

Mécanisme de contribution volontaire et effet de contexte: théorie et évidence expérimentale¹

J.-C. VERGNAUD, M. WILLINGER*, A. ZIEGELMEYER

Bureau d'Economie Théorique et Appliquée, Université Louis Pasteur, 38 boulevard d'Anvers,
67000 Strasbourg, France.

Les raisons sous-jacentes au phénomène de sur-contribution à un bien public - phénomène observé dans les expériences relatives aux mécanismes de contribution volontaire - restent largement inexplorées. Au cours de ces expériences, les sujets sont trop « coopératifs » par rapport à ce que prédit la théorie économique. En cela, ces résultats contrastent avec ceux des expériences sur les oligopoles et sur les ressources communes qui produisent généralement des équilibres de Nash. Or, les externalités créées au cours de ces deux types d'expériences sont de signe opposé; dans les expériences sur les oligopoles et celles sur les ressources communes les choix des sujets créent des externalités négatives, alors que dans les expériences relatives aux mécanismes de contribution volontaire les choix des sujets créent des externalités positives. Dans cet article nous comparons les effets d'un « contexte positif » et ceux d'un « contexte négatif » sur le degré de coopération des sujets placés dans un mécanisme de contribution volontaire. Le résultat de notre expérience est que les sujets participant au contexte positif se montrent significativement plus coopératifs que ceux participant au contexte négatif. Un modèle dynamique de flux de comportements est ensuite construit afin d'établir une équivalence entre des comportements, caractérisés empiriquement à partir de nos données expérimentales, et le niveau de contribution moyen dans les deux contextes.

VOLUNTARY CONTRIBUTIONS AND FRAMING EFFECT: THEORY AND EVIDENCE.

Overcontribution in experiments on voluntary contribution to a public good has not yet received a fully satisfactory explanation. In those experiments, subjects cooperate, in contradiction with the game-theoretic prediction. In contrast, in oligopoly-experiments and common pool resource experiments, the Nash equilibrium is usually observed. The externalities generated by the two types of experiments are of the opposite sign: in oligopoly and common pool resources experiments the subjects' choices generate a negative externality, while in the public good experiment, the actions of the other subjects create a positive externality. In this paper we designed an experiment which allows us to compare the level of overcontribution of the two contexts. Our experiment shows that subjects in the positive context contribute (significantly) more to the public good than the subjects in the negative context. We design a model that relates the difference between the two levels of average contribution, to behavioral differences between two types of players. We provide an empirical characterization of the underlying behaviors.

Mots clés: Biens publics; Mécanisme de Contribution Volontaire; Effet de Contexte; Réciprocité

Classification JEL: C92, H41

1. Introduction

De nombreux résultats expérimentaux montrent qu'en matière de contribution volontaire à un bien public, le comportement de passager clandestin n'est pas systématique. Les prédictions théoriques paraissent donc pêcher par excès de « pessimisme » quant à la production de biens publics. En effet, les sujets de ces expériences sur-contribuent au bien public par rapport à la contribution d'équilibre du jeu sous-jacent. De fait, en termes de bien être, les gains perçus sont plus importants que ceux

¹ Nous remercions, pour sa lecture critique, Francesco De Palma ainsi que Claudia Keser pour ses précieux conseils.

* Tel. : +33 41 52 09, fax : +33 388 61 37 66, email : willma@cournot.u-strasbg.fr.

qu'obtiendraient un groupe de sujets adoptant tous la stratégie dominante du jeu. Si ces résultats présagent de la réalité, ils indiquent alors que les contributions volontaires à un bien public sont en réalité plus importantes que ne le suggère la théorie et qu'une coopération spontanée émerge. Les comportements observés en laboratoire étant en contradiction avec la théorie des jeux standard, il convient d'essayer de comprendre ce qui contribue à cette coopération spontanée.

Les hypothèses avancées pour expliquer cette sur-contribution sont multiples. La première invoque la possibilité d'erreurs de la part des sujets qui n'identifient pas immédiatement quelle est la stratégie dominante. Comme dans les expériences le jeu est répété un nombre fini de fois, par apprentissage, ce phénomène de sur-contribution devrait s'amenuiser et disparaître. Une seconde explication repose sur l'idée de considérations stratégiques liées à l'information incomplète quant aux types des joueurs. Cette hypothèse reprend dans un cadre de jeu public l'argument avancé par Kreps et al. [1982] pour expliquer la coopération rationnelle dans un dilemme du prisonnier répété un nombre fini de fois. Il suffit que chaque joueur pense qu'il y a une probabilité non nulle que parmi les autres joueurs certains soient irrationnels², pour qu'il y ait un intérêt stratégique à se faire passer pour irrationnel. Cette explication sera appelée hypothèse stratégique par la suite. Au contraire de ces deux premières explications, d'autres hypothèses remettent plus fortement en cause les présupposés de la théorie des jeux. Les sociologues évoquent la motivation sociale des individus telle que l'altruisme (Dawes, McTavish et Shaklee [1977] et Marwell et Ames [1979]): dans ce cas les joueurs se préoccupent à la fois de leurs gains mais aussi des gains des autres joueurs. Andreoni [1990] distingue l'altruisme pur, évoqué à l'instant, de l'altruisme impur (appelé « warm-glow giving »): les joueurs se préoccupent de leurs gains mais ils sont aussi attentifs à l'acte de faire du bien aux autres joueurs. Enfin, Levati [1997] évoque la prise de conscience des joueurs du bénéfice mutuel qu'engendre la coopération et leur tentative de persuader les autres de coopérer. Nous verrons dans la revue de littérature (section 2) ainsi que sur la base de nos propres résultats, qu'aucune de ces hypothèses n'est pleinement satisfaisante et suffisante à elle seule.

Dans cet article nous proposons un modèle dynamique de flux de comportements intégrant plusieurs des facteurs explicatifs présentés précédemment. Bien que nous supposions que les agents ont reconnu leur stratégie dominante, d'un point de vue individualiste, nous introduisons trois hypothèses complémentaires pour expliquer leur motivation à sur-contribuer. Tout d'abord, les agents reconnaissent leur intérêt mutuel à contribuer au bien public. En effet, lorsqu'ils estiment que les autres sont bienveillants envers eux-mêmes alors, par réciprocité, ils répondent en sur-contribuant également. La réciprocité est le moteur de fond du phénomène de coopération. On peut lui donner des bases formelles en recourant par exemple à un modèle introduisant l'équité (« fairness » de Bolton [1991], Rabin [1993], ..). Une seconde motivation est celle de persuader et d'instiguer un cercle « vertueux » de coopération, de donner l'exemple en quelque sorte. Nous reprenons ici l'argument avancé par Levati [1997]. Enfin, puisque ce souci de réciprocité existe, il peut être stratégiquement intéressant de sur-contribuer pour récolter ensuite les fruits de cette incitation. Avec cette dernière motivation, nous reprenons l'explication de Kreps et al. [1982]. Nous avons développé ce modèle pour expliquer les profils parfois erratiques des contributions des sujets en laboratoire. Dans notre modèle, de tels changements s'expliquent par une modification de la perception qu'un sujet a des autres joueurs ou bien encore, comme une tentative d'incitation.

Le second apport de notre travail réside dans de nouveaux résultats expérimentaux permettant de comparer les contributions volontaires à un bien public dans deux contextes différents. Le premier traitement, que nous appelons contexte positif du fait que les agents perçoivent la contribution des autres agents comme une externalité positive, correspond à une situation de bien public. Le second traitement, appelé contexte négatif car la non-contribution des autres agents apparaît comme une externalité négative, correspond quant à lui plutôt à un jeu de ressource commune. Formellement, les jeux pour chacun de ces deux traitements sont identiques et les deux équilibres de Nash sont strictement les mêmes. Ils ne diffèrent que par le mode de présentation des gains. Andreoni [1995a]

² Nous qualifions d'irrationnel un joueur qui n'adopte pas la stratégie dominante c'est-à-dire qui sur-contribue au bien public eu égard à l'équilibre.

qui a initié cette voie de recherche, a observé que les contributions volontaires dans le contexte négatif sont significativement plus proches des prédictions de l'équilibre de Nash. Nous observons également ce phénomène dans notre expérience.

Cet effet de contexte a des implications importantes pour l'économie des biens publics. Ces résultats suggèrent en effet que la coopération spontanée sera plus forte pour la production d'un bien public que pour sa conservation (la problématique de la ressource commune). Pour expliquer cet effet de contexte, Andreoni [1995a] y voit un parallèle avec les effets de présentation de la théorie de la décision (« framing effect ») et suggère de reprendre l'hypothèse proposée par Kahneman et Tversky [1979] d'une asymétrie des comportements individuels entre une situation de gain et une situation de perte. Dans une situation d'interaction stratégique, cette voie ne nous semble pas prometteuse (voir Vergnaud et Ziegelmeyer [1998]) et nous avançons plutôt l'idée d'une différence de perception par rapport aux comportements des autres: les sujets jugent avec moins de bienveillance les contributions des autres au bien public dans le contexte négatif. Introduite dans notre modèle, cette hypothèse nous permet d'expliquer un niveau plus faible de contribution dans le contexte négatif. D'autre part, nous proposons une caractérisation empirique des types de comportements développés dans le modèle, ce qui nous permet de valider celui-ci plus complètement sur la base de nos résultats expérimentaux.

L'article est organisé comme suit. Dans la section 2, nous faisons un survol de la littérature expérimentale en nous attachant plus particulièrement à montrer l'insuffisance des hypothèses traditionnelles proposées pour expliquer le phénomène de sur-contribution. Dans la deuxième partie de cette seconde section, nous évoquons l'expérience de J. Andreoni [1995a] qui fut, à notre connaissance, la première expérience caractérisant l'effet de contexte dans un mécanisme de contribution volontaire à un bien public. Notre protocole expérimental est explicité dans la troisième section et nous présentons les résultats obtenus dans la section suivante. La section 5 développe un modèle traduisant l'évolution des comportements des agents en réaction à leur environnement et montre la validité empirique de ce modèle en proposant une caractérisation des comportements observés dans notre expérience.

2. Un aperçu de la littérature expérimentale sur les biens publics

2.1 Phénomène de sur-contribution et explications avancées

Il existe une importante littérature expérimentale sur les biens publics car de nombreuses variantes ont été proposées de façon à apprécier la robustesse du phénomène de sur-contribution et à déterminer les facteurs qui participent à ce phénomène. Dans la plupart de ces travaux, le jeu de base était tel que la stratégie dominante était de ne rien contribuer au bien public tandis que pour le maximum de surplus collectif, il s'agissait pour tous les joueurs d'investir la totalité de leur dotation en bien public. Quatre faits stylisés majeurs se dégagent de l'ensemble des résultats expérimentaux sur les biens publics (voir Ledyard [1995] pour un survol exhaustif de la littérature):

- lors de la première période du jeu répété, les sujets contribuent au bien public, en moyenne, entre 40 et 60 % de leur dotation;
- la sur-contribution moyenne diminue au fur et à mesure des répétitions du jeu;
- le niveau de contribution au bien public est minimum lors de la dernière période du jeu répété;
- l'adoption de la stratégie dominante par l'ensemble des sujets est rarement observée et ce même en dernière période.

Keser [1996] propose un protocole original, où la stratégie dominante est d'investir partiellement dans le bien public. L'auteur observe un niveau de contribution au bien public de 33 % supérieur au niveau de l'équilibre lors de la première période ainsi qu'une décroissance de ce niveau de contribution sans que celui-ci n'atteigne le niveau correspondant à la stratégie dominante même en dernière période. Les quatre résultats expérimentaux majeurs concernant les biens publics avec équilibre en coin sont donc également confirmés dans le cas d'un équilibre intérieur.

