



Le comportement coopératif à la lumière de deux conceptions de la moralité : devoir et choix

By/Par | **Alejandro Caparros¹, Jean-Christophe Pereau² et Tarik Tazdaït³**

RÉSUMÉ

Nous nous proposons dans ce papier d'étudier la question de la coopération à la lumière de la moralité. Cela nous conduit à discerner deux formes de moralité, celle conçue comme un devoir (impératif catégorique de Kant) et celle conçue comme un choix (l'altruisme). Néanmoins, ces deux représentations se recoupent dans leur traitement de la coopération. Dans la mesure où elles sont formalisées dans le cadre d'un jeu statique, les résultats qui en ressortent ne nous disent pas plus qu'il est dans l'intérêt d'un individu de s'y soumettre lorsque les autres en ont fait autant. Toutefois, il n'est pas expliqué comment les autres en sont venus à adopter un comportement moral. Cela soulève en substance le problème de l'outil utilisé, celui-ci définissant les comportements à l'équilibre sans nous livrer la moindre information sur la situation qui prédominait avant d'arriver à l'équilibre.

¹ Department of Economics, Institute of Economics and Geography (IEG), Spanish Council for Scientific Research (CSIC), Pinar 25, 28006 Madrid.

² Université de Marne La Vallée - O.E.P, Cité Descartes, 5 boulevard Descartes, Champs Sur Marne, 77454, Marne-La-Vallée cedex 2

³ CNRS - EHESS – CIRED, Jardin Tropical, 45 bis, avenue de la Belle Gabrielle, 94736 Nogent sur Marne cedex.

INTRODUCTION

La coopération au sein de l'action commune a constitué un sujet d'étude majeur⁴, et continue encore à alimenter la réflexion tant la logique de l'action individuelle semble être incompatible avec toute initiative collective. Précisons qu'il est très généralement admis que le type de situation auquel ont à faire face des agents confrontés au besoin de coopérer peut être représenté par des jeux de la famille du dilemme du prisonnier (DP)⁵. Il décrit une situation où chaque joueur a le choix entre deux stratégies : coopérer (C) ou trahir (T). L'histoire qui se profile derrière ce jeu est lié à l'arrestation de deux complices. La police ne dispose pas de preuves suffisantes pour identifier le responsable du méfait, c'est pourquoi le juge d'instruction a besoin que l'un des suspects témoigne contre l'autre. Il les interroge séparément de façon à confondre le responsable. Le marché qu'il propose à chacun est tel que la paix (cas de figure où les deux suspects coopèrent) est préférable à la guerre ouverte (cas de figure où les deux suspects se dénoncent mutuellement), mais où l'attaque surprise, c'est à dire que l'un parle pendant que l'autre se tait, est payante. Par conséquent, le conflit d'intérêt entre les joueurs est tel que la meilleure solution pour l'un est la pire pour l'autre, ce qui conduit à une solution non-coopérative stable (T,T), mais sous-optimale au sens de Pareto, car dominée par une solution coopérative (C,C) qui ne peut être atteinte de façon spontanée. Ce résultat permet de souligner l'opposition entre intérêt individuel et intérêt collectif, lorsqu'il y a indépendance des prises de décision.

De manière générale, si le DP se traduit par une solution inefficace c'est que *"pour un joueur qui n'est pas entièrement assuré des intentions pacifiques de son partenaire, l'usage de la stratégie agressive s'impose au nom des intérêts individuels"* (Moulin, 1981), et ce bien que l'intérêt collectif favorise une issue plus intéressante pour chacun des joueurs. Il est toujours possible d'objecter qu'en communiquant, les agents pourraient s'accorder sur l'issue coopérative. Certes, mais c'est oublier que rien ne les empêche de revenir sur leur parole.

Une autre façon d'entrevoir une sortie du paradoxe serait d'admettre l'intervention d'un tiers-arbitre à qui il reviendrait la mission de contraindre les individus à agir en fonction de l'intérêt commun. Il suffit, pour cela, de supposer qu'avant que le jeu ne prenne forme, les acteurs passent entre eux un contrat dont l'exécution serait contrôlée et rendue obligatoire par un tiers. Nous pouvons supposer que la "loi du milieu", qui punit les "délateurs", jouera ce rôle. En effet, si l'on songe qu'à notre sortie de prison, un tueur du gang nous attendra pour nous assassiner, la perspective d'une moindre durée d'emprisonnement ne saurait nous réjouir. Nous voyons dans cet exemple que la présence d'un tiers, ici une institution en marge du système légal, permet à nos prisonniers de trouver une issue préférable à celle auquel ils aboutiraient s'ils étaient livrés à eux-mêmes. Mais cette

⁴ Pour une synthèse introductive, voir Olson (1965).

⁵ Ce jeu a été introduit par Merrill Flood et Melvin Dresher au sein de la RAND Corporation en 1950 ou 1951, pour ensuite être nommé dilemme du prisonnier par Albert Tucker lors d'un séminaire de théorie des jeux organisé au département de psychologie de l'Université de Stanford (Straffin, 1980 ; Raiffa, 1992).

explication reste beaucoup trop partielle pour que l'on puisse s'en contenter : qui met en mouvement l'échelon hiérarchique ? La peur que les dénonciations s'étendent aux commanditaires ? On est en droit de s'interroger sur le statut du tiers, la manière dont il s'impose et surtout son motif d'intervention.

La situation idéale serait celle où les agents respecteraient spontanément la norme de coopération une fois celle-ci énoncée. Chacun pourrait en effet réaliser combien il serait plus raisonnable pour tout le monde de s'en remettre à une attitude coopérative plutôt que de faire primer l'intérêt personnel. L'argument censé conduire tout le monde à coopérer est classique. Lorsqu'un individu envisage de faire cavalier seul, il doit se demander : *"Et si tout le monde en faisait autant ?"*.

Soulignons que du point de vue des théories du contrat social (Machiavel, 1532 ; Hobbes, 1651 ; Locke, 1690 ; Rousseau, 1762) ce questionnement est totalement éclipsé. Comme le note à juste titre Rosanvallon (1989), les théories du contrat social ne font que déplacer le problème de l'interaction du niveau individuel au niveau inter-étatique :

— dans l'état de nature, il s'agit d'étudier l'instabilité des relations humaines et la forme de l'autorité permettant de la dépasser dans l'Etat, entendu comme un corps politique, une communauté unifiée dont l'une des principales attributions est la protection de ses membres vis-à-vis de toute menace extérieure ;

— mais, admettre la constitution contractuelle de sociétés civiles ne permet pas de régler la question des relations entre ces sociétés, qui demeurent inscrites, en l'absence d'une autorité qui leur serait supérieure, dans un état de nature similaire à celui dans lequel se trouvaient initialement les individus. L'interdépendance entre les sociétés civiles repose donc la question de l'interaction et permet la réaffirmation de la pensée stratégique.

De nombreux auteurs, pour leur part, ont défendu l'idée que le développement d'une moralité individuelle pouvait permettre la résolution des problèmes pour lesquels des actions individuellement rationnelles conduisent à une issue collectivement inefficace. L'analyse de la moralité en économie s'effectue en partie par le recours à des références philosophiques, à l'image de l'impératif catégorique de Kant (Laffont, 1975 ; Gauthier, 1986 ; Etzioni, 1987 ; Bordignon, 1990 ; Roemer, 1996, et bien d'autres). La morale y est alors décrite comme un devoir qui s'imposerait à tous. Toutefois la moralité ne se réduit pas au seul élément philosophique. Certains auteurs expliquent la coopération en supposant les individus altruistes (Becker, 1974, 1981). L'altruisme, sans être explicitement évoqué sous l'angle de la moralité, n'en est pas moins traité comme un problème de moralité : dans un cas comme dans l'autre les motivations se confondent aux comportements (Ballet, 2000), c'est pourquoi les travaux analytiques concernés se sont développés autour du comportement, délaissant la liaison entre motivations et comportement. L'altruisme est , en quelque sorte, perçu comme un choix moral.

