires : les offres du premier joueur au second sont rejetées avec une probabilité comprise entre 40 et 60% lorsqu'elles représentent moins de 20% de la somme initiale.

³¹ Dans un jeu d'échange de don, le *proposant* fait une offre w , qui peut être interprétée comme un salaire au *répondant*. Le *répondant* décide alors d'accepter ou de rejeter cette offre. S'il refuse, les deux joueurs ont des gains nuls. En cas d'acceptation, le répondant doit décider de son effort e . Selon les prédictions du jeu, le répondant ne rejettera jamais l'offre du proposant et choisira toujours l'effort le plus faible possible. Donc le proposant choisira w le plus bas possible. Toutes les études sur ce jeu montrent que les efforts et les niveaux de salaires sont corrélés positivement.

³² Dans un jeu de confiance, un *proposant* reçoit une somme de monnaie y de la part de l'expérimentaliste. Il envoie alors une somme z entre 0 et y à un *répondant*. Le répondant reçoit alors le triple de la somme, soit $3z$. Ce dernier décide alors d'une somme comprise entre 0 et $3z$ à envoyer au *proposant*. Les résultats expérimentaux montrent que les proposants envoient des sommes non nulles aux répondants qui renvoient en retour également des sommes non nulles.

³³ Le jeu du dictateur diffère du jeu d'ultimatum dans la mesure où le répondant ne peut, dans ce jeu, qu'accepter l'offre du proposant. Forsythe et al (1994) ont montré que les offres dans un jeu du dictateur sont plus faibles que celles d'un jeu d'ultimatum.

uniquement de ses propres gains, mais également de ceux des agents avec lesquels il interagit. Dans le modèle de Bolton et Ockenfels, les agents comparent leur gain matériel au gain moyen du groupe. Ils ressentent alors de l'aversion à l'inégalité si leur revenu s'éloigne du revenu moyen. Ce qui importe est donc de rapprocher son revenu au plus près du revenu moyen du groupe. Au contraire, dans le modèle de Fehr et Schmidt (1999), les agents comparent leur gain à celui de chacun des autres agents.

Ce chapitre vise à proposer une modélisation de la pression des pairs fondée sur une modification des fonctions d'utilité des agents permettant d'incorporer la composante sociale de leurs motivations. La thèse défendue ici est que la modélisation de la pression des pairs gagnerait en pertinence si elle enrichissait sa représentation, en considérant que les agents incorporent des considérations sociales dans leur fonction d'utilité. La fonction d'utilité proposée s'appuie sur les considérations distributives d'aversion à l'inégalité auto-centrée. L'utilité des sujets ne dépend alors plus uniquement de leurs gains pécuniaires mais également des différences de gains avec les autres sujets. L'idée est que si les agents sont suffisamment averses à l'inégalité, ils peuvent être incités à sanctionner ceux qui ne coopèrent pas afin de réduire les dissonances de gain. La menace de sanction devient alors crédible et les sujets sont incités à coopérer. L'utilisation d'une telle démarche se devant d'être prudente, un soin particulier sera apporté à justifier en début de chapitre la nature des arguments introduits dans la modélisation par le biais d'observations robustes et stables.

Le chapitre est organisé de la manière suivante. La deuxième section s'interroge sur l'intérêt de modifier les fonctions d'utilité en incorporant les préférences sociales des individus. La troisième section est consacrée à la présentation du modèle séquentiel de pression des pairs sous l'hypothèse d'aversion à l'inégalité. Dans la quatrième section, le modèle de pression des pairs est testé à partir d'un jeu de bien public. La cinquième section présente les résultats expérimentaux qui testent l'hypothèse d'aversion à l'inégalité. Enfin la sixième section conclut le chapitre.

2. Le poids théorique des préférences sociales

Un certain nombre d'auteurs ont émis quelques réserves quant à l'introduction de nouveaux arguments dans la fonction d'utilité (Güth, 1995 ; Lévy-Garboua et Rapoport, 2000). Güth (1995) dénonce ainsi l'aspect artificiel d'une telle démarche. Selon ces auteurs, le comportement des agents serait plutôt le résultat d'un processus dynamique de prise de décision. L'adjonction de nouveaux arguments dans la fonction d'utilité ne permettrait pas d'expliquer les comportements individuels. Cette interprétation constitue une voie intéressante afin d'expliquer les comportements individuels. Toutefois, elle pose un certain nombre de difficultés analytiques. En effet, il paraît difficile de dégager, pour un jeu donné, une tendance générale des comportements dans la mesure où chaque joueur diffère dans ses propres expériences. Par ailleurs, cette théorie ne permet pas d'expliquer les comportements des joueurs dans un contexte de jeu non répété. Enfin une telle approche conduit, comme le soulignent Fehr et Schmidt (1999) et Bolton et Ockenfels

(2000), à relâcher l'hypothèse fondamentale de rationalité des agents.

Les approches fondées sur les considérations sociales des agents se posent en complémentarité des travaux qui considèrent le comportement des agents comme un processus dynamique de prise de décision. En effet, alors ces travaux visent à étudier l'aspect *stratégique* des comportements, les approches fondées sur les considérations sociales cherchent à analyser les aspects *non stratégiques* des comportements individuels.

Tableau IV.1 : Littérature sur les fonctions d'utilité modifiées

Auteurs	Fonctions d'utilité (forme générale)
<i>Rabin (1993)</i>	avec et
<i>Dufwenberg et Kirchsteiger (1998)</i>	
<i>Charness et Rabin (2000)</i>	
<i>Bolton (1991)</i>	
<i>Bolton et Ockenfels (2000)</i>	
<i>Fehr et Schmidt (1999)</i>	

Falk & Fischbacher (1998)	avec f : nœud terminal du jeu : terme de réciprocité (réaction) , où représente la comparaison des gains et une pondération dont la valeur est déterminée par l'intention attribuée à l'autre
---------------------------	---

Il est possible de regrouper ces différentes approches en deux grandes catégories. La première catégorie met l'accent sur les intentions des joueurs (Rabin, 1993 ; Dufwenberg et Kirchsteiger, 1998). Elle repose sur la notion de réciprocité. La deuxième catégorie est essentiellement axée sur les considérations distributives du jeu (Charness et Rabin, 2000 ; Bolton, 1991 ; Bolton et Ockenfels, 2000 ; Fehr et Schmidt, 1999). L'approche de Falk et Fischbacher (1998) tente, quant à elle, de réconcilier les intentions et les considérations distributives.

2.1. Le rôle des intentions

Rabin (1993) propose une formulation de la fonction d'utilité des joueurs qui incorpore des éléments psychologiques comme les croyances dans le cadre de jeux simultanés. L'utilité des joueurs dépend alors à la fois des gains monétaires du joueur mais également des intentions qu'il attribue aux actions des autres. Rabin utilise pour cela un jeu

psychologique construit par Geanakoplos, Pearce et Stacchetti (1989). Dans un jeu psychologique, l'utilité des sujets ne dépend pas uniquement de leurs gains mais également de leurs croyances et des croyances des autres sur leurs croyances. Notons $a_i \in A_i$ la stratégie du joueur i . Lorsque i décide de sa stratégie, il a des croyances quant à la stratégie choisie par le joueur j . Notons b_j la croyance du joueur i sur la stratégie que j va suivre. Le joueur i a également des croyances sur les croyances du joueur j quant à la stratégie du joueur i . Cette croyance sur les croyances de j est notée c_j . La fonction d'utilité d'un joueur est composée de ses gains matériels et de ses gains psychologiques. Le joueur i maximise donc une fonction d'utilité définie par:

(2)

D'après cette fonction, il apparaît que la bienveillance repose sur une comparaison entre le gain attendu du jeu (sachant que j est supposé choisir la stratégie b_j) et le gain jugé équitable qui correspond à la moyenne arithmétique entre le plus élevé et le plus faible des gains Pareto-efficients susceptibles d'être attribués à j par l'action de i . Cette comparaison est pondérée par la différence entre le gain potentiel le plus élevé de j et son gain le plus faible du jeu. Si i revendique une part plus importante, reflète une action malveillante vis à vis de j . La perception de la bienveillance de j envers i est construite de façon similaire à la fonction précédente. Elle est donnée par :

(3)

Le joueur maximise sa fonction d'utilité définie par (1). A partir de cette fonction, il apparaît que la meilleure réponse du joueur i à une action anticipée comme malveillante (respectivement bienveillante) est représentée par la punition (respectivement par la récompense), même si ces pratiques sont coûteuses pour le joueur qui les utilise.

Cette fonction d'utilité offre l'avantage de rendre rationnelles et crédibles de telles pratiques coûteuses. Toutefois, le modèle de Rabin comporte un certain nombre de limites. D'une part, la multiplicité des équilibres est une caractéristique générale de ce modèle. En particulier, il y a un équilibre dans lequel les deux joueurs sont bienveillants et un équilibre où ils sont malveillants. Il est alors difficile de prévoir quel équilibre va être joué. D'autre part, cette fonction d'utilité est définie pour des jeux simultanés et non pour des jeux séquentiels. Dufwenberg et Kirchsteiger (1998) ont appliqué le modèle de Rabin (1993) à des jeux séquentiels. Les auteurs introduisent la notion d'équilibre séquentiel réciproque (SRE). En jeu séquentiel, les joueurs sont amenés à réviser leurs croyances à chaque nœud du jeu. Cela suppose donc de distinguer les croyances initiales des joueurs de leurs croyances révisées. En effet, la nature de ces croyances dépend fortement du sous-jeu dans lequel les joueurs se trouvent au moment de leur choix. Cependant, même dans les jeux séquentiels extrêmement simples, l'analyse d'équilibre devient rapidement complexe.

Les travaux expérimentaux qui se sont attachés à tester les intentions donnent des résultats contradictoires. Blount (1995) a cherché à vérifier l'hypothèse fondée sur les intentions à partir d'un jeu d'ultimatum joué sous deux traitements différents. Dans un premier traitement, appelé traitement aléatoire, l'offre du proposant est déterminée de façon aléatoire. Le répondant sait comment l'offre est déterminée et que le proposant n'est pas responsable de cette offre. Dans un deuxième traitement, appelé traitement

avec intentions, le proposant détermine lui-même l'offre. La comparaison des deux traitements permet de vérifier le rôle joué par les intentions des joueurs. Alors que dans le premier traitement, les intentions ne jouent pas, elles jouent pleinement dans le deuxième traitement. Blount observe que les répondants rejettent les offres non équitables dans les deux traitements. Toutefois, il observe que le seuil d'acceptation est sensiblement plus élevé dans le traitement avec intentions. Bolle et Kritikos (1998) obtiennent des résultats en faveur des *intentions*. Bolton, Brandts et Katok (1997) ; Bolton, Brandts et Ockenfels (2000) ne trouvent quant à eux aucune évidence en faveur du principe des intentions. Charness (à paraître) obtient des conclusions similaires à partir d'un jeu de don-contre don.

2.2. Le poids des considérations distributives

La deuxième approche de fonction d'utilité modifiée dans la littérature suppose que les agents ont des préférences sociales, c'est à dire que leur fonction d'utilité dépend non seulement de leur propre intérêt pécuniaire mais également des gains perçus par les autres agents. Etant données ces préférences sociales, tous les agents sont supposés se comporter de façon parfaitement rationnelle. Soit $\{1, 2, \dots, n\}$ un ensemble d'individus et $x = (x_1, x_2, \dots, x_n)$ l'allocation des ressources où x_i représente les ressources matérielles allouées à l'agent i . Selon l'hypothèse de maximisation des gains monétaires habituelle, l'utilité du joueur i ne dépend que de x_i . Lorsque les agents ont également des préférences sociales, l'utilité du joueur i est également affectée par des variations de $x_j, j \neq i$.

2.2.1. Altruisme et envie

Une personne est altruiste si les dérivées partielles de $u(x_1, \dots, x_n)$ par rapport à x_1, \dots, x_n sont strictement positives, c'est à dire si son utilité augmente avec le bien être des autres (Andreoni et Miller, 2000 ; Charness et Rabin, 2000)³⁴. L'approche fondée sur l'altruisme permet d'expliquer un certain nombre de résultats expérimentaux tels que les résultats des jeux du dictateur ou les comportements observés dans les jeux de don contre don. Toutefois, une telle approche ne peut expliquer pourquoi les sujets s'engagent dans des comportements de sanction qui sont souvent coûteux pour eux-mêmes. Ainsi, l'altruisme ne peut expliquer, ni les niveaux élevés de refus d'offres dans le jeu d'ultimatum, ni les résultats observés sur les jeux de bien public avec sanction. Contrairement à l'altruisme, l'envie suppose que les agents ressentent de la désutilité s'ils ont des gains inférieurs aux autres $\partial u(x_1, \dots, x_n) / \partial x_j < 0, j \neq i$ mais sont indifférents au fait que les autres puissent avoir des gains inférieurs (Bolton, 1991). Si l'envie permet d'expliquer les résultats des jeux de négociation et de bien public avec sanction, elle ne peut pas expliquer les résultats des jeux du dictateur, de don contre don ou de bien public sans opportunité de sanction.

³⁴ L'encyclopédie Britannica (1998, 15^{ème} édition) définit une personne altruiste comme quelqu'un qui ressent l'obligation de « mettre en avant les plaisirs et soulager les souffrances des autres ».

2.2.2. L'aversion à l'inégalité

Les travaux sur l'aversion à l'inégalité proposent d'intégrer dans une même modélisation les idées d'altruisme et d'envie. Ces modèles supposent qu'un sujet est altruiste envers ceux dont les gains sont inférieurs à un certain niveau jugé équitable et est envieux à l'égard de ceux dont les gains sont au-dessus de ce niveau. Plus précisément, un individu est averse à l'inégalité³⁵ s'il ressent de la désutilité à percevoir des revenus qu'il juge inéquitables (Fehr et Schmidt, 1999). Cette définition pose alors la question de savoir comment les individus jugent qu'un revenu est équitable ou non.

« L'iniquité existe pour une personne chaque fois que celle-ci perçoit comme inégal le rapport entre sa contribution et ses résultats par comparaison avec les contributions et les résultats d'une autre personne » Adams 1965, *Inequity in Social exchange*, *Advances in Experimental Psychology*, Academic Press

Le revenu de référence, généralement utilisé pour évaluer une situation donnée, fait lui-même l'objet d'un processus complexe de comparaisons sociales. La prise en compte des comparaisons sociales comme élément fondamental pour expliquer le comportement des individus a fait l'objet d'un certain nombre de travaux en psychologie sociale (Homan, 1961, Adam, 1963, 1965) et en sociologie (Davis, 1959, Pollis, 1968). Adams (1963, 1965) propose une théorie de l'équité qui s'appuie sur les travaux d'Homan (1961). La théorie de l'équité repose sur l'hypothèse que chaque agent rationnel s'investit dans une relation sociale tant que les avantages qu'il en retire sont proportionnels à ses investissements. Tout déséquilibre se traduit par un sentiment d'iniquité. L'insatisfaction, née d'une iniquité relative ressentie, doit alors se traduire par une modification des attitudes ou des comportements. Selon Adams, un salarié i apprécie les avantages A_i qu'il retire de son emploi en fonction des contributions C_i qu'il fournit, relativement à ceux des autres salariés. Ainsi compare-t-il son ratio A_i/C_i à celui de ses collègues de travail j A_j/C_j .

Des études empiriques ont également mis en évidence l'importance des considérations sur les gains relatifs. Clark et Oswald (1996) montrent que les comparaisons de revenu ont un impact significatif sur la satisfaction dans l'emploi. Loewenstein, Thompson et Bazerman (1989) ont étudié l'importance des gains relatifs à partir d'une étude en psychologie expérimentale. Les sujets reçoivent un salaire fictif et observent les salaires des autres sujets. Ils doivent alors classer par rang les salaires différents de leur propre salaire. À partir de cette étude, les auteurs estiment l'importance qu'attachent les individus aux gains relatifs. Ils observent que les sujets ont une aversion forte pour l'inégalité à leur désavantage mais également dans une moindre mesure de l'aversion pour l'inégalité à leur avantage.

2.2.2.1. Fehr et Schmidt (1999)

Fehr et Schmidt (1999) définissent l'aversion à l'inégalité autocentrée de la façon suivante : une personne ressent de l'aversion à l'inégalité lorsqu'elle ressent de la désutilité à percevoir des gains inférieurs à ceux des autres (inégalité à son désavantage) et/ou à

³⁵ Fehr et Schmidt utilisent également le terme d'aversion pour l'iniquité.

percevoir des gains supérieurs à ceux des autres (inégalité à son avantage). Les auteurs supposent que les agents sont davantage sensibles à l'inégalité désavantageuse qu'à l'inégalité avantageuse. L'aversion à l'inégalité est auto-centrée si une personne est uniquement sensible à l'iniquité qu'elle ressent vis à vis des autres mais ne se préoccupe pas de l'iniquité qui existe entre les autres. Contrairement aux modèles d'apprentissage qui relâchent l'hypothèse de rationalité et supposent que tous les sujets sont uniquement motivés par leurs stricts gains pécuniaires (Roth et Erev, 1995), Fehr et Schmidt introduisent de l'hétérogénéité en supposant que les individus pussent être motivés par des considérations sociales tout en conservant l'hypothèse de rationalité.

Soient n joueurs indexés par $i \in \{1, \dots, n\}$ et $x = (x_1, \dots, x_n)$ le vecteur des gains monétaires. La fonction d'utilité du joueur $i \in \{1, \dots, n\}$ est donnée par :

Le premier terme de l'équation (5) représente le gain du joueur i . Le deuxième terme traduit la perte d'utilité inhérente à l'inégalité désavantageuse ressentie ($x_j > x_i$) tandis que le troisième terme mesure la perte d'utilité inhérente à l'inégalité j avantageuse ressentie ($x_j < x_i$). Le gain de l'individu i est alors maximisé pour $x_i = x_j$. La perte d'utilité inhérente à l'inégalité désavantageuse est supérieure à celle inhérente à l'inégalité avantageuse. Pour simplifier, les auteurs supposent que la fonction d'utilité est linéaire en gain et en aversion à l'inégalité. Lorsque les fonctions de gain des agents sont connaissance commune, ils peuvent alors anticiper les conséquences de chacune des stratégies possibles et l'aversion à l'inégalité peut alors affecter leurs décisions.

2.2.2.2. Bolton et Ockenfels (2000)

Bolton et Ockenfels (2000) ont développé un modèle d'aversion à l'inégalité proche de celui de Fehr et Schmidt. Ils utilisent une formulation spécifique de la fonction de préférence appelée ERC³⁶ :

c est la somme totale des gains. y_i représente la part du gain relatif pour le joueur i . Les préférences sont conformes aux hypothèses suivantes :

A0 : v_i est continue et deux fois différentiable par rapport à y_i et \bar{y}_i .

L'hypothèse A1 traduit que la fonction est monotone croissante à taux décroissant des gains monétaires. C'est au travers de l'hypothèse A2 que se manifeste l'aversion pour l'inégalité du joueur. Toutes choses étant égales par ailleurs, l'utilité est maximale quand la part de chaque individu est égale à la moyenne des gains.

$r_i(c)$ est une valeur plafond correspondant à la plus grosse part du partage social de référence que le joueur souhaite s'approprier pour un gain moyen donné k . $s_i(c)$ est le seuil minimal d'acceptation du joueur i représentant la part de c qui le rend indifférent entre accepter ou refuser une offre. Bolton et Ockenfels font l'hypothèse d'hétérogénéité des joueurs en supposant qu'ils se distinguent par la valeur de leurs seuils $r_i(c)$ et $s_i(c)$. La distribution de ces seuils est décrite par les fonctions de distribution f et g avec $f(r_i/k) > 0$ sur l'intervalle $[1, \bar{y}]$ et $g(s_i/k) > 0$ sur l'intervalle $(0, 1]$ pour un gain absolu moyen $k > 0$.

³⁶ ERC correspond à Equité Reciprocité et Concurrence.

Les joueurs se trouvent dans une situation d'information incomplète puisqu'ils ne sont pas en mesure de connaître les seuils de rejet qui caractérisent les agents avec qui ils jouent mais où les fonctions de distribution de ces seuils (f et g) sont de connaissance commune. Ils peuvent donc évaluer la probabilité qu'une de leur offre soit rejetée. Trouver les équilibres ERC d'un jeu revient donc à maximiser la fonction de motivation dans le cadre d'un jeu bayésien. Afin d'obtenir des résultats dans les différentes configurations de jeu, les auteurs définissent une fonction de motivation particulière compatible avec les hypothèses présentées ci-dessus, soit :

L'intérêt de cette formulation est que le type d'un joueur est caractérisé par son TMS entre les gains matériels et l'aversion pour l'inégalité, ici égale à a/b .

Bolton et Ockenfels appliquent leur modèle à différents jeux expérimentaux. Ils obtiennent des résultats similaires à ceux de Fehr et Schmidt dans les jeux avec deux joueurs. Le point commun des modèles de Fehr et Schmidt (1999) et de Bolton et Ockenfels (2000) est que dans les deux modèles les joueurs ne se préoccupent pas de la distribution des gains entre les autres joueurs. Lorsque les jeux ont plus de deux joueurs, il existe au contraire des différences significatives entre les prédictions des deux modèles. En effet, dans ce cas, Fehr et Schmidt supposent qu'un sujet se compare à chacun des autres sujets séparément. Contrairement à Fehr et Schmidt, dans l'approche de Bolton et Ockenfels, les joueurs comparent leur gain matériel au gain moyen du groupe et cherchent alors par leurs décisions à se rapprocher de la moyenne du groupe. Ainsi, si le gain du sujet i est inférieur à la moyenne, alors il cherchera à réduire le gain d'un autre joueur afin de faire baisser la moyenne du groupe. Il est alors indifférent entre réduire le gain d'un sujet qui a un gain supérieur ou inférieur au sien.

Les travaux expérimentaux qui se sont attachés à comparer les modèles de Bolton et Ockenfels et Fehr et Schmidt ont montré que les individus ne se comparent pas à la moyenne du groupe mais plutôt aux autres membres du groupe pris séparément (Fehr et Fischbacher, 2000 ; Engelmann et Strobel, 2000; Charness et Rabin, 2000 ; Zizzo et Oswald, 2000). Dans une relation bilatérale, il n'y a pas d'ambiguïté sur la référence de comparaison. Dans un jeu avec plusieurs joueurs, le choix de la référence est moins évident. Fehr et Fischbacher (2000) ont testé le modèle de Bolton et Ockenfels (2000) à l'aide du jeu suivant. Soit un jeu avec trois joueurs, respectivement les joueurs A, B et C. Dans un premier temps, le joueur A reçoit une dotation initiale de 100 ECU et doit décider du montant de cette dotation qu'il attribue au joueur B. Dans un deuxième temps, le joueur B reçoit la somme attribuée par le joueur A. Le joueur B n'a, par ailleurs, aucune décision à prendre. Enfin, dans un troisième temps, le joueur C reçoit une dotation initiale de 50 ECU. Il peut alors utiliser cette dotation afin de punir le joueur A après avoir observé combien ce dernier a attribué au joueur B. Pour une unité monétaire dépensée par C en punition, le gain de A est réduit de 3 unités. Selon les prédictions du modèle de Bolton et Ockenfels (2000), C ne devrait jamais punir. Fehr et Fischbacher observent au contraire, que les joueurs A sont fortement sanctionnés lorsqu'ils attribuent au joueurs B un montant relativement faible de la dotation. Cette observation indique donc que les sujets sont averses aux inégalités entre les joueurs.