Parmi les quatre faits stylisés évoqués précédemment, les trois premiers sont compatibles avec deux des explications que nous avons évoquées: l'hypothèse que les sujets font des erreurs et que par apprentissage ils finissent par découvrir leur stratégie dominante et l'hypothèse stratégique. Bien que l'apprentissage et le comportement stratégique soient deux hypothèses séduisantes, un examen attentif des résultats expérimentaux montre que ces deux hypothèses sont insuffisantes à elles seules pour expliquer la sur-contribution.

Tout d'abord, l'idée que le phénomène de sur-contribution est un artefact expérimental dû aux erreurs des sujets est satisfaisante pour expliquer les résultats observés avec un équilibre de Nash en coin, mais elle l'est beaucoup moins pour l'expérience de Keser [1996]. En effet, si erreurs il y a, elles devraient être réparties symétriquement autour de l'équilibre, ce qui n'est pas le cas. Par ailleurs, Andreoni [1995b] a élaboré un protocole expérimental permettant de séparer la sur-contribution due aux erreurs de celle due à d'autres facteurs explicatifs: l'auteur conclut que le phénomène de sur-contribution ne peut être réduit à des erreurs et que la décroissance du niveau de contribution est plutôt à mettre sur le compte d'une bienveillance avortée. Enfin, des sujets expérimentés, c'est-à-dire ayant déjà participé à une expérience de biens publics, redémarrent une seconde expérience du même type en sur-contribuant à nouveau. De plus, lors de la seconde itération, le niveau de contribution de première période est supérieur au niveau de contribution de dernière période obtenu lors de la première itération (Isaac, Walker et Thomas [1984], Palfrey et Prisbrey [1997]). Ces résultats limitent la portée de l'hypothèse d'apprentissage.

L'hypothèse stratégique, bien que plus convaincante, n'est pas non plus pleinement satisfaisante. Kreps et al. [1982] ont montré, dans un dilemme du prisonnier répété un nombre fini de fois, que si les agents avaient une croyance commune sur le fait que l'autre soit irrationnel alors coopérer devient rationnel: en adoptant un comportement coopératif le premier agent « passe » pour irrationnel et incite le second agent à être coopératif. Cette attitude coopérative du second agent peut perdurer même si le premier agent a cessé de coopérer. Ainsi, l'émergence de comportements coopératifs repose sur l'abandon de la connaissance commune de la rationalité. Toutefois ce comportement stratégique cesse d'avoir un intérêt pour un agent rationnel dans les dernières périodes du jeu. Si tous les agents étaient rationnels, et qu'il existait une croyance commune d'une irrationalité, alors on devrait observer une sur-contribution dans les premières périodes et un brusque effondrement vers la stratégie dominante quelques périodes avant la fin du jeu. Or bien que l'on observe une convergence vers l'équilibre de Nash, celui-ci est très rarement atteint en fin de jeu. Une sur-contribution significativement différente de zéro est encore observée en dernière période dans la très grande majorité des expériences. De plus, l'hypothèse stratégique est acceptable seulement pour un protocole expérimental où les sujets appartiennent à un seul et même groupe tout au long de l'expérience. Andreoni [1988] a élaboré un protocole permettant de tester l'hypothèse stratégique. Au traitement habituel, où chaque sujet appartient à un seul et même groupe tout au long de la session expérimentale (traitement appelé traitement des « partners »), il compare un traitement où les groupes sont redistribués de période en période; l'intérêt de construire une réputation d'altruiste est réduit dans ce second traitement appelé traitement des « strangers ». Andreoni obtient le résultat surprenant que les « strangers » contribuent plus au bien public que les « partners », en contradiction évidente avec l'hypothèse stratégique³.

Puisque l'hypothèse d'apprentissage et l'hypothèse stratégique ne sont pas pleinement satisfaisantes, il faut envisager l'hypothèse que certains agents contribuent au bien public pour d'autres raisons. Faut-il pour autant accepter l'idée d'individus altruistes avancée par les sociologues ? La réponse est clairement non. En effet, une fois de plus, de nombreux travaux expérimentaux sur les biens publics remettent en cause cette explication. Tout d'abord, on observe une variation du niveau de contribution lorsque le rapport entre le rendement marginal du bien privé et le rendement marginal du bien public varie. De plus, la décroissance du niveau de contribution au cours des répétitions est en désaccord avec les préférences stables dans le temps d'un individu altruiste. Enfin, nos résultats expérimentaux remettent en cause de manière radicale cet argument. En revanche, le cadre alternatif

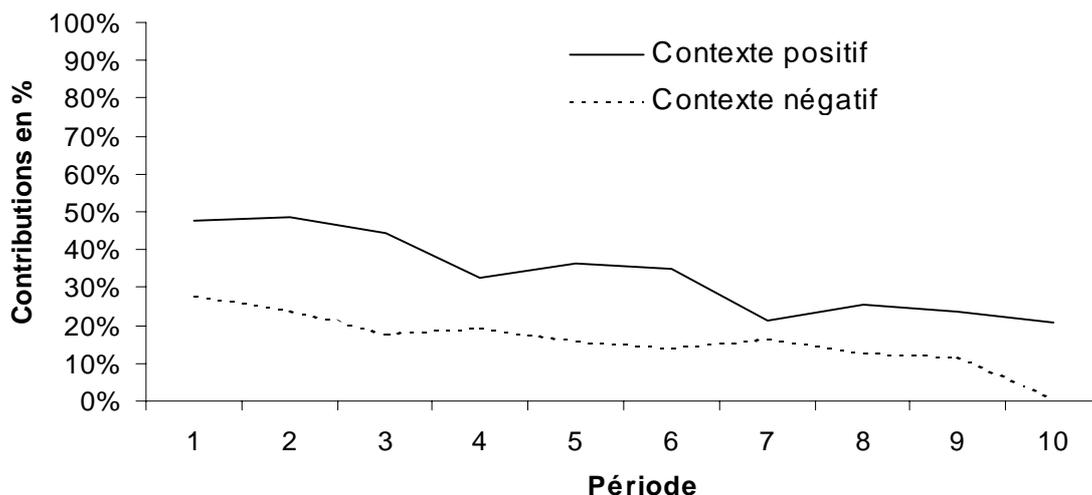
³ Malgré tout, l'expérience de Andreoni [1988] a été répliquée par Croson [1996] et le résultat contraire a été obtenu, à savoir que les partners ont sur-contribué, en moyenne, plus fortement que les strangers.

proposé par Levati [1997] semble mieux à même d'être corroboré par les résultats expérimentaux. Son explication repose sur l'argument de la persuasion. Les sujets comprennent que leur intérêt mutuel est de contribuer l'intégralité de leur dotation à la fourniture du bien public afin d'atteindre l'optimum de Pareto. Etant donné qu'ils ne peuvent communiquer entre eux pour s'accorder sur la « bonne stratégie », ils adoptent une stratégie de persuasion qui consiste à contribuer très fortement durant les premières périodes, mêmes si les autres se comportent en passagers clandestins. Cet effort de sur-contribution est destiné à faire comprendre aux autres leur volonté de coopérer⁴. Toutefois ce comportement de persuasion possède ses propres limites. Au fur et à mesure que le jeu approche de la fin, l'intérêt de persuader les autres se réduit. La convergence vers l'équilibre de Nash est alors expliquée par la « déception » résultant de tentatives de persuasion infructueuses. C'est dans ce sens que vont les explications non standard que nous reprenons dans notre modèle. Plutôt que de s'en tenir uniquement à l'hypothèse d'agents « irrationnels » d'un certain type, nous proposons de définir plus précisément les fondements comportementaux de ces agents, en nous basant notamment sur le principe de réciprocité, qui transparait dans les commentaires post-expérimentaux des sujets. L'avantage est de permettre une formalisation plus riche et susceptible de prédictions théoriques plus fines, et d'introduire une hypothèse de perception expliquant l'effet de contexte.

2.2 L'effet de contexte: une nouvelle variable d'étude dans le cadre des biens publics expérimentaux

Nous présentons maintenant les résultats de l'expérience de J. Andreoni [1995a] sur l'effet de contexte dans un environnement de bien public avec équilibre en coin. Nous rappelons que dans ce cadre expérimental, la stratégie dominante est de rien contribuer au bien public tandis que l'optimum de Pareto est atteint si chaque joueur investit la totalité de sa dotation en bien public.

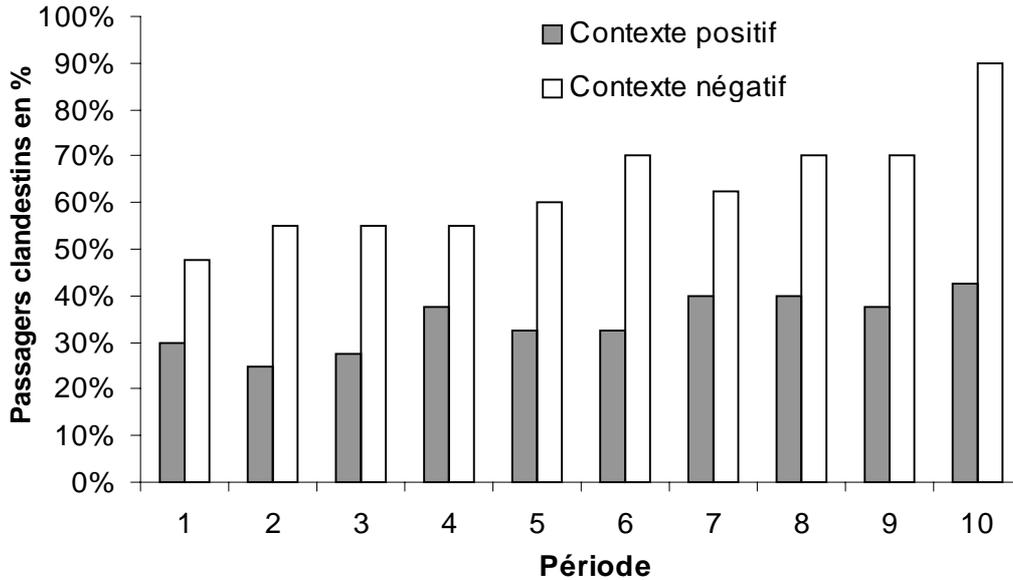
Andreoni constitue des groupes de 5 sujets. Les sujets sont confrontés à 10 répétitions du jeu de base, mais changent de groupe à chaque période pour éviter les effets de réputation. Par ailleurs, l'effet de contexte est produit en écrivant les contraintes budgétaires des agents de deux manières différentes⁵. Ainsi, le contexte négatif n'est qu'une présentation alternative du contexte positif où les incitations financières restent identiques. Le contexte est donc neutre d'un point de vue théorique ce qui implique qu'à la fois la stratégie dominante et l'optimum social sont inchangés. Les résultats de cette expérience sont résumés par le graphique 1 et l'histogramme 1.



Graphique 1: Contributions au bien public par période (en %) pour les deux traitements dans l'expérience de J. Andreoni [1995a].

⁴ Il est à noter que cette explication est en contradiction avec l'analyse standard de la théorie des jeux puisque les déviations hors équilibre doivent être considérées comme corrélées.

⁵ Notre procédure de construction du contexte, calquée sur celle de Andreoni, est présentée en section 3.



Histogramme 1: Pourcentage de sujets ne contribuant rien au bien public par période pour les deux traitements dans l'expérience de J. Andreoni [1995a].

Le graphique 1 indique le pourcentage de sur-contribution au bien public pour les deux contextes, période par période. La différence entre les taux de sur-contribution indique qu'en moyenne la sur-contribution est supérieure de plus de 17 % dans le contexte positif. Un test du rang de Mann-Whitney permet à l'auteur de conclure que cette différence est significative au seuil de 1 % pour l'ensemble des périodes. L'histogramme 1 indique le pourcentage de sujets qui ont adopté la stratégie dominante, période par période. L'histogramme montre clairement qu'à chaque période le pourcentage de passagers clandestins est significativement supérieur dans le contexte négatif. Incidemment, cet histogramme montre également que dans le contexte négatif les observations sont beaucoup plus en accord avec les prédictions théoriques que dans le contexte positif.