C'est à une analyse critique de ce programme de recherche que nous consacrons cet article. Il s'agit, plus précisément, de montrer que l'analyse économique reste en mesure de s'enrichir de concepts provenant d'autres disciplines, à l'image de ceux empruntés à la

philosophie. Cependant cette interdisciplinarité reste confinée à une problématique bien définie : comment s'établit la coopération lorsqu'elle est appuyée par des considérations morales ? Comme nous le montrons, il est impossible d'en cerner le pourquoi, ou encore le processus d'émergence de la morale.

Le papier s'articule autour du plan suivant. La section (2) porte sur l'impératif catégorique de Kant et l'un de ses prolongements possibles qui consiste en la construction par Margolis (1982) d'un modèle où les individus seraient animés à la fois d'un intérêt individuel et de groupe. En section (3), nous revenons sur une autre forme de morale, celle du comportement altruiste, pour en souligner les différentes formes recensées tout en s'attachant à la représentation que l'économie en offre. La section (4) prolonge l'étude de l'altruisme compris comme un acte de générosité. Il s'agit en fait d'associer l'acte aux intentions qui le génèrent. Enfin, la section (5) dresse un bilan de cette analyse critique.

1. LE SENS DU DEVOIR MORAL

L'impératif catégorique de Kant : Un principe d'universalité

Comment se fait-il, se demande Laffont (1975), qu'au moins dans certains pays, les gens ne jettent pas négligemment leurs canettes de soda sur les plages et daignent parfois faire un détour pour les déposer dans une poubelle ? Il n'est pourtant pas dans leur intérêt immédiat d'adopter un tel comportement. Sans doute n'apprécieraient-ils pas de voir la plage où ils se promènent jonchée de déchets, mais l'économie d'un déplacement aussi pénible vers la poubelle vaut bien un accroissement marginal infime de la saleté de la plage. La contribution réelle de chaque individu à la dégradation du lieu est minime au regard du déplacement qu'il doit effectuer vers la poubelle. D'après les hypothèses standards il faudrait donc s'attendre à ce que les plages soient de véritables décharges à ciel ouvert. Or, au moins dans certains pays précise Laffont, il n'en est rien : les gens ne se comportent pas nécessairement comme le prévoit la théorie. *"La nature humaine est capable d'autres comportements que le schéma égoïste imputé à "l'homo oeconomicus" par la théorie économique"* (Laffont, 1975). En particulier, chacun peut prendre en compte les "contraintes macroéconomiques" liées à une certaine manière d'agir. Cela revient, selon l'auteur, à supposer que les individus sont guidés par un comportement moral conforme à "l'impératif catégorique" de Kant⁶. La règle de décision des agents s'en trouve modifiée. Chacun fait l'hypothèse que tous les autres se comportent comme lui, et sous cette nouvelle contrainte macroéconomique maximise sa

⁶ Laffont n'est pas le seul à qualifier de kantien ce type de règle de conduite. D'une manière générale, la littérature économique parle de principe kantien dès lors qu'un acteur suppose raisonnable de rendre son action soutenable (au regard de ses préférences) et compatible avec le fait que les autres pourraient adopter la même attitude que lui. C'est ce que font, par exemple, Collard (1978, 1983) et Harsanyi (1980) qui parle à ce sujet *"d'engagement rationnel"*. Sugden (1984), pour sa part, assimile cette attitude à un engagement inconditionnel.

fonction d'utilité. Le calcul aboutit à la conclusion assez évidente que chacun a intérêt à limiter ses actions lorsqu'elles impliquent une externalité négative.

Pour y voir plus clair, attachons nous à spécifier le concept "d'impératif catégorique". Dans son ouvrage *Fondements de la Métaphysique des Mœurs*, Kant (1798) défend le point de vue selon lequel il est possible de déterminer des principes moraux qui soient en mesure de gouverner notre conduite à l'égard des autres. Pour cela, chaque individu doit se mettre à la place des autres afin d'identifier ses devoirs envers eux et, a fortiori, envers lui-même. Etant indépendante des buts particuliers de chacun, l'éthique kantienne est de ce fait une éthique de la raison pure. Quel que soit l'objectif d'un individu, quel que soient ses désirs, ses penchants, s'il veut agir moralement, il lui faut agir en se conformant à l'impératif catégorique, c'est à dire en se prononçant suivant une règle de conduite qu'il souhaiterait voir s'imposer comme un principe universel : "*Agis uniquement d'après la maxime qui fait que tu peux vouloir en même temps qu'elle devienne une loi universelle*" (Kant, 1798). Un individu pourrait, par exemple, ne pas tenir ses promesses. Il en tirerait profit en trompant autrui. Mais l'individu en question ne peut considérer son attitude comme digne d'approbation. S'il le faisait, il ferait d'une loi universelle le fait de ne pas tenir ses promesses. La notion de promesse n'aurait plus aucun sens, et rien ne pourrait fonctionner dans le respect de tout un chacun. Ce même individu ne saurait plus comment vivre dans un tel monde. Une attitude n'est donc morale que si elle revêt un caractère d'universalité. Dans une certaine mesure, nous retrouvons ce type d'argument dans les résultats d'une expérience menée par Quattrone et Tversky (1986). En interrogeant un ensemble d'individus sur leur comportement électoral, ces auteurs aboutissent à une explication qui peut être résumée dans les termes suivants. Les individus pensent être représentatifs d'un certain groupe, c'est pourquoi ils en viennent à adopter un raisonnement allant dans ce sens : si j'agis de la façon dont j'agis, il y a alors de fortes chances que ceux qui me ressemblent en fassent autant. L'analogie avec Kant réside dans le principe d'universalité, à la différence près qu'il n'est pas motivé par le souci de rendre le comportement universel mais de considérer que les autres, vous ressemblant, choisiront la même voie : l'individu se voit comme agent universel, du moins pour le groupe auquel il se réfère. La dimension universelle ne touche plus à l'acte, comme chez Kant, mais à l'individu, ce qui est loin d'être la même chose car plus rien ne garantit la mise en œuvre du comportement coopératif plus qu'un autre.

Ce que l'on est en droit de se demander, dans le contexte révélé par Laffont, c'est pourquoi les individus adopteraient-ils spontanément cette règle de conduite kantienne ? Les hommes ont en commun la raison, précise Kant, c'est ce qui les rend identiques les uns aux autres. Or c'est à travers la raison qu'ils pensent et qu'ils arrivent aux mêmes résultats. Plus encore, en se mettant ensembles pour raisonner ils devraient aboutir aux mêmes conclusions, et ce quel que soit leur sexe, leur race ou leur origine sociale. La morale prend dans cette optique la forme d'un consensus entre des êtres doués de raison. Ce consensus fruit de la raison est ce que Kant appelle "*le règne des fins*". Si chacun se considère comme "*auteur d'une législation universelle*" en rapportant sa vision à chacune des autres visions et même à "*chacune des actions envers soi-même*", il en découle un système de devoir se fondant sur le principe de la liberté de choix de chacun. Les principes ainsi établis par des personnes morales guidés par le devoir seront alors des principes de justice et d'équité.