Charness et Rabin (2000) ont également testé les prédictions des modèles de Bolton et Ockenfels (2000) et Fehr et Schmidt (1999). Ils considèrent le jeu suivant : trois joueurs,

A, B et C. Le joueur C doit choisir entre deux allocations de gains pour les trois joueurs : (575,575,575) ou (900,300,600). Puisque les deux allocations procurent au joueur C un gain exactement identique au gain moyen des deux autres joueurs (c'est à dire $1/3$ du surplus), le joueur C devrait choisir d'après le modèle de Bolton et Ockenfels (2000) la première allocation qui lui procure un gain supérieur. Au contraire, si les prédictions du modèle de Fehr et Schmidt (1999) sont vérifiées, alors le joueur C devrait choisir la première allocation dès lors que le gain de C est identique à chacun des gains individuels de chaque joueur. Charness et Rabin observent que plus de la moitié des sujets choisissent la première allocation des gains, ce qui s'oppose aux prédictions du modèle de Bolton et Ockenfels (2000) et semble confirmer les prédictions du modèle de Fehr et Schmidt.

Une objection souvent faite aux modèles axés sur les considérations distributives est qu'ils négligent d'incorporer un élément prépondérant, à savoir l'influence des intentions. En effet, les modèles de Bolton et Ockenfels et Fehr et Schmidt supposent que les joueurs ne se préoccupent pas des intentions des autres joueurs. L'absence de modélisation explicite des intentions dans les modèles d'aversion à l'inégalité ne signifie cependant pas qu'ils soient incompatibles avec une interprétation des comportements réiproques fondée sur les intentions. En effet, les paramètres α et β du modèle de Fehr et Schmidt peuvent être interprétés, à partir d'une approche fondée sur les intentions, puisque les intentions bonnes ou mauvaises sont en générales inférées par ce qu'implique une action en termes d'équité. Ils peuvent alors varier selon l'importance accordée aux intentions. Ainsi, les individus qui ont le désir de punir une mauvaise intention se comportent comme s'ils détestaient avoir des gains inférieurs à ce qu'ils jugent être équitable ; ceux qui récompensent de bonnes intentions se comportent comme s'ils préféraient avoir des gains supérieurs à cette même référence d'équité. Par conséquent, les paramètres de préférence dans le modèle de Fehr et Schmidt sont compatibles avec l'interprétation de la réciprocité qui repose sur les intentions.

Il s'avère que, même si aucune modélisation du comportement n'est entièrement satisfaisante, le succès des modèles de Bolton et Ockenfels et Fehr et Schmidt repose d'une part sur leur simplicité et d'autre part sur leur capacité à expliquer une grande variété d'observations expérimentales. C'est pour ces raisons que le choix de la fonction d'utilité modifiée utilisée par la suite pour modéliser la pression des pairs porte sur l'approche d'aversion à l'inégalité de Fehr et Schmidt (1999).

3. La tentation du passage clandestin contrariée par les considérations distributives

L'objet de cette section est de modéliser la pression des pairs, en considérant que l'utilité des joueurs dépend non seulement des profits du jeu standard original mais également de l'aversion à l'inégalité auto-centrée. L'approche d'aversion à l'inégalité retenue est celle de Fehr et Schmidt (1999). En effet, l'approche de Bolton et Ockenfels (2000) semble peu

adéquate puisqu'elle ne permet pas d'expliquer pourquoi les coopérateurs désirent punir uniquement les *passagers clandestins* et ne sont pas indifférents à punir un coopérateur ou un passager clandestin³⁷. On montre que sous certaines conditions relatives aux paramètres d'aversion à l'inégalité, la coopération peut émerger lorsque les sujets ont l'opportunité de sanctionner leurs pairs. L'idée défendue ici est que l'incorporation des considérations sociales d'aversion à l'inégalité dans la fonction d'utilité des agents permet de rendre rationnel tout comportement de sanction. Si les comportements de sanction deviennent crédibles, alors les sujets sont incités à coopérer à la production de l'output.

Dans un premier temps un modèle sans pression des pairs sous l'hypothèse d'aversion à l'inégalité auto-centrée est présenté. Ce modèle permet de comparer les niveaux de coopération obtenus avec et sans opportunité de sanction. Dans un deuxième temps est présenté un modèle avec pression des pairs. Enfin, dans un troisième temps, certaines hypothèses concernant le coût des sanctions et leur amplitude sont relâchées.

3.1. Modèle sans pression des pairs

Au début du jeu, les joueurs forment un groupe afin de produire ensemble un bien. S'ils coopèrent, davantage d'output est produit que s'ils ne coopèrent pas. Dans un tel jeu, sans l'hypothèse d'aversion à l'inégalité, il n'existe aucune ambiguïté quant à l'issue du jeu impliquant des individus égoïstes et rationnels : chaque sujet est incité à adopter un comportement de passager clandestin. Les prédictions théoriques de ce jeu restent-elles inchangées sous l'hypothèse d'aversion à l'inégalité ?

3.1.1. Hypothèses et fonctions de gain

Soit $n \geq 2$ le nombre de membres du groupe et m , $m \leq n$, le nombre de joueurs qui coopèrent. Une fois réalisé, l'output est partagé de façon égale entre tous les membres du groupe. Soit $f(m,n)$ l'output moyen par membre et X la désutilité pour un joueur à coopérer à la production de l'output. Lorsque tous les membres du groupe coopèrent, chaque joueur reçoit $f(n,n)-X$. Si certains joueurs ne coopèrent pas, ils reçoivent $f(m,n)$ alors que ceux qui coopèrent reçoivent $f(m,n)-X$. Lorsque personne ne coopère à la production du bien, aucun bien n'est produit. La fonction de gain de chaque joueur i peut donc s'écrire de la façon suivante :

(i)

Les hypothèses du modèle sont les suivantes :

(1) Un sujet qui adopte une stratégie de passager clandestin obtient un gain plus élevé qu'un coopérateur. Adopter un comportement de passager clandestin est donc une stratégie dominante telle que pour tout $m > 0$,

³⁷ Il est nécessaire ici de distinguer ici la norme à partir de laquelle les sujets décident ou pas de sanctionner un comportement (g dans le modèle de FS) de la référence utilisée pour comparer ses gains à ceux des autres joueurs. Dans le modèle de FS (1999), les sujets comparent leur gain à celui de chacun des autres sujets. Dans le modèle de BO (2000), les sujets comparent leur gain au gain moyen du groupe.

(ii)

Il est supposé par ailleurs que

, (iii)

ce qui traduit l'existence d'économies d'agrégation : la présence d'un coopérateur supplémentaire accroît le niveau moyen de l'output. En notant $m=n$, l'inégalité ci-dessus implique que les profits individuels sont plus grands lorsque le nombre de coopérateurs augmente. Normalisons l'output dans le cas où il n'y a aucun coopérateur tel que : $f(0,n)=0$; ceci traduit l'expression selon laquelle il n'y a pas d'économie d'agrégation sans coopération. On suppose par ailleurs que la présence d'un passager clandestin n'accroît pas le niveau moyen d'output. Par conséquent,

(iv)

(2) La coopération est socialement valable, de sorte que le gain net de coût est plus élevé si tous les joueurs coopèrent que s'ils ne coopèrent pas.

(v)

(3) Conformément au modèle de Fehr et Schmidt (1999), les individus ne sont pas uniquement motivés par leurs gains pécuniaires mais également par l'aversion à l'inégalité auto-centrée. La fonction d'utilité de chaque joueur $i \in \{1, \dots, n\}$ est alors donnée par :

(1)

Les paramètres $\alpha_i > 0$ et $\beta_i, 0 \leq \beta_i < 1$ traduisent respectivement l'aversion à l'inégalité que ressent un agent i à son désavantage et à son avantage. Le premier terme de l'équation (1) représente les gains pécuniaires du joueur i . Le deuxième terme de (1) mesure la perte d'utilité inhérente à l'inégalité désavantageuse ressentie. Ce terme est positif si $x_i > x_j$, et nul autrement. Le troisième terme représente la perte d'utilité en cas d'inégalité désavantageuse. On reprend les hypothèses de Fehr et Schmidt selon lesquelles la fonction d'utilité est linéaire en gain et en aversion pour l'inégalité et que $\alpha_i > \beta_i$.

3.1.2. Résolution du modèle

Il s'agit d'examiner ici l'équilibre de Nash symétrique dans lequel tous les joueurs choisissent de ne pas coopérer. Lorsque les agents n'ont pas l'opportunité de punir leurs pairs, la coopération ne peut pas émerger. Ceci provient du fait que la seule possibilité pour un coopérateur de réduire l'inégalité désavantageuse ressentie vis à vis de ceux qui ne coopèrent pas consiste à réduire son propre effort de production.

Proposition 1 : *lorsque les sujets n'ont pas l'opportunité de se sanctionner mutuellement, alors personne n'est incité à coopérer à la production de l'output.*

Preuve: Afin de montrer que la combinaison des stratégies où tous les sujets adoptent un comportement de passager clandestin est un équilibre de Nash, il faut vérifier que personne n'est incité à dévier de cet équilibre. Considérons la fonction d'utilité du joueur i s'il adopte un comportement de passager clandestin tandis que les autres adoptent également un comportement de passager clandestin. Le profit de chaque joueur

i s'écrit alors:

(2)

Supposons que le joueur *i* dévie de l'équilibre et choisit de coopérer. Alors son gain est donné par:

(3)

Le premier terme de la fonction d'utilité du joueur *i* représente l'output moyen que reçoit le joueur *i*. Le deuxième terme représente le coût pour l'agent *i* de coopérer. Enfin, le troisième terme traduit l'inégalité désavantageuse ressentie par *i* vis à vis des (*n*-1) autres joueurs qui adoptent un comportement de passager clandestin. Le gain net d'être passager clandestin est *X*.

Considérons l'incitation pour l'agent *i* à dévier. L'agent *i* est incité à coopérer si et seulement si la condition suivante est vérifiée:

(4)

D'après l'hypothèse (ii), on sait que : $f(m,n)-X < f(m-1,n)$. Il apparaît donc que $f(0,n) > f(1,n)-X$. Par ailleurs, lorsque *i* coopère alors que tous les autres sujets adoptent un comportement de passager clandestin, il ressent une inégalité désavantageuse vis à vis des (*n*-1) autres joueurs. Puisque $\alpha_i > 0$, alors la condition ci-dessus n'est jamais vérifiée. Personne n'est donc incité à dévier de l'équilibre. Par conséquent, les sujets ne sont pas incités à coopérer même lorsqu'ils sont averses à l'inégalité. Les prédictions théoriques restent donc identiques à celles du modèle standard sans aversion à l'inégalité.³⁸

3.2. Modèle avec pression des pairs

Considérons maintenant un jeu séquentiel où les agents ont l'opportunité de punir leurs pairs en leur attribuant des sanctions. Dans une première étape, les sujets décident simultanément de coopérer ou pas à la production de l'output. Dans une deuxième étape, ils décident simultanément de sanctionner ou pas leurs pairs. A l'instar d'Hirshleifer et Rasmusen (1989), on considère que les décisions de coopération et de sanction sont discrètes (coopérer ou ne pas coopérer ; sanctionner ou ne pas sanctionner). Comme on l'a vu dans le chapitre 1, dans les modèles de pression des pairs sans aversion à l'inégalité, la coopération ne peut pas théoriquement émerger lorsque sanctionner est coûteux car les menaces de sanction ne sont pas crédibles. En effet, en deuxième étape du jeu personne n'est incité à sanctionner puisque cela est coûteux. En première étape du jeu, les joueurs ne coopèrent pas puisqu'ils anticipent qu'ils ne seront pas sanctionnés à l'étape suivante. La stratégie d'équilibre parfait en sous jeu consiste à ne jamais coopérer et à ne jamais sanctionner.

En revanche, si l'on considère que les joueurs sont averses à l'inégalité, alors la coopération peut émerger sous certaines conditions. La coopération repose sur le fait que la sanction devient crédible si les sujets sont suffisamment averses à l'inégalité. L'idée est la suivante : en deuxième étape du jeu, il existe un gain pour un coopérateur

³⁸ Une preuve formelle de l'unicité de l'équilibre sans coopération est présentée en annexe A.

suffisamment averses à l'inégalité, à sanctionner un passager clandestin afin de réduire l'inégalité désavantageuse ressentie.

3.2.1. Hypothèses et fonctions de gain

(4) En plus des hypothèses générales avancées plus haut, une hypothèse est faite sur l'ampleur des paramètres telle que :

(vi)

Quand un joueur est puni par un pair, il reçoit une sanction Y qui réduit son gain. La sanction doit être suffisamment élevée afin que ceux qui dévient n'obtiennent pas un profit plus élevé que ceux qui ne dévient pas de l'équilibre. Autrement dit, la sanction doit être plus élevée que le gain que peut retirer un agent en adoptant un comportement de passager clandestin. Sanctionner est coûteux pour celui qui inflige la punition. Ainsi un sujet qui sanctionne un autre membre de son groupe subit un coût, cY ; $0 < c \leq 1$.

3.2.2. Résolution du modèle

Dans cette section, on étudie la combinaison de stratégies appelée stratégie de *coopération* qui supporte un équilibre parfait en sous jeu symétrique qui consiste à coopérer en première étape du jeu et à sanctionner tous ceux qui n'ont pas coopéré, sauf si l'on a dévié de la stratégie d'équilibre à la première étape ; dans ce cas, on dévie également en deuxième étape. A l'équilibre, tous les sujets coopèrent à la production de l'output et personne n'est donc sanctionné en deuxième étape du jeu. Si un joueur dévie de l'équilibre, les autres continuent de coopérer mais le sanctionnent.

Proposition 2 : *lorsque les agents sont suffisamment averses à l'inégalité, l'opportunité de sanctionner ses pairs permet d'atteindre la coopération.*

Preuve : Contrairement au modèle sans pression des pairs, l'opportunité de sanctionner permet aux sujets de réduire l'inégalité ressentie sans avoir à réduire son propre effort de production. L'induction à rebours est utilisée pour vérifier que la combinaison des stratégies consistant à coopérer et à sanctionner tous ceux qui ne coopèrent pas est un équilibre parfait en sous jeu.

(1) En deuxième étape du jeu, les sujets décident de punir ou pas les autres membres de leur groupe. Soit $n \geq 2$, le nombre de membres dans le groupe et m , le nombre de coopérateurs à l'étape précédente. Vérifions qu'aucun sujet i ne soit incité à dévier à cette étape du jeu. (i) Considérons tout d'abord le cas où le sujet i a dévié à l'étape précédente, c'est à dire s'il n'a pas coopéré. Si le joueur i décide de punir un passager clandestin, il subit alors la désutilité à sanctionner. Or, comme il n'y a pas de différence de gain avec les autres passager clandestin, le joueur i n'est donc pas incité à sanctionner un autre passager clandestin. On vérifie donc que si un sujet dévie de la stratégie d'équilibre en première étape, il dévie également en deuxième étape. (ii) Considérons maintenant le cas où le joueur i n'a pas dévié à l'étape précédente, c'est à dire qu'il a coopéré. S'il ne dévie pas de la stratégie d'équilibre à cette étape et qu'il punit ceux qui ne coopèrent pas, sa fonction d'utilité s'écrit :

Le premier terme de la fonction d'utilité ci-dessus représente l'output moyen que

l'agent i reçoit. Le deuxième terme représente le coût pour l'agent i de coopérer. Le troisième terme représente le coût de sanctionner les $(n-m)$ passagers clandestins. Le quatrième terme traduit l'inégalité désavantageuse que l'agent i éprouve vis à vis des passagers clandestins. Cette inégalité s'accroît du fait de l'existence du coût à sanctionner mais est réduite par les sanctions que reçoivent les passagers clandestins.

Considérons maintenant la fonction d'utilité de l'agent i s'il choisit de ne pas punir les *passagers clandestins* telle que :

Si l'agent i choisit de ne pas punir les autres membres de son groupe, il économise alors le coût $cY(n-m)$ mais ressent plus d'inégalité désavantageuse vis à vis des passagers clandestins, ce qui génère une perte d'utilité de $[\square (n-m)Y(1-c(n-m))]/(n-1)$. Il subit par ailleurs de l'inégalité avantageuse vis à vis des $(m-1)$ joueurs qui subissent le coût de sanctionner, ce qui provoque une perte d'utilité de $[(n-m) \square (m-1)cY]/(n-1)$. La comparaison de (5) et (6) montre qu'un joueur décide de punir un passager clandestin si et seulement si la condition suivante est vérifiée³⁹ :

Quelques manipulations algébriques montrent que la condition ci-dessus est équivalente à la condition suivante:

(2) En première étape du jeu, les agents décident de coopérer ou non à la production du bien. Supposons que la condition (8) soit vérifiée. Considérons les incitations des agents à ne pas coopérer, c'est à dire à dévier de l'équilibre dans la première étape du jeu. Le gain immédiat obtenu en déviant est $f(n-1, n)$ au lieu de $f(n, n)-X$. Cependant, s'il dévie, l'agent i sera puni en deuxième étape, ce qui lui procure une perte de $(n-1)Y$. Les hypothèses (iii) et (iv) impliquent que $f(m-1, n) < f(m, n)$ (l'output s'accroît avec le nombre de coopérateurs) et d'après l'hypothèse (vi) $X < (n-1)Y$. Les joueurs ne sont donc pas incités à dévier de l'équilibre en première étape du jeu. En deuxième étape, personne ne sanctionne puisque tous coopèrent en première étape. En comparant les modèles avec et sans pression des pairs, il apparaît que l'opportunité de sanction peut faire émerger la coopération car la menace de sanction est crédible, du fait du gain à sanctionner découlant de l'aversion à l'inégalité.

Remarque 1 : Lorsque la condition (8) n'est pas vérifiée, aucun sujet n'est alors incité à sanctionner les *passagers clandestins* en deuxième étape et personne ne coopère en première étape.

Remarque 2 : Il existe un autre équilibre symétrique parfait en sous jeu où personne ne coopère et personne ne sanctionne puisque cela procure les mêmes gains pour chaque joueur et ne génère donc pas d'inégalité. En effet, lorsque tous les sujets adoptent un comportement de passager clandestin, il n'existe alors pas d'inégalité de gain et donc aucune incitation à sanctionner en deuxième étape du jeu. L'équilibre avec coopération E1 domine au sens de Pareto l'équilibre sans coopération E2. Supposons que le jeu soit répété deux fois et que les issues du premier jeu soient observées avant que le deuxième jeu commence. Alors les joueurs peuvent jouer l'équilibre Pareto dominé comme une menace en cas de non-coopération. L'idée est que les joueurs anticipent différentes issues au deuxième jeu selon les issues du premier jeu (Gibbons, 1992, pp82-88). Si le

³⁹ On retrouve la condition donnée par Fehr et Schmidt (1997), dans leur proposition 5 p19.

jeu n'est pas répété, il est difficile à priori de déterminer quel équilibre sera choisi. Cependant, si l'on applique le critère de Pareto dominance sans perfection, E1 sera l'équilibre sélectionné par les joueurs puisque tous le préfèrent (Rasmusen, 1984). Ce raffinement est toutefois arbitraire comme on l'a déjà précisé dans le chapitre précédent puisqu'il suppose que les agents se coordonnent de façon implicite sur l'équilibre pareto dominant.

3.3. Relâchement de certaines hypothèses

3.3.1. Cas particulier : $c=1$

On a supposé jusqu'à présent que la désutilité de punir est donnée par cY , $0 < c \leq 1$. On s'intéresse dans cette section au cas particulier où $c=1$, c'est à dire que le coût de punir est égal au coût d'être puni. Sous cette condition, la menace de sanction n'est plus crédible et personne ne contribue au bien public. L'intuition est la suivante : lorsque $c=1$, alors les sanctions ne permettent pas de réduire l'inégalité de gains. Il n'existe donc plus d'incitation à sanctionner les autres membres du groupe.

Proposition 3 : *lorsque la désutilité de sanctionner est égale à la désutilité d'être sanctionné, alors la coopération ne peut pas apparaître.*

Preuve :

Récrivons la condition (8) sous l'hypothèse $c=1$:

Puisque $0 \leq c \leq 1$ alors la condition (9) n'est jamais vérifiée. Donc personne n'est incité à sanctionner en deuxième étape du jeu. En première étape, la stratégie dominante est de ne pas coopérer puisque personne ne sera sanctionné en deuxième étape.

3.3.2. Application des sanctions

On considère dans cette sous-section le cas où la pénalité Y reçue par un passager clandestin est indépendante du nombre de sanctionneurs⁴⁰. Le jeu d'ostracisme présenté dans le chapitre précédent est un exemple d'une telle sanction puisqu'il suffit qu'un seul joueur décide d'exclure un autre joueur pour que la décision soit effective. Un problème de passager clandestin affecte alors les décisions des sanctions d'exclusion. En effet, sanctionner les *passagers clandestins* est efficient au sens de Pareto puisque les inégalités de gains avec les *passagers clandestins* sont alors réduites. Cependant, comme les sanctions sont indépendantes du nombre de sanctionneurs, chaque joueur espère que les autres excluent les *passagers clandestins* à sa place. Chaque joueur pris individuellement a donc intérêt à adopter un comportement de passager clandestin, c'est à dire à ne pas sanctionner. En conséquence, personne ne devrait coopérer en première

⁴⁰ Coleman, J. (1990) distingue un comportement de *sanction incrémental* d'un comportement de sanction héroïque. Un comportement de sanction incrémental suppose un coût de sanctionner faible et l'effet d'une sanction également faible mais qui s'additionne à l'effet des sanctions imposées par les autres sanctionneurs. Au contraire, dans le cas des comportements de sanction héroïque, l'effet total de la sanction émane d'un seul acteur (par exemple, la fable d'Aesop's, "The mice in council"). La différence principale entre ces deux types de sanction repose sur l'amplitude des sanctions.

étape puisque personne ne sanctionne en deuxième étape.

Proposition 4 : *lorsque les sanctions sont indépendantes du nombre de sanctionneurs, les menaces de punition ne sont pas crédibles.*

Preuve : réécrivons la condition (8) sous l'hypothèse d'indépendance des sanctions vis à vis du nombre de sanctionneurs :

Il apparaît clairement que cette condition n'est jamais vérifiée.