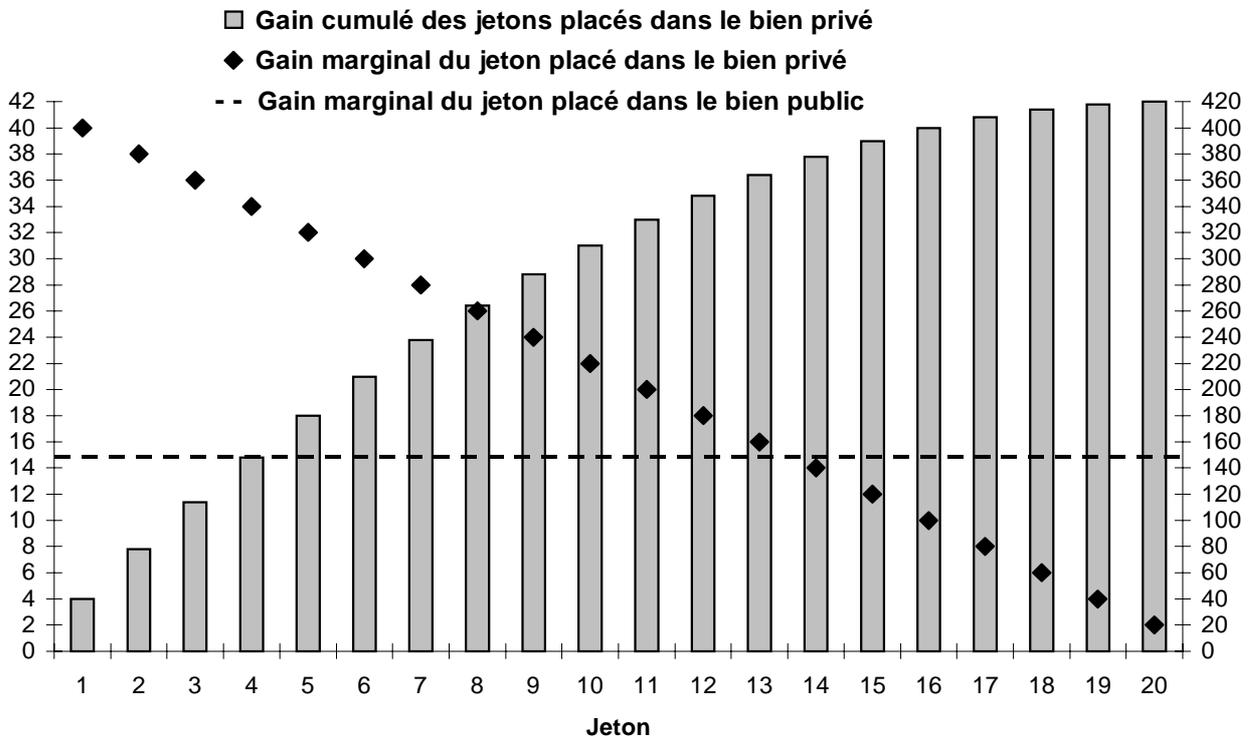
3. Protocole expérimental

3.1 Le jeu de base

Pour déterminer l'effet du contexte sur le comportement individuel de souscription à un bien public, nous comparons deux traitements: l'un avec un contexte positif et l'autre avec un contexte négatif. Le traitement de référence, identique à celui de C. Keser [1996], est le contexte positif. Il possède la propriété d'avoir un équilibre de Nash intérieur unique du jeu. Dans ce traitement la contribution d'un joueur au bien public crée une externalité positive. Le second traitement, correspondant au « contexte négatif », est une présentation différente du contexte positif obtenu en réécrivant les contraintes budgétaires des agents. Cette variation du contexte implique que la non-contribution d'un sujet au bien public crée une externalité négative pour les autres membres du groupe.

Le jeu de base de notre expérience est un jeu symétrique et non coopératif auquel quatre joueurs participent. L'utilité du joueur i est définie par: $u(x_i, y) = 41 x_i - x_i^2 + y$. Les deux premiers termes $41 x_i - x_i^2$ représentent l'utilité associée au bien privé et y l'utilité associée au bien public. L'utilité associée au bien public est donc simplement égale au niveau de production du bien public qui est donné par $y = 15 t$ où $t = \sum_{k=1}^4 t_k$ est la contribution totale du groupe au bien public. A chaque période,

chaque joueur reçoit une dotation de 20 jetons qui doivent être répartis, en totalité, entre le bien privé et le bien public. Le graphique 2 représente le gain marginal de chaque jeton investi dans le bien privé, le gain marginal de chaque jeton investi dans le bien public ainsi que le gain cumulé des jetons placés dans le bien privé.



Graphique 2: Gains marginaux et gains cumulés du jeu de base.

Cette présentation du jeu est celle qui correspond au contexte positif⁶. Lorsqu'un joueur contribue au bien public, il crée une externalité positive pour les autres membres du groupe.

Sous l'hypothèse que le jeu de base est à information complète et que les joueurs ont un comportement rationnel de maximisation de leurs gains, l'unique stratégie dominante d'un joueur est de placer 13 jetons dans le bien privé et 7 jetons dans le bien public. Ce résultat découle de la maximisation de la fonction objectif:

$$G(x_i, \sum_{j \neq i} t_j) = 41x_i - x_i^2 + 15 \left[20 - x_i + \sum_{j \neq i} t_j \right].$$

Ainsi 13 jetons placés dans le bien privé apportent un gain maximum au joueur i , quel que soit l'investissement des autres personnes du groupe. Le graphique 2 montre que les treize premiers jetons placés dans le bien privé ont un gain marginal supérieur à 15, donc supérieur au gain marginal du bien public. Mais l'optimum social requiert que chaque membre du groupe place tous ses jetons dans le bien public puisque le rendement de quatre jetons placés dans le bien public est supérieur au rendement marginal du premier jeton placé dans le bien privé ($4 \cdot 15 = 60 > 40$).

Grâce à la fonction d'utilité quadratique retenue on obtient une solution intérieure pour l'équilibre du jeu de base. Cette paramétrisation présente deux avantages: premièrement ce choix nous permet de contrebalancer le caractère médian de la contribution de première période⁷, et deuxièmement ce choix

⁶ Le bien public que nous considérons ici est un bien public pur à obligation d'usage, il peut donc être assimilé à une externalité.

⁷ En effet, la contribution moyenne au bien public observée en première période se situe, dans la majorité des expériences portant sur des mécanismes de contributions volontaires, approximativement au milieu de l'espace de choix des sujets.

évite que les erreurs des sujets soient toujours commises dans la même direction. En effet, dans le cas où la stratégie dominante est de ne rien contribuer au bien public (solution en coin), une contribution peut résulter soit d'un acte coopératif soit d'une erreur. Avec une solution intérieure, on peut séparer ces deux effets si l'on fait l'hypothèse que les erreurs sont distribuées symétriquement autour de l'équilibre. Cette hypothèse est acceptable puisque la fonction de gain des sujets est symétrique autour de la stratégie dominante; de par sa seule décision, un sujet est autant pénalisé s'il contribue trop au bien public que s'il n'y contribue pas assez⁸. Il paraît donc raisonnable de supposer que les erreurs se compensent et que la fonction de gain retenue permet d'éviter le biais expérimental de la solution en coin (voir Andreoni [1995b]).

Reproduisant la démarche adoptée par Andreoni [1995a], nous réécrivons les contraintes budgétaires des autres joueurs afin de faire apparaître le contexte négatif. Pour chaque joueur i ,

$\sum_{k=1}^4 t_k = t_i + \sum_{j \neq i} t_j$ où $t_j = 20 - x_j$. En remplaçant cette dernière écriture dans la fonction de gain du joueur i nous obtenons:

$$G(x_i, \sum_{j \neq i} x_j) = 41x_i - x_i^2 + 15(20 - x_i) + 900 - 15 \sum_{j \neq i} x_j.$$

Cette présentation alternative du jeu est celle que nous nommerons le contexte négatif. Le coût d'opportunité d'investir dans le bien privé est le défaut d'investissement dans le bien public. Ainsi, dire que la contribution d'un individu au bien public crée une externalité positive pour les autres membres du groupe revient à dire que l'investissement d'un individu dans le bien privé crée une externalité négative pour les autres membres du groupe. La différence entre les deux contextes ne peut donc être qu'une différence de perception. Un sujet placé dans le contexte positif aurait tendance à percevoir les contributions au bien public des autres sujets de son groupe comme une source supplémentaire de gains. Au contraire, un sujet placé dans le contexte négatif aurait tendance à percevoir les non-contributions au bien public des autres sujets de son groupe comme une source de pertes. Par ailleurs, dans le contexte négatif, les sujets perçoivent un capital de 900 de manière automatique et ce à chaque période.

On peut résumer les deux contextes du point de vue du joueur i :

- ⇒ dans le contexte positif, au début de chaque période, l'ensemble des jetons des autres sujets de son groupe est placé dans le bien privé. Lorsque les autres sujets de son groupe déplacent leurs jetons vers le bien public ils améliorent la satisfaction du sujet i ;
- ⇒ dans le contexte négatif, au début de chaque période, l'ensemble des jetons des autres sujets de son groupe est placé dans le bien public. Le sujet i dispose d'un capital de départ de 900. Ce capital est érodé lorsque les autres sujets de son groupe déplacent leurs jetons vers le bien privé; ce faisant ils diminuent la satisfaction du sujet i .

3.2 Le jeu de l'expérience

Le jeu de base est répété durant 15 périodes par le même groupe de quatre joueurs. Au début d'une période, chaque joueur décide individuellement combien de jetons il alloue à chacun des deux biens. A la fin de la période, chaque joueur est informé de son gain pour chacun des biens ainsi que du nombre total de jetons investis par l'ensemble du groupe dans le bien public. Ensuite la période suivante commence. La situation de décision est exactement la même à chaque période. Chaque joueur sait que le jeu se termine au bout de 15 périodes. Enfin, à tout moment de l'expérience, chaque joueur dispose de l'historique du jeu⁹.

⁸ Ceci bien entendu ne reste vrai que tant que le nombre de jetons contribué par le joueur est supérieur ou égal à zéro.

⁹ Les jeux d'instructions - disponibles auprès des auteurs - sont une écriture précise, et pratiquement vide de références économiques, des deux contextes. Leur consultation permet une meilleure compréhension de la distinction qui est faite entre les deux contextes.

Sous l'hypothèse d'information complète et de rationalité des joueurs le jeu répété possède un unique équilibre de Nash parfait en sous-jeux; chaque joueur alloue 7 jetons au bien public à chaque période. Cette prédiction théorique s'applique aussi au contexte négatif. En d'autres termes, la solution théorique du jeu répété dans un contexte positif est identique à celle du jeu répété dans un contexte négatif; le contexte est neutre d'un point de vue théorique.

3.3 Organisation de l'expérience

48 étudiants ont participé à l'expérience, organisée en quatre sessions. Trois groupes de 4 personnes ont été constitués lors de chaque session; ceci porte le nombre total d'observations indépendantes à 6 par traitement. Les sujets ayant participé à l'expérience étaient majoritairement des étudiants de premier cycle et plusieurs disciplines étaient représentées à savoir l'économie, la gestion des entreprises, la biologie appliquée, etc. La proportion de garçons et de filles était presque identique à l'intérieur de chaque traitement. Les participants ont été recrutés par téléphone au sein d'un pool d'environ 500 sujets, volontaires pour participer à des expériences.