On parvient tout de même mal, à ce stade de la réflexion, à s'expliquer ce surgissement d'une préoccupation collective chez les agents, sinon dans leurs fins, du moins dans la limitation de l'usage de leurs moyens (puisqu'ils s'imposent une contrainte). En effet la mise en avant d'une morale kantienne n'a de véritable signification, pour nous, que si elle ressort d'une interaction stratégique entre les individus, ce qui est loin d'être le cas. Laffont définit cette morale en internalisant le comportement de groupe par l'introduction d'une contrainte d'ordre macroéconomique, qui s'exprime en fonction de la consommation agrégée du bien à l'origine de la pollution.

On peut toujours, à l'image de Sugden (1991), affirmer qu'aucun joueur ne désirerait que la trahison devienne une loi universelle. En effet, si un agent avait à choisir une règle qui s'appliquerait universellement, et si en faisant ce choix il est guidé par ses préférences, il paraît naturel qu'il choisisse une règle qui, si elle était suivie par chaque joueur, les conduirait à la coopération mutuelle. Un tel raisonnement, il est vrai, fournit à chacun des joueurs une raison de coopérer. Cette ligne de défense à l'appui de l'argument kantien est reprise d'Ulmann-Margalit (1977). Pourquoi une telle norme peut-elle entraîner l'adhésion des individus, se demande l'auteur ? La réponse proposée ne s'adresse pas simplement à la force efficace de l'argument de généralisation : *"But suppose everyone did the same?"*. Elle est censée valoir pour d'autres types de normes sociales. Ces normes sont générées, s'imposent et sont respectées parce qu'elles sont des solutions à des problèmes posés par certains types d'interactions sociales. C'est en ce sens qu'il faut comprendre la notion d'émergence, s'agissant des normes : l'argument kantien peut s'imposer et régler la conduite des personnes parce qu'il est parfaitement adapté à la situation qu'il est susceptible de gérer : *"Pour résumer nous nous sommes posés la question : en quoi consiste la force de persuasion de l'argument de généralisation, en dépit de son statut logique douteux ? Ma réponse est que cet argument est applicable aux situations structurées par un DP généralisé à ce type de situations seulement. Son pouvoir tient au fait qu'il pointe le véritable dilemme inhérent à ces situations et qu'il est un appel à des moyens de le résoudre, c'est à dire à des normes interdisant de faire l'action du type A (celle qui est collectivement désastreuse)"* (Ulmann-Margalit, 1977).

Cette explication souffre d'une importante limite : elle se fonde sur le seul socle fonctionnaliste. Si la règle kantienne s'impose aux acteurs, c'est parce qu'elle permet de résoudre, de manière adéquate, leur problème. Outre que cette sorte d'explication n'est jamais à même d'apporter une condition suffisante à l'institution d'une règle, elle se referme en un parfait diallèle dans le contexte présent. Le DP est décrit ici comme une situation émettant un appel fort pour qu'intervienne une norme régulatrice, et la norme kantienne s'impose parce qu'elle répond à cet appel. Ulmann-Margalit complète son raisonnement par le recours à une argumentation évolutionniste : *"Je propose, ajoute-t-il, de considérer l'argument soutenant ce livre, dans un sens importé et quelque peu métaphorique, comme une théorie de la sélection naturelle du développement des normes"*. Cette juxtaposition d'arguments de type économique/fonctionnaliste avec des institutions mal assurées de type évolutionniste se retrouve également chez Schotter (1981). Bien que l'outil d'analyse de *The Economic Theory of Social Institutions* soit exclusivement la théorie des jeux, une large place est accordée en introduction à une argumentation de type

darwinienne. En effet, dans la vision de cet auteur les systèmes économiques et sociaux évoluent comme le font les espèces, pour assurer leur survie et leur croissance. Chaque problème économique lié à cette évolution requiert une institution sociale pour la résoudre, et les sociétés qui arrivent à créer l'ensemble adéquat d'institutions sociales survivront et prospéreront. En revanche, celles qui n'y arriveront pas, seront amenées à périr. En d'autres termes, les institutions sociales sont nos outils adaptatifs et on ne peut survivre sans elles.

On constate à travers nos commentaires que ces développements font figure de solution alternative lorsque l'approche centrale est jugée implicitement défailante pour rationaliser l'émergence effective du comportement coopératif ainsi que sa généralisation. Cependant, le rôle que l'on entend faire jouer à l'évolution sélective n'est pas toujours le même. Chez Ulmann-Margalit, il s'agit seulement de donner une indication sur le "comment la norme est arrivée là", tandis que chez Schotter le problème semble être plus profond : il s'agit de compléter la méthode analytique, qui ne procure que des arguments en faveur de la stabilité locale, par un évolutionnisme intuitif censé fournir les arguments en faveur de la stabilité globale. Cela ne fait que relancer un peu plus le débat car il nous faut bien fournir une explication à l'émergence de la coopération en tant que résultat direct de la rationalité des agents.

L'arbitrage intérêt de groupe / intérêt individuel

Une autre façon d'appréhender cette question serait de modifier le modèle du choix rationnel afin d'y intégrer un arbitrage entre intérêt de groupe et intérêt individuel, cette possibilité d'arbitrage pouvant faire ressortir l'acte coopératif comme issue d'équilibre. C'est la voie empruntée par Margolis (1982).

Selon cet auteur, en réduisant le comportement de l'agent à celui de l'intérêt individuel, la théorie du choix rationnel a été conduite à négliger l'entité à laquelle elle se réfère. Chaque fois qu'un agent est pris comme unité de revenu et de consommation, cela revient implicitement à supposer que l'agent est cette entité dont l'intérêt est d'être maximisé. L'attrait majeur du modèle de choix rationnel réside dans son aspect pratique. En effet, sa simplicité a le mérite d'apporter plus par les développements théoriques qu'elle autorise, qu'elle ne retire par son caractère simplificateur. Il n'en reste pas moins que Margolis le considère sans fondement théorique car en pratique les individus agissent comme si leur intérêt se référait à des entités plus grandes qu'eux-mêmes. Ils apparaissent agir, non pas en fonction de leur propre intérêt, mais en fonction de l'intérêt d'une collectivité ; ce serait à cette seule condition que les bases d'une société pourraient apparaître.

Ainsi, en chaque agent il y aurait deux individus : un S-individu qui valorise l'intérêt individuel, et un G-individu qui valorise l'intérêt de groupe. Une règle d'allocation des ressources censée assurer la cohérence du comportement doit être également introduite. Elle repose sur deux principes dont l'énoncé est le suivant : plus le montant de ressources que j'ai dépensé à des fins non égoïste est élevé, plus le poids que j'accorde à mon intérêt personnel en allouant une part marginale des ressources est important. En outre, plus élevé

est le bénéfice que je confère au groupe comparativement à celui qui résulte d'une dépense marginale des ressources à mes propres fins, plus je tendrai à agir de façon altruiste.

Le premier principe composant cette règle porte sur l'importance de la participation de l'agent dans une activité d'ordre social, tandis que le second concerne la valeur de l'effort marginal de l'agent dans la poursuite de l'intérêt de groupe.