Donc la coopération est soutenue si les agents sont suffisamment sensibles à l'inégalité des gains relatifs et s'ils désirent réduire cette dissonance de gains en punissant les autres joueurs. En effet, même s'il est coûteux de punir les autres agents, les sujets peuvent gagner en utilité à punir les *passagers clandestins*. En comparant les cas avec et sans opportunité de sanctionner, il est aisé de voir que la coopération est considérablement améliorée s'il existe une opportunité de sanctionner les *passagers clandestins*. Sans punition, aucun joueur ne coopérera. A l'inverse, avec punition les mêmes joueurs coopéreront.

4. Le pouvoir explicatif de l'aversion à l'inégalité

Cette section vise à tester les prédictions théoriques des modèles de pression des pairs sous l'hypothèse d'aversion à l'inégalité. Pour être effectif, le pouvoir prédictif des modèles présentés, se doit d'être supérieur à celui du modèle standard fondé sur la seule hypothèse d'indifférence aux autres. Afin de tester les modèles de pression des pairs de la section précédentes, on utilise les jeux de bien public avec et sans sanction présentés dans les chapitres 2 et 3. On vérifie si les observations permettent de tester si les comportements de sanction reflètent l'aversion pour l'inégalité exprimée par les joueurs. L'intérêt de se focaliser uniquement sur les comportements de sanction est double. D'une part, l'analyse des comportements de contributions découle directement des comportements de sanction puisque les contributions sont conditionnées par l'existence de sanctions. D'autre part, l'étude des comportements de contribution a déjà fait l'objet d'une analyse détaillée dans les chapitres précédents.

Dans une première sous-section, les prédictions théoriques des différents jeux de bien public avec sanction sous l'hypothèse d'aversion pour l'inégalité sont présentées⁴¹. La deuxième sous-section présente les résultats expérimentaux dans une perspective d'aversion à l'inégalité. Enfin la troisième sous-section fait l'objet d'une discussion sur l'interprétation des résultats obtenus.

4.1. Prédictions théoriques des jeux de bien public avec aversion à

⁴¹ L'intérêt d'une telle présentation résulte du fait que les prédictions théoriques de ces jeux diffèrent légèrement du modèle de pression des pairs présenté plus haut. En effet, dans un jeu de bien public avec sanction, les décisions des joueurs en termes de coopération et de sanctions sont continues ($g_i \in [0, y]$ et $P_{ij} \in [0, 10]$) et non discrètes ($X > 0 / X = 0$) et $P > 0 / P = 0$).

l'inégalité

Les jeux présentés dans les chapitres précédents sont étudiés ici sous l'hypothèse d'aversion à l'inégalité. Les tableaux 2a et 2b résument la structure ainsi que les prédictions des différents jeux.

Tableau IV.2b : Descriptif des expériences de bien public avec/sans opportunité de sanction

Traitements	1 ^{ère} étape : contribution	2 ^{ème} étape : sanction
Traitement SP	Contribution au financement du bien public $n=4$; $\alpha=0.4$	
Traitement MP	Contribution au financement du bien public $n=4$; $\alpha=0.4$	Attribution de points de sanction monétaire
Traitement NP	Contribution au financement du bien public $n=4$; $\alpha=0.4$	Attribution de points de sanction non monétaire
Traitement ESC	Contribution au financement du bien public $n=4$; $\alpha=0.4$	Exclusion non coûteuse ⁴² à la participation d'un second projet
Traitement EC	Contribution au financement du bien public $n=4$; $\alpha=0.4$	Exclusion coûteuse ⁴³ à la participation d'un second projet

Tableau IV.2b : Prédictions théoriques des différents traitements

Traitements	Modèle standard	Modèle sous l'hypothèse d'aversion à l'inégalité
SP	$g_i=0, i=1, \dots, n$	$g_i=0, i=1, \dots, n$
MP	$g_i=0; p_{ij}=0, i, j=1, \dots, n; i \neq j$	$g_i=g; g \in [0, y]$ et $p_{ji}=(g-g_i)/(m-c)$ si $g_i < g; i, j=1, \dots, n; i \neq j$
NP	$g_i=0; p_{ij}=0, i, j=1, \dots, n; i \neq j$	$g_i=0; p_{ij}=0, i, j=1, \dots, n; i \neq j$
EC	$g_{i1}=0; \text{personne n'exclut}; g_{i2}=0; i, j=1, \dots, n; i \neq j$	$g_{i1}=0; \text{personne n'exclut}; g_{i2}=0; i, j=1, \dots, n; i \neq j$

4.1.1. Jeu de bien public sans opportunité de sanction

Dans le jeu de bien public sans opportunité de sanction (SP), les sujets décident simultanément de leur niveau de contribution $g_i \in [0, y], i=1, \dots, n$ à un bien public. Chaque joueur a une dotation initiale y . Sous l'hypothèse d'aversion à l'inégalité, l'utilité des joueurs dépend non seulement des profits du jeu standard original mais également de l'aversion à l'inégalité. La fonction d'utilité de chaque joueur $i \in 1, \dots, n$ est alors

⁴² et ⁵³ : L'exclusion est coûteuse ou non coûteuse pour celui qui exclut.

⁴³

donnée par :

(11)

avec

où a représente le retour marginal constant du bien public $G = g_j$.

Considérons l'équilibre où personne ne contribue au bien public. Afin de s'assurer que la stratégie qui consiste à ne pas contribuer au bien public supporte un équilibre de Nash, il suffit de vérifier que personne n'est incité à dévier de cet équilibre proposé. Soit la fonction d'utilité du joueur i s'il contribue $g_i=0$, tandis que les autres, à l'équilibre, ne contribuent pas non plus au bien public.

Considérons la déviation pour le sujet i qui consiste à contribuer un montant positif au bien public:

Si un sujet i accroît sa contribution, il réduit son gain monétaire de $1-a$ et subit par ailleurs de l'inégalité désavantageuse vis à vis des $n-1$ autres sujets qui ne contribuent pas au bien public. Donc le sujet i n'est pas incité à dévier de l'équilibre (voir annexe B pour l'unicité de l'équilibre). L'introduction de l'aversion à l'inégalité dans un jeu de bien public sans opportunité de sanction ne change donc pas les prédictions théoriques du jeu : la stratégie dominante consiste à ne pas contribuer au bien public⁴⁴.

4.1.2. Jeu de bien public avec opportunité de sanction monétaire

Dans le jeu avec opportunité de sanction monétaire (MP), les sujets ont l'opportunité d'assigner des points de sanction monétaire après avoir observé le niveau de contribution des autres agents. Sous l'hypothèse d'aversion à l'inégalité, la fonction d'utilité de chaque joueur i est redéfinie par :

avec

Si les sujets sont suffisamment averses à l'inégalité, ils seront incités à sanctionner ceux qui adoptent un comportement de passager clandestin même si punir est coûteux. Les coopérateurs ont alors la possibilité de réduire les dissonances de gains avec les *passagers clandestins* sans avoir à réduire leur propre niveau de contribution. La menace de sanction est alors crédible et incite les sujets à contribuer au bien public.

La sanction optimale p_{ij}^* égalise les gains des sujets i et j tel que :

L'induction à rebours est utilisée une fois de plus pour vérifier que la combinaison de stratégies ci-dessus est vérifiée. Vérifions qu'en deuxième étape du jeu, aucun des m coopérateurs (ayant pour préférences $\alpha_i, \alpha_j \geq 1-a$) n'est incité à réduire son niveau de

⁴⁴ Fehr et Schmidt (1999) ont appliqué leur modèle d'aversion à l'inégalité à différents jeux expérimentaux et notamment à un jeu de contribution volontaire au financement d'un bien public sans opportunité de sanction. Ils observent que lorsque tous les agents ont les mêmes préférences relatives à l'aversion à l'inégalité ($1-a > \alpha_i$), alors il existe un seul équilibre où personne ne contribue au bien public ($g_i=0$). Fehr et Schmidt montrent qu'il existe un autre équilibre avec des niveaux positifs de contribution si l'on suppose que les sujets ont des préférences relatives à l'aversion à l'inégalité différentes de sorte qu'il existe k sujets avec $\alpha_i + \alpha_j > 1$, (où a représente le rendement marginal du bien public), tandis que les autres sujets j ont des préférences tel que $\alpha_i + \alpha_j < 1$. Si k est tel que $k/(n-1) < (\alpha_i + \alpha_j - 1)/(\alpha_i + \alpha_j)$, les k sujets ne contribuent pas au bien public ($g_i=0$) tandis que les autres sujets j contribuent $g_j = g \in [0, y]$.

sanction. Si un sujet décide de réduire son niveau de sanction p_{ij} , il économise le coût de sanctionner mais fait désormais l'expérience de plus d'inégalité à son désavantage par rapport aux *passagers clandestins*. Il subit par ailleurs de l'inégalité avantageuse vis à vis de ceux qui sanctionnent et subissent le coût de sanctionner. Un coopérateur n'est pas incité à réduire son niveau de sanction si et seulement si la condition suivante est vérifiée :

4.1.3. Jeu de bien public avec opportunité de sanction non monétaire

Dans le jeu avec opportunité de sanction non monétaire (NP), les agents ont l'opportunité de sanctionner leurs pairs en leur infligeant des points de sanction non monétaire. Ces points n'affectent ni les gains de ceux qui les reçoivent ni de ceux qui les donnent. Ils indiquent seulement le niveau de désapprobation ressentie vis à vis du comportement d'un autre sujet. Dans ce jeu, les points non monétaires ne permettent donc pas de réduire les dissonances de gains puisque qu'il n'y a pas de coût à sanctionner et/ou à être sanctionné (voir condition (9) de la section 3.3). Il n'existe alors pas d'incitation à sanctionner. En première étape, la stratégie dominante consiste à ne pas coopérer puisque personne ne sera sanctionné en deuxième étape.

4.1.4. Jeu de bien public avec possibilité d'exclusion coûteuse

Dans le jeu avec possibilité d'exclusion (EC), les sujets ont la possibilité d'exclure d'un deuxième projet les autres joueurs après avoir observé leur effort de contribution à un premier projet. Il suffit qu'un seul joueur décide d'exclure un autre membre du groupe pour que ce dernier soit exclu. Lorsqu'un joueur est exclu, il ne peut pas participer à un deuxième projet.

Réolvons le jeu d'ostracisme par induction à rebours. En troisième étape du jeu, les $(n - n_{os})$ sujets qui n'ont pas été exclus à l'étape précédente décident de leur niveau de contribution au deuxième projet. La stratégie dominante consiste donc à ne pas contribuer au bien public. En deuxième étape, les sujets décident d'exclure ou non les autres joueurs. Dans le jeu précédent avec sanctions monétaires, un sujet pouvait être incité à sanctionner un passager clandestin afin de réduire la dissonance des gains. Dans ce jeu d'ostracisme, il suffit qu'un sujet décide d'exclure un autre membre de son groupe pour que sa décision soit appliquée. Un coopérateur n'a donc plus d'incitation à sanctionner un autre membre de son groupe s'il pense que les autres vont le sanctionner à sa place (voir condition (10) de section 3.3). En effet, un joueur peut espérer que les autres réduisent à sa place les dissonances de gains en sanctionnant ceux qui ne coopèrent pas au premier projet. Si tous les sujets sanctionnent ceux qui ne contribuent pas au premier projet, les dissonances de gains sont alors réduites, ce qui procure une utilité pour tous les joueurs en termes de réduction de l'inégalité. Cependant, chaque joueur pris isolément maximise son utilité s'il ne sanctionne pas du fait d'un problème de passage clandestin sur les décisions de sanction. La stratégie dominante consiste donc à ne pas sanctionner en deuxième étape. En première étape, personne ne devrait contribuer au bien public puisque personne ne sanctionne en deuxième étape.

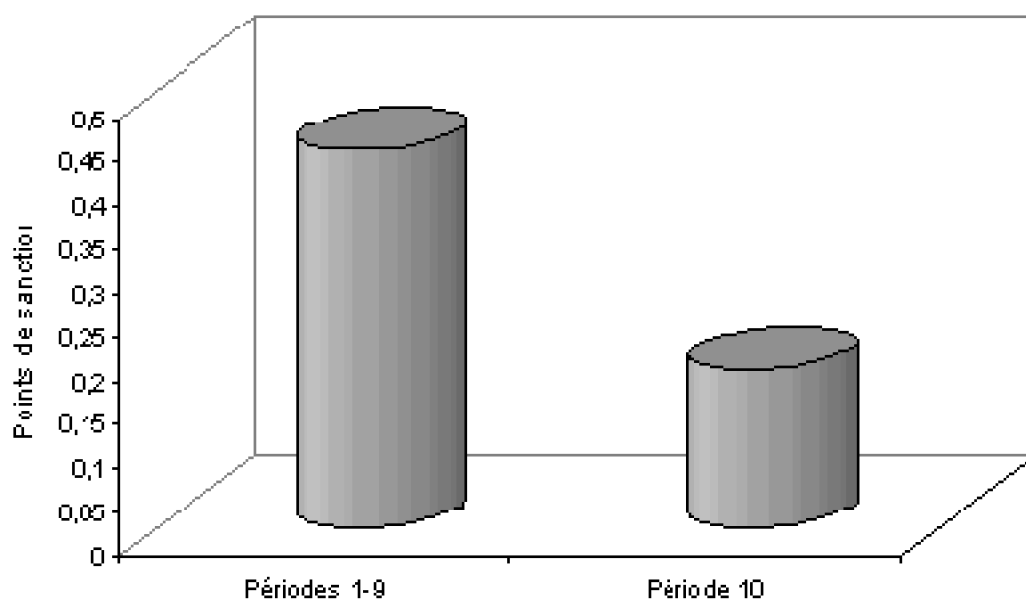
4.2. Résultats expérimentaux

On s'intéresse dans cette sous-section à l'étude des comportements individuels de sanction à partir des données expérimentales présentées dans les chapitres précédents. Dans un premier temps, on cherche à distinguer les comportements de sanction d'ordre stratégique des comportements non stratégiques. En effet, le chapitre précédent a montré que les agents sanctionnent à la fois pour inciter les autres à davantage coopérer dans le futur (raison stratégique) mais également pour des considérations sociales (raisons non stratégiques). Dans un deuxième temps, on montre que ces motivations d'ordre non stratégique sont essentiellement des motivations fondées sur l'hypothèse d'aversion à l'inégalité. Pour cela, on étudie la relation existant entre les sanctions attribuées à un joueur par un autre et les différence de gains entre ces deux joueurs.

4.2.1. Comportements de sanction

Afin de dissocier les raisons *stratégiques* à sanctionner des raisons *non-stratégiques*, on compare le nombre de sanctions attribuées dans la période finale du jeu et dans les 9 autres périodes, et ce, dans les différents jeux. En effet, en dernière période, seul l'effet non stratégique permet d'expliquer pourquoi les sujets sanctionnent les autres membres du groupe. Au contraire, dans les 9 premières périodes, un sujet peut décider de sanctionner les autres membres de son groupe à la fois pour des raisons non stratégiques et pour des raisons stratégiques. En dernière période, l'effet stratégique disparaît.

Si les raisons stratégiques ne sont pas les seuls motifs des agents à sanctionner leurs pairs et si l'hypothèse d'aversion à l'inégalité est vérifiée, on devrait observer que les sujets continuent de sanctionner en dernière période du traitement MP. Au contraire, aucun sujet ne devrait sanctionner pour des raisons non stratégiques liées à l'aversion à l'inégalité dans les traitements NP et EC. En effet, dans le traitement NP, les points de sanctions ne permettent pas de réduire les différences de gains. Dans le traitement EC personne n'est incité à sanctionner du fait du problème de passager clandestin sur les sanctions. Le graphique 1 décrit la distribution des sanctions monétaires attribuées à un individu selon la période du jeu.



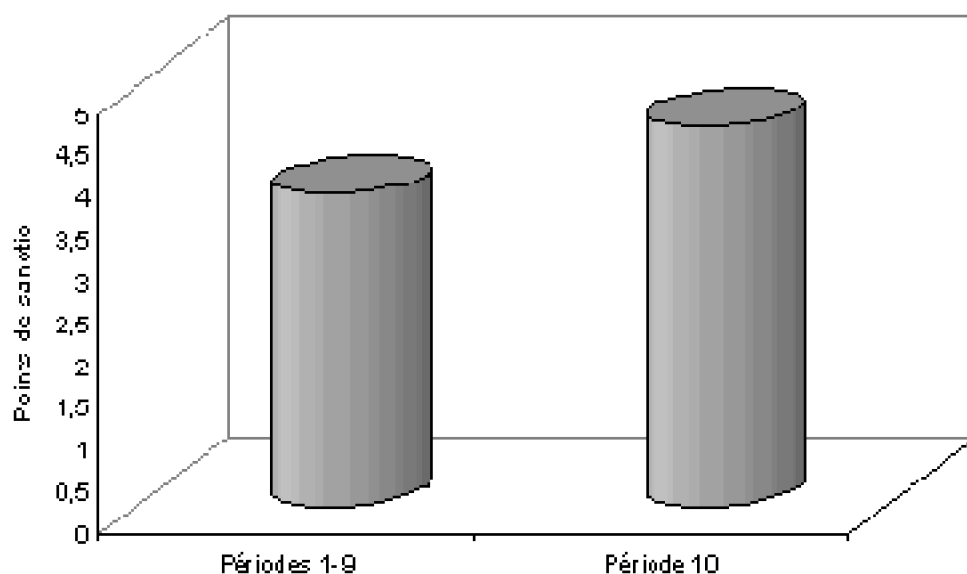
Graphique VI.1 : Nombre moyen de sanction monétaire distribués par individu et par période (Traitement MP)

On observe qu'un sujet attribue en moyenne 1/2 point de sanction monétaire à un autre membre de son groupe à chaque période lors des 9 premières périodes et environ 0.3 point à la dernière période. Ainsi, les sujets attribuent moins de points de sanction en dernière période du traitement MP que lors des 9 périodes précédentes. Cette observation semble montrer que les sujets sanctionnent essentiellement pour des raisons d'ordre non stratégique, ce qui semble infirmer les prédictions théoriques sous l'hypothèse d'aversion à l'inégalité. Toutefois, cette différence n'est pas significative d'après le test de *Wilcoxon signed rank* ($p > 0.05$). Par ailleurs, la diminution des sanctions en dernière période peut s'expliquer en partie par le fait que plus de 80% des sujets contribuent toute leur dotation de 20 ECU en dernière période. Le niveau relativement important des sanctions en dernière période, malgré cette explication, souligne donc au contraire l'existence des motivations de sanction d'ordre *non stratégique*.

RESULTAT 1 : *En dernière période du traitement MP, les sujets continuent à sanctionner les autres membres de leur groupe.*

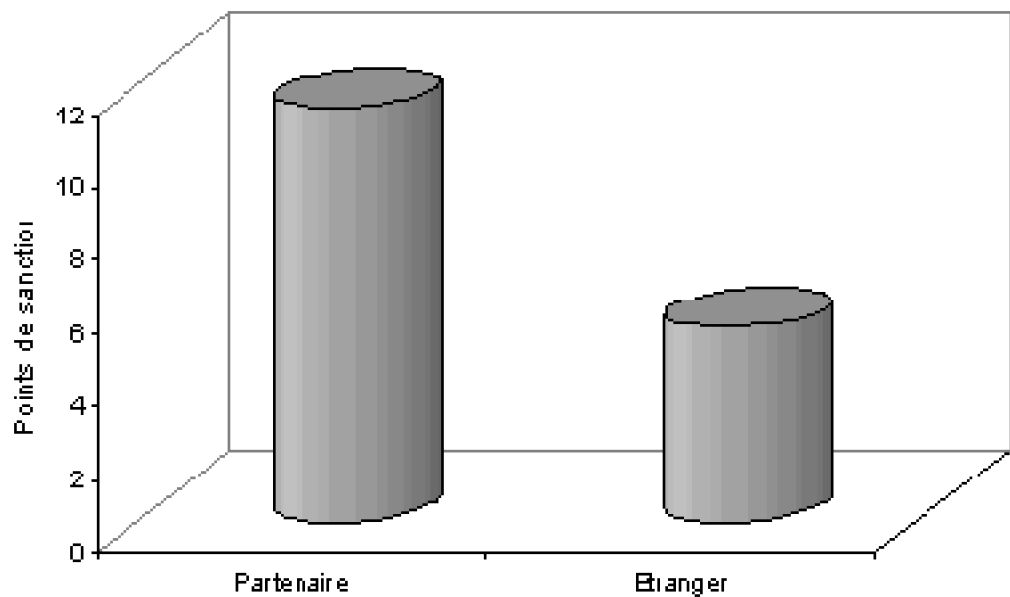
Le résultat 1 met en évidence le fait que les sujets continuent de sanctionner en dernière période. Les agents sont donc également motivés par des raisons non stratégiques, ce qui est conforme avec l'hypothèse d'aversion à l'inégalité.

Le graphique 2 décrit la distribution du niveau moyen des sanctions non monétaires distribuées par un individu à chaque période dans les 9 premières périodes et en dernière période du traitement NP.



Graphique IV.2 : Points de sanction non monétaire distribués par individu et par période (Traitement NP)

Le graphique 2 indique des niveaux élevés de sanction pour l'ensemble des périodes. En moyenne, un sujet affecte à chaque période environ 4 points de sanction non monétaire à un individu lors des 9 premières périodes et 4.5 points en dernière période. Une comparaison du nombre moyen de points de sanction distribués en périodes 1-9 et dans la période 10 indique que les sujets sont davantage incités à sanctionner en dernière période. D'après le test de *Wilcoxon signed rank* sur le nombre moyen de points non monétaires attribués par un individu, cette différence est significative ($p < 0.05$). Afin de confirmer les observations du graphique 10, comparons les sessions conduites en partenaire (NP) et en étranger (NS). Cette comparaison permet de dissocier davantage les motifs de sanction d'ordre stratégique et non stratégique.



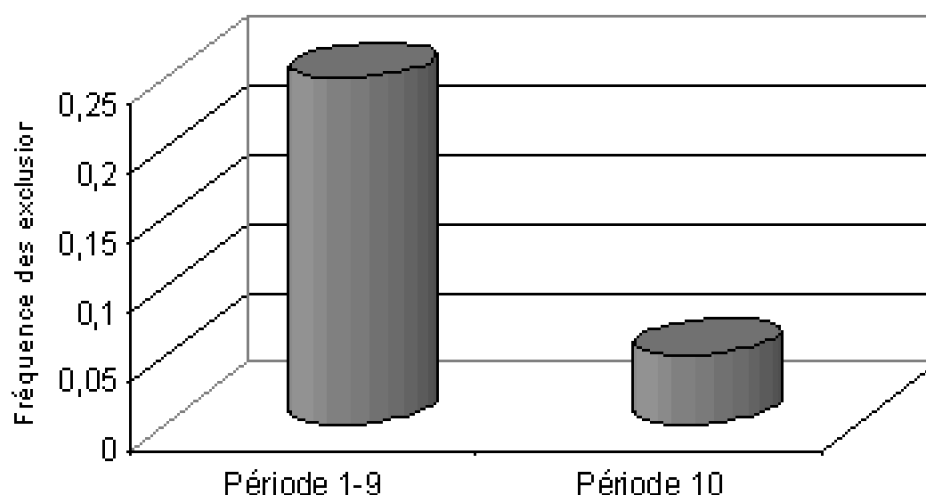
Graphique IV.3 : Points de sanction non monétaire distribués par individu et par période (Traitement NP/NS)

Le graphique 3 compare les comportements de sanction selon que les joueurs sont appariés ou non avec les mêmes personnes à chaque période du jeu. Il apparaît que les sujets attribuent davantage de points de sanction non monétaire lorsque les groupes sont fixes. En moyenne, un sujet attribue 11 points de sanction non monétaire à un individu en condition partenaire et seulement 5 points en condition étranger. D'après le test de *Wilcoxon Mann Whitney*, le nombre moyen de points de sanction non monétaire attribué par un individu à un autre individu est significativement plus élevé dans le traitement MP que dans le traitement NP ($z=3.06$, $p=0.0011$).