Chaque sujet était affecté de façon aléatoire à un groupe de quatre personnes. Afin de préserver l'anonymat et d'éviter toute pression morale, l'expérience s'est déroulée à travers un réseau informatique. L'application informatique nécessaire à la gestion de l'expérience a été développée à l'aide de la boîte à outils «RatImage» [1995]. Les sujets disposaient à leur arrivée d'un jeu d'instructions sur papier.

Dans chaque salle deux assistants étaient chargés de lire les instructions et de vérifier la bonne compréhension de celles-ci par les sujets. Avant que ne commence la première période, les sujets devaient répondre à un questionnaire proposé par l'ordinateur. Le jeu a commencé après que tous les sujets aient répondu correctement à toutes les questions. Pour repérer un éventuel effet de « mode de réponse », la moitié des groupes devaient indiquer le montant de jetons placés dans le bien privé et l'autre moitié le montant de jetons placés dans le bien public¹⁰. A la fin de chaque période, les sujets étaient informés de leur gain (ou perte) relatif à chaque bien, de leur contribution à chaque bien, de la contribution totale du groupe au bien public (ou privé) et de leur gain total sur l'ensemble du jeu¹¹. Les sujets capitalisaient au fur et à mesure du déroulement de la session leurs gains en points; 1000 points valant 5 francs. Une fois les quinze périodes achevées, les sujets pouvaient expliquer brièvement leur stratégie par écrit. Il est important de noter que les sujets ont très majoritairement répondu correctement au questionnaire précédant le jeu et qu'il est apparu, au travers des explications relatives aux stratégies adoptées, qu'une majorité avait identifié la stratégie dominante.

A la fin de chaque session, les étudiants ont été payés individuellement en fonction des gains réalisés sur l'ensemble du jeu. La moyenne des gains perçus a été de 89 francs. Chaque session a duré environ une heure et trente minutes. Aucune limite de temps n'était imposée aux sujets pour prendre leur décision.

4. Résultats expérimentaux: l'effet de contexte

4.1 Résultats de l'expérience

Le contexte positif

¹⁰ Dans un même contexte, aucune différence significative n'a été observée entre les contributions au bien public des groupes plaçant leurs jetons dans le bien public et celles des groupes plaçant leurs jetons dans le bien privé.

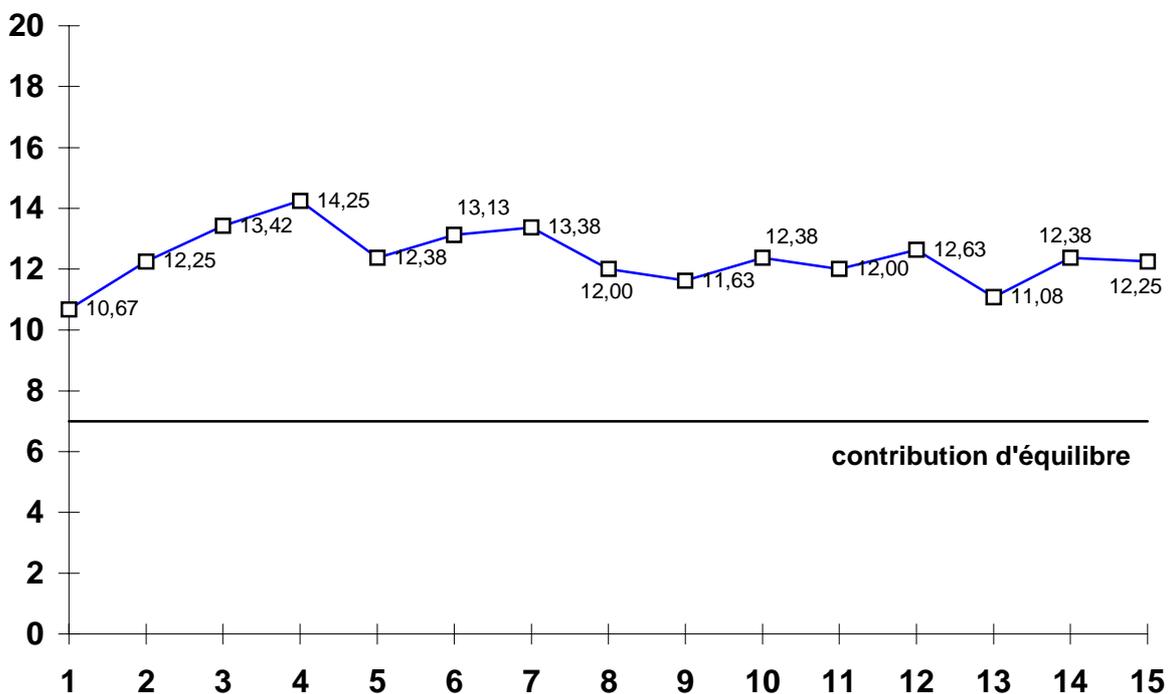
¹¹ Dans le cadre de l'expérience, les termes bien privé ont été remplacés par investissement individuel et les termes bien public ont été remplacés par investissement collectif.

Les vingt-quatre sujets ayant pris part au premier traitement, ont investi en moyenne 12,4 jetons dans le bien public, ce qui correspond à un taux de sur-contribution¹² de 41,5 % par rapport à la contribution d'équilibre (7 jetons). Le tableau 1 détaille l'ensemble de ces résultats.

Périodes	De 1 à 5	De 6 à 10	De 11 à 15	De 1 à 15
contribution moyenne (en jetons)	12,6	12,5	12,1	12,4
taux de sur-contribution	43,1%	42,3%	39,2%	41,5%

Tableau 1: Contributions moyennes et taux de sur-contribution pour le contexte positif.

Un test du Khi-deux confirme que les individus ont une contribution au bien public supérieure à 7 jetons ($p = 0,00007$). Le graphique 3 rend compte de l'évolution de la contribution moyenne sur les quinze périodes et nous permet de comparer celle-ci à la contribution d'équilibre.



Graphique 3: Sentier temporel d'évolution de la contribution moyenne pour le contexte positif.

Les données observées dans le contexte positif sont similaires aux données produites par l'expérience de Keser [1996]. Il semble donc que le protocole que nous avons utilisé soit robuste. Par ailleurs, les résultats obtenus sont en adéquation avec les résultats traditionnels d'un mécanisme de contribution volontaire à un bien public: un degré de coopération significativement non nul est observé. Contrairement à l'expérience de Keser [1996], pour laquelle on observe une décroissance significative du niveau de contribution au bien public au cours des périodes, nous n'observons qu'une faible décroissance non-significative dans le cadre de notre expérience. On peut supposer que ce phénomène de décroissance marquée eut été observé si nous avions adopté un nombre identique de périodes à celui adopté par Keser [1996] à savoir 25 périodes.

Le contexte négatif

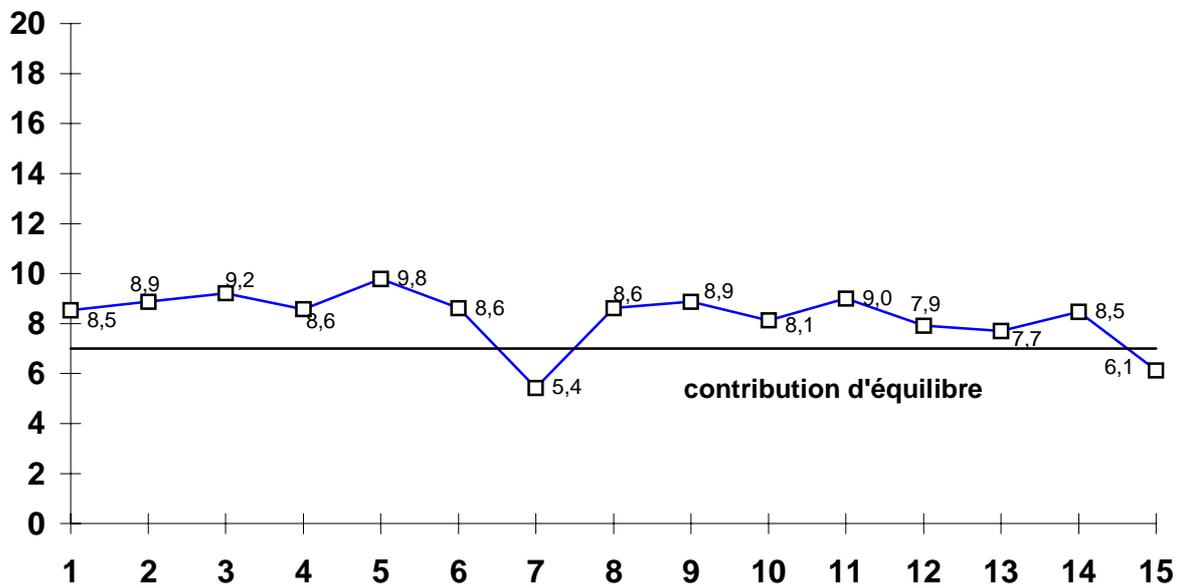
¹² On définit le taux de surcontribution comme le taux de contribution où 7 jetons correspond au fait de ne rien contribuer à l'activité publique, puisqu'il s'agit d'un comportement de passager clandestin.

Les vingt-quatre sujets ayant pris part au second traitement, ont investi en moyenne 8,3 jetons investis dans le bien public, ce qui correspond à un taux de sur-contribution de 9,7 % par rapport à la contribution d'équilibre. Le tableau 2 présente de manière détaillée l'ensemble de ces résultats.

Périodes	De 1 à 5	De 6 à 10	De 11 à 15	De 1 à 15
contribution moyenne (en jetons)	9	7,9	7,8	8,3
taux de sur-contribution	15,4 %	7,1 %	6,5 %	9,7 %

Tableau 2: Contributions moyennes et taux de sur-contribution pour le contexte négatif.

Un test du Khi-deux accepte l'hypothèse nulle selon laquelle le niveau moyen de contribution observé est égal au niveau de contribution d'équilibre ($p = 0,71$). Ainsi, le contexte négatif est un cadre expérimental capable de produire, en moyenne, le comportement de passager clandestin prédit par la théorie standard des jeux. Le graphique 4 rend compte de l'évolution de la contribution moyenne sur les quinze périodes et nous permet de comparer celle-ci à la contribution d'équilibre.