Afin d'en éclaircir la portée, introduisons quelques notations. Soit S' (resp. G') l'utilité marginale associée à l'intérêt individuel (resp. de groupe). De même, soit s (resp. g) le montant qui a été alloué au S-individu (resp. G-individu), c'est à dire le montant qui aura été dépensé pour maximiser l'utilité S (resp. G). On appellera g/s le ratio de participation, celui-ci donnant les proportions dans lesquelles l'individu a réparti ses ressources. Enfin, soit W l'importance que l'agent accorde à l'utilité S , compte tenu que cela résulte d'un mécanisme interne à l'individu et dont il est complètement inconscient. Etant donné ces conventions, la règle d'allocation exige que :

1- plus le ratio G'/S' est élevé, plus grande doit être la tendance pour l'agent d'allouer une unité de ressource marginale au G-individu plutôt qu'au S-individu;

2- la fonction W doit varier positivement avec le ratio de participation, g/s , dans la mesure où la probabilité que l'agent allouera une unité de ressource marginale au S-individu plutôt qu'au G-individu doit décroître avec le ratio de participation;

3- l'agent sera à l'équilibre lorsque la propension à favoriser l'intérêt individuel compense légèrement la propension à agir dans l'intérêt du groupe.

Formellement, l'agent est en situation d'équilibre si et seulement si :

$$W = G'/S' ,$$

$$\text{où : } W = f(g/s) ; W^{\circ} \geq 1, W' > 0 .$$

Cette dernière équation stipule que dès l'instant où l'agent a dépensé une somme dans l'intérêt de groupe (i.e. aussi longtemps que $g = 0$; il s'ensuit : $g/s = 0$), alors $W = W^{\circ}$. Ainsi, dans le cas le plus simple où $W^{\circ} = 1$, il ne donne pas plus d'importance à son intérêt personnel qu'à l'intérêt de groupe. Cependant, tout accroissement dans le ratio de participation, g/s , accroît W :

$$W' = dW/d(g/s) > 0 .$$

Tant que l'hypothèse de diminution de l'utilité marginale tient, alors si l'agent se trouve hors équilibre avec $W > G'/S'$, il s'en rapprochera en modifiant son allocation des ressources d'intérêt de groupe en intérêt individuel (diminuant g , augmentant s). En effet, en modifiant les ressources de g vers s , g/s diminue entraînant simultanément une diminution de W et un accroissement de G'/S' . En revanche, si l'agent se trouve hors équilibre avec $W < G'/S'$, nous prêtons le raisonnement inverse. La figure (1) illustre la

situation d'équilibre. La courbe en pointillée décrit l'augmentation de W selon l'accroissement de la part du revenu alloué à l'intérêt de groupe. La courbe en gras décrit le ratio G'/S' , dont la pente est négative (i.e $S'' < 0$, $G'' \leq 0$). L'intersection des deux courbes donne l'allocation d'équilibre, défini en E .

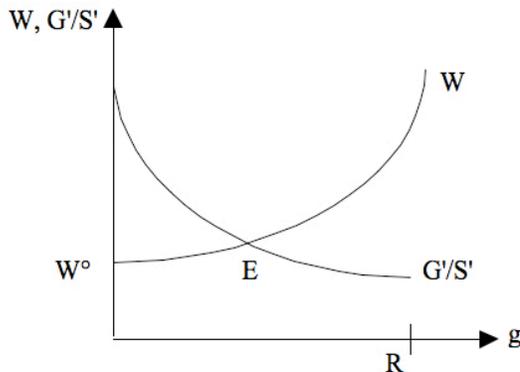


figure 1

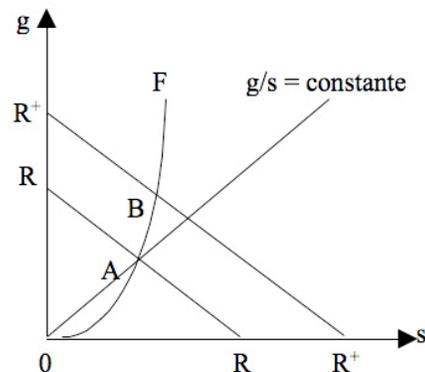


figure 2

La figure (2) illustre une propriété importante de l'équilibre dans la mesure où elle décrit l'ensemble des allocations d'équilibre lorsque l'agent observe un accroissement de son revenu. L'allocation relative au G-individu (g) est représentée sur l'axe vertical et celle relative au S-individu (s) l'est sur l'axe horizontal. La somme de g et s est égale à R et désigne le revenu de l'agent. L'agent étant soumis à sa contrainte de revenu, ses allocations possibles sont données par la droite RR . Lorsque nous nous déplaçons au nord-ouest le long de RR (ce qui correspond sur la figure (1) à se déplacer vers la droite), W augmente tandis que G'/S' diminue. Quelque part le long de cette droite nous devons aboutir à l'unique point d'équilibre tel que $W = G'/S'$. Ce point se situe en A sur (2).

Traçons maintenant un vecteur partant de l'origine et passant par A . Celui-ci nous donne l'ensemble des points pour lesquels le ratio de participation, g/s , et donc W , est le même qu'en A : la droite est donc relative à W constant. Lorsque l'on se déplace le long de cette droite, le ratio G'/S' croît (S' décroissant et G' restant constant). A se trouve alors être le seul point où l'agent est à l'équilibre tout comme il est le seul point le long de la droite associée à la contrainte budgétaire.

Pour l'ensemble des points le long de la droite à W constant qui sont plus proche de l'origine que de A , nous devons avoir $W > G'/S'$: G' sera le même qu'en A ainsi que W , mais S' sera plus important. Pour ceux qui sont plus éloignés de l'origine que de A , la condition sera inversée. Concrètement, cela signifie que pour un revenu inférieur à R , le point d'équilibre doit se trouver au sud-est de la droite relative à W constant ; en revanche, pour un revenu supérieur à R , le point d'équilibre se situera au nord-ouest de la même droite.

Ce raisonnement s'appliquant à chacun des points d'équilibre, le chemin revenu- dépenses d'équilibre de l'agent doit avoir une forme inclinée vers le haut à l'image de la courbe OF en figure (2). Ainsi, pour un revenu $R^+ > R$, l'allocation d'équilibre de l'agent doit passer par un point, ici B, au nord-ouest de l'intersection entre la droite à W constant et la contrainte budgétaire élargie $R+R^+$. A un niveau de revenu encore plus important, le nouvel équilibre passerait au nord-ouest de la ligne partant de l'origine et passant par B, et ainsi de suite.

La règle d'allocation introduite par Margolis implique donc que les dépenses allant dans le sens de l'intérêt de groupe sont un bien supérieur : plus le revenu augmente plus la fraction des ressources allouées au G-individu augmente.

Que nous apprend ce modèle ? Considérons la situation pour laquelle l'agent détermine une unique allocation entre les dépenses du S- et du G-individu. Plus son revenu s'élève, plus son intérêt personnel peut être satisfait pour une fraction donnée du revenu affectée au S- individu. D'autre part, pour l'agent les dépenses d'un autre membre de la société n'auront pas d'effet perceptible sur les opportunités de dépenses agrégées dans l'intérêt de groupe. Ainsi, dans le cas statique, S' diminue avec un accroissement des ressources allouées au S-individu, G' restant constant. Il en résulte qu'une hausse du revenu suppose que le ratio de participation de l'agent, g/s , doit augmenter. Dans la mesure où le revenu croît, la part du revenu allouée à l'intérêt de groupe croît également même si le poids de l'intérêt personnel évolue dans un sens identique.

Mais, bien qu'il existe un large éventail de situations où l'hypothèse d'un contexte statique reste une hypothèse raisonnable, l'inverse est tout aussi vrai, c'est-à-dire qu'il existe un large éventail de situations où elle se trouve être empiriquement indéfendable.