RESULTAT 2 : *Le nombre de sanctions non monétaires distribuées en dernière période du traitement NP est significativement plus élevé que dans les périodes précédentes. Par ailleurs, le nombre de sanctions attribuées dans le traitement NS est significativement plus faible que dans le traitement NP.*

Ce résultat met en exergue l'importance, dans le traitement NP, des motivations stratégiques à sanctionner par rapport aux motivations non stratégiques. Toutefois, les sujets continuent de sanctionner en dernière période du traitement NP. Cela semble donc infirmer l'hypothèse d'aversion à l'inégalité.

Le graphique 4 compare le niveau moyen des décisions d'exclusion par individu et par période dans les 9 premières périodes et en dernière période.



Graphique VI.4 : Exclusions par individu et par période (traitement EC)

On observe un niveau élevé d'exclusion lors des 9 premières périodes (25% des sujets excluent). Toutefois, en dernière période, seulement 5% des sujets excluent. Cette différence est significative au seuil de 5%.

RESULTAT 3 : *Les sujets excluent peu leurs pairs en dernière période, où seules les raisons non stratégiques expliquent les comportements de sanction*

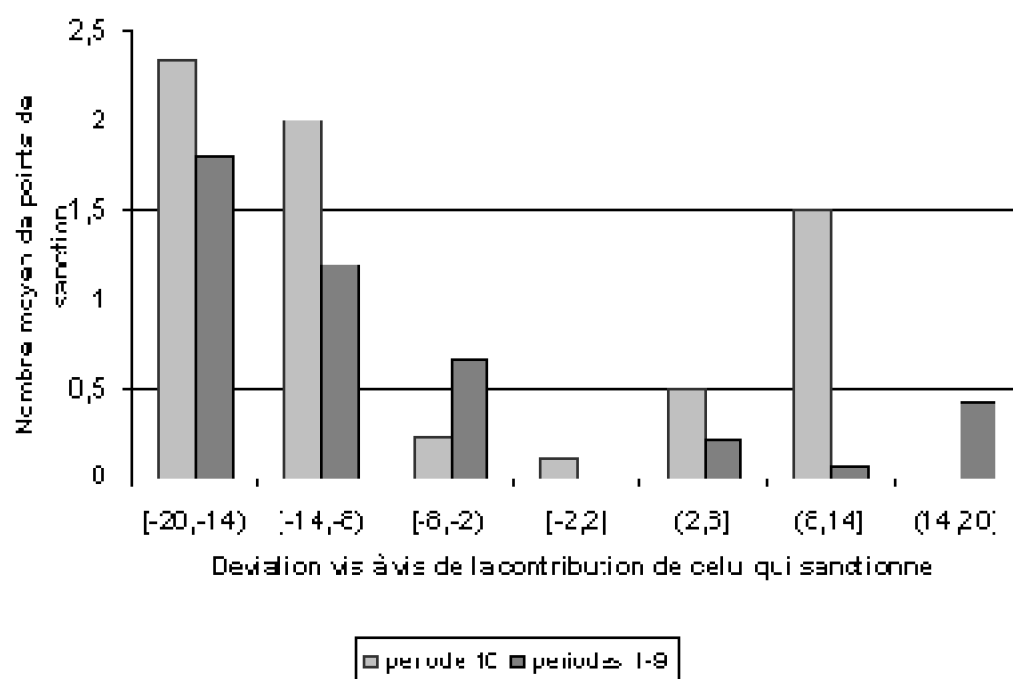
Conformément à ces prédictions théoriques, le résultat 3 montre que les sujets excluent peu en dernière période du traitement EC.

4.2.2. Sanctions et niveaux de contribution

On a montré jusqu'à présent que les agents étaient motivés par des raisons stratégiques mais également non stratégiques à sanctionner leurs pairs. L'objet de cette sous-section est de montrer que les raisons d'ordre non stratégique sont essentiellement relatives à l'aversion pour l'inégalité auto-centrée. Pour cela, on étudie la relation existant entre les sanctions attribuées par un sujet à un autre membre du groupe et la différence de gain entre ces deux joueurs. L'idée est de vérifier si les agents sont incités à sanctionner leurs pairs afin de réduire les inégalités de gain.

4.2.2.1. Attribution des points de sanction monétaire et non monétaire

Le graphique 5 illustre la façon dont les sanctions monétaires sont affectées selon les déviations de contribution par rapport à la contribution de celui qui sanctionne. Il montre combien de points de sanction monétaire un joueur affecte à un autre joueur lorsque ce dernier contribue plus ou moins que lui.



Graphique VI.5: Points de sanction monétaire donnés pour une déviation vis-à-vis de la contribution de celui qui sanctionne

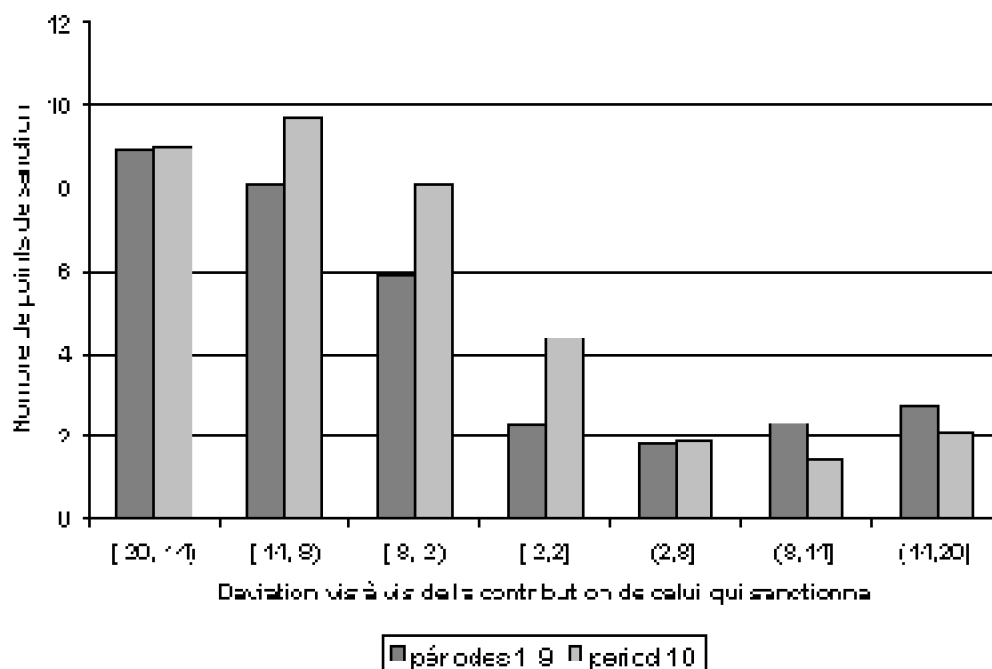
On observe qu'un joueur est d'autant plus sanctionné que sa contribution est inférieure à celle du joueur qui sanctionne. Lorsque les joueurs contribuent davantage que celui qui sanctionne, ils sont également sanctionnés mais dans une moindre mesure. Enfin, les contributions qui se rapprochent de la contribution de celui qui sanctionne, c'est à dire dans l'intervalle $[-2,2]$ sont très peu, voire pas sanctionnées. Au regard de ces observations, il semblerait que les raisons d'ordre non stratégique relatives à l'aversion à l'inégalité génèrent des motivations à sanctionner les autres membres du groupe. En effet, les agents seraient incités à sanctionner ceux qui contribuent moins qu'eux afin de réduire les différences de gains.

Le graphique 5 compare également les comportements de sanction de dernière période aux périodes précédentes. Les sanctions sont plus sévères en dernière période que dans les périodes précédentes sur certains intervalles. Ainsi, un joueur qui contribue entre 14 et 20 ECU de moins que la contribution de celui qui sanctionne reçoit en moyenne 1.8 point de sanction monétaire lors des 9 premières périodes et 2.4 points en dernière période. De même, lorsqu'un joueur contribue entre 8 et 14 ECU de plus que celui qui sanctionne, il reçoit 1.5 points en dernière période contre moins de 0.5 points lors des 9 premières périodes. Toutefois, dans le traitement MP, les comportements de sanction entre les 9 premières périodes et la dernière période ne sont pas significativement différents ($p > 0.05$). Les raisons d'ordre stratégique ne semblent donc pas générer de motivations supplémentaires à sanctionner par rapport aux raisons d'ordre non stratégique de la dernière période.

Le résultat 4 résume la façon dont sont affectées les sanctions monétaires.

RESULTAT 4 : Dans le traitement MP, les joueurs sanctionnent essentiellement ceux qui contribuent moins qu'eux.

Le graphique 6 illustre la relation entre le nombre de points non monétaires attribués et la déviation de la contribution du joueur sanctionné par rapport à la contribution de celui qui sanctionne.

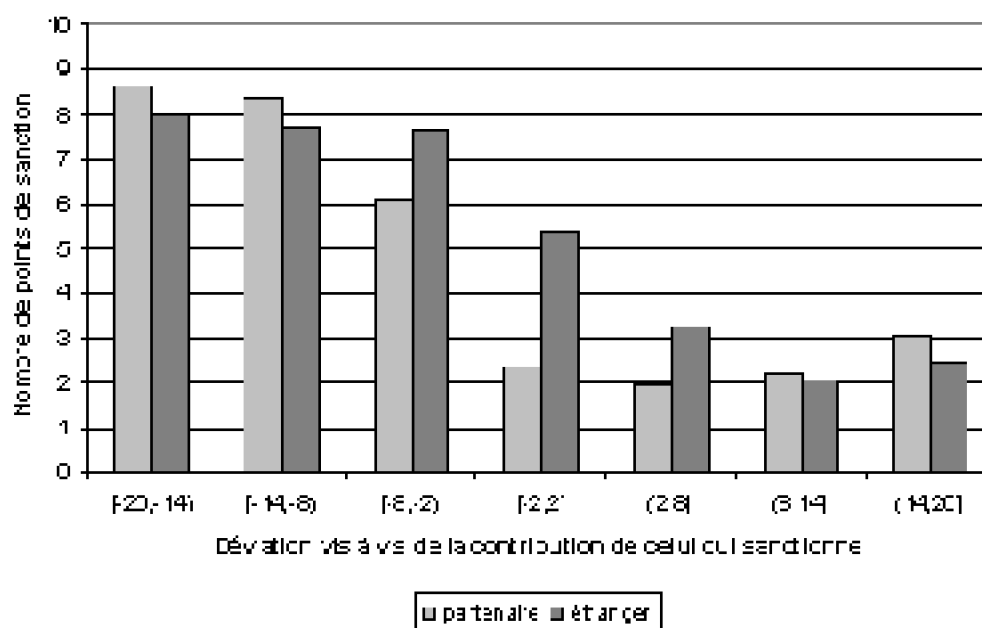


Graphique IV.6: points de sanction non monétaire donnés pour une déviation vis-à-vis de la contribution de celui qui sanctionne

Comme dans les sessions avec sanction monétaire, les joueurs attribuent davantage de points de sanction à ceux qui contribuent moins qu'eux. Les déviations positives sont également punies mais dans une moindre mesure. La comparaison des 9 premières périodes et de la dernière période montre que les déviations négatives sont davantage punies en dernière période. Ainsi, un sujet qui contribue entre 8 et 14 ECU de moins que celui qui sanctionne reçoit presque 10 points de sanction en dernière période contre 8 points lors des 9 premières périodes. Toutefois, ces différences ne sont pas significatives ($p > 0.05$).

RESULTAT 5 : Les comportements de sanction dans le traitement NP sont similaires à ceux observés dans le traitement MP.

Le graphique 7 compare les comportements de sanction par attribution de points non monétaires selon les conditions partenaire et étranger.



Graphique IV.7: points de sanction non monétaire donnés pour une déviation vis-à-vis de la contribution de celui qui sanctionne (Partenaire/étranger)

Le graphique 7 montre que les déviations négatives sont plus fortement sanctionnées que les déviations positives. La comparaison des traitements partenaire et étranger ne fait pas apparaître de différences significatives sur les comportements de sanction. L'étude de la régression ci-dessous permet d'analyser de façon plus détaillée les comportements de sanction :

La variable expliquée est la sanction (respectivement nombre de points monétaires, non monétaires) que le joueur i donne au joueur j . Les variables explicatives sont respectivement *Déviator négative en valeur absolue vis à vis de la contribution du joueur i* et *Déviator positive vis à vis de la contribution de i* . Le coefficient traduit la relation entre le fait que le joueur j contribue moins que i et la sanction que i impose à j . A l'inverse, β_2 traduit le fait que j contribue plus que i . Le tableau 3 donne les estimations de la régression ci-dessus.

Tableau IV.3 : Relation entre les points de sanction monétaire et non monétaire attribués et les déviations de contribution

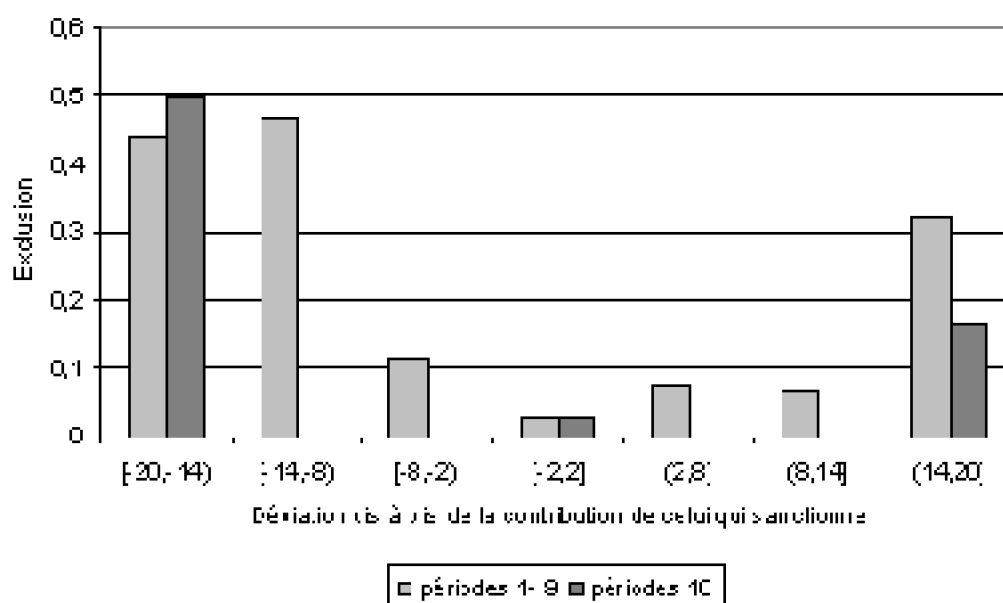
Variable expliquée : points de sanction donnés						
	MP \square	MP \square Dernière période	NP	NP Dernière période	NS	NS Dernière période
Constante	-2.6457*** (0.19289)	-4.08*** (1.01)	2.551*** (0.1482)	4.53*** (0.5103)	5.505*** (0.157)	6.083*** (0.515)
\square_0	0.3024*** (0.0216)	0.368*** (0.100)	0.4533*** (0.0223)	0.2982*** (0.0645)	0.223*** (0.02567)	0.199*** (0.0741)
\square_1	0.09362*** (0.02196)	0.2261** (0.104)	-0.01756*** (0.0221)	-0.1915*** (0.0642)	-0.288*** (0.0256)	-0.374*** (0.0781)
\square_2	N=1310 Log L -889.98	N=131 Log L -73.06	N=1200 F[2,1197]= 246.85*** R ² 0.29	N=120 F[2,117]= 22.49*** R ² 0.265	N=1176 F[2,1173]= 153.65*** R ² 0.207	N=95 F[2,92]= 21.9*** R ² 0.322
Note : écart type en parenthèse. *** 1% de degré de significativité, ** 5%, de degré de significativité * 10% de degré de significativité ; \square des estimations TOBIT ont été réalisées par le traitement MP. Les autres régressions sont des estimations OLS						

Les résultats du tableau 3 confirment les observations des graphiques 5, 6 et 7. Les coefficients \square_1 sont positifs et très significatifs pour les trois traitements (MP NP et NS). Ainsi, les joueurs sont sanctionnés lorsque leur contribution est plus faible que la contribution de ceux qui sanctionnent. Dans le traitement MP, \square_2 est également positif mais moins significatif que \square_1 . Donc les déviations positives sont également sanctionnées par l'attribution de points de sanction monétaire. Toutefois, les déviations négatives sont plus fréquemment et plus fortement sanctionnées que les déviations positives. \square_2 est négatif et significatif dans les traitements NP et NS, ce qui signifie que les déviations positives ne sont pas sanctionnées par l'attribution de points non monétaires.

Si l'on compare les coefficients \square_1 des données sur l'ensemble des périodes et en dernière période, il apparaît que ces coefficients ne sont pas très différents selon les périodes pour les trois traitements. Cela indique que les motivations d'ordre stratégique ne génèrent pas de sanctions supplémentaires par rapport aux motivations non stratégiques.

4.2.2.2. Attribution des sanctions d'exclusion

Le graphique 8 illustre la relation entre les déviations de contribution par rapport à la contribution de celui qui exclut et les comportements d'exclusion.



Graphique VI.8: Décisions d'exclusion pour une déviation vis-à-vis de la contribution de celui qui sanctionne

Le graphique 8 montre que les déviations négatives importantes sont les plus fortement sanctionnées par exclusion. Ainsi, sur l'ensemble des périodes, plus de 40% des déviations négatives dans l'intervalle $[-20, -14)$ sont sanctionnées d'exclusion. Les déviations positives sont moins sévèrement sanctionnées. Le graphique 8 montre également que les sujets excluent davantage ceux qui contribuent moins qu'eux que ceux qui contribuent comme eux. Ainsi, un sujet est rarement puni si sa contribution individuelle est proche de la contribution du sujet qui sanctionne, c'est à dire dans l'intervalle $[-2, 2]$. Un joueur qui contribue de façon similaire à celui qui inflige les sanctions a une probabilité inférieure à 5% d'être exclu.

A l'exception de l'intervalle $[-20, -14)$, les sujets sanctionnent moins sévèrement les déviations négatives et positives en dernière période que lors des périodes précédentes. Ces observations mettent en évidence la prédominance des motivations d'ordre stratégique dans les 9 premières périodes du traitement EC comparées aux motivations non stratégiques de dernière période.

Le tableau 4 contient les estimations du modèle Probit suivant :

La variable expliquée peut prendre deux valeurs : $Y_{ij} = 1$ si le joueur i décide d'exclure j et $Y_{ij} = 0$ autrement.

Tableau 4: relation entre les décisions d'exclusion et les déviations de contribution en dernière période (modèle PROBIT)

Variable expliquée : décision d'exclusion Y_{ij}		
	Traitement EC	Traitement EC dernière période
Constante α_0	-1.70*** (0.6838)	-2.008*** (0.2567)
α_1	0.0973*** (0.0114)	0.0751* (0.0303)
α_2	-0.02620 (0.0269)	0.0497 (0.03409)
	N=1320 Log L -242.35 LogLR-279.77 Pseudo R^2 (Mac Fadden) ⁴⁵ 0.133 Pourcent. de var. correct. prédites 94.02%	N=132 Log L -21.02 LogLr -24.1 Pseudo R^2 (Mac Fadden) 0.0769 Pourcent. de var. correct. prédites 95.45%
Note : écart type en parenthèse. *** 1% de degré de significativité, ** 5%, de degré de significativité * 10% de degré de significativité		

Le coefficient α_1 est significatif pour l'ensemble des périodes, ce qui confirme le fait que les sujets excluent ceux qui contribuent moins qu'eux. En dernière période, ce coefficient est moins élevé et également moins significatif. Le coefficient α_2 n'est pas significatif. La comparaison des comportements de sanction selon les périodes montre que les motivations à exclure sont essentiellement d'ordre stratégique puisque les exclusions sont moins fréquentes en dernière période. Le résultat 6 décrit la relation entre la décision d'un joueur d'exclure un autre membre du groupe et les différences de gain entre ces deux joueurs.

RESULTAT 6 : *les sujets excluent essentiellement pour des raisons d'ordre stratégiques. Lorsqu'ils sanctionnent pour des raisons non stratégiques, ils punissent essentiellement ceux qui contribuent moins qu'eux.*

4.2.3. Pertinence de l'analyse des comportements sous l'hypothèse d'aversion à l'inégalité

Les observations des sous-sections précédentes confirment largement les prédictions du modèle théorique de pression des pairs sous l'hypothèse d'aversion à l'inégalité. En effet, conformément aux prédictions théoriques sous l'hypothèse d'aversion à l'inégalité, dans le traitement MP, les sujets sanctionnent les autres membres de leur groupe (résultat 1). L'étude de la relation entre sanction et contribution montre par ailleurs qu'ils sanctionnent essentiellement ceux qui contribuent moins qu'eux (résultat 4).

Les résultats obtenus à partir du traitement EC sont-ils également conformes avec l'hypothèse d'aversion à l'inégalité ? Le résultat 6 montre que les sujets sont effectivement averses à l'inégalité mais qu'ils sanctionnent peu en dernière période (résultat 3). Ces observations confirment les prédictions théoriques sous l'hypothèse d'aversion à l'inégalité selon lesquelles les sujets averses à l'inégalité ne devraient pas exclure leurs pairs, du fait de l'existence d'un problème de passager clandestin sur les sanctions.

Les résultats du traitement NP remettent-ils en cause l'hypothèse d'aversion à

⁴⁵ Selon la méthode de Mac Fadden, $\text{pseudo } R^2 = 1 - \log L / \log L_R$

l'inégalité auto-centrée ? Il semble que oui. La comparaison des traitements NP et NS montre que lorsque les groupes sont variables et donc, que seules les raisons non stratégiques peuvent jouer, le nombre de sanction imposées est significativement plus faible (résultat 2). Toutefois, la comparaison des périodes 1 à 9 et de la dernière période du traitement NP, montre que les sanctions non monétaires sont toujours appliquées en dernière période. Ce résultat infirme donc les prédictions théoriques des modèles d'aversion à l'inégalité.