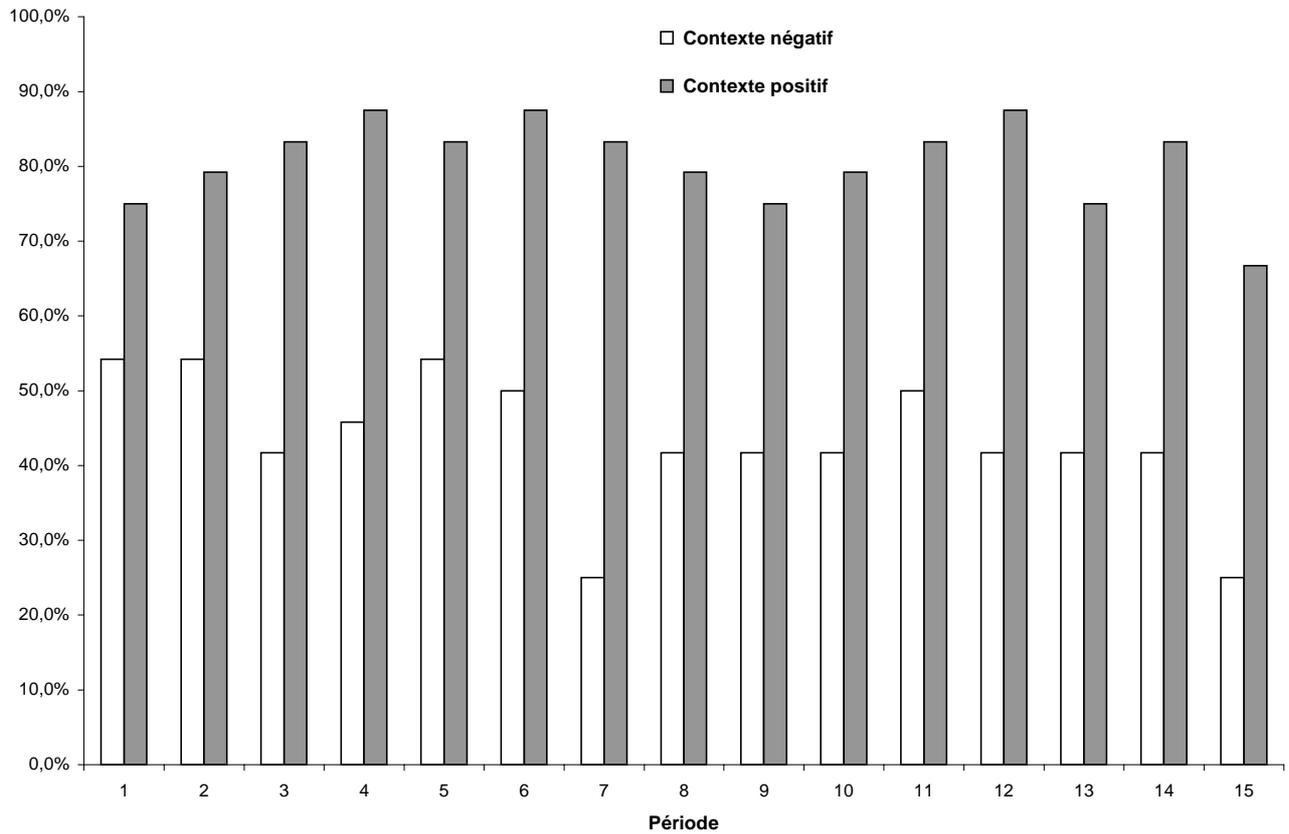


Graphique 4: Sentier temporel d'évolution de la contribution moyenne pour le contexte négatif.

Comme pour le contexte positif on peut noter une légère décroissance, non-significative, du niveau de contribution au bien public.

4.2 Comparaison entre les deux traitements: observation de l'effet de contexte et mise en parallèle avec des résultats expérimentaux passés

Un test unilatéral de permutation, test non-paramétrique, confirme que la contribution moyenne à l'activité publique pour le contexte positif est supérieure à la contribution moyenne à l'activité publique pour le contexte négatif; l'hypothèse nulle selon laquelle les contributions moyennes sont identiques est rejetée au seuil de 3 % ($p = 2,86$ %). Il ressort de la comparaison entre les deux traitements que les individus ne prennent pas des décisions d'investissement identiques suivant le contexte dans lequel ils sont placés. L'histogramme 2 illustre la différence entre les deux contextes quant au nombre de sujets sur-contributeurs.



Histogramme 2: Pourcentage de sujets ayant une contribution à l'activité publique strictement supérieure à 7 jetons par période pour les deux contextes.

Un test unilatéral de permutation indique que le nombre de sujets sur-contributeurs dans le contexte positif est supérieur au nombre de sujets sur-contributeurs dans le contexte négatif; l'hypothèse nulle selon laquelle le nombre de sujets sur-contributeurs est identique dans les deux contextes est rejetée au seuil de 1 % ($p = 0,006$). Bien que nous considérons le comportement de sur-contribution, à la différence de J. Andreoni [1995a] qui établit sa comparaison à partir du comportement de passager clandestin, nos résultats sont en accord avec les siens puisque près du double des sujets adoptent un comportement coopératif dans le contexte positif par rapport au contexte négatif.

La comparaison entre la contribution moyenne au bien public dans le contexte positif et la contribution moyenne au bien public dans le contexte négatif établit donc l'existence d'un effet de contexte qui influence le comportement des individus quant à leur décision d'investissement. Par ailleurs, pour le contexte négatif, nos résultats sont similaires à ceux obtenus pour les expériences sur les ressources communes. Le fait que le niveau de contribution observé soit conforme à la prédiction théorique, semble donc résulter de la présence d'une externalité négative. La question est donc de savoir quelles sont les particularités propres au contexte positif qui induisent une sur-contribution. La section 5 propose de donner une réponse en termes de comportements.

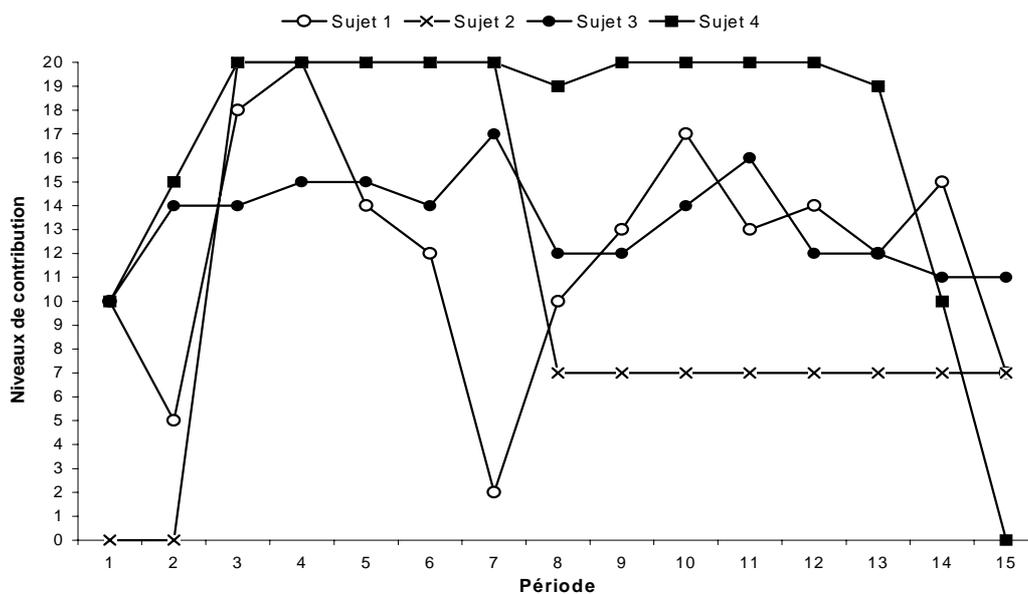
5. Un modèle de flux

L'objectif de cette section est d'élaborer un modèle permettant de rendre compte de la différence de niveau de contribution entre les deux contextes. Le modèle proposé est fondé sur l'hypothèse de coexistence de deux types d'agents: des agents de « type Stratégique », dont le comportement s'accorde avec le comportement prédit par l'équilibre en stratégie dominante, et des agents de « type Réciproque », dont le comportement est animé par un désir de réciprocité en réponse aux actions des

autres. Plusieurs travaux récents contribuent à donner un fondement théorique au comportement de réciprocité, notamment Rabin [1993], Bolton [1997], et Fehr et al. [1997]. L'hypothèse de la coexistence d'agents de type Réciproque et d'agents de type Stratégique se justifie au vu des résultats de notre expérience, comme nous le verrons par la suite. Notre modèle montre que le différentiel de contribution entre les deux contextes peut être expliqué à partir d'une différence de perception quant aux comportements des autres. Ce résultat est obtenu dans un cadre théorique éloigné du cadre expérimental. En effet, notre modèle n'explique pas les sentiers temporels de contribution observés dans l'expérience. De plus, nous n'explicitons pas entièrement les fondements décisionnels des différents comportements adoptés, nous raisonnerons sur la base d'une population large d'agents et l'horizon temporel du modèle ne se limite pas obligatoirement à 15 périodes.

Notre modèle se distingue d'autres modèles fondés sur l'hypothèse de diversité des comportements, par deux caractéristiques. La première caractéristique est que nous autorisons les agents à changer de type en cours de jeu. Ainsi, un agent Réciproque peut devenir Stratégique, et vice versa. Les causes de ces changements de type sont exogènes au modèle (on peut les assimiler à des « mutations »). La seconde caractéristique est que pour un type donné, le comportement d'un agent n'est pas figé, il peut évoluer. Contrairement au changement de type, le changement de comportement est endogène au modèle. Les agents peuvent donc décider d'adopter un comportement différent en fonction du déroulement du jeu. Plus spécifiquement, nous supposons que chaque type est susceptible d'adopter plusieurs comportements. L'adoption d'un comportement spécifique par un agent dépend soit du niveau de la contribution moyenne observée, soit de l'évolution de cette contribution.

Le graphique 5 illustre le sentier temporel des contributions individuelles au sein de l'un des 12 groupes constitutifs de notre expérience. Comme le révèle le graphique, il est difficile d'imaginer que les sujets au sein de ce groupe ont toujours adopté le même comportement. Le sujet 2 par exemple, après avoir commencé à contribuer zéro s'est montré très sur-contributeur jusqu'à la période 7, après quoi, il a changé de comportement et a joué la stratégie d'équilibre pour le restant du jeu. De même, les autres sujets ont des contributions élevées à certaines périodes, alors qu'à d'autres périodes ces contributions se rapprochent de la contribution d'équilibre (quelques contributions sont même inférieures à 7); l'hypothèse de stabilité des comportements nous paraît dès lors peu pertinente pour rendre compte des données observées.



Graphique 5: Un exemple d'évolution des contributions individuelles au sein d'un groupe.

L'effet de contexte est expliqué dans ce modèle par une différence de perception de la contribution des agents, cette perception influençant le comportement de réciprocité des agents. L'idée est que les

agents sont moins tolérants dans le contexte négatif et par conséquent jugent l'action des autres avec moins de bienveillance. Le modèle proposé cherche dès lors à expliquer la différence du niveau de contribution dans les deux contextes, par une différence de changement de comportement des agents de type Réciproque. La traduction formelle de cette hypothèse nous permet de dégager des éléments de statique comparative qui sont testés sur nos données.

5.1 Types et comportements

5.1.1 Type Réciproque

Le souci de réciprocité que nous modélisons est plus complexe que de l'altruisme simple. Quand l'agent se sent redevable des autres – par exemple, parce que ceux-ci ont largement investi en bien public – il investira lui aussi beaucoup. Inversement, s'il juge que les autres n'ont pas fait preuve à son égard d'une bienveillance suffisante, il investira peu en bien public cherchant avec malveillance à faire subir des pertes aux autres, quitte à en subir lui même. C'est en ce sens qu'il ne s'agit pas d'un altruisme simple puisque l'agent ne tient compte positivement des autres que s'il se sent redevable¹³. Pour saisir cette réciprocité, nous supposons qu'il existe deux comportements: Bienveillant (noté B) et Malveillant¹⁴ (noté M). L'agent de type Réciproque peut également adopter, à une période, un comportement d'incitation, c'est-à-dire effectuer un investissement plus important en bien public que l'investissement moyen observé chez le reste de la population à la période précédente. Ce comportement correspond à l'incitation-persuasion évoquée par Levati [1997]; l'agent de type Réciproque montre son bon-vouloir pour essayer de relancer le cercle vertueux de la réciprocité. En résumé, l'agent de type Réciproque peut adopter un comportement Bienveillant, Malveillant ou un comportement d'Incitateur Réciproque (noté IR). Nous noterons b_{IR} le taux de passage du comportement Bienveillant au comportement Incitateur Réciproque (i_{B} le taux de passage en sens contraire) et b_M le taux de passage du comportement Bienveillant au comportement Malveillant (m_B le taux de passage en sens contraire).

Le passage entre Bienveillant et Incitateur Réciproque est fonction de l'évolution de la contribution moyenne du reste de la population. Si celle-ci se dégrade, un agent Bienveillant sera davantage enclin à tenter de relancer la coopération en contribuant fortement. Par contre, un agent Malveillant n'a pas a priori de raison de devenir Incitateur. Il pourrait néanmoins devenir Incitateur avec un taux de changement de comportement très faible. Dans la version du modèle que nous présentons ici nous supposons que le passage de Malveillant à Incitateur Réciproque est impossible. Toutefois, les prédictions théoriques du modèle restent inchangées si l'on suppose que le taux de changement de comportement vers Incitateur Réciproque est plus fort pour un Bienveillant que pour un Malveillant.