Pendant longtemps, il a semblé que le groupe auquel l'espèce humaine appartenait coulait de source. Mais au fur et à mesure que les sociétés humaines se sont organisées, la question de l'identité du groupe est devenu plus complexe. Cela est encore plus vrai dans le monde moderne où différentes formes d'associations peuvent être envisagées : en termes de parentés, de traditions, de religions. A travers la nature et les déterminants de l'identité de groupe, nous abordons des thèmes propres à la sociologie. Cela constitue l'un des aspects les plus intéressants du modèle : fournir un lien potentiel entre les deux modes d'analyse que sont l'économie et la sociologie.

Adoptons maintenant ce schéma au problème des biens publics⁷ afin d'en étudier le pouvoir explicatif. Tout d'abord, il est important de rappeler que la théorie du choix rationnel n'a pas encore été en mesure de produire une véritable définition de la demande en biens publics qui soit cohérente avec les observations empiriques. Selon ces dernières, les agents ont des préférences pour les biens publics, ce qui reste difficilement explicable en termes d'intérêt individuel. Une multitude de travaux en ont été

⁷ Un bien public est un bien dont la consommation (ou l'utilisation) par un agent n'empêche pas d'autres agents de le consommer (ou l'utiliser) simultanément (Bénard, 1985). On dit aussi qu'il y a non rivalité entre les individus à propos de leur consommation.

consacrés (Cornes et Sandler, 1996). Plus simplement, la littérature retient comme cadre de référence une approche qui consiste à spécifier des préférences aux formes restreintes.

Margolis, pour sa part, soutient que son modèle est suffisamment général pour définir des demandes en biens publics. Pourtant il est facile de vérifier sur la base de celui-ci que la S-utilité de l'agent perd toute sa pertinence lorsqu'il en vient à déterminer sa contribution en bien public. Dans la mesure où sa S-utilité le conduira à adopter un comportement de passager clandestin, il ne contribuera au financement d'aucun bien public. Toute contribution à laquelle il se soumettra découlera donc des ressources allouées à sa G-utilité. Mais sa G-utilité l'amène à privilégier le bien-être social. En conséquence, il devra allouer ses ressources entre les biens publics et les autres activités favorables au groupe. Selon la logique de cette théorie, il choisira le bien qui génère le bien-être social le plus élevé pour une contribution marginale. Mais rien ne prouve que cela profitera nécessairement au financement d'un bien public. Nous pouvons même ajouter que l'activité la plus à même de répondre à ce comportement est la charité. En considérant un individu motivé par le bien-être des autres agents, Margolis a en fait formulé plus une théorie de l'altruisme que de la coopération.

Cette limite ne doit pas faire oublier que l'arbitrage entre S- et G-individu a le mérite d'introduire une représentation de l'agent qui se veut plus proche de la réalité.

En effet, le cadre retenu conduit Margolis à remarquer que nous avons en l'agent un G-individu qui semble avoir les propriétés du "*surmoi*" de Freud, un S-individu qui ressemble au "*ça*", et une troisième entité qui s'interpose entre les deux et qui semble avoir la fonction intégrante si essentielle à la notion freudienne du "*moi*". Cette référence à la psychologie est d'autant plus importante qu'elle montre combien cette discipline est susceptible d'enrichir la réflexion économique⁸, notamment par les possibilités d'expérimentations qu'elle permet. C'est d'ailleurs sur ce principe que se sont développés les jeux expérimentaux.

2. LE SENS DU CHOIX MORAL

Différentes approches de l'altruisme

La prise en compte dans l'analyse économique de l'impératif catégorique de Kant pose problème. Ce principe correspondant à un devoir, la préoccupation morale de l'individu se réduit à une contrainte qui limite ses choix et ne s'analyse pas comme un choix en tant que tel. Une explication des conditions qui conduisent un individu à se conformer délibérément à une contrainte morale doit établir que celle-ci est choisie par un individu dépourvu pour les intérêts d'autrui. Cela ne signifie aucunement que l'individu ne prenne pas en

⁸ Sur ce thème citons l'excellente synthèse de Rabin (1998).

considération les intérêts d'autrui, il ne le fait qu'au nom de son intérêt propre. De ce point de vue l'approche de Margolis ne peut elle aussi être retenue.

Le comportement altruiste, en revanche, vérifie la condition précédente puisqu'il consiste en une action ; action qui résulterait du sentiment de sympathie que l'individu éprouve pour autrui. Le sentiment de sympathie et l'altruisme ont été évoqués de façon concomitante par Smith (1759) dans *Théorie des Sentiments Moraux*. Cet ouvrage recense une large gamme de sentiments propres à l'homme et la sympathie, comme les autres, constitue un mobile d'ordre non économique qui induirait un besoin accru de bien-être. En effet, à l'image d'un certain nombre d'auteurs du 17^{ème} siècle, Smith reconnaît que l'estime de soi est l'un des traits fondamentaux de l'homme. Elle le pousse à rechercher considération et approbation d'autrui, c'est pourquoi elle ne peut qu'être orientée vers des objectifs matériels : une meilleure réussite économique s'accompagnant pour l'individu de l'approbation et l'admiration de ses semblables. L'altruisme, sous-jacent à la sympathie, se concilierait donc avec l'intérêt individuel. Pour reprendre Maître (2000), chez Smith "*l'intérêt individuel comprend le self-love [l'amour de soi] et la sympathie*". Les deux œuvres maîtresses de cet auteur, *Théorie des Sentiments Moraux* et *Recherches sur la Nature et les Causes de la Richesse des Nations*, se complètent en ce sens que dans la première Smith s'intéressait aux ressorts non économiques du comportement individuel, tandis que dans la deuxième il examinait les conditions sous lesquelles le bien-être individuel pouvait être atteint, attestant ainsi que les motivations non économiques jouent un rôle moteur dans la recherche d'une amélioration des conditions matérielles.

Auguste Comte (1852) a développé une représentation alternative dans *Catéchisme Positiviste*. Frappé par la misère et l'oppression des hommes, Comte s'est alors engagé à œuvrer pour le bien de l'humanité en cherchant à penser une nouvelle forme d'organisation de la société. C'est dans ce contexte qu'il définit l'altruisme comme la capacité "*de vivre pour autrui*", cette attitude permettant aux hommes d'être en rapport les uns avec les autres et de travailler à l'amélioration de leur sort⁹. L'altruisme est évoqué en réaction aux transformations sociales et naturelles induites par la révolution industrielle ; son introduction s'insère dans l'élaboration de la nouvelle religion préconisée par Comte, la "*religion de l'humanité*". Bien que la réorganisation sociale escomptée se doit d'être politique, elle ne réussira que si elle s'opère au niveau de l'individu : le renouvellement proposé touche donc l'individu et les institutions dans lesquelles il évolue. Il ne s'agit pas de comprendre les ressorts du comportement coopératif tel qu'il est observé dans les faits, mais de concevoir une morale qui pourrait voir le jour si l'altruisme était privilégié : l'altruisme deviendrait obligation au même titre que l'impératif de Kant. Nous retrouvons ici la limite liée à tout raisonnement normatif. En déterminant les principes auxquels les individus devraient se soumettre pour améliorer leur situation, Comte fait, en grande partie, abstraction de la réalité. Le prince anarchiste Kropotkine (1902) ouvre, pour sa part, une troisième voie. En procédant d'une étude historique dans *L'Entraide, un Facteur*

⁹ Comme le précise Pierre Arnould en préface de l'édition de 1966 de *Catéchisme Positiviste* : "*Comte [...] ouvre à la morale des perspectives immenses par la formule : vivre pour autrui, afin de survivre par et dans autrui*".