5. Conclusion

Ce chapitre a permis de montrer l'importance des considérations distributives. Parmi celles-ci, l'aversion à l'inégalité semble jouer un rôle majeur. Même s'il est coûteux de punir, les sujets peuvent être incités à punir les *passagers clandestins* lorsqu'ils sont suffisamment averses à l'inégalité. En effet, les sanctions permettent de réduire les dissonances de gains entre les joueurs. La coopération peut alors apparaître lorsque les agents ont la possibilité de se sanctionner mutuellement. A l'inverse, sans opportunité de sanction, les sujets ne peuvent pas soutenir la coopération même s'ils sont averses à l'inégalité. La seule façon de réduire les différences de gain avec un passager clandestin consiste pour un coopérateur, à réduire son propre effort.

Les résultats expérimentaux tirés des jeux de bien public avec et sans opportunité de sanction confirment en partie les prédictions théoriques du modèle de pression des pairs sous l'hypothèse d'aversion à l'inégalité. On observe que les motivations non stratégiques liées à l'aversion à l'inégalité jouent un rôle considérable dans l'explication des comportements de sanction. Ainsi, les joueurs sanctionnent essentiellement ceux qui contribuent moins qu'eux. Cela semble indiquer que les coopérateurs cherchent à réduire les différences de gains avec les passagers clandestins. Toutefois, si l'hypothèse d'aversion à l'inégalité apporte une explication pertinente aux comportements de sanction, elle ne permet pas d'expliquer toutes les observations. En effet, les résultats expérimentaux montrent que les sujets ne sanctionnent pas uniquement pour des raisons non stratégiques mais également afin d'inciter à la coopération dans le futur (raison stratégique). Par ailleurs, comment expliquer que les sujets sanctionnent leurs pairs alors que les points de sanction non monétaire n'affectent pas les gains de ceux qui les reçoivent et ne peuvent donc pas réduire l'inégalité des gains ?

Une première explication est consiste à mettre en avant le caractère non coûteux des sanction, ce qui nécessite alors d'être prudent quant à une remise en cause trop hâtive de l'hypothèse d'aversion à l'inégalité. Toutefois, Falk, Fehr et Fischbacher (2000) ⁴⁶ observent un résultat similaire avec un jeu de dilemme du prisonnier avec possibilité de sanction coûteuse. Ils étudient deux traitements différents. Dans un premier traitement, les points de sanction permettent de réduire la différence de gains entre les joueurs car le coût d'être sanctionné est supérieur au coût de sanctionner. Dans un deuxième

⁴⁶ FFF par la suite.

traitement, les points de sanction ne permettent pas de réduire cette différence de gain car le coût de sanctionner est identique au coût d'être sanctionné. Un point de sanction coûte une unité monétaire pour celui qui la reçoit mais également pour celui qui l'attribue. Donc, les points de sanction ne modifient pas les différences de gains entre celui qui sanctionne et celui qui reçoit les points⁴⁷. Les auteurs observent que 51% des sujets coopèrent et que 25% des sujets punissent les comportements de passager clandestin dans le second traitement. Ils comparent ce résultat avec celui obtenu dans le premier traitement. Dans ce traitement, seulement 61% des sujets coopèrent et 36% des sujets punissent (en attribuant 5.7 points de sanction en moyenne). Les observations de Falk, Fehr et Fischbacher (2000) suggèrent donc que les sujets qui sanctionnent les autres membres de leur groupe n'ont pas le désir de changer l'équité de l'allocation des gains.

Doit-on, au regard de ces résultats, rejeter l'hypothèse d'aversion à l'inégalité ? Deux arguments peuvent être avancés en faveur des modèles d'aversion à l'inégalité. Le premier argument consiste, tout en conservant l'hypothèse d'aversion à l'inégalité, à supposer que les comparaisons interpersonnelles ne portent pas sur les gains individuels mais sur les utilités des joueurs. L'idée est que si l'attribution des points de sanction non monétaire ne réduit pas les différences de gain, elle peut toutefois affecter l'utilité des sujets. En effet, on a mis en évidence dans le chapitre 2 que les sujets étaient sensibles à la désapprobation de leurs pairs et qu'ils contribuaient davantage au bien public afin d'éviter la désutilité de recevoir des sanctions non monétaires qui reflètent la désapprobation de leurs pairs. Si les sanctions non monétaires affectent l'utilité des sujets sanctionnés, alors, il peut être dans l'intérêt d'un agent de punir ses pairs afin de réduire les dissonances d'utilité des sujets. En considérant une définition élargie de l'aversion à l'inégalité où un individu compare non plus ses gains à ceux des autres individus mais son niveau d'utilité, alors les observations réalisées dans le traitement NP ne permettraient pas forcément d'invalidier l'hypothèse d'aversion à l'inégalité.

Un deuxième argument consiste à mettre en exergue la multiplicité des motivations des sujets. L'aversion à l'inégalité n'est alors qu'une explication parmi d'autres des comportements individuels observés. A l'instar de Falk et Fischbacher (1998), on peut alors envisager de réconcilier dans un même modèle les approches fondées sur les intentions et les considérations distributives des agents. L'intérêt d'une telle démarche est d'augmenter le pouvoir explicatif du modèle, toutefois, au risque de complexifier considérablement la modélisation des comportements.

⁴⁷ Le traitement NP diffère du deuxième traitement de FFF (2000) dans la mesure où il n'existe pas de coût à sanctionner ni à être sanctionné.

Conclusion générale

Dans la lignée de Kandel et Lazear (1992), de nombreux travaux comme ceux de Barron et Gjerde (1997) ou de Dong et Dow (1993) se sont intéressés au contrôle mutuel au sein des équipes de production. L'intérêt fondamental de ces modèles a été de mettre en lumière l'efficacité du contrôle mutuel. Cependant, ces travaux reposent sur une hypothèse forte selon laquelle les sanctions ne font pas l'objet d'une décision mais sont imposées automatiquement en cas de détection d'un comportement non coopératif. Or, les agents ne se contentent généralement pas de contrôler leurs pairs mais décident également des sanctions à imposer à d'éventuels passagers clandestins (Miller, 1992).

L'objet de cette thèse était de dépasser les modèles de contrôle mutuel existants et d'enrichir la modélisation de la pression des pairs en prenant en considération les décisions des agents concernant les sanctions à prendre envers d'éventuels passagers clandestins. La prise en compte des décisions de sanction confère ainsi davantage de réalisme au modèle de pression des pairs. Les principaux résultats théoriques des modèles de pression des pairs montrent que la discipline par les pairs est efficace dès lors que les agents se contentent de décider de leur effort de contrôle. Au contraire, lorsque les décisions portent sur les sanctions à imposer et que sanctionner est coûteux, la pression des pairs devient totalement inexistante. Les agents anticipent alors qu'ils ne seront pas sanctionnés et adoptent donc un comportement de passager clandestin.

La soumission à réfutation par la méthode expérimentale du modèle de pression des pairs avec choix des sanctions a montré que les agents ne se comportent pas comme le prédit la théorie. En effet, ils n'hésitent pas à sanctionner les autres membres du groupe

même si cela est coûteux pour eux. Par ailleurs, l'opportunité de sanctionner accroît considérablement le niveau de coopération. La divergence entre les prédictions théoriques et les résultats expérimentaux conduit inévitablement à s'interroger sur les motivations qui guident les comportements de contribution et de sanction des agents. La comparaison des traitements expérimentaux avec sanctions monétaires et sanctions non monétaires a permis de dissocier deux raisons principales pour expliquer pourquoi les sujets contribuent davantage au bien public lorsqu'il existe une opportunité de sanction. Les sujets seraient incités à davantage coopérer afin d'éviter non seulement les conséquences directes des sanctions qui se traduisent par une réduction de leurs gains pécuniaires mais également indirectes, en termes de désapprobation des pairs.

Deux raisons principales ont également été avancées pour expliquer les comportements de sanction des sujets. La première explication est d'ordre *stratégique* et repose sur l'idée que les sujets punissent leurs pairs, même si sanctionner est coûteux dans l'immédiat car ils attendent des gains futurs des sanctions. La deuxième explication est *non stratégique* et suggère que les sujets sanctionnent leurs pairs car ils sont motivés par des considérations d'ordre social. L'hypothèse retenue est que les agents sanctionnent leurs pairs afin de réduire les inégalités de gains observées avec les autres joueurs. L'incorporation de l'hypothèse d'aversion à l'inégalité auto-centrée dans les fonctions d'utilité permet d'expliquer en grande partie les comportements de sanction des sujets.

Les expérimentations réalisées ont par ailleurs mis en exergue que l'efficacité de la pression des pairs était fortement conditionnée par la nature des sanctions. Ainsi, les sanctions monétaires qui affectent les gains de ceux qui les reçoivent mais également de ceux qui les infligent sont plus efficaces mais aussi plus persistantes dans le temps que les sanctions non monétaires qui reposent uniquement sur la désapprobation des pairs. Par ailleurs, lorsque les sanctions sont suffisamment fortes, comme par exemple l'ostracisme, alors, la menace de sanction est à elle seule suffisante pour inciter les agents à coopérer davantage. Dès lors que la sanction est rarement appliquée, elle permet au groupe d'économiser les coûts inhérents à la pression des pairs.

Dans le prolongement de cette thèse, deux pistes de recherche peuvent être avancées. La première voie de recherche possible serait d'étudier la possibilité pour le principal d'influencer ou même de créer la pression des pairs. En effet, puisque la pression des pairs est efficace, on peut alors s'interroger sur l'intérêt pour l'entreprise d'influencer et de canaliser la pression des pairs. A ce titre, les entreprises japonaises dépensent beaucoup de temps et d'argent à mettre en place les conditions qui faciliteront l'émergence de la pression des pairs au sein des équipes de travail (Aoki, 1988). Ainsi, elles organisent des rencontres entre les employés et leur famille en dehors du cadre du travail (voyage, clubs...) afin qu'ils tissent des liens sociaux étroits et que cela facilite leur aptitude à coopérer sur leur lieu de travail. Cet esprit d'équipe canalise les comportements des individus et peut réduire les comportements opportunistes puisque chaque membre de l'équipe doit être loyal envers l'autre et ne pas se comporter en passager clandestin⁴⁸.

La deuxième voie de recherche consiste à s'interroger sur l'opportunité pour le principal à déléguer directement aux agents l'autorité sur le contrôle. A ce titre, Aghion et Tirole (1997) distinguent l'autorité formelle de l'autorité réelle. Alors que l'autorité formelle

est définie comme le droit de choisir des actions affectant entièrement ou en partie une organisation, l'autorité réelle correspond quant à elle au contrôle effectif des décisions ⁴⁹. Dès lors que les agents ont une autorité réelle sur leurs pairs, c'est à dire une information privilégiée du fait de leur proximité et qu'ils n'hésitent pas à sanctionner les éventuels passagers clandestins, il peut être dans l'intérêt du principal de déléguer directement l'autorité formelle sur les décisions de contrôle et de sanction directement aux agents eux-mêmes.

⁴⁸ Selon Varian (1990) l'efficacité des cercles de qualité au japon trouve ses fondements dans les caractéristiques mêmes de la société japonaise et de son système éducatif qui privilégie les valeurs de conformité, de loyauté et d'empathie. Il apparaît alors moins coûteux d'instaurer un « esprit d'équipe » au sein des entreprises dans la société japonaise que dans la société américaine davantage individualiste (Hollander, 1990).

⁴⁹ Cette distinction entre autorité formelle et autorité réelle se rapproche de la description de l'autorité «rationnelle» ou «légale» de Max Weber (1968). Weber montre que les employés peuvent exercer un pouvoir considérable sur la «machine bureaucratique».

Annexes

ANNEXES CHAPITRE I

ANNEXE I.A 1 : Le modèle d'Holmström (1982)

Comment alors inciter les agents à ne pas adopter un comportement de passager clandestin coûteux pour l'équipe dans son ensemble? Selon Holmström (1982), le problème du passager clandestin n'est pas seulement la conséquence de l'impossibilité d'observer les actions mais surtout la conséquence d'un schéma d'incitation inapproprié. L'auteur s'interroge alors sur l'existence d'un schéma de partage des profits entre les membres d'une équipe, satisfaisant les conditions de Pareto optimalité. En effet, une règle de partage où l'output est entièrement partagé conduit inévitablement au problème du passager clandestin. Selon lui, ce problème peut être résolu grâce à un schéma de répartition où l'output n'est pas partagé uniquement entre les agents. La solution avancée par Holmström consiste à faire intervenir une tierce personne comme *créancier résiduel*. Le problème du passager clandestin est résolu selon Holmström par l'introduction d'un principal qui sanctionnerait les agents, en prélevant la totalité de l'output si le niveau optimal d'output n'est pas atteint. Holmström considère une équipe de production composée de n agents qui fournissent chacun un effort « a » non observable

avec un coût v_i , afin de produire un bien x qui sera partagé entre les agents après sa production. Soit $s_i(x)$, la part de l'output que reçoit chaque agent i . Holmström s'interroge sur l'existence d'une règle de partage de l'output tel que :

Soit la fonction de gain de chaque agent i :

avec la condition de premier ordre par rapport à a :

Soit a^* l'équilibre de Nash qui satisfait la condition de Pareto optimalité telle que :

La Pareto optimalité suppose que :

L'égalité des équations (3) et (5) suppose que $s_i' = 1$, ce qui est incompatible avec l'équation (1) puisque la différenciation de (1) implique que :

ANNEXE I.A 2 : Le modèle d'Esvaran et Kotwal (1984)

Si le schéma d'incitation proposé par Holmström a le mérite d'apporter une solution au problème de passager clandestin, il a toutefois suscité un certain nombre de critiques. Selon Esvaran et Kotwal (1984), le schéma de rémunération proposé par Holmström (1982) ne tient pas compte d'un éventuel aléa moral de la part du principal. En effet, après avoir imposé un seuil minimal de production, le principal peut être incité à corrompre un membre de l'équipe. Il peut demander à un agent de réduire son effort pour que le niveau d'output de l'équipe descende en dessous d'un certain seuil dans la mesure où les gains du principal sont plus élevés lorsque les agents n'atteignent pas le niveau d'output fixé dans le schéma de rémunération. Le principal peut ainsi proposer à l'agent qu'il souhaite corrompre, le schéma de rémunération suivant :

L'agent i reçoit un gain supérieur ($bn + \Delta > b$), s'il accepte de réduire son effort de production afin que l'output descende en dessous de $x(a^*)$ mais reste toutefois au-dessus du niveau .

ANNEXE I.A 3 : Le modèle de Rasmusen (1987)

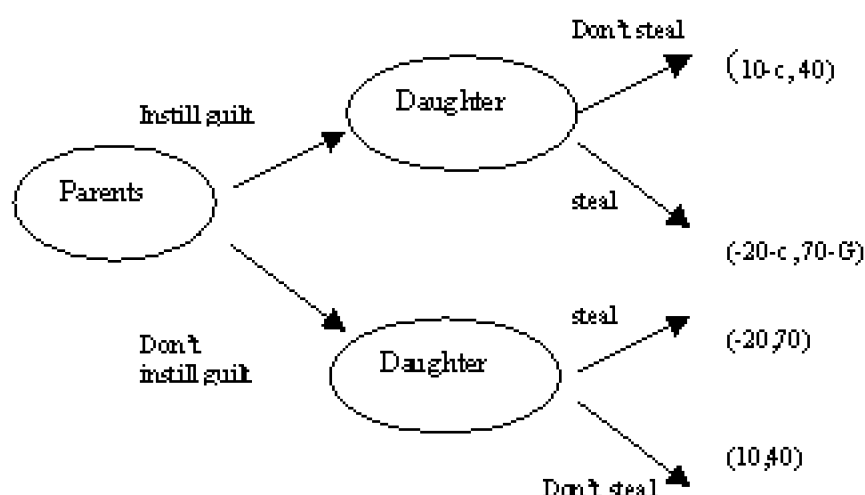
Rasmusen (1987) a montré qu'il est possible d'obtenir un résultat optimal avec une équipe de production sans nécessairement avoir recours à un principal. En effet, un contrat où tout l'output est partagé entre les agents peut être efficient si les agents sont suffisamment averses au risque. Le problème du passager clandestin dans les équipes peut être éliminé en introduisant une loterie où un agent, choisi de façon aléatoire, est récompensé tandis que les autres sont sanctionnés. Rasmusen propose le schéma de rémunération suivant :

Lorsque le niveau optimal d'output est atteint, chaque agent reçoit une part $s_i(x) = x/n$ de l'output. Quand ce niveau d'output n'est pas atteint, un agent, choisi au hasard, reçoit un paiement $s_i(x) = x + (n-1)w$, tandis que les $n-1$ autres agents reçoivent un paiement négatif $s_i(x) = -w$. Donc selon Rasmusen (1987), l'intervention du principal n'est pas justifiée puisqu'un contrat où tout l'output est partagé entre les agents peut être efficient si les agents sont suffisamment averses au risque. Andolfatto et El Nosal (1997) ont toutefois montré que l'argumentation proposée par Rasmusen (1987) dépend de

conditions extrêmes quant à la vérifiabilité des actions individuelles pour ce qui est des tentatives de renégociation. Plus précisément, les auteurs montrent que l'équilibre est supporté par une menace non crédible. Autrement dit, l'équilibre de Nash n'est pas parfait en sous jeu. En effet, si un output bas est observé, alors tous les membres du groupe sont incités à renégocier le contrat initial afin d'éviter le risque encouru par la loterie.

Le contrôle mutuel peut-il être une alternative au schéma d'incitation d'Holmström (1982) ? L'idée du contrôle mutuel est que la famille est l'entité la mieux placée pour contrôler ses membres (Stiglitz, 1990 ; Varian, 1990 ; Arnot et Stiglitz, 1991 ; Dong et Dow, 1993). En effet, alors que le partage des profits fournit aux agents des incitations fortes à adopter un comportement opportuniste, un tel mode de rémunération peut également les inciter à se contrôler mutuellement. S'il est indéniable qu'il existe des incompatibilités potentielles entre les intérêts du groupe et ceux de ses membres, ces intérêts n'entrent pas forcément en conflit.

ANNEXE I.B : Internalisation de la pression des pairs



Graphique : mécanismes d'internalisation de la culpabilité

Le jeu décrit par le graphique ci-dessus se présente de la façon suivante. En deuxième étape du jeu, la fille décide de dérober ou pas un objet. Dans le cas où ses parents ne lui ont pas inculqué le sentiment de culpabilité, son utilité à dérober l'objet est supérieure à celle de ne pas le voler. Elle choisit donc de dérober l'objet. Dans le cas où ses parents lui ont inculqué un sentiment de culpabilité, dérober l'objet lui procure une utilité de $(70-G)$ tandis que ne pas le voler lui procure une utilité de (40) . G représente ici la désutilité inhérente au sentiment de culpabilité ressenti. Elle décidera donc de ne pas voler l'objet si et seulement si son utilité de ne pas voler l'objet est supérieure à celle de le dérober $(40) > (70-G)$. Donc lorsque le sentiment de culpabilité est suffisamment fort, $(G > 30)$, elle décidera de ne pas dérober l'objet. En première étape, les parents décident d'investir dans la culpabilité. S'ils décident d'investir dans la culpabilité, ils subissent un coût c mais ils s'attendent à ce que leur fille ne vole pas, ce qui leur procure un gain

d'utilité. Au contraire, s'ils décident de ne pas investir dans la culpabilité, ils font l'économie du coût c mais subissent alors la désutilité inhérente au fait que leur fille dérobera l'objet. Ils choisiront d'investir en culpabilité pour des valeurs faibles de c (c'est à dire dans l'exemple lorsque $c < 30$). L'internalisation de la pression des pairs peut aider à comprendre les phénomènes de conformité⁵⁰. Ainsi, les sentiments de culpabilité, les coutumes renforcent la participation des membres du groupe.

ANNEXE I.C Statique comparative avec système de deux équations et deux inconnues

Afin de déterminer comment les choix optimaux d'effort de production varient en fonction des choix d'effort de contrôle, procédons à une analyse de statique comparative :

On sait que les choix optimaux d'effort de production $e_1^* = l(a_{21}, a_{12})$, et $e_2^* = Q(a_{12}, a_{21})$ doivent satisfaire les conditions du premier ordre :

En différenciant ces deux expressions par rapport à a_1 on obtient :

Comme les agents sont identiques, un équilibre symétrique suppose que $\partial l / \partial a_1 = \partial Q / \partial a_2$ et $\partial l / \partial a_2 = \partial Q / \partial a_1$. La résolution du système à deux équations et deux inconnues ci-dessus conduit aux solutions suivantes :

soit :

Puisque les agents sont identiques, un équilibre symétrique suppose que $l'_{a_{21}} = Q'_{a_{12}}$ et $l'_{a_{12}} = Q'_{a_{21}}$. Par ailleurs, $f''_{e_{1e}} = f''_{e_{2e}}$, $C''_{e_{1e}} = C''_{e_{2e}}$, $C''_{e_{1e}}(e_i, a_i) = C''_{e_{2e}}(e_j, a_j)$ et $P''_{e_{1e}}(e_i, a_i) = P''_{e_{2e}}(e_j, a_j)$. Il est supposé que :

ANNEXES CHAPITRE II

ANNEXE II.A : Instructions pour l'expérience

⁵⁰ Une norme internalisée est une norme qui n'est pas considérée comme une contrainte. Ainsi on se conforme à une norme internalisée non parce qu'on craint les sanctions que susciterait une déviation mais parce qu'on souhaite activement s'y conformer. Par exemple, considérons la norme qui consiste à aider quelqu'un en détresse. Je peux choisir d'aider cette personne car j'en attends de la reconnaissance du groupe social auquel j'appartiens, ou au contraire je m'attends à être sanctionné socialement si je n'aide pas cette personne. Si la norme est internalisée, je peux aider la personne en détresse car je souhaite réellement le faire et me sentirais coupable de ne pas le faire. Iannaccone (1992) a montré que les interdits religieux jouent souvent le rôle d'écran protecteur en décourageant les membres les moins engagés et en augmentant le niveau moyen de participation du groupe. En effet, en interdisant par exemple les activités alternatives à la pratique religieuse, les interdictions favorisent une substitution de ces activités à la pratique religieuse et augmentent donc la participation. Ces interdictions et autres sentiments représentent un coût pour ceux qui ne se conforment pas aux standards. Ce coût peut décourager les membres du groupe les moins engagés et les plus enclins à adopter des comportements de passagers clandestins. A l'inverse, ces interdictions renforcent la participation des autres membres du groupe religieux.

Instructions générales

Vous participez à une expérimentation en économie, à l'issue de laquelle vous pourrez gagner une certaine somme d'argent. Vos gains dépendent de vos décisions et des décisions des autres participants à l'expérience. Il est donc important de lire avec attention ces instructions. Les instructions que nous vous avons distribuées sont pour votre information privée. Il est interdit de communiquer avec les autres participants pendant l'expérience, sous peine d'exclusion de la session et des gains. Toutes les transactions menées dans l'expérimentation et vos gains seront calculés en ECU (Experimental Currency Units). Vos gains en ECU pour cette session expérimentale seront convertis en francs et payés en liquide à la fin de la session expérimentale sur la base suivante : votre gain final en ECU est la somme de vos gains réalisés lors de chacune des parties composant cette session. Ce gain final en ECU sera converti en Francs français sur la base suivante : 1 ECU vaut 0.08 Franc. Vous recevrez en plus un forfait de participation de 25 Francs. Cette session expérimentale est composée de parties successives. Dans chaque partie, les participants sont divisés en groupes de 4. Vous jouerez donc avec 3 autres sujets. Durant toute cette session, vous serez en relation avec les mêmes sujets. Vous ne connaîtrez pas l'identité de ces personnes.