5.1.2 Type Stratégique

Le type Stratégique cherche uniquement à maximiser ses gains sans souci de réciprocité; en cela il adopte un comportement Individualiste qui se confond avec la stratégie dominante. Mais étant confronté à des agents potentiellement Réciproques, sa meilleure stratégie n'est plus nécessairement de s'en tenir à la stratégie dominante puisqu'il a les moyens d'influencer par son propre investissement les investissements en bien public des autres agents. Un comportement d'incitateur est donc possible, mais pour des raisons différentes de celle de l'Incitateur de type Réciproque. L'incitation par une forte contribution à une période du jeu peut s'avérer stratégiquement payante pour les périodes suivantes. En résumé, le type Stratégique peut adopter un comportement Individualiste

¹³ Rabin [1993] a proposé un modèle de prise en compte par l'agent de la bienveillance qu'un autre agent peut avoir à son égard. Ainsi, l'utilité générale de l'agent est obtenue en ajoutant un terme représentatif de la bienveillance réciprocque des agents.

¹⁴ Nous définissons plus loin la stratégie d'investissement qui est alors réalisée.

(noté I) ou un comportement d'Incitateur Stratégique (noté IS). Nous noterons i_{IS} le taux de passage du comportement Individualiste au comportement Incitateur Stratégique (i_{SI} le taux de passage en sens contraire).

5.1.3 Changement de type

Nous supposons qu'en plus des changements de comportement pour chacun des deux types d'agents, des changements de type peuvent se produire. Ce changement de type peut s'effectuer dans les deux sens suivants. Un agent Bienveillant ou Malveillant peut adopter un comportement Individualiste et devenir alors un agent de type Stratégique. A l'inverse, un agent Individualiste peut adopter un comportement Bienveillant ou Malveillant (selon un taux de changement de type qui correspond aux proportions existantes). De plus, nous supposerons que le taux de changement de type d'un agent Réciproque vers un agent Stratégique (et vice versa) est constant et identique dans les deux contextes. En effet, nous pensons que ni la contribution moyenne observée ni son évolution ne peuvent expliquer le changement de type puisque les motivations pour l'existence d'une coopération au sein de la population sont à l'opposé pour les deux types. Nous noterons r_S le taux de changement du type Réciproque vers le type Stratégique dans les deux contextes (s_R le taux de changement en sens contraire) et nous posons $0 < r_S, s_R < 1$. Nous faisons l'hypothèse que le changement de type ne s'effectue pas quand l'agent est incitateur, car le comportement d'incitation doit se comprendre comme un pari sur l'avenir. L'agent de type Réciproque espère convaincre les autres de contribuer plus largement afin de maintenir la contribution du groupe à un niveau élevé par la suite¹⁵. Un agent de type Stratégique, en décidant d'être incitateur, tentera également de convaincre le reste de la population d'augmenter sa contribution, mais pour profiter de cette générosité éventuelle par la suite.

La figure 1 résume l'ensemble des évolutions possibles en termes de flux de population. Les flux de changement de comportement sont symbolisés par des flèches blanches et les flux de changement de type par des flèches grisées.

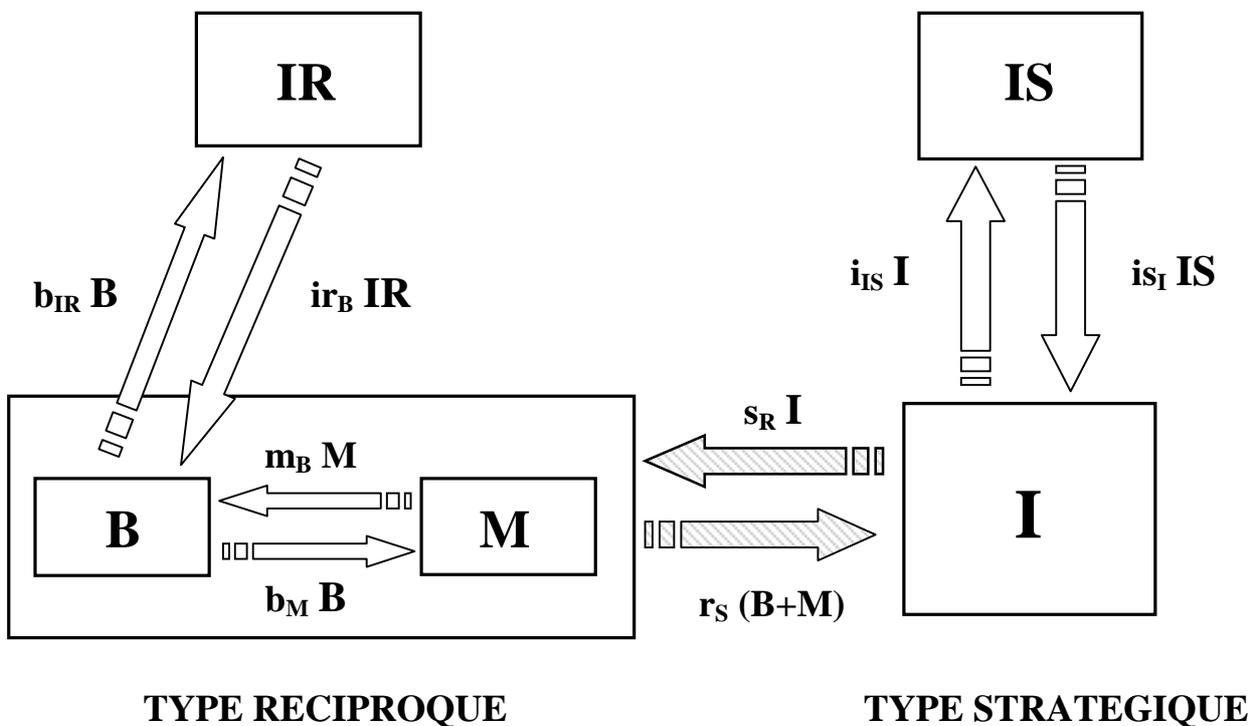


Figure 1: Les flux « inter-type » et « intra-type ».

¹⁵ Nous rappelons que, à l'intérieur du type Réciproque, seul l'agent ayant adopté le comportement Bienveillant a la possibilité d'inciter.

5.2 Résolution du modèle

Le modèle est basé sur des comportements en termes de niveau de contribution et sur des hypothèses d'évolution de ces comportements. La résolution du modèle nous permettra de comparer les tailles relatives des différents comportements à l'état stationnaire dans les deux contextes et d'établir une équivalence entre ces tailles relatives et les niveaux de contribution moyens à l'état stationnaire dans les deux contextes.

5.2.1 Caractérisation des comportements

Nous caractérisons les comportements de contribution dans un but essentiellement descriptif, afin de rendre compte du différentiel de contribution moyen dans les deux contextes. Notons t_n^i la contribution au bien public de l'agent i à la période n et \bar{t}_{n-1}^i la contribution moyenne au bien public des autres agents à la période $n-1$.

L'agent Individualiste a un comportement unique, il adopte la stratégie d'équilibre. Par contre, l'agent de type Réciproque modifie son comportement suivant le niveau moyen de contribution observé chez le reste de la population à la période précédente. Soit il juge de manière bienveillante ce niveau moyen observé et dans ce cas il y répond par un niveau de contribution identique. En ce sens, il se montre bienveillant à l'égard des autres en réponse à la bienveillance passée des autres agents à son propre égard. Mais si l'agent de type Réciproque juge de manière malveillante ce niveau de contribution observé alors il répond à la malveillance des autres à son égard en étant malveillant à son tour. Dans ce cas, il est disposé à sacrifier une partie de son gain individuel afin de « punir » les autres de leur malveillance. Finalement, nous caractérisons le comportement d'un agent Incitateur, comme un agent qui contribue davantage au bien public que la contribution moyenne observée chez les autres agents à la période précédente. En résumé, les comportements sont caractérisés comme suit:

- l'agent Individualiste investit systématiquement 7 jetons dans le bien public;
- l'agent Bienveillant investit dans l'activité publique l'investissement moyen observé à la période précédente chez le reste de la population;
- l'agent Malveillant investit en bien public moins que les 7 jetons de la stratégie dominante, soit un montant égal à $7 - a$, avec $a \in \mathbb{IN} \cap [1, 7]$;
- l'agent Incitateur investit en bien public plus que l'investissement moyen de ses partenaires à la période précédente, soit $\bar{t}_{n-1}^i + b$, avec $b \in \mathbb{IN} \cap [1, 20 - \bar{t}_{n-1}^i]$.

5.2.2 Hypothèses sur les taux de changement de comportement

L'adoption par l'agent i d'un comportement à la période n dépend soit du niveau moyen de contribution du reste de la population à la période précédente, \bar{t}_{n-1}^i , soit de l'évolution de ce niveau moyen observé entre les deux périodes précédentes, $\dot{t}_{n-1}^i = \bar{t}_{n-1}^i - \bar{t}_{n-2}^i$. Nous rappelons que le type de l'agent i à la période n est lui déterminé de manière exogène.

Hypothèse 1: sens de variation des taux de changement de comportement

- a) Le taux m_B est croissant avec \bar{t}_{n-1}^i et le taux b_M est décroissant avec \bar{t}_{n-1}^i . De plus, $\forall \bar{t}_{n-1}^i$, nous posons $0 < m_B, b_M < 1$;
- b) Les taux b_{IR} et i_{IS} sont décroissants avec \dot{t}_{n-1}^i et les taux i_{rB} et i_{sI} sont croissants avec \dot{t}_{n-1}^i . De plus, $\forall \dot{t}_{n-1}^i$, nous posons $0 < b_{IR}, i_{IS}, i_{rB}, i_{sI} < 1$.

La population d'agents étant supposée large, la contribution moyenne de tous les agents, agent i compris, à la période précédente est identique à la contribution moyenne des $(-i)$ agents à la période précédente. Nous omettrons donc par la suite l'indice supérieur $(-i)$.

D'après l'hypothèse 1, le passage de Bienveillant à Malveillant dépend simplement du niveau de coopération entre les agents au sein de la population. Si le niveau de coopération, mesuré par la contribution moyenne de la période précédente, augmente, le nombre d'agents qui passe de Malveillant à Bienveillant augmente. En revanche, la décision d'adopter un comportement Incitateur est fonction de l'évolution du niveau de contribution. Ainsi, dans le cas où la contribution moyenne au bien public diminue, $\dot{t}_{n-1} < 0$, alors davantage d'agents vont adopter un comportement d'incitateur afin de relancer la coopération qui se dégrade, et ce indépendamment de leur type.

Hypothèse 2: taux de changement de comportement et contexte

- a) Pour une même contribution moyenne, $(m_B)^+ > (m_B)^-$ et $(b_M)^+ < (b_M)^-$;
 b) Pour une même variation de la contribution moyenne, $(b_{IR})^+ = (b_{IR})^-$, $(i_{IS})^+ = (i_{IS})^-$, $(i_{IR})^+ = (i_{IR})^-$ et $(i_{SI})^+ = (i_{SI})^-$.

Le symbole + renvoie au contexte positif et le symbole - au contexte négatif.

L'hypothèse 2 exprime l'idée d'une différence de perception, suivant le contexte, d'un même niveau moyen de contribution observé; cette perception différente influençant le comportement de réciprocité de l'agent. Par contre, une variation de la contribution moyenne est perçue de manière identique dans les deux contextes; l'adoption du comportement incitateur n'est donc pas influencée, de manière directe, par le contexte.