d'*Evolution*, cet auteur examine la part qu'a eu l'entraide¹⁰ dans l'évolution progressive de l'humanité. A la suite du succès de l'ouvrage de Darwin (1859) *l'Origine des Espèces*, il devenait commun, pour certains de ses contemporains, de parler de "*l'âpre et impitoyable lutte pour la vie*" soutenue par chaque animal contre tous les autres animaux mais aussi par chaque "*sauvage*" contre tous les autres "*sauvages*", et par chaque homme civilisé contre tous ses concitoyens¹¹. Or dans le monde animal, la solidarité tient une place beaucoup plus significative que ne le laisse entendre "*la lutte pour la vie*". Kropotkine constate, par exemple, que les problèmes de surpopulation face aux ressources trouvent souvent leur issue dans des alternatives autres que le conflit. La migration d'une partie du groupe constitue l'une de ces options¹². En d'autres termes, la vie en société s'avère être la meilleure arme dans la "*lutte pour la vie*" comprise au sens d'une lutte contre toutes les conditions naturelles défavorables à l'espèce, et non comme une lutte pour acquérir les moyens de subsistance nécessaires à la vie. La protection mutuelle obtenue par le biais de la vie en société garantit la conservation de l'espèce et son développement : les espèces qui ont réduit la lutte individuelle à ses limites les plus étroites sont devenues à la fois les plus nombreuses et les plus prospères. S'intéressant aux tribus sibériennes, aux insulaires polynésiens, aux corporations médiévales et à la société industrielle naissante, Kropotkine retrouve cet aspect chez l'homme. Dès l'âge de pierre, l'entraide s'est imposée comme une nécessité, aussi bien de survivance que de progrès, conduisant les hommes à vivre en clans, en tribus. Ils ont ainsi développé un grand nombre des institutions sociales qui détermineront plus tard les lignes principales du progrès. Ce fut le cas avec la chasse en commun, la défense en commun ou la possession en commun d'un territoire. Quand de nouveaux besoins se sont affirmés, alors sont apparus les communes, les villages, les cités et les Etats. De même, le succès d'une guerre est proportionné à l'influence de l'entraide au sein des nations, cités ou tribus qui entrent en conflit. Point qui n'est pas souvent mentionné mais Darwin (1871) avait défendu une position analogue dans *La Descendance de l'Homme*¹³. Il soulignait que l'homme est un être moral en ce sens qu'il est capable de comparer ses actes passés ou futurs, et de les approuver ou de les désapprouver. Au fur et à mesure que l'homme avançait en civilisation, les tribus en sont

¹⁰ Dans *La Morale Anarchiste*, Kropotkine (1891) parle indifféremment de *solidarité et d'altruisme* en opposition à l'égoïsme, alors que dans *l'Entraide, un Facteur d'Evolution*, il opte pour le terme d'entraide en opposition à l'égoïsme.

¹¹ Dans une lettre adressée à Engels le 18 juin 1862, Marx décrivait la théorie darwinienne en ces termes : "*Il est curieux de voir comment Darwin retrouve chez les bêtes et les végétaux sa société anglaise avec la division du travail, la concurrence, l'ouverture de nouveaux marchés, les inventions, et la lutte pour la vie de Malthus. C'est le bellum omnium contra omnes [la guerre de tous contre tous] de Hobbes*". Dans un courrier daté du 12 novembre 1875 et adressé à Lavrov, Engels ajoutait : "*L'interaction des corps naturels -vivants et morts- implique aussi bien l'harmonie que le conflit, aussi bien la lutte que la coopération. Si, par conséquent, un soi- disant naturaliste se permet de résumer toute la richesse, toute la variété de l'évolution historique en une formule étroite et unilatérale, celle de la lutte pour la vie, formule qui ne peut être admise même dans le domaine de la nature qu'avec quelques réserves, ce procédé contient déjà sa propre condamnation*".

¹² Cette stratégie est celle adoptée par les castors.

¹³ Kropotkine (1891) écrivait à ce sujet : "*N'en déplaise aux vulgarisateurs de Darwin, ignorant chez lui tout ce qu'il n'avait pas emprunté à Malthus, le sentiment de solidarité est le trait dominant de la vie de tous les animaux qui vivent en sociétés*".

venues à se réunir en communautés plus importantes, et la raison a poussé chaque individu à étendre ses instincts sociaux à l'ensemble des membres de la nation, puis à tous les hommes de toutes les nations et de toutes les races.

Kropotkine complète cette appréciation en ajoutant que l'instinct d'entraide transmis à l'humanité par les héritages d'une très longue évolution est tel que le développement de l'Etat sous sa forme la plus autoritaire n'a pu briser les institutions protectrices d'entraide. En se fondant sur de "*vagues agrégations d'individus*", l'Etat se donne pour objectif de parler d'une unique voix au nom de tous. Pour être le seul lien d'union des individus, il lui faut préserver la prédominance du "*moi individuel*" au détriment du collectif, effaçant par là même le facteur d'entraide. La Rome Impériale, par exemple, a adopté cette attitude en cherchant, notamment, à gommer les particularités culturelles garantissant l'unité des différents peuples de l'empire. C'est toutefois l'affirmation de l'entraide qui poussa les tribus barbares à combattre Rome, plus pour préserver leurs institutions d'entraide que pour la conquérir. N'est-ce pas également le besoin d'entraide qui a trouvé refuge dans les associations secrètes d'ouvriers du 19^{ème} siècle alors que l'Etat leur refusait le droit d'être ? Aujourd'hui l'entraide se manifeste encore dans la multitude d'associations qui englobent toutes les manifestations de la vie sociale.

L'auteur ne nie nullement l'importance de la revendication du "*moi individuel*" dans la communauté, mais soutient l'argument que le progrès ne peut être attribué au triomphe de l'individualisme et de la concurrence. Tout progrès résulte de la capacité de l'entraide à contenir l'affirmation du "*moi individuel*". Selye (1974) ne rejoint qu'en partie cette conclusion, et réactualise par là même l'argument de Smith. Pour ce chercheur en médecine, chez tous les êtres vivants s'occuper de soi est un instinct tout ce qu'il y a de plus naturel, et aucun principe moral ne peut modifier cet état de faits. Mais faire du bien aux autres est parti intégrante de l'égoïsme, car les hommes sont des êtres sociaux qui ont besoin de respect et de gratitude : personne ne peut vouloir du mal à quelqu'un qui lui est utile, c'est là une condition nécessaire à la sécurité de chacun dans la vie en société. Sur le strict plan biologique, il n'existerait donc pas d'opposition entre égoïsme et altruisme. Cette observation en vient à suggérer la pertinence d'un "*égoïsme altruiste*".

Bien que la représentation d'Adam Smith a été celle que les économistes ont retenu, il a fallu attendre les travaux de Marglin (1963), Kolm (1966), De Alessi (1969), Hochman et Rodgers (1969) pour qu'une conceptualisation analytique soit proposée. De ces travaux, il est ressorti une définition neutralisée (Koulibaly, 1998) : l'analyse économique ne s'est pas attachée à expliquer ce qui motive l'individu altruiste, mais plutôt à conceptualiser l'altruisme sur une base comportementale. Ainsi un individu altruiste est décrit comme incorporant l'utilité d'autrui dans sa propre utilité en la considérant comme un bien ; on dit que l'altruisme se caractérise par une interdépendance de l'utilité. Plus simplement, l'utilité d'un altruiste dépend positivement de celle des autres.