Déroulement des 10 premières parties

Les 4 sujets composant un groupe ont pour rôle de participer à un projet, sous la forme d'une contribution à la constitution d'une somme globale qui sera ensuite partagée entre eux. Cette somme résulte des contributions individuelles des 4 sujets. Chaque partie se déroule de la manière suivante. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation qu'il affecte au projet. Après avoir choisi votre contribution au projet, en indiquant un nombre compris entre 0 et 20, vous devez cliquer sur le bouton OK. Une fois que cette décision est prise, vous ne pouvez plus en changer pour la partie en cours. Une fois que tous les membres de votre groupe auront pris leur décision, votre écran vous indiquera le montant total en ECU affecté au projet par tous les membres du groupe (incluant votre contribution). L'écran vous indiquera aussi combien vous avez gagné pour cette partie. Votre gain est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain total est donc calculé par l'ordinateur de la manière suivante :

$(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet)

Le gain de chaque membre du groupe est calculé de la même manière, ce qui signifie que chaque membre du groupe reçoit le même revenu du projet. Chaque ECU de votre dotation que vous conservez pour vous-même représente pour vous un gain de 1 ECU. Si à la place, vous avez décidé d'affecter cet ECU au projet, alors la contribution totale au projet augmente de 1 ECU. Votre revenu issu du projet augmente alors de 40% ($1 \text{ ECU} = 0.4 \text{ ECU}$). Le revenu des autres membres du groupe augmentera aussi de 0.4 ECU par personne, et donc le revenu total du projet pour le groupe augmentera de 1.6 ECU. Votre

contribution au projet augmente donc également le revenu des autres membres du groupe. De la même manière, vous retirez un revenu de chaque ECU affecté par les autres membres du groupe au projet. Pour chaque point contribué par un autre membre du groupe, vous gagnez 40% (1 ECU) = 0.4 ECU.

Déroulement des 10 parties suivantes (MP)

Chaque partie se déroule désormais en deux étapes. Vous continuez à jouer avec les mêmes membres du groupe que lors des 10 premières parties. Lors de la première étape, vous devez décider combien d'ECU vous désirez contribuer à un projet. Dans la seconde étape, vous êtes informé des contributions de chacun des 3 autres membres de votre groupe au projet. Vous pouvez alors décider de diminuer ou non leurs gains de la première étape en leur distribuant des points. Les paragraphes suivants décrivent le jeu en détail. La première étape est identique aux 10 premières parties de cette session. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette première étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain de première étape qui est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire 20 – votre contribution au projet), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain de 1^{ère} étape est donc calculé par l'ordinateur de la manière suivante :

(20 – votre contribution au projet)

+ 40% (contributions totales au projet)

Le gain de chaque membre du groupe est calculé de la même manière, ce qui signifie que chaque membre du groupe reçoit le même revenu du projet.

Lors de la seconde étape, votre écran vous indiquera le montant que chacun des autres membres de votre groupe a contribué au projet. Chacun de vous a maintenant la possibilité de réduire ou de laisser inchangé le gain de chacun des autres sujets en distribuant des points. Vous pouvez distribuer un certain nombre de points pour un membre de votre groupe pour manifester une certaine désapprobation (10 points pour la plus grande désapprobation, 0 point pour l'absence de désapprobation). Chaque point distribué à un sujet particulier diminue son gain de 1^{ère} étape de 10%. De la même manière, votre gain peut être modifié si les autres sujets le souhaitent. Vous êtes d'abord informé du montant que chacun des 3 autres sujets a contribué au projet lors de la 1^{ère} étape du jeu. Attention : l'ordre dans lequel vous sont indiquées les décisions des 3 autres sujets est modifié aléatoirement à chaque partie (autrement dit, par exemple, le chiffre qui apparaîtra en 1^{er} sur votre écran ne correspondra pas toujours à la décision du même joueur). Vous devez décider ensuite du nombre de points que vous distribuez à chacun des 3 autres membres de votre groupe pour réduire ou maintenir leur gain. Vous devez entrer une valeur pour chaque sujet, comprise entre 0 et 10 points. Si vous ne désirez pas diminuer le gain d'un autre sujet, vous devez entrer la valeur 0. Si vous distribuez des points, vous subissez un coût qui dépend du nombre de points distribués à chaque sujet.

Plus vous distribuez de points, plus votre coût est élevé. Votre coût total est égal à la somme des coûts des points distribués à chacun des 3 autres membres de votre groupe. Le tableau suivant illustre la relation entre le nombre de points distribués à un sujet et leur coût associé :

# points	0	1	2	3	4	5	6	7	8	9	10
Coût	0	1	2	4	6	9	12	16	20	25	30

Si vous donnez 2 points à 1 membre du groupe, cela vous coûte 2 ECU ; si vous donnez 9 points à un autre membre, cela vous coûte 25 ECU de plus ; si vous donnez 0 point au dernier membre du groupe, cela ne vous coûte rien de plus. Dans ce cas, le coût total des points que vous avez distribués est de 27 ECU (2+25+0). Ces coûts seront indiqués sur votre écran. Tant que vous n'avez pas cliqué sur le bouton OK sur votre écran, vous pouvez modifier votre décision. Si vous donnez 0 point à un autre membre du groupe, vous ne changez pas son gain. Mais si vous distribuez 1 point à un sujet, vous réduisez son gain issu de la première étape de la partie de 10% ; si vous lui distribuez 2 points, vous réduisez son gain de 20% ; etc. Le nombre de points que vous distribuez définit donc de combien vous voulez diminuer son gain de la 1^{ère} étape.

Votre gain final pour chaque partie est désormais calculé par l'ordinateur de la manière suivante :

Si vous avez reçu moins de 10 points

= (gain de la 1^{ère} étape)*[(10-nombre de points reçus)/10] – coût des points distribués

Si vous avez reçu 10 points ou +

gain final pour la partie = - coût des points que vous avez distribués

Il est à noter que, dans le calcul des gains, un maximum de 10 points reçus est pris en compte. Par exemple, si vous avez reçu 3 points de la part de tous les autres membres du groupe, votre revenu de la 1^{ère} étape est réduit de 30%. Si vous avez reçu 4 points, votre revenu de 1^{ère} étape est réduit de 40%. Si vous avez reçu 10 points ou plus, vous avez perdu 100% de votre gain de 1^{ère} étape. Dans ce cas, vous réalisez une perte si vous avez distribué des points. Le montant de cette perte correspond au coût des points que vous avez vous-même distribué aux autres. Votre gain de deuxième étape peut donc être négatif, si le coût des points distribués excède votre gain de 1^{re} étape. Mais de manière générale, vous pouvez toujours éviter les pertes avec certitude à travers vos décisions.

Déroulement des 10 parties suivantes (NP)

Chaque partie se déroule désormais en deux étapes. Vous continuez à jouer avec les mêmes membres du groupe que lors des 10 premières parties. Lors de la première étape, vous devez décider combien d'ECU vous désirez contribuer à un projet. Dans la seconde étape, vous êtes informé des contributions de chacun des 3 autres membres de votre groupe au projet. Vous pouvez alors signifier votre approbation ou votre désapprobation à l'égard des décisions de chacun des autres membres du groupe en leur distribuant des

points. Les paragraphes suivants décrivent le jeu en détail.

La première étape est identique aux 10 premières parties de cette session. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette première étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain individuel. Votre gain est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : Ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain est donc calculé par l'ordinateur de la manière suivante :

$(20 - \text{votre contribution au projet})$

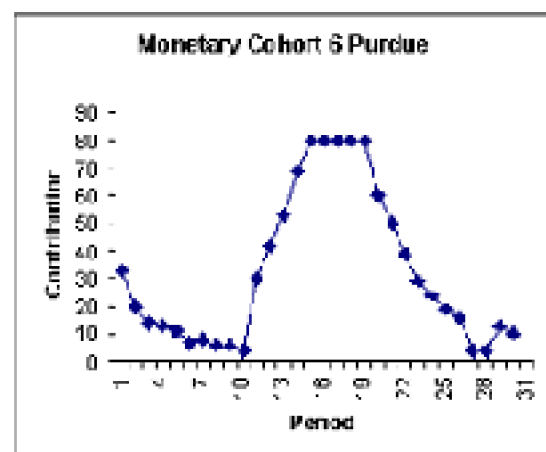
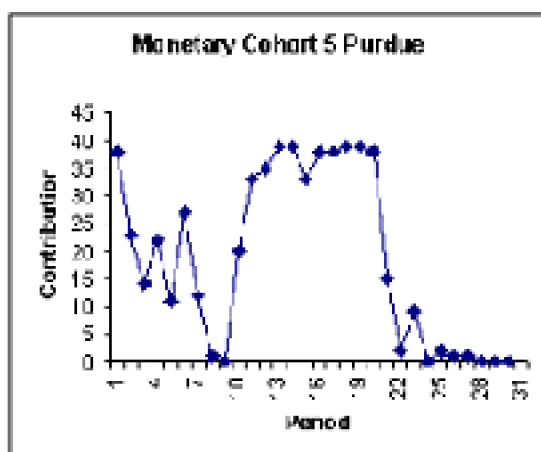
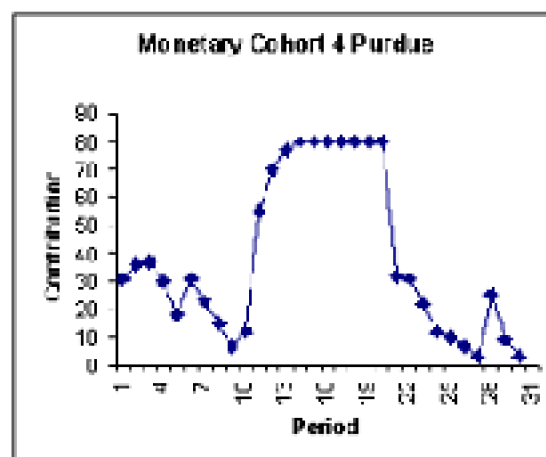
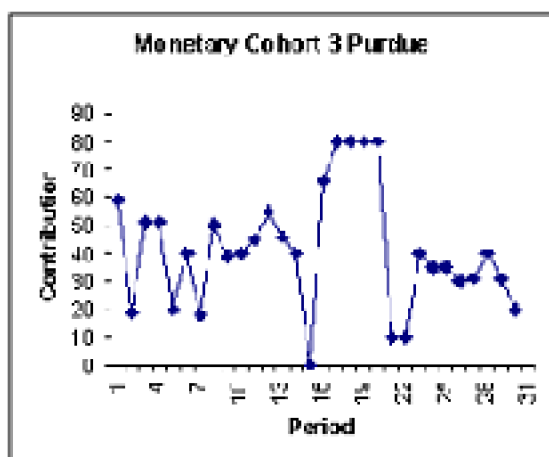
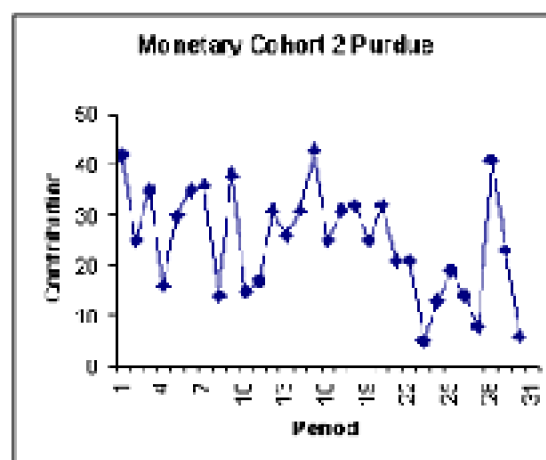
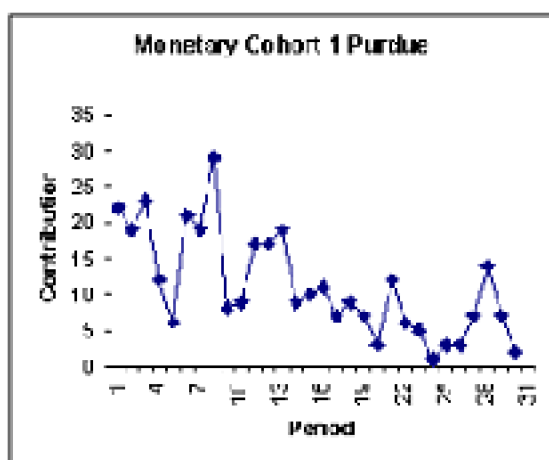
+ 40% (contributions totales au projet)

Le gain de chaque membre du groupe est calculé de la même manière, ce qui signifie que chaque membre du groupe reçoit le même revenu du projet. Lors de la seconde étape, votre écran vous indiquera le montant que chacun des autres membres de votre groupe a contribué au projet. Chacun de vous a maintenant la possibilité de signifier votre approbation ou votre désapprobation à chacun des autres sujets en distribuant des points. Vous pouvez distribuer un grand nombre de points pour un membre de votre groupe pour manifester une certaine désapprobation (10 points pour la plus grande désapprobation, 0 point pour l'absence de désapprobation). De la même manière, les autres membres de votre groupe peuvent vous attribuer des points de désapprobation. Vous êtes d'abord informé du montant que chacun des 3 autres sujets a contribué au projet lors de la 1^{ère} étape du jeu. Attention : l'ordre dans lequel vous sont indiquées les décisions des 3 autres sujets est modifié aléatoirement à chaque partie (autrement dit, par exemple, le chiffre qui apparaîtra en 1^{er} sur votre écran ne correspondra pas toujours à la décision du même joueur). Vous devez décider ensuite du nombre de points que vous distribuez à chacun des 3 autres membres de votre groupe. Vous devez entrer une valeur pour chaque sujet, comprise entre 0 et 10 points. Si vous ne désirez pas signifier de désapprobation à l'égard d'un autre sujet, vous devez entrer la valeur 0. Une fois que tous les membres du groupe auront pris leur décision, votre écran affichera vos gains de la partie.

Déroulement des 10 dernières parties

Ces 10 dernières parties sont identiques aux 10 premières et nous appliquons donc les instructions des 10 premières parties de cette session. Dans ces dernières parties, vous êtes toujours en relation avec les mêmes sujets que lors des jeux précédents. A la fin de ces 10 parties, je vous remercie de rester à votre place jusqu'à ce que je vous invite à la quitter et de rester silencieux pendant ce temps.

ANNEXE II.B Niveaux de contribution de groupe (sessions MP à Purdue et à Lyon GATE)



ANNEXES CHAPITRE III

ANNEXE III.A : Instructions de l'expérience

Instructions générales

Vous participez à une expérimentation en économie, à l'issue de laquelle vous pourrez gagner une certaine somme d'argent. Vos gains dépendent de vos décisions et des décisions des autres participants à l'expérience. Il est donc important de lire avec attention ces instructions. Toutes les transactions menées dans l'expérimentation et vos gains seront calculés en ECU (Experimental Currency Units). Vos gains en ECU pour cette session expérimentale seront convertis en francs et payés en liquide à la fin de la session expérimentale sur la base suivante : votre gain final en ECU est la somme de vos gains réalisés lors de chacune des parties composant cette session. Ce gain final en ECU sera converti en Francs français sur la base suivante : 100 ECU vaut 4.5 Franc. Vous recevrez en plus un forfait de participation de 15 Francs. Cette session expérimentale est composée de parties successives. Dans chaque partie, les participants sont divisés en groupes de 4. Vous jouerez donc avec 3 autres sujets. Durant toute cette session, vous serez en relation avec les mêmes sujets. Vous ne connaîtrez pas l'identité de ces personnes.

Déroulement des 10 premières parties

Les 4 sujets composant un groupe ont pour rôle de participer à un projet, sous la forme d'une contribution à la constitution d'une somme globale qui sera ensuite partagée entre eux. Cette somme résulte des contributions individuelles des 4 sujets. Chaque partie se déroule de la manière suivante. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation qu'il affecte au projet. Après avoir choisi votre contribution au projet, en indiquant un nombre compris entre 0 et 20, vous devez cliquer sur le bouton OK. Une fois cette décision prise, vous ne pouvez plus en changer pour la partie en cours. Une fois que tous les membres de votre groupe auront pris leur décision, votre écran vous indiquera le montant total en ECU affecté au projet par tous les membres du groupe (incluant votre contribution). L'écran vous indiquera aussi combien vous avez gagné pour cette partie.

Votre gain est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain total est donc calculé par l'ordinateur de la manière suivante :

$(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet)

Chaque ECU de votre dotation que vous conservez pour vous-même représente pour vous un gain de 1 ECU. Si à la place, vous avez décidé d'affecter cet ECU au projet, alors la contribution totale au projet augmente de 1 ECU. Votre revenu issu du projet augmente alors de 40% ($1 \text{ ECU} = 0.40 \text{ ECU}$). Le revenu des autres membres du groupe augmentera aussi de 0.40 ECU par personne. Cela signifie que votre contribution au projet augmente donc également le revenu des autres membres du groupe. De la même manière, vous retirez un revenu de chaque ECU affecté par les autres membres du groupe au projet. Pour chaque ECU contribué par un autre membre du groupe, vous gagnez 40% ($1 \text{ ECU} = 0.40 \text{ ECU}$).

* * *

Si vous avez des questions concernant ce que vous venez de lire, merci de lever la main. Nous viendrons répondre immédiatement à vos questions. Les instructions que nous vous avons distribuées sont pour votre information privée. Il est interdit de communiquer avec les autres participants pendant l'expérience, sous peine d'exclusion de la session et des gains.

Déroulement des 10 parties suivantes*

Chaque partie se déroule désormais en trois étapes. Vous continuez à jouer avec les mêmes membres du groupe que lors des 10 premières parties. Lors de la première étape, vous participez à un premier projet où vous devez décider combien d'ECU vous désirez contribuer à ce premier projet. Dans la seconde étape, vous êtes informé des contributions de chacun des 3 autres membres de votre groupe au projet. Vous pouvez alors décider d'exclure ou pas un ou plusieurs membres de votre groupe. Dans la troisième étape, seuls les joueurs non exclus peuvent participer à un deuxième projet où ils décident combien d'ECU ils désirent contribuer à ce deuxième projet. Les joueurs exclus ne participent pas au deuxième projet. Les paragraphes suivants décrivent le jeu en détail.

La première étape est identique aux 10 premières parties de cette session. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette première étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain de première étape qui est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain issu du premier projet est donc calculé par l'ordinateur de la manière suivante :

Gain issu du premier projet:

$(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet)

Lors de la seconde étape, votre écran vous indiquera le montant que chacun des 3

autres membres de votre groupe a contribué au projet lors de la 1^{ère} étape du jeu. Attention : l'ordre dans lequel vous sont indiquées les décisions des 3 autres sujets est modifié aléatoirement à chaque partie (autrement dit, par exemple, le chiffre qui apparaîtra en 1^{er} sur votre écran ne correspondra pas toujours à la décision du même joueur). Ayant été informé de la contribution individuelle de chacun des autres membres du groupe, chacun de vous a maintenant la possibilité d'exclure un ou plusieurs autres sujets. De la même manière, vous pouvez être exclu si au moins un autre membre du groupe le souhaite. Si vous excluez un membre du groupe, il ne peut pas participer au deuxième projet et ses gains pour ce deuxième projet sont nuls. Il suffit qu'un seul participant au moins décide d'exclure un autre joueur pour que ce dernier soit exclu. L'exclusion ne dure que la période en cours, c'est à dire qu'un joueur exclu à une période revient à la période suivante. Si vous désirez exclure un participant vous devez cocher la case d'exclusion qui se rapporte à sa contribution. A l'inverse, si vous ne désirez pas exclure un participant, vous choisissez de ne pas cocher cette case. Lorsque vous excluez un ou plusieurs participants, vous réduisez la taille du groupe du nombre de participants exclus et donc le nouveau projet ne prendra en compte que les contributions des participants non exclus. En fin de cette deuxième étape, vous êtes informé de votre exclusion ou de votre non-exclusion et du nombre de joueurs exclus.

En troisième étape, seuls les joueurs n'ayant pas été exclus participent à un nouveau projet, sous la forme d'une contribution à la constitution d'une somme globale qui sera ensuite partagée entre eux. Cette somme résulte des contributions individuelles des joueurs non exclus. Chaque sujet reçoit à nouveau une dotation de 20 ECU. Chacun des joueurs décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain issu du deuxième projet qui est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des contributions individuelles au projet.

Votre gain issu du deuxième projet est donc calculé par l'ordinateur de la manière suivante :

Votre gain issu du deuxième projet:

Si vous n'êtes pas exclu: $(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet des joueurs non exclus)

Si vous êtes exclu : gain nul pour ce deuxième projet

Votre gain total pour chaque période est calculé par l'ordinateur de la manière suivante :

Votre gain total pour chaque période :

Gain issu du premier projet + Gain issu du deuxième projet

Déroulement des 10 parties suivantes**

Chaque partie se déroule désormais en trois étapes. Vous continuez à jouer avec les mêmes membres du groupe que lors des 10 premières parties. Lors de la première étape, vous participez à un premier projet où vous devez décider combien d'ECU vous désirez contribuer à ce premier projet. Dans la seconde étape, vous êtes informé des contributions de chacun des 3 autres membres de votre groupe au projet. Vous pouvez alors décider d'exclure ou pas un ou plusieurs membres de votre groupe. Dans la troisième étape, seuls les joueurs non exclus peuvent participer à un deuxième projet où ils décident combien d'ECU ils désirent contribuer à ce deuxième projet. Les joueurs exclus ne participent pas au deuxième projet. Les paragraphes suivants décrivent le jeu en détail.

La première étape est identique aux 10 premières parties de cette session. Au début de chaque partie, chaque sujet reçoit une dotation de 20 ECU. Chacun des 4 sujets décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette première étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain de première étape avant décision d'exclusion qui est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des 4 contributions individuelles au projet.