5.2.3 Solution d'équilibre du modèle

Nous sommes en mesure maintenant de définir l'état stationnaire de la population. A l'état stationnaire, les niveaux relatifs des comportements étant constants, le niveau de contribution moyen ainsi que les taux de passage d'un comportement à un autre sont constants. Nous supposons que la taille de la population est constante et, sans perte de généralité, nous normalisons cette taille à 1: $B + M + I + IS + IR = 1$. Ainsi, l'état stationnaire est défini par les équations suivantes:

$$(1) (i_{IR})^* IR^* + (m_B)^* M^* + s_R \frac{B^*}{B^* + M^*} I^* = B^* [(b_M)^* + r_S + (b_{IR})^*];$$

$$(2) s_R I^* = r_S [M^* + B^*];$$

$$(3) (b_{IR})^* B^* = (i_{IR})^* IR^* ;$$

$$(4) (i_{IS})^* I^* = (i_{SI})^* IS^* ;$$

$$(5) B^* + M^* + IR^* + IS^* + I^* = 1 ;$$

$$(6) \bar{t}^* = \bar{t}^* B^* + (7-a) M^* + (\bar{t}^* + b) IN^* + 7 I^*.$$

où X^* est la taille relative du comportement X à l'état stationnaire ($X = B, M, I, IS$ et IR), $(x_X)^*$ est le taux de changement de comportement à l'état stationnaire, \bar{t}^* est le niveau de contribution moyen à l'état stationnaire et $IN^* = IR^* + IS^*$ désigne le niveau relatif du comportement Incitateur à l'état stationnaire.

La prédiction théorique de notre modèle consiste en une équivalence entre les niveaux relatifs des différents comportements et le niveau de contribution moyen de la population à l'état stationnaire.

Proposition 1.

Il existe une relation d'équivalence entre l'ensemble des inégalités $\{(M^*)^+ < (M^*)^-, (B^*)^+ > (B^*)^-, (I^*)^+ < (I^*)^-, (IN^*)^+ > (IN^*)^-\}$ et l'inégalité $(\bar{t}^*)^+ > (\bar{t}^*)^-$.

Preuve

Voir Annexe 1.

L'ensemble des inégalités constitutives de la prédiction théorique de notre modèle peut être testé à partir de la caractérisation empirique des comportements. Nous rappelons que l'effet de contexte, $(\bar{t}^*)^+ > (\bar{t}^*)^-$, a déjà été testé avec succès. La proposition 1 peut être complétée par deux inégalités non testables empiriquement: $(IR^*)^+ > (IR^*)^-$ et $(IS^*)^+ < (IS^*)^-$.

5.3 Test du modèle

Dans cette section, nous validons empiriquement la prédiction théorique de notre modèle, à savoir la proposition 1, de manière indirecte. En effet, notre principal résultat expérimental est l'effet de contexte. Bien entendu, notre cadre théorique a été élaboré de manière à expliquer cette effet de contexte. Mais notre modèle prédit aussi un ensemble d'inégalités sur les niveaux relatifs des différents comportements. Si nous montrons que ces inégalités sont vérifiées par nos données expérimentales alors notre modèle sera validé empiriquement. Pour ce faire, nous commençons tout d'abord par recenser la prévalence des comportements en relation avec les stratégies définies dans la sous-section 5.2.1.

5.3.1 Définition des caractérisations empiriques

Nous proposons ici des critères de classification des sujets qui correspondent aux hypothèses du modèle, mais qui tolèrent certaines variations par rapport aux comportements qui y sont décrits. Ceci dans le souci de pouvoir caractériser chaque sujet à chaque période du jeu expérimental et ce quel que soit son niveau de contribution à l'activité publique. Ainsi, nous définissons des zones, et non plus des niveaux comme dans notre modèle, d'appartenance de la contribution de chaque sujet à chaque période du jeu expérimental. Par ailleurs, la caractérisation retenue définit les comportements du sujet Incitateur, du sujet Malveillant et du sujet Individualiste ; le comportement Bienveillant est donc défini par défaut. Pour caractériser un sujet en période courante, nous comparons sa contribution (t_n^i) à la contribution moyenne des trois autres membres de son groupe à la période précédente (\bar{t}_{n-1}^{-i}). Une telle caractérisation ne peut donc se faire que sur les 14 dernières périodes du jeu expérimental.

Une distinction est faite suivant que le niveau de contribution moyen observé est strictement supérieur au niveau de contribution d'équilibre ou inférieur. Nous justifierons cette distinction lors de l'évocation de la caractérisation de chaque comportement.

Nous n'avons pas établi de distinction empirique entre un Incitateur Réciproque et un Incitateur Stratégique. Cette distinction nécessite de prendre en compte l'évolution du comportement du sujet sur au moins trois périodes, avec une caractérisation plus fine des comportements. Nous nous contenterons donc d'une comparaison de l'ensemble de tous les Incitateurs dans les deux contextes pour valider empiriquement le modèle, et ce bien que celui-ci précise les inégalités existantes suivant le type de l'incitateur.

La caractérisation empirique des comportements est présentée dans le tableau 3 ci-dessous.

Contribution moyenne des trois autres membres du groupe à la période précédente	$\bar{t}_{n-1}^{-i} \leq 7$	$\bar{t}_{n-1}^{-i} > 7$
	INCITATEUR	$t_n^i \geq 7 + b$
BIENVEILLANT	-	$6 + b < t_n^i < \bar{t}_{n-1}^{-i} + b$
MALVEILLANT	$0 \leq t_n^i \leq 7 - a$	$0 \leq t_n^i \leq 7 - a$
INDIVIDUALISTE	$8 - a \leq t_n^i \leq 6 + b$	$8 - a \leq t_n^i \leq 6 + b$

Tableau 3: Caractérisation empirique des comportements.

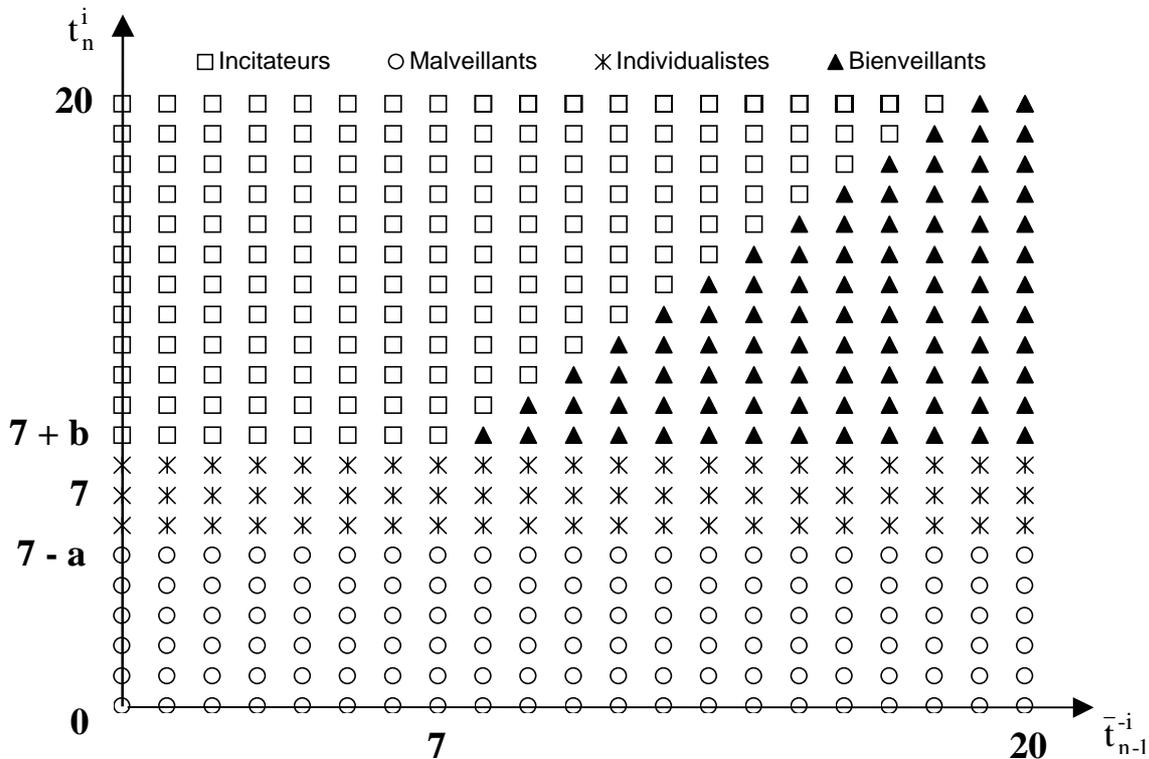
La caractérisation d'un sujet Incitateur est distincte suivant le niveau moyen de contribution qu'il observe. En effet, un sujet est considéré comme Incitateur à la période courante s'il investit plus que la contribution moyenne des autres membres du groupe observée à la période précédente, soit $t_n^i \geq \bar{t}_{n-1}^{-i} + b$ ($b > 0$), et si cette contribution moyenne observée est strictement supérieure à la contribution d'équilibre. Mais, dans le cas où la contribution moyenne observée est inférieure à la contribution d'équilibre, un sujet est considéré comme Incitateur s'il investit plus que la contribution d'équilibre, soit $t_n^i \geq 7 + b$. Cette distinction est nécessaire afin d'éviter que des contributions inférieures à 7 puissent être considérés comme incitatrices. Ainsi, nous restreignons l'ensemble des contributions considérées comme incitatrices aux contributions strictement supérieures à 7. Par ailleurs, nous avons retenu trois valeurs possibles pour b , $b \in \{1, 2, 3\}$. Nous souhaitons par là obtenir une large validation empirique de notre modèle sans toutefois assigner des valeurs « extrêmes » au paramètre relatif au degré d'incitation.

Le modèle postule qu'un agent Individualiste contribue exactement 7 jetons au bien public, ce qui correspond à la stratégie d'équilibre. Pour la caractérisation empirique nous relâchons légèrement cette hypothèse, en considérant comme Individualiste un sujet qui contribue un montant « proche » de 7. Cet « élargissement » du comportement Individualiste est nécessaire pour obtenir une taille relative de sujets considérés comme Individualistes significative. De plus, l'élargissement du zonage correspondant au comportement Individualiste autour de la contribution d'équilibre permet de tenir compte d'erreurs de faible amplitude de la part des sujets. En effet, il paraît raisonnable de supposer que des erreurs peu coûteuses, correspondantes à des contributions voisines de 7, puissent être commises par un sujet foncièrement individualiste.

Dans le modèle proposé, un agent Malveillant investit un montant fixe strictement inférieur à 7 jetons dans le bien public. Pour la caractérisation empirique nous considérons comme Malveillant un sujet qui investit $7 - a$ jetons et moins dans le bien public. Là encore, pour vérifier la robustesse des prédictions de notre modèle nous avons testé deux valeurs pour a , $a \in \{1, 2\}$.

La caractérisation du comportement Bienveillant est définie par défaut. Ainsi, ce comportement ne peut être caractérisé dans le cas où la contribution moyenne observée est inférieure à 7. En effet, nous considérons qu'un sujet ayant une contribution du même ordre que la contribution moyenne observée alors que cette contribution moyenne est inférieure à la contribution d'équilibre ne peut être caractérisé comme Bienveillant.

Nous avons donc retenu 6 jeux de paramètres afin de confirmer le plus largement possible la validation empirique de notre modèle: $a \in \{1, 2\}$ et $b \in \{1, 2, 3\}$. A titre d'exemple, sur le graphique 6, nous avons reporté le zonage correspondant à la caractérisation empirique des comportements pour le jeu de paramètres ($a = 2, b = 2$).



Graphique 6: Représentation des différents comportements pour le jeu de paramètres ($a = 2, b = 2$).

5.3.2 Résultats

Le tableau 4 résume les résultats obtenus pour la caractérisation empirique des comportements et ce pour l'ensemble des 6 jeux de paramètres possibles. Comme le montre le tableau 4, la prédiction théorique du modèle concernant les inégalités sur les tailles relatives des comportements dans les deux contextes est validée si l'on considère une taille relative moyenne sur les 6 jeux de paramètres pour chacun des comportements.