La conceptualisation offerte par l'économie

Afin de cerner les effets de l'altruisme, Becker (1981) a appliqué la définition neutralisée au problème de l'allocation des ressources au sein de la famille. L'altruisme comme élément explicatif des relations intrafamiliales avait été introduit sous forme d'intuition par Smith (1776) : *"après lui même, [tout homme] consacre son affection aux membres de sa famille, à ceux qui vivent dans la même maison que lui, ses parents, ses enfants, ses frères et ses sœurs. Ils sont naturellement les personnes dont le bonheur ou le malheur dépendent de sa conduite"*. Mais il n'a pas considéré ce qui pouvait se passer lorsque les comportements altruiste et égoïste se trouvent confrontés l'un à l'autre, aspect que Becker s'est proposé d'expliquer. Son analyse va nous servir de point d'appui pour l'étude des spécificités du concept.

Pour que l'altruisme prenne tout son sens, l'individu altruiste doit disposer d'une parfaite connaissance des fonctions d'utilité d'autrui. Cela nécessite des interactions fréquentes entre les agents, chose qui semble plus réaliste au sein de la famille qu'en dehors. Dans une certaine mesure Steedman (1995) confirme ce point de vue en reprenant l'exemple suggéré par Wicksteed. Ce dernier affirmait que cela n'avait aucun sens de soutenir qu'une ménagère adoptait une attitude égoïste en effectuant ses achats sur la place du marché, pour ensuite se comporter de façon altruiste en les servant dans son foyer. Il existerait selon Wicksteed, un paradoxe entre concevoir un égoïsme sur le marché et un altruisme au sein de la famille. Il y a bien une opposition entre les deux attitudes mais pas au niveau où le situe Wicksteed. Pour Steedman (1995), l'opposition porte sur la nature des relations économiques qui caractérisent les situations de marché et familiales, ainsi que sur leurs motifs : la relation marchande ne s'appuie pas sur la sympathie tandis que celle-ci s'exprime plus facilement avec les personnes que l'on connaît le mieux, que ce soit les membres de la famille, les amis ou les proches voisins.

Ces précisions faites considérons, dans le prolongement de Becker (1981), une famille à deux personnes dont l'un des membres est altruiste. Formellement, l'altruisme du membre 1 signifie que sa fonction d'utilité s'écrit :

$$U_1 = U(z_1, \varphi(U_2)), \quad \text{avec } \frac{\partial U_1}{\partial U_2} > 0,$$

où U_1 et U_2 représentent les niveaux d'utilité respectifs des membres du couple, φ est une fonction positive de U_2 et z_1 (resp. z_2) le montant de consommation en bien z supporté par 1 (resp. 2). Dans la mesure où l'altruiste consacre une partie y de son revenu (noté R_1) en faveur du second membre, sa contrainte budgétaire a pour forme :

$$z_1 + y = R_1.$$

Compte tenu du supplément de revenu dont il bénéficie, l'agent 2 a alors une consommation supérieure à celle de son propre revenu, R_2 :

$$z_2 = R_2 + y.$$

En tirant y de la dernière équation pour reporter son expression dans la précédente équation, nous obtenons :

$$z_1 + z_2 = R_1 + R_2 = S_1,$$

S_1 désignant le revenu familial de l'altruiste. L'allocation optimale des ressources déterminé par ce dernier est déduite de la maximisation de son utilité sous la contrainte de revenu familial. Il en découle la condition d'équilibre :

$$\frac{\partial U/\partial z_1}{\partial U/\partial z_2} = 1.$$

De celle-ci se déduit les fonctions de demande de 1 pour z_1 et z_2 :

$$z_1 = z_1(S_1) \quad \text{et} \quad z_2 = z_2(S_2),$$

où $dz_i/dS_i > 0$, pour $i = 1, 2$, si 1 souhaite accroître z_1 et z_2 quand son revenu familial s'accroît, et si $y > 0$. Ainsi la consommation des deux membres de la famille dépend du revenu familial, et non pas seulement du revenu propre à chacun d'eux.

Dans la mesure où le revenu familial est égal à la somme des revenus de l'altruiste et du bénéficiaire, l'altruiste ne s'engagera pas dans une action qui diminuerait davantage le revenu de son conjoint qu'elle n'augmenterait le sien. Mais il s'engagera dans une action qui diminuera son revenu moins qu'elle n'augmentera celui de son conjoint. Notons que si le bénéficiaire adoptait un comportement purement égoïste, cela ne le pousserait pas pour autant à s'engager dans une action qui nuirait à son bienfaiteur. Supposons, en effet, qu'une telle action augmente le revenu du bénéficiaire tout en réduisant encore plus celui de l'altruiste. L'altruiste, tenant compte de la réduction constatée de son revenu familial, réduira sa contribution plus que ce que gagnerait le bénéficiaire. En d'autres termes, il ne faut pas oublier que c'est l'altruiste qui détermine la consommation totale du bénéficiaire ; et aussi égoïste que soit le bénéficiaire, il n'en internalise pas moins le bénéfice qu'il peut retirer de son bienfaiteur dans le choix de ses propres actions. Tout se passe comme si la présence d'un membre altruiste dans le couple poussait l'autre membre à maximiser aussi le revenu familial de l'altruiste. C'est ce que Becker (1974) appelle le théorème du "*rotten kid*"¹⁴. Par conséquent, chacun des membres d'une famille altruiste se trouve protégé contre les aléas de l'incertitude. Toute modification extérieure susceptible de détériorer la situation de l'un des

¹⁴ Ou encore le théorème de l'enfant gâté.

membres sera alors compensé par les autres membres : une redistribution des ressources sera observée, de sorte que chacun supporte une part du désagrément.

En élargissant le panel des caractéristiques relatives à autrui dans les préférences d'un agent, il devient possible de décrire aussi bien le comportement altruiste que celui d'envie (Rabin, 1998). La fonction d'utilité globale retenue aurait pour expression :

$$U_1(z_1, z_2) = U_1(z_1) + r_1 U_2(z_2).$$

Suivant les valeurs de r_1 , différents comportements peuvent être discernés. Si $r_1 > 0$, l'agent se réjouit de l'accroissement de gain d'autrui, il est alors altruiste. Si $r_1 < 0$, les gains d'autrui affectent négativement l'agent, il est donc envieux. Si $r_1 = 0$, l'agent ne s'en tient qu'à son seul intérêt, il est donc égoïste.

Quels que soient les développements introduits quant à la forme des fonctions d'utilité, l'analyse se réduit à nous présenter les seules conséquences du comportement altruiste, reprenant par là même le point de vue de Marshall (1890) selon lequel les motivations de l'altruisme, aussi riches et variées quelles soient, ne peuvent se mesurer. Becker (1981) complète ce commentaire par le suivant : *"J'avance une définition de l'altruisme qui se veut en cohérence avec le comportement, plutôt que de développer une discussion philosophique quant à ce qui motive vraiment les individus"*. Même s'il existe une multitude de raisons à l'altruisme, il aurait toujours été possible de les dénombrer et de les analyser une à une pour, au moins, en déduire le degré d'interaction.