Votre gain en fin de première étape est donc calculé par l'ordinateur de la manière suivante :

Gain en fin de première étape :

$(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet)

Lors de la seconde étape, votre écran vous indiquera le montant que chacun des 3 autres membres de votre groupe a contribué au projet lors de la 1^{ère} étape du jeu. Attention : l'ordre dans lequel vous sont indiquées les décisions des 3 autres sujets est modifié aléatoirement à chaque partie (autrement dit, par exemple, le chiffre qui apparaîtra en 1^{er} sur votre écran ne correspondra pas toujours à la décision du même joueur). Ayant été informé de la contribution individuelle de chacun des autres membres du groupe, chacun de vous a maintenant la possibilité d'exclure un ou plusieurs autres sujets. De la même manière, vous pouvez être exclu si au moins un autre membre du groupe le souhaite. Si vous excluez un membre du groupe, il ne peut pas participer au deuxième projet et ses gains pour ce deuxième projet sont nuls. Il suffit qu'un seul participant au moins décide d'exclure un autre joueur pour que ce dernier soit exclu. L'exclusion ne dure que la période en cours, c'est à dire qu'un joueur exclu à une période revient à la période suivante. Si vous désirez exclure un participant vous devez cocher la case d'exclusion qui se rapporte à sa contribution. A l'inverse, si vous ne désirez pas exclure un participant, vous choisissez de ne pas cocher cette case. Lorsque vous excluez un ou plusieurs participants, vous réduisez la taille du groupe du nombre de participants exclus et donc le nouveau projet ne prendra en compte que les contributions

des participants non exclus. Si vous excluez un ou plusieurs participants, vous subissez un coût en ECU proportionnel au nombre de joueurs que vous décidez d'exclure. Le tableau ci-dessous illustre les coûts associés à chaque exclusion

Nombre de joueurs que je décide d'exclure	Aucun joueur	1 joueur	2 joueurs	3 joueurs
Coût d'exclure	0 ECU	5 ECU	10 ECU	15 ECU

En fin de cette deuxième étape, vous êtes informé de votre exclusion ou de votre non-exclusion, du nombre de joueurs exclus et de votre gain issu du premier projet.

Votre gain en fin de deuxième étape qui est **le gain issu du premier projet** est donc calculé par l'ordinateur de la manière suivante :

Gain issu du premier projet:

Gain de fin de première étape-coût d'exclure

En troisième étape, seuls les joueurs n'ayant pas été exclus participent à un nouveau projet, sous la forme d'une contribution à la constitution d'une somme globale qui sera ensuite partagée entre eux. Cette somme résulte des contributions individuelles des joueurs non exclus. Chaque sujet reçoit à nouveau une dotation de 20 ECU. Chacun des joueurs décide séparément du montant de cette dotation, compris entre 0 et 20, qu'il souhaite affecter au projet. Une fois cette décision prise, vous devez cliquer sur le bouton OK. Ensuite, cette décision n'est plus modifiable pour la partie en cours. A la fin de cette étape, vous serez informé individuellement du montant de la somme effectivement constituée et de votre gain issu du deuxième projet qui est composé de deux éléments : d'une part, le montant de votre dotation que vous avez gardé pour vous-même (c'est-à-dire $20 - \text{votre contribution au projet}$), d'autre part, votre revenu issu du projet : ce revenu représente 40% de la somme des contributions individuelles au projet.

Votre gain issu du deuxième projet est donc calculé par l'ordinateur de la manière suivante :

Votre gain issu du deuxième projet:

Si vous n'êtes pas exclu: $(20 - \text{votre contribution au projet})$

+ 40% (contributions totales au projet des joueurs non exclus)

Si vous êtes exclu : gain nul pour ce deuxième projet

Votre gain total pour chaque période est calculé par l'ordinateur de la manière suivante :

Votre gain total pour chaque période :

Gain issu du premier projet + Gain issu du deuxième projet

* * *

Si vous avez des questions concernant ce que vous venez de lire, merci de lever la main. Il sera répondu immédiatement à vos questions. A la fin de la session expérimentale, je vous remercie de rester à votre place jusqu'à ce que je vous invite à la

quitter et de rester silencieux pendant ce temps.

ANNEXE III.B1

Encadré III.B1.théorème (Fudenberg et Tirole (1991))

S'il existe dans un sous jeu une multiplicité d'équilibre, alors il est désormais possible de punir un joueur qui a dévié à l'étape précédente en choisissant l'équilibre pareto dominé. Au contraire l'équilibre préféré est choisi s'il ne dévie pas.

Considérons ici une version simplifiée du jeu d'ostracisme avec uniquement deux sujets qui prennent trois décisions successives : dans un premier temps, les deux sujets choisissent de contribuer (C) ou de ne pas contribuer (NC) à un premier projet commun ($g_{i,1}=y_1$ ou $g_{i,1}=0$, avec $y_1=20$). Dans un deuxième temps, chaque sujet observe la contribution individuelle de l'autre sujet au projet commun et peut décider de sanctionner ou pas l'autre sujet en l'excluant. Dans une troisième étape, seuls les sujets non exclus peuvent participer à un deuxième projet. Les sujets non exclus décident alors de contribuer (C) ou de ne pas contribuer (NC) à un deuxième projet commun ($g_{i,2}=y_1$ ou $g_{i,2}=0$, avec $y_2=20$).

On cherche à montrer que la solution de coopération (C1,C2) qui n'est pas *à priori* un équilibre de Nash peut être atteinte en première étape du fait de l'existence de deux équilibres de Nash au deuxième jeu. Afin de résoudre le jeu, procédons par récurrence à rebours. En dernière période, les agents non exclus jouent de façon simultanée un jeu de contribution.

Jeu de troisième étape

		Joueur 1	
		C1	NC1
Joueur 2	C1	$\square_{1,1}+32, \square_{2,1}+32$	$\square_{1,1}+16, \square_{2,1}+36$
	NC1	$\square_{1,1}+36, \square_{2,1}+16$	$\square_{1,1}+20, \square_{2,1}+20$

Où $\square_{i,1}$ représente le gain du sujet 1 issu du premier projet. Donc la stratégie dominante est de ne pas contribuer au deuxième projet. Chaque sujet garde alors sa dotation initiale $y_2=20$ qui s'ajoute à ses gains de première étape. En deuxième étape du jeu, les sujets décident de façon simultanée de sanctionner (S) ou de ne pas sanctionner (NS) l'autre sujet. Soit la représentation du jeu sous forme stratégique :

Jeu de deuxième étape

		Joueur 1	
		S	NS
Joueur 2	S	$\square_{1,1}, \square_{2,1}$	$\square_{1,1}+20, \square_{2,1}$
	NS	$\square_{1,1}, \square_{2,1}+20$	$\square_{1,1}+20, \square_{2,1}+20$

Il existe *au moins* deux équilibres de Nash en stratégie pure correspondant au sous jeu de deuxième étape du jeu : (S_1, S_2) et (NS_1, NS_2) . L'équilibre (NS_1, NS_2) domine (S_1, S_2) (au sens de Pareto). Soient les gains associés à chacun des équilibres: $p_i(NS_1, NS_2) = \frac{1}{2} + 20 > p_i(S_1, S_2) = \frac{1}{2}$, $i=1,2$. Puisqu'il existe deux équilibres de Nash dans le jeu de deuxième étape, les sujets peuvent anticiper que les différentes issues du premier jeu seront suivies par différents équilibres en sous jeu dans la deuxième étape (cf van Damme, 1991, Gibbons, 1992). Supposons que chaque sujet i , $i=1,2$, anticipe que l'autre sujet joue S_j en deuxième étape si le sujet i joue NC_i et qu'au contraire que j joue NS_j en deuxième étape si le sujet i joue C_i au premier jeu. En notant b^* la stratégie d'équilibre pour les deux sujets, on peut écrire :

□ En première étape, les sujets doivent décider de leur effort de contribution au premier projet. On peut alors représenter le jeu de première étape de la façon suivante :

Jeu de première étape

		Joueur 1	
		C2	NC2
Joueur 2	C2	52,52	36,36
	NC2	36,36	20,20

Il existe un seul équilibre de Nash : $(C1, C2)$. Donc la coopération peut être atteinte du fait de l'existence de multiples équilibres parfaits en sous de deuxième étape du jeu. On vérifie que la meilleure déviation possible pour un sujet consiste à jouer NC en première étape et S en deuxième étape. Soit les gains pour le sujet 1 :

Alors que la stratégie d'équilibre conduit à :

Donc la stratégie b_1^* est la meilleure réponse à b_2^* dans un jeu séquentiel à deux étapes avec multiplicité d'équilibre en deuxième étape. Par symétrie, le même argument tient pour le sujet 2. Il est donc possible d'atteindre l'issue $(C1, C2)$ en première étape du fait de la menace qui pèse que l'équilibre Pareto dominé soit joué en deuxième étape au lieu de l'équilibre Pareto dominant. Donc la menace d'exclusion est une menace crédible pour les sujets puisqu'elle n'est pas directement coûteuse pour ceux qui l'applique. La coopération peut alors être soutenue en première étape.

ANNEXE III.B2 : Définition de la coopération

On a considéré qu'un sujet était coopérateur s'il contribuait un montant positif au bien public. Une autre définition de la coopération serait de considérer par exemple qu'un sujet est coopérateur uniquement lorsqu'il contribue toute sa dotation au bien public. Il existerait donc autant de niveau de contribution d'équilibre que de définition de la coopération. On montre dans cette annexe que la condition (7) est vérifiée pour toute définition de la coopération. A l'instar de Fehr et Schmidt (1999)⁵¹, on suppose qu'un sujet coopère s'il

⁵¹ Voir Fehr et Schmidt (1998) , "A Theory of Fairness Competition and Cooperation", working paper, pp16-20.

contribue un montant positif $g \in]0, y]$ au bien public. D'après ces auteurs, l'intérêt d'une telle démarche est d'offrir une solution théorique générale qui englobe toutes les définitions de la coopération. Ainsi, au sens strict de définition de la coopération un sujet coopère s'il contribue $g=y$ tandis qu'au sens large de la définition, un sujet coopère s'il contribue un montant non nul au bien public.

Vérifions si la combinaison de stratégies appelée *bannissement*, qui consiste à ce que chaque joueur i contribue un montant positif $g_i = g \in]0, y]$ au premier projet, exclure tous ceux qui ne coopèrent pas au premier projet et enfin ne pas contribuer au deuxième projet, forme un équilibre parfait en sous jeu. A l'équilibre, chacun coopère au premier projet. Si un sujet dévie de l'équilibre, c'est à dire s'il contribue moins que g , alors les autres continuent à coopérer mais excluent le sujet déviant. Dans le deuxième projet, la stratégie de passager clandestin est une stratégie dominante puisque les décisions de contribution à ce deuxième projet ne peuvent faire l'objet de représailles.

Procédons par induction à rebours afin de montrer que la combinaison de stratégies où tous les sujets suivent les stratégies de *bannissement* est une stratégie d'équilibre. Pour cela, montrons qu'aucun sujet n'est incité à dévier de cet équilibre à chacune des étapes du jeu.

- En dernière étape du jeu, les sujets qui n'ont pas été exclus décident de leur niveau de contribution au deuxième projet. Aucun sujet n'est incité à contribuer au projet dans la mesure où la décision à cette étape du jeu ne fait pas l'objet de représailles.
- En deuxième étape, chaque sujet observe la contribution individuelle de chacun des autres membres du groupe et décide d'exclure ou pas les autres sujets. Puisque les sujets s'attendent à ce que le deuxième projet ne soit pas produit quelle que soit la taille du groupe, le coût indirect à exclure un membre est donc nul. Chaque sujet i est donc incité à exclure tous les sujets qui ne coopèrent pas dans le premier projet

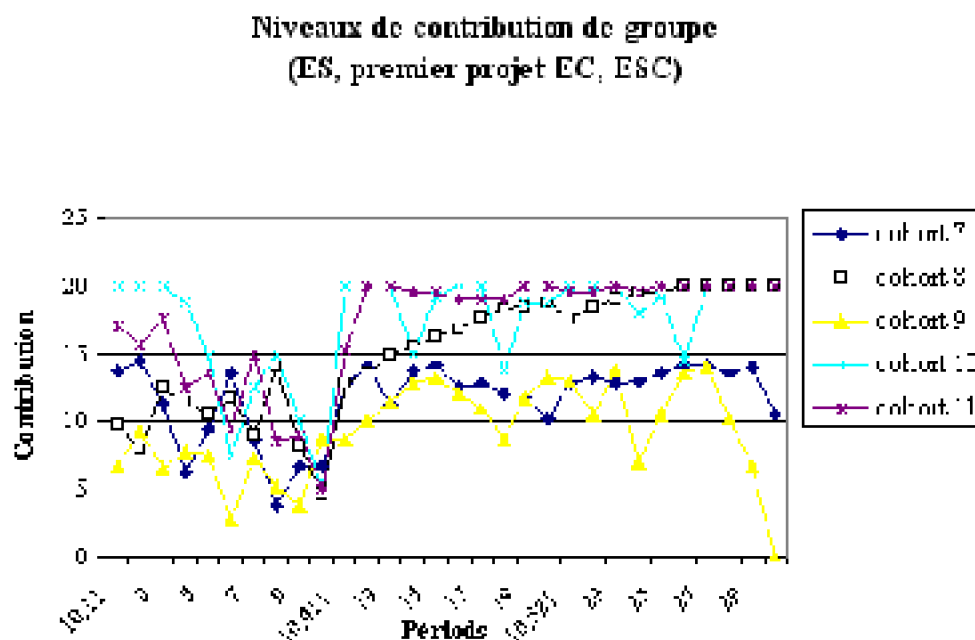
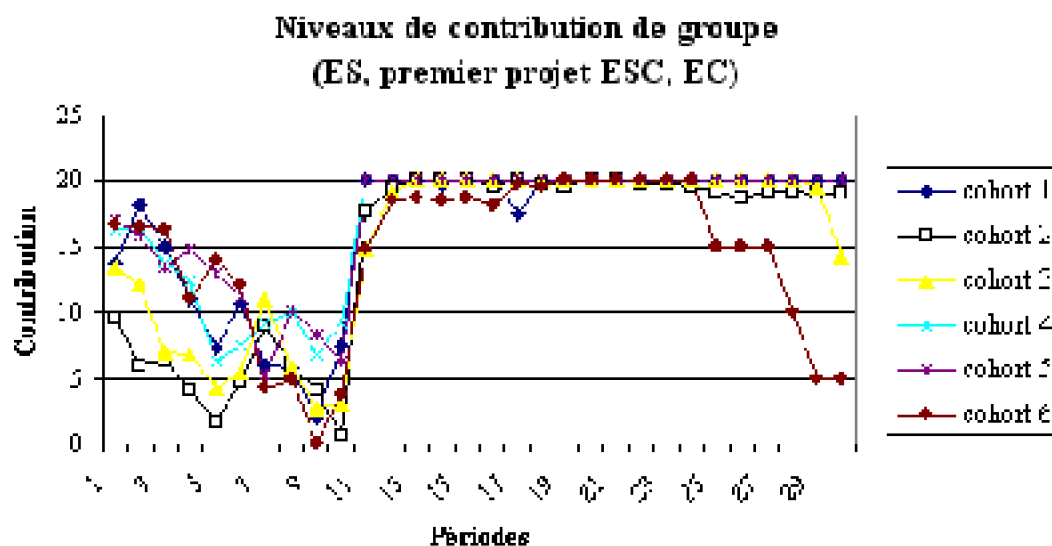
Le premier terme de l'inégalité représente le profit du sujet i s'il contribue g au premier projet et adopte un comportement de passager clandestin au deuxième projet. Le deuxième terme de l'inégalité ci-dessus traduit le gain du sujet i s'il contribue moins que g au premier projet et est exclu du deuxième projet. Ses gains pour la période se résument donc aux gains issus du premier projet. La condition ci-dessus est-elle toujours vérifiée? Selon l'hypothèse sur les économies d'agrégation, on sait que :

De plus, on sait que $y \geq g_i$. Par conséquent la condition ci-dessus est toujours vérifiée. Ainsi les sujets choisissent de coopérer au premier projet. La menace d'exclusion est donc une menace crédible qui permet d'atteindre la coopération dans le premier projet. A l'équilibre, chaque sujet i coopère au premier projet et personne n'est exclu. En troisième étape du jeu, personne ne contribue au deuxième projet. La coopération peut ainsi être atteinte dans le premier projet au coût d'une non-coopération au deuxième projet.

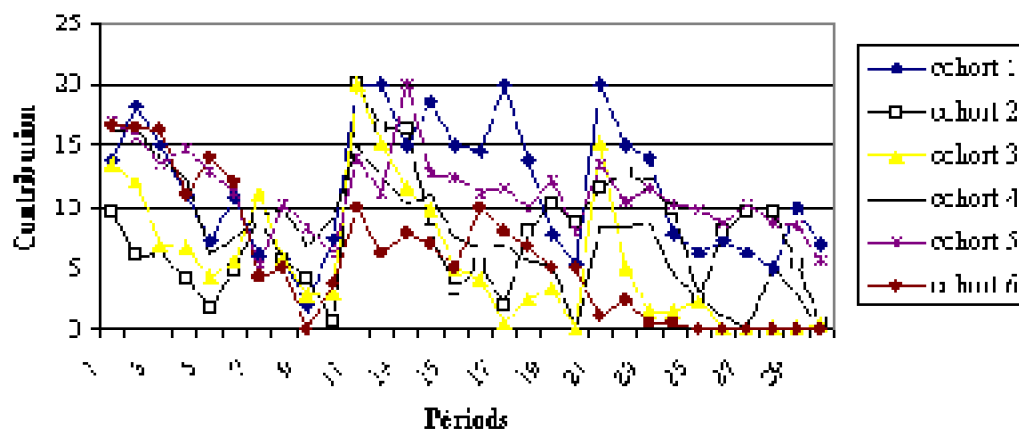
Il existe donc un continuum de revenus d'équilibre selon la définition d'un coopérateur, c'est à dire selon la valeur prise par g . Ainsi lorsque $g=20$, tous ceux qui ne contribuent pas toute leur dotation sont exclus par leurs pairs. Au contraire, lorsque $g=1$, alors seuls ceux qui contribuent aucun ECU au bien public font l'objet d'exclusion. Fehr et Schmidt (1999) avancent le raffinement d'équilibre qui consiste à éliminer tous les

équilibres Pareto dominés pour montrer que l'équilibre sur lequel se coordonnent les joueurs (même s'ils choisissent de façon indépendante leurs stratégies) est celui où tous les joueurs contribuent toute leur dotation au bien public.

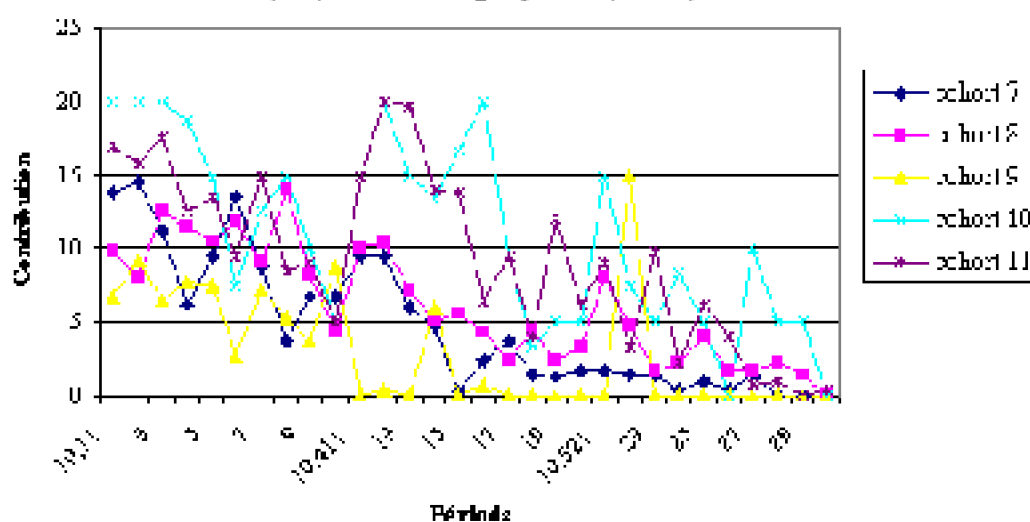
ANNEXE III.C : Contributions de groupe



**Niveaux de contribution de groupe
(ES, deuxième projet ESC, EC)**



**Niveaux de contribution de groupe
(ES, deuxième projet EC, ESC)**



ANNEXE III.D: tests non paramétriques

Tests Wilcoxon signed rank sur la contribution moyenne de groupe

Sessions 1-6

Premier projet p A(ccepte)/ R(ejet) Ho			
(1-10)=(11-20)	0.0156	R	(1-10)<(11-20)
(1-10)=(21-30)	0.0156	R	(1-10)<(21-30)
(21-30)=(11-20)	0.3125	A	(21-30)<(11-20)

en vertu de la loi du droit d'auteur.

Sessions 7-11

Deuxième projet p A(ccepte)/ R(ejet) Ho			
(1-10)=(11-20)	0.2813	A	(1-10)>(11-20)
(1-10)=(21-30)	0.1094	A	(1-10)>(21-30)

Sessions 7-11

Premier projet p A(ccepte)/ R(ejet) Ho			
(1-10)=(11-20)	0.0313	R	(1-10)<(11-20)
(1-10)=(21-30)	0.0313	R	(1-10)<(21-30)
(21-30)=(11-20)	0.2188	A	(21-30)>(11-20)

Deuxième projet p A(ccepte)/ R(eject) Ho			
(1-10)=(11-20)	0.2813	R	(1-10)>(11-20)
(1-10)=(21-30)	0.0313	R	(1-10)>(21-30)

Tests Wilcoxon Mann-Whitney sur la contribution moyenne de groupe

Comparaisons des contributions moyennes : sessions (1-6) / (7-11)

p A(ccepte)/ R(ejet) Ho			
(1-10),T1=(1-10),T2	0.5368	A	(1-10,T1)>(1-10,T2)
(11-20,T1)=(11-20,T2)	0.0043	R	(11-20,T1)>(11-20,T2)
(21-30,T1)=(21-30,T2)	0.0628	A	(21-30,T1)>(11-20,T2)

Comparaisons des gains moyens : sessions (1-6) / (7-11)

p A(ccepte)/ R(ejet) Ho			
(1-10),T1=(1-10),T2	0.7922	A	(1-10,T1)>(1-10,T2)
(11-20,T1)=(11-20,T2)	0.0087	R	(11-20,T1)>(11-20,T2)
(21-30,T1)=(21-30,T2)	0.0887	A	(21-30,T1)>(11-20,T2)

ANNEXES CHAPITRE IV

ANNEXE IV.A Equilibre coopératif sans pression des pairs

Il existe un autre équilibre de Nash où tous les sujets coopèrent à la production de l'output. Afin de montrer que la combinaison des stratégies où tous les sujets coopèrent

en vertu de la loi du droit d'auteur.

est un équilibre de Nash, il faut vérifier que personne n'est incité à dévier de cet équilibre. Le profit de chaque joueur i si tous les sujets coopèrent s'écrit :

ANNEXE IV.B Equilibre coopératif sans pression des pairs dans un jeu de bien public

Soit la fonction d'utilité du joueur i s'il contribue g , c'est à dire s'il ne dévie pas de l'équilibre :

Il est évident qu'aucun sujet n'a intérêt à contribuer davantage que g . En effet, si un sujet augmente sa contribution, il réduit son gain pécuniaire de $(1-a)$ et augmente celui des autres sujets, ce qui produit une inégalité à son désavantage. Considérons maintenant la déviation qui consiste à réduire sa contribution de \square :

Si le sujet i réduit sa contribution au bien public, il obtient un gain pécuniaire de $(1-a)$ mais ressent de l'inégalité à son avantage vis-à-vis des $n-1$ autres sujets qui contribuent g . Le sujet i n'est donc pas incité à réduire sa contribution si $(1-a) < \square_i$. Si $(1-a) < \square_i$, il existe alors un équilibre unique où personne ne contribue au bien public. L'introduction de l'aversion à l'inégalité dans un jeu de bien public sans opportunité de sanction ne change alors pas les prédictions théoriques du jeu.