	Contexte négatif	Contexte positif
Malveillant (M)	29,61 %	10,71 %
Incitateur (IN = IR + IS)	27,08 %	40,08 %
Individualiste (I)	34,08 %	11,90 %
Bienveillant (B)	9,23%	37,30 %

Tableau 4: Tailles relatives moyennes des comportements dans les deux contextes.

Afin de tester la significativité des tailles relatives moyennes, nous avons comparé les tailles relatives des 6 groupes constitutifs du contexte positif aux tailles relatives des 6 groupes constitutifs du contexte négatif et ce pour chaque jeu de paramètres (pour chaque comportement, nous avons

établi une moyenne de sa taille relative dans chaque groupe sur les 14 dernières périodes du jeu expérimental). Nous avons donc pu tester les prédictions du modèle concernant les différences des niveaux relatifs de comportements entre le contexte positif et le contexte négatif pour 6 jeux de paramètres distincts. Les inégalités $B^+ > B^-$, $M^+ < M^-$, $I^+ < I^-$ et $IN^+ > IN^-$ ont été testées contre l'hypothèse nulle d'égalité des niveaux relatifs de comportements dans les deux contextes à l'aide du test U de Mann-Whitney, test non-paramétrique. A l'exception de l'inégalité $I^+ < I^-$ qui n'est pas significative pour les deux jeux de paramètres $\{a = 1, b = 1\}$ et $\{a = 1, b = 2\}$, les autres inégalités sont acceptées au seuil de 5% quel que soit le jeu de paramètres. Par ailleurs, l'inégalité $(M+I)^+ < (M+I)^-$, qui combinée avec les inégalités $IN^+ > IN^-$ et $M^+ < M^-$ est suffisante pour établir l'effet de contexte, est elle aussi acceptée au seuil de 5% quel que soit le jeu de paramètres. Ainsi, les inégalités sur les tailles relatives des comportements, ces inégalités étant selon notre modèle équivalentes à l'effet de contexte, sont statistiquement significatives. Nous pouvons donc affirmer que la prédiction théorique de notre modèle descriptif est empiriquement validée par les tests effectués sur nos données expérimentales.

6. Conclusion

L'apport de ce papier est double. D'une part, nous avons montré expérimentalement que le phénomène de sur-contribution à un bien public est spécifique au contexte positif, alors que dans le contexte négatif le niveau de contribution moyen n'est pas significativement différent du niveau de contribution d'équilibre. D'autre part, nous avons proposé un modèle de flux de comportements qui rend compte du différentiel observé de niveau de contribution entre les deux contextes. Le premier apport permet d'étendre le résultat antérieur de Andreoni [1995a] à un cadre plus général pour lequel l'équilibre de contribution n'est pas caractérisé par une contribution nulle, mais par un niveau de contribution individuel strictement positif.

Le modèle que nous avons développé est fondé sur l'idée centrale que les comportements des agents ne sont pas figés une fois pour toutes, mais sont susceptibles d'évoluer. Pour modéliser cette évolution, nous avons défini deux types de comportements de référence, le type Stratégique et le type Réciproque. Alors que le comportement de type Stratégique est fondamentalement en accord avec la logique du comportement égoïste de la théorie des jeux, le comportement de type Réciproque est mû par une logique différente. L'agent de type Réciproque se montrera bienveillant envers le groupe, s'il juge que le groupe s'est montré bienveillant à son égard en contribuant « suffisamment » au bien public. Dans le cas contraire, il pourra se montrer malveillant, en choisissant un niveau de contribution individuel suffisamment faible pour punir les autres, quitte à sacrifier une partie de ses gains potentiels. L'évolution des comportements dans ce modèle se fait à deux niveaux: à un niveau « inter-type » et à un niveau « intra-type ». Le changement inter-type est réalisé par mutation. Un agent de type Stratégique peut devenir de type Réciproque et vice versa, avec une probabilité qui est exogène au modèle. En revanche, une évolution intra-type est également possible, sous la forme de changements de comportements. Un agent de type Réciproque peut adopter un comportement Bienveillant, un comportement Malveillant ou un comportement Incitateur, selon le niveau ou l'évolution de la contribution moyenne des autres agents de la population. Le comportement Incitateur, qui consiste à contribuer plus que les autres, peut se justifier par la volonté de relancer la coopération lorsque celle-ci s'est dégradée. De la même façon un agent de type Stratégique peut adopter un comportement Individualiste ou un comportement Incitateur, mais dans ce dernier cas ses motivations se distinguent de celle de l'agent de type Réciproque. A l'état stationnaire de ce modèle, on peut alors identifier la prévalence des deux types d'agents et des différents comportements de ceux-ci dans les deux contextes. Nous montrons qu'il existe un lien entre les niveaux relatifs des différents comportements et le niveau moyen de contribution au bien public. Nous avons ainsi une explication en termes de comportement de l'effet de contexte. Les comportements, contrairement aux contributions, ne sont pas directement observables. Par contre, il est possible de les caractériser empiriquement, à partir de leur description théorique. Cette caractérisation empirique, que nous avons réalisée à partir de plusieurs paramétrisations, nous a permis de valider globalement le modèle.

Le modèle proposé peut bien sûr s'appliquer à d'autres expérimentations sur la contribution à un bien public. Mais il offre surtout l'avantage de proposer une représentation des comportements plausible pour ce type de jeux. Étant donné que le phénomène de sur-contribution, observé expérimentalement dans de très nombreuses expériences, reste en grande partie inexplicé, notre modèle contribue au développement d'un cadre explicatif permettant de rendre compte du phénomène de sur-contribution, en termes de comportements. Néanmoins, si l'idée de changements de comportements paraît séduisante, notre analyse reste essentiellement statique. Certes notre modélisation est capable d'expliquer le différentiel des niveaux de contribution mais pas leur évolution au cours du temps. Il est donc nécessaire, si l'on veut mieux comprendre le phénomène de sur-contribution, de proposer des analyses dynamiques.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANDREONI J. [1988], « Why free ride ? Strategies and learning in public goods experiments », *Journal of Public Economics*, 37, p. 291-304.
- [1990], « Impure Altruism and Donations to Public Goods : A theory of Warm-Glow giving ? », *Economic Journal*, 100, p. 464-477.
- [1995a], « Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments », *The Quarterly Journal of Economics*, 110, p. 1-17.
- [1995b], « Cooperation in Public-Goods Experiments: Kindness or Confusion ? », *American Economic Review*, sept. 1995, p.891-904.
- BOLTON G. [1991], « A Comparative Model of Bargaining : Theory and Evidence », *American Economic Review*, 81, p. 1096-1136.
- CROSON R. T. A. [1996], « Partners and Strangers revisited », *Economics Letters*, 53, p. 25-32.
- DAWES R., MCTAVISH J. ET SHAKLEE H. [1977], « Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation », *Journal of Personality and Social Psychology*, 35, p. 1-11.
- FEHR E., GÄCHTER S. ET KIRCHSTEIGER G. [1997], « Reciprocity as a contract enforcement device: Experimental evidence », *Econometrica*, 65, p. 833-860.
- ISAAC R. M., WALKER J. M. ET THOMAS S. H. [1984], « Divergent Evidence on Free Riding : An Experimental Examination of Possible Explanations », *Public Choice*, 43, p. 113-149.
- KAHNEMAN D. ET TVERSKY A. [1979], « PROSPECT THEORY: An Analysis of Decision Under Risk. », *Econometrica*, 47, p. 263-291.
- KESER C. [1996], « Voluntary Contributions to a Public Good when Partial Contribution is a Dominant Strategy », *Economics Letters*, 50, p. 359-366.
- KREPS D. M., MILGROM P., ROBERTS J. ET R. WILSON. [1982], « Rational cooperation in the finitely repeated prisoners' dilemma », *Journal of Economic Theory*, 27, p. 245-252.
- LEDYARD J. O. [1995], « Public Goods: A Survey of Experimental Research », dans KAGEL et ROTH, *The Handbook of Experimental Economics*, Princeton University Press, p. 111-194.
- LEVATI M. V. [1997], « An Experimental Study on the Role Played by Persuasive Behaviour in Public Goods Experiments », Miméo.

- MARWELL G. ET AMES R. [1979], « Experiments on the provision of public goods I: Resources, interest, group size, and the free-rider problem. », *American Journal of Sociology*, 84, p. 1335-1360.
- PALFREY T. ET PRISBREY J. [1997], « Anomalous Behavior in linear public goods experiments: how much and why ? », *American Economic Review*, 87, p. 829-846.
- RABIN M. [1993], « Incorporating Fairness into Game Theory and Economics », *American Economic Review*, Décembre 1993, p. 1281-1302.
- RATIMAGE [1995], « Research Assistance Toolbox for Computer-Aided Human Behavior Experiments », Klaus Abbink and Abdolkarim Sadrieh, University of Bonn.
- VERGNAUD J.-C. ET ZIEGELMEYER A. [1998], « Prospect Theory et « effet miroir »: une incapacité à expliquer l'effet de contexte », Miméo.

ANNEXE 1

i) La résolution du système constitué des équations (1) à (5) nous permet d'écrire la taille relative de chaque comportement à l'état stationnaire en fonction des différents taux de passage à l'état stationnaire. Ainsi,

$$M^* = \frac{K_1}{\frac{(m_B)^*}{(b_M)^*}(K_2 + K_3) + K_2}, B^* = \frac{K_1}{K_2 \left(\frac{(b_M)^*}{(m_B)^*} + 1 \right) + K_3},$$

$$I^* = \frac{K_1 \frac{r_S}{s_R}}{K_2 + K_3 \left(\frac{(m_B)^*}{(b_M)^* + (m_B)^*} \right)}, IS^* = \frac{K_1 \frac{r_S}{s_R} \frac{(i_{IS})^*}{(i_{SI})^*}}{K_2 + K_3 \left(\frac{(m_B)^*}{(b_M)^* + (m_B)^*} \right)},$$

$$IR^* = \frac{K_1 \frac{(b_{IR})^*}{(i_{r_B})^*}}{K_2 \left(\frac{(b_M)^*}{(m_B)^*} + 1 \right) + K_3} \text{ et } IN^* = 1 - \frac{K_1 \left(1 + \frac{r_S}{s_R} \right)}{K_2 + K_3 \left(\frac{1}{1 + \frac{(b_M)^*}{(m_B)^*}} \right)},$$

où $K_1 = s_R (i_{r_B})^* (i_{SI})^*$, $K_2 = (i_{r_B})^* (r_S (i_{IS})^* + (i_{SI})^* (r_S + s_R))$ et $K_3 = s_R (b_{IR})^* (i_{SI})^*$. Les hypothèses 1 et 2 nous permettent de conclure que si $(\bar{t}^*)^+ > (\bar{t}^*)^-$ alors $(M^*)^+ < (M^*)^-$, $(B^*)^+ > (B^*)^-$, $(I^*)^+ < (I^*)^-$ et $(IN^*)^+ > (IN^*)^-$. De plus, les inégalités suivantes sont vérifiées: $0 < IN^*$, M^* , B^* , I^* , IS^* , $IR^* < 1$ et $0 < (\bar{t}^*)^+$, $(\bar{t}^*)^- < 20$.

ii) L'équation (6) se réécrit $\bar{t}^* = 7 + \frac{b IN^* - a M^*}{M^* + I^*}$. De manière immédiate, nous déduisons des inégalités $(M^*)^+ < (M^*)^-$, $(B^*)^+ > (B^*)^-$, $(I^*)^+ < (I^*)^-$ et $(IN^*)^+ > (IN^*)^-$ l'inégalité $(\bar{t}^*)^+ > (\bar{t}^*)^-$.