Bibliographie

- ADDESSALEM T. (1997), "Biens publics avec exclusion : allocations efficaces, production décentralisée", *CNRS Editions*.
- Aghion, P., Tirole, J. (1997), "Formal and Real Authority in Organization", *Journal of Political Economics*, vol. 105, no.1, pp. 1-29.
- ANDOLFATTO D., NOSAL E. (1997), "Optimal team contracts", *Canadian Journal of Economics*, no2, pp.385-396.
- Andreoni, J. (1988), "Why free ride? Strategies and learning in public goods experiments." *Journal of Public Economics* 37:291-304.
- Andreoni, J. (1996), "Cooperation in public-goods experiments: Kindness or Confusion?", *American Economic Review* 85(4):891-904.
- ANDREONI J., MILLER J. (2000), "Giving According to GARP : An experimental Test of the rationality of Altruism", *mimeo, University of Wisconsin and Carnegie Mello University*.
- AOKI M. (1986), "Horizontal versus Vertical Information Structure of the Firm", *American Economic Review*, 76 (5), pp971-83.
- ARNOT R., STIGLITZ J. (1991), "Moral hazard and non market institutions : dysfunctional crowding out or peer monitoring?", *American Economic Review*, vol. 81, n° 1, pp.179-190.
- AXELROD R. (1986), "An Evolutionary Approach to Norms", *American political science*

review, vol.80,no.4, déc.

- BARR A. (2001), "Social dilemmas and shame-based Sanctions : Experimental results from rural Zimbabwe", *Center for the study of African Economies, University of oxford*.
- Barron J. M., Gjerde K.P. (1997), "Peer pressure in an Agency relationship", *Journal of Labor economics*, vol. 15, no. 2, pp. 234-254.
- BECKER G.S. (1968), "Crime and Punishment : An economic Approach"; *Journal of Political Economy*, 76, pp169-217.
- BERG J., DICKHAUT J., McCABE K. (1995), "Trust, Reciprocity and Social History", *Games and Economic Behavior* x, 122-142.
- BETTENHAUSEN K., MURNIGHAN J. (1985), "The Emergence of Norms in Competitive Decision-Making Groups.", *Administrative Science Quarterly*, vol.30, pp350-72.
- BHARGAVA S. (1991), "Profit sharing and profitability : evidence from UK firm level data", 'Draft St John's College, Oxford
- BINMORE K., KIRMAN a., TANI P. (1993), "Frontiers of Game Theory", *MIT Press*
- BINMORE K. (1994), "Game Theory and the Social Contract", *Volume 1 : Playing fair*, Cambridge, MASS: MIT Press.
- Blau, P. (1964), "Exchange and Power in social life", *John Wiley and Sons, Inc. New York, USA*.
- BOLLE F., KRITIKOS a. (1998), "Self Centered Inequality Aversion versus Reciprocity and Altruism", *discussion paper, europea-universität Viadrina, Frankfurt*.
- BOLTON G., OCKENFELS A. (2000), "A theory of equity, reciprocity and competition.", *American economic Review*, vol.100, pp.166-193.
- Bolton G., Ockenfels A., (1997), "ERC, A theory of Equity, Reciprocity and Competition", *Working Paper, august.; American Economic Review (2000)*, 90, 1, mars, pp. 166-193.
- BOWLES S. carpenter j., GINTIS H. (2001), "Mutual Monitoring in Teams : Theory and Evidence on the Importance of Residual Claimancy and Reciprocity", *working paper*.
- Bowles S., Gintis H. (2001), "The Economics of Shame and Punishment" *working paper, university of mass*.
- BOYD R., RICHERSON P. (1985), "Culture and the Evolutionary Process, University Press of Chicago
- BRENNAN G., PETTIT P. (1993), "Hand Invisible and Intangible", *Synthese*, 94, 191-225.
- BUCHANAN J.M. (1965), "An Economic Theory of Clubs", *Economica*, 32, pp1-14
- CABLE J. AND WILSON N. (1990) "Profit sharing and productivity : some further evidence", *The economic Journal*, vol. 100, pp550-5555.
- CABLE J. AND WILSON N. (1990) "Profit sharing and productivity : an analysis of UK engineering firms", *The economic Journal*, vol. 99, pp366-75
- CAHUC P., DORMONT B. (1992), "L'intéressement en France : allègement du coût salarial ou incitation à l'effort ?", *Economie et statistiques*, no257, pp35-43.
- Cahuc, P., Kramarz, F. (1997), "Voice and Loyalty as a Delegation of Authority : a

- Model and a Test on Matched Worker-firm Panels", *Journal of Labor Economics*, vol. 15, no. 4, pp. 658-688.
- CALVO G. A., WELLISZ S. (1978), "Supervision, Loss of Control, and the Optimum Size of the Firm", *Journal of Political Economy*, vol. 86. n° 5
- CAMERER C. (1997), "Progress in Behavioral Game Theory", *Journal of Economic Perspective*, vol. 11, no4, pp167-188.
- CARMICHAEL L. (1988), "Incentive in Academics : Why Is There tenure?", *Journal of political Economy*, vol. 96, no3, pp453-472.
- CARPENTER J. (1999), "Mutual monitoring in teams : theory and experimental evidence", *working paper Middlebury College*.
- CAVALLI-SFORZA L., FELDMAN M. (1981), "Cultural Transmissions and Evolution", *Princeton University Press*.
- CHARNESS G. (1997), "Attribution and Reciprocity in a Simulated Labor Market: An Experimental Investigation", *Working Paper, Department of Economics, UC-Berkeley, March 1998*.
- CHARNESS G, RABIN M (2000a), "Social preferences : Some Simple Tests and a New Model", *working paper, University of California at Berkeley*.
- Charness G., rabin M (2000b), "Responsibility and effort in an experimental labor market", *journal of economic behavior and organization*, 42, 375-384.
- CLARK A., OSWALD J. (1996), "Satisfaction and Comparison Income", *Journal of Public Economics*, LXI, pp359-381.
- Coleman J. (1990), "Foundations of social theory", *Harvard University Press, Cambridge, MA, USA*.
- Coutrot T. (1996), "Relations sociales et performance économique, une première analyse empirique du cas français", *Travail et Emploi*, no.66 pp.39-58.
- CRAIG J., PENCAVEL J. (1995), "Participation and productivity : A comparison of Worker Cooperatives and Conventional Firms in the Plywood industry", *Brookings Papers : Microeconomics*, pp121-160.
- DAVIS J. (1959), "A Formal interpretation of the Theory of relative Deprivation", *Sociometry*, XXII, pp280-296.
- Dawes R. (1975), "Formal Models of dilemmas in social decision-making", *Human Judgment and Decision Processes*, NY : Academic Press. 87-107.
- DEMOUGIN D., FLUET C. (1999), "Costly Sanctions and the Maximum Penalty Principle", *working paper No. 100, Université du Québec à Montréal*.
- Dickinson D.(2000), "The carrot vs. stick in work team Motivation", *working paper ERI 2000-06 Utah State University*.
- DONG X., DOW G. (1993), "Monitoring Costs in Chinese agricultural Teams", *Journal of political Economy*, vol. 101, no3, pp.539-553.
- DuFWENBERG M., KIRCHSTEIGER G. (1998), "A theory of sequential Reciprocity", *Discussion Paper. CenTER, Tilburg University*.
- Durkheim E. (1951) "Suicide, a study in sociology", *NY, free press*
- ECKMAN P. (1992), "An argument for basic emotions", *Cognition and emotion*,

pp169-200

- Eichberger J. (1993) "Game Theory for Economists" *Academic Press*
- Elster J. (1989), "Social Norms and Economic Theory", *Journal of Economic Perspectives*, vol.3, no. 4, pp. 99-117.
- ELSTER J. (1998), "Emotions and Economic Theory", *Journal of economic Literature*, vol. XXXVI, pp47-74.
- engelmann d., STROBEL M. (2000), "An Experimental Comparison of the Fairness Models by Bolton and Ockenfels and by Fehr and Schmidt", *Discussion paper, Humboldt University, Berlin*.
- ESWARAN M., KOTAL A. (1984), "The Moral Hazard of Budget-breaking", *Rand journal of Economics*, vol.15, no.4, pp.578-581.
- FALK A, FEHR E., FISCHBACHER U. (2000), "Informal sanctions", *working paper, University of Zurich*.
- FALK A., FISCHBACHER U. (1998), "A Theory of Reciprocity", *University of Zurich, Working paper*
- Falkinger J., Fehr E, Gächter S., Winter-Ebmer R. (2000), "A simple Mechanism for the efficient provision of Public goods-Experimental evidence", *American economic review*, 90, 247-264.
- Fehr, E., S. Gächter.(1998), "Cooperation and punishment in public goods experiments.", *University of Zurich working paper*.
- Fehr, E., S. Gächter.(1999), "Collective action as a Social Exchange", *Journal of Economic Behavior and Organization*, vol. 39, pp341-369.
- Fehr, E., S. Gächter.(2000), "Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4): 980-994.
- FEHR E., SCHMIDT K. (1997), "A Theory of Fairness and reciprocity - Evidence and economic Applications", *Working paper, university of Zurich*.
- FEHR E., SCHMIDT K. (1999), "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics*, pp. 114, 817-868.
- FEHR E., TYRAN J.R. (2001), "Does Money Illusion Matter?", *American Economic Review* 91(5), pp1239-62.
- FELDMAN A., KIRMAN A. (1974), "Fairness and Envy", *American Economic Review* (1974), vol. 64, issue 6, pp. 995-105.
- Fischbacher, U., S. Gächter, E. Fehr. 2001, "Are people conditionally cooperative? Evidence from a public goods experiment", *Economics Letters* 71(3): 397-404.
- FITZROY F., KRAFT K. (1987), "Cooperation, productivity and Profit sharing", *The Journal of Industrial Economics*, vol. XXXV, dec. 1986, no2.
- FONTAINE P. (1997), "Identification and economic Behavior Sympathy and Empathy in Historical Perspective", *Economics and Philosophy*, 13, pp261-280.
- FRANCIS H. (1985), "The law, Oral tradition and the Mining Community", *Journal of Law and Society* 12, 267-271.
- FRANK R. (1994), "Microeconomics and Behavior", *New York: W.W. Norton & Company*.

- FUDENBERG D., MASKIN E. (1986), "The Folk Theorem in repeated Games with Discounting or with Incomplete Information", *Econometrica*, 54, 533-556.
- FUDENBERG D., TIROLE J. (1991), "Game theory", *Cambridge, Massachusetts: MIT press*.
- Geanakoplos J., Pearce D., Stacchetti E. (1989), "Psychological games and sequential rationality", *Games and Economic Behavior*, 1, pp 60-79.
- GIBBONS R. (1992), "Game Theory for Applied Economists", *Princeton University press*.
- GREEMAN N., GUELLEC D. (1994) , "Organisation du travail, technologie et performances : une étude empirique", *Economie et Prévision*, no 113-114.
- GUTH W. (1995), "An Evolutionary Approach to explaining Cooperative Behavior by Reciprocal Incentives", *International Journal of game Theory*, 24, 323-344.
- GUTH W., SCHMITTBERGER R., SCHWARZE B. (1982), "An experimental Analysis of Ultimatum Bargaining", *Journal of Economic Behavior and Organization*, III, pp. 367-88.
- HANSEN D. (1997), "Individual Responses to a Group Incentive", *Industrial and Labor Relations Review*, 51,1, pp37-49
- Hirschman A.(1970), "Exit, Voice and Loyalty", *Cambridge, MA : Harvard University Press*.
- Hirschman s. (1980), "La passion et les intérêts". , *PUF, paris*.
- HIRSHLEIFER D., RASMUSEN E. (1989), "Cooperation in a repeated prisoners' Dilemma with ostracism", *Journal of economic Behavior and Organization*, 12 (1989) 87-106.
- Hollaender, H. (1990), "A Social Exchange Approach to Voluntary Cooperation", *American Economic Review* 80(5): 1157-1167.
- HOLMSTROM B. (1982), "Moral hazard in teams", *The Bell Journal of Economics*, pp. 324-340.
- HOLMSTROM B., MILGROM P. (1991), "Multitask principal-Agent Analysis : Incentives Contracts, asset Ownership, and Job design", *Journal of Law, Economics, and organization*, 7 special issue, pp24-52.
- Homans, G. (1961), "Social behavior : its elementary forms", *Harcourt Brace & World, New York, NY, USA*.
- IANNACCONE L. (1992), "Sacrifice and Stigma : Reducing Free riding in Cults, Communes and Other Collectives", *Journal of Political economy*, 100, 271-91.
- Isaac, R.M, Walker J., Thomas (1984), "Divergent evidence on free riding : an experimental examination of possible explanations.", *Public Choice* 43 (1) :pp113-49.
- Isaac, M., J. Walker. (1988), "Group size effects in public goods provision: The voluntary contributions mechanism.", *Quarterly Journal of Economics* 103:179-99.
- Isaac M.R, McCue K., Plott C. (1985), "Public goods provision in an experimental environment", *Journal of Public economics* 26 :51-74.
- isaac m.r., walker j.m., williams a.w. (1994), "Group Size and the voluntary provision of Public Goods", *Journal of public economics*, 54, 1-36.

- JONES S. R. (1984) "the Economics of Conformism". *Oxford: Blackwell*
- Jones D.C., Svejnar J. (1985), « Participation, Profit Sharing, Worker Ownership and Efficiency in Italian Producer Cooperatives », *Economica*, 52, pp.449-465.
- JONES D.C., KATO T. (1995), "The Productivity effects of Employee Stock-Ownership Plans and Bonuses : Evidence from Japanese Panel Data", *The American Economic review*, vol.85, no3, pp391-412.
- Kandel, E., Lazear, P.(1992), "Peer Pressure and Partnerships", *Journal of Political Economics*, ,vol. 100, no. 4, pp.801-817.
- KANDORI M. (1992), "Social Norms and Community Enforcement", *Review of Economic Studies* , 59, 63-80.
- KERN H., SCHUMANN M. (1989), "La fin de la division du travail", *Maison des Sciences de l'Homme, Paris*.
- Keser, C. (1996), "Voluntary contributions to a Public Good when partial contribution is a dominant strategy.", *Economic Letters* 50:359-66.
- KESER C., VAN WINDEN F. (1996), "Partners contribute more to public goods than strangers : conditional cooperation", *Working paper, University of Karlsruhe and University of Amsterdam*.
- Keser, C., F. van Winden.(2000), "Conditional cooperation in voluntary contributions to public goods.", *Scandinavian Journal of Economics* 102:23-39.
- Knack, S. (1992), "Civic norms, social sanctions and voter turnout", *Rationality and Society*, 4: 133-156.
- KREPS D. (1997), "The interaction between norms and economic incentives : intrinsic motivation and extrinsic incentives", *AEA Papers and proceeding*, vol.87no2.
- KREPS D., WILSON R. MILGROM P., ROBERTS J. (1982), "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", *Journal of economic Theory*, 27, 245-252.
- KRUSE D.(1992), "Profit Sharing and productivity : Microeconomic Evidence from The united States", *The Economic Journal*, 102, pp24-36.
- LAZEAR P. E. (1989), "Pay Equality and Industrial Politics", *Journal of Political Economy*, 97,pp.561-580.
- LAZEAR P. E. (2000), "Performance, Pay and Productivity", *The American Economic Review*, vol.90, pp.1346-1361.
- Ledyard, J. (1995), "Public goods: A survey of experimental research", in: J. Kagel and R. Roth, eds. *Handbook Of Experimental Economics*, Princeton University Press, Princeton, NJ.
- LEVINE D. (1998), "Modeling Altruism and spitefulness in experiments", *Review of Economic Dynamics* 1, 593-622.
- LEVY-GARBOUA L, RAPOPORT B. (2000), "A Theory of Social Norms, Fairness and Competition", *working paper, Université Paris 1, Panthéon-Sorbonne*
- LIN J.Y. (1990), "Collectivization and china's agricultural crisis in 1959-1961", *Journal of Political Economics*, ,vol. 98, , pp.1228-52.
- LINDBECK A. (1997), "Incentives and social norms in household behavior ", *AEA papers and proceeding*, vol87,no2.

- Loewenstein G. (1996), "Out of control : Visceral Influences on Behavior", *Organizational Behavior and Human decision processes*, vol. 65 (3) , pp 272-92
- Loewenstein G. (2000), "Emotions in Economic Theory and Economic Behavior", *American Economic Review*, 90 (2), 426-432.
- Loewenstein G., Thompson, Bazerman (1989), "Social Utility and Decision Making in Interpersonal Contexts", *Journal of Personality and social Psychology*, LVII, no3., pp426-441.
- MARMAROS D., SACERDOTE B.(2002), "Peer and social networks in job search", *European Economic Review* 46 pp.870-879
- MASCLET D., NOUSSAIR C., TUCKER S., VILLEVAL M.C. (2002), "Monetary and Non Monetary Punishment in the Voluntary Contributions Mechanism", *forthcoming in the American Economic Review*
- MASCLET D. (2001), "Peer Pressure in Work Teams : The Effects of Inequity Aversion", mimeo, *Workshop 2002 of the European Association of Labor Economist, Paris, Journées de l'Association française d'Economie, Lyon*
- MEIDINGER C., RULLIERE J.L., VILLEVAL M.C.V. (2001), "Free riding and Fairness in Principal-Multi-agent relationships : experimental Evidence, *working paper, GATE, Université Lyon2*
- MIDLER (1988) "De l'automatisation à la modernisation, les transformations de l'industrie automobile. Premier épisode : une expérience novatrice chez Renault », *Annale des mines*, no13.
- MILGROM, ROBERT, (1992), "Economics, organization and Management", *Englewood cliffs, NJ : Prencetice Hall*.
- MILLER W. (1992), "Managerial dilemmas : the Political Economy of Hierarchy", *Cambridge University Press*
- MILLER W. (1993), "Humiliation", *ithaca, NY, Cornell U. press*.
- MOIR R. (1998), "Spies and Swords : Costly Monitoring and sanctioning in a Common Pool resource Environment", pp42, *Department of Economics, Saint Johns : University of New Brunswick*.
- Ostrom, E., (1990), "Governing the Commons : The evolution of Institutions for Collective Action", *Cambridge, UK: Cambridge University Press*.
- Ostrom, E., J. Walker, R. Gardner. (1992), "Covenants with and without a sword: Self-governance is possible.", *The American Political Science Review* 86(2):404-17.
- Palfrey, T., J. Prisbrey. (1996), "Altruism, reputation and noise in linear public goods experiments.", *Journal of Public Economics* 61:409-27.
- Palfrey, T., J. Prisbrey. (1997), "Anomalous behavior in linear public goods experiments: How much and why?", *American Economic Review* 87(5):829-46.
- PARSON T. (1954), "Essay in Sociological theory", *NY Free Press*.
- PLUTCHIK R. (1980), "Emotions : A psycho evolutionary synthesis, NY : Harper and Row.
- POLLIS N, (1968), "Reference Groups Re-examined", *British Journal of Sociology*, XIX, pp300-307.

- POSNER A.R., RASMUSEN E.R. (1999), "Creating and enforcing Norms, with special Reference to Sanctions", *working paper, university of Indiana*.
- RABIN M. (1993), "Incorporating Fairness into game Theory and economics", *American economic Review*, 83, 1281-1302.
- RASMUSEN E. (1987), "Moral Hazard in Risk –Averse Teams", *Rand Journal of economics*, vol.18, no3, pp. 428-436.
- RASMUSEN E. (1994), "Games and information : an introduction to game theory", *Blackwell, Cambridge MA and Oxford UK*
- REHDER R. (1990), "Japanese Transplants : After the Honeymoon", *Business Horizons*, 87-98.
- RIELD A., ULE A. (2002) , "Exclusion and Cooperation in Social Network Experiments", *Working paper, Amsterdam*.
- ROETHLISBERGER F.J., DICKSON V. (1939) "Management and the worker", *Cambridge, Harvard University Press*.
- ROTH, A., EREV, I. (1995), "Bargaining Experiments", *in Handbook of Experimental Economics, KAGEL and ROTH, Princeton University Press*.
- SACERDOTE B. (2000),"Peer effects with Random Assignment : results for Dartmouth Roomates", *working paper, NBER*
- SAMUELSON P.(1977),"Thoughts on profit sharing', *Zeitschrift fur die Gesamte staadtswissenschaft (Special Issue)*.
- SAMUELSON L. (1993) "Does Evolution eliminate Dominated Strategies", pp 213-236, *In Binmore K., Kirman A, tani P (eds) Frontiers of Game Theory, MIT*.
- Sefton M., Shupp R., walker J. (2000), "The effect of Rewards and Sanctions in Provision of Public Goods", *Indiana university, working paper*.
- Sefton, M., Steinberg. R. (1996), "Reward structures in public good experiments.", *Journal of Public Economics* 61:263-87.
- SCHELLING, T. (1960), "The Strategy of Conflict", *Cambridge, Harvard University press*.
- SPAGNOLO G. (1999), "Social relations in the workplace : a "linked games" approach", *working paper, Stockholm School of Economics*.
- STIGLITZ (1990), "Peer monitoring and credit markets", *The world bank economic review*, vol 4, no.3, pp 351-366.
- TYLER T.R. (1990),"Why people Obey the Law?", *New-Haven, CT: Yale, University Press*.
- TYRAN J.R., FELD L. (2002),"Why people Obey the law : Experimental evidence from the provision of Public Good", *Working paper, University of St Gallen*
- VAN DAMME E.(1991), "Stability and Perfection of Nash Equilibria", *Springer Verlag; 2Nd/Rev edition*
- VARIAN H. R. (1990), "Monitoring agents with other agents", *journal of institutional and theoretical economics*, pp153-174.
- Weimann, J. (1994), "Individual behavior in a free riding experiment.", *Journal of Public Economics* 54:185-200.

- WEITZMAN M., KRUSE D. (1990) "profit sharing and productivity" *In paying for Productivity* (ed A. Blinder) Washington DC : The brookings Institutions.
- Yamagishi (1986), "The provision of a sanctioning system as a public good.", *journal of Personality and social Psychology*, 51 (1):110-16.
- YANG Y. (1995), "Degree of Supervision , Moral Hazard, and Hierarchical Control", *Journal of Economic Behavior and Organization*, vol. 26, pp93-105.
- ZIZZO D., OSWALD A. (1999), "Are People Willing to Pay to reduce Others' Incomes ?", *working paper*, Oxford University.