L'absence du facteur motivation rend caduque le traitement analytique proposé par Becker : le pouvoir explicatif du modèle ne résiste pas à la confrontation avec les faits. On le voit si l'on s'appuie sur les travaux de Monroe (1996) : cette enseignante en sciences politiques a réalisé entre 1988 et 1990 une série d'interviews auprès de personnes qui avaient porté assistance à des juifs lors de la seconde guerre mondiale. Bien que beaucoup de ces personnes avaient conscience qu'elles mettaient en jeu leur propre vie, ainsi que celle de leur famille, cela ne les a pas empêché de sauver des individus qui leur étaient totalement étrangers. Devant une telle situation, l'économiste se contenterait de souligner que le coût est certes très élevé, mais le gain que devrait retirer le sauveur pourrait l'être encore plus. Or, les interviews de Monroe révèlent que certains d'entre eux ont été jusqu'à refuser toute récompense matérielle ou autre. L'individu altruiste, au sens de l'économie, est prêt à s'engager en faveur d'autrui mais pas au point d'y perdre. Aussi, le comportement des sauveurs de juifs met-il en défaut l'approche de Becker. Selon Monroe, l'explication de ce choix d'aide découlerait d'une vision du monde à laquelle les sauveurs adhèreraient. Par leur engagement, c'est toute l'humanité qu'ils engageraient ; leur vision du monde s'accompagnant d'une forme d'empathie.

En fait, la limite de la théorie du choix rationnel dans son traitement de l'altruisme provient de ce qu'elle l'envisage par ses seules conséquences. S'intéresser aux motivations des individus requiert la définition d'un cadre d'analyse plus riche, qui induit l'introduction de

concepts d'équilibre adéquats, les notions d'équilibre retenues ne devant plus se restreindre aux seules stratégies adoptées mais tenir compte également des motivations (Rabin, 1993).

L'intentionnalité à l'origine de l'altruisme

On ne peut pas vraiment dire que la moralité appréhendée par l'économie nous ait été d'une aide précieuse pour comprendre le comportement coopératif. Etant donné cet état de fait, il paraîtrait raisonnable d'en conclure que le modèle du choix rationnel ne permet pas d'expliquer la coopération, justifiant par là même certaines des critiques qui lui sont formulées¹⁵. Pourtant, tel n'est pas le cas. En y regardant de plus près, nous pouvons dire qu'à l'aide d'une représentation de l'individu qui se veut la plus simple possible, nous avons étudié, sous l'angle de la coopération, les possibilités qu'offraient cette représentation. Le tour ayant été réalisé sans grand succès, quelle reformulation de *l'homo œconomicus* pourrait être envisagée afin de l'enrichir ? Les travaux initiés dans le champs de l'économie comportementale (ou psychologie économique) s'inscrivent dans cette démarche. Cette branche de l'économie se propose de fournir une meilleure compréhension du système cognitif humain et de ses conséquences, *l'homo œconomicus* y étant enrichi de motivations, d'intentions. Il devient alors possible, à la lumière des intentions, d'analyser certaines propriétés de l'altruisme.

Le principe de la psychologie économique¹⁶ est de s'attacher à la diversité des motivations de l'individu. Cette approche suppose que les individus ne forment pas une population homogène. Cette diversité peut être appréhendée de deux façons. Soit elle provient de l'existence d'individus différents qui forment une société dont la composition est stable, soit elle est le fait de l'individu lui-même qui porte en lui des sentiments contradictoires. Ces deux optiques se rejoignent toutefois dans le questionnement qu'elle soulève : le problème de la cohérence même de l'individu.

Le modèle élaboré par Rabin (1993), qui traite justement de bienveillance, s'inscrit dans cette démarche. Cet auteur a développé l'idée selon laquelle l'intentionnalité d'une action peut être en mesure de favoriser l'émergence d'un comportement coopératif. L'intention repose ici sur un souci d'équité (de générosité) dans le partage. Chaque agent jugeant des intentions prêtées à l'action des autres, le gain qu'un individu i retire de son action (a_i) sera non seulement fonction de celle-ci mais également de ses croyances sur les stratégies des autres (b_i) et de ses croyances sur les croyances des autres sur sa propre stratégie (c_i). Dans la mesure où l'intention met en avant l'influence des croyances des joueurs, elle nécessite alors la définition d'une nouvelle classe d'équilibre : l'équilibre de bienveillance réciproque¹⁷.

¹⁵ Voir, par exemple, Etzioni (1990).

¹⁶ C'est à la suite de l'article fondateur de Geanakoplos, Pearce et Stacchetti (1989) que plusieurs auteurs ont développé cette voie de recherche.

¹⁷ Nous reprenons la traduction offerte par Meidinger, Robin et Ruffieux (1999) du "fairness equilibrium" de Rabin.

Afin d'explicitier cette notion, il nous faut introduire quelques précisions quant aux caractéristiques des agents. Pour ce faire, considérons un jeu à deux joueurs, 1 et 2. Chaque individu est supposé doté d'une fonction de bienveillance, $f_i(a_i, b_j)$, $i \neq j$, et la bienveillance de i envers j est donnée par :

$$f_i(a_i, b_j) = \frac{U_j(a_i, b_j) - U_j^e(b_j)}{U_j^{\max}(b_j) - U_j^{\min}(b_j)}.$$

$U_j^{\max}(b_j) = \text{Max } U_j(a_i, b_j)$ et $U_j^{\min}(b_j) = \text{Min } U_j(a_i, b_j)$ désignent, respectivement, le gain le plus élevé et le moins élevé que l'agent j est en mesure d'obtenir parmi l'ensemble des gains possibles. Le terme $U_j^e(b_j)$ correspond au niveau d'utilité équitable :

$$U_j^e(b_j) = [U_j^h(b_j) + U_j^l(b_j)]/2,$$

sachant que $U_j^l(b_j)$ (resp. $U_j^h(b_j)$) représente le gain le plus faible (resp. le plus élevé) que j puisse obtenir parmi l'ensemble des gains associés à une situation pareto efficace. Ce point de référence permet d'évaluer la générosité de i envers j .

Les croyances du joueur i sur les intentions du joueur j sont définies par la fonction $\tilde{f}_j(b_j, c_i)$ qui, par construction, est formellement équivalente à la fonction $f_j(a_j, b_i)$:

$$\tilde{f}_j(b_j, c_i) = \frac{U_i(c_i, b_j) - U_i^e(c_i)}{U_i^{\max}(c_i) - U_i^{\min}(c_i)}.$$

Les préférences de chacun des individus sont alors caractérisées par une fonction d'utilité $U_i(a_i, b_j, c_i)$ qui comprend les gains matériels, $U_i(a_i, b_j)$, en plus des deux fonctions précédentes :

$$U_i(a_i, b_j, c_i) = U_i(a_i, b_j) + \tilde{f}_j(b_j, c_i) [1 + f_i(a_i, b_j)].$$

Cette fonction exprime l'arbitrage qu'effectue l'agent i entre la satisfaction qu'il retire de son gain matériel et celle qu'il retire de l'appréciation qu'il se fait du comportement de l'agent j . Les préférences ainsi définies reflètent bien certaines spécificités comportementales. En effet, si le joueur i pense que j le traite négativement, cas où $\tilde{f}_j(\cdot) < 0$, i souhaitera alors le lui rendre en choisissant une action a_i telle que $f_i(\cdot)$ soit suffisamment faible ou tout simplement négative. En revanche, si le joueur j traite i avec générosité, $\tilde{f}_j(\cdot)$ sera alors positive et i souhaitera le rendre à j en se comportant tout aussi généreusement. Cela

n'empêche pas chaque agent de se trouver confronté à un arbitrage entre souci de réciprocité et recherche de bien-être matériel.

Rabin définit un équilibre de bienveillance réciproque par analogie à l'équilibre de Nash (condition (1)), à la seule différence qu'il impose la condition supplémentaire selon laquelle les croyances doivent correspondre au comportement effectif (condition (2)). D'après la condition (1), étant donnée la croyance de i quant à la bienveillance de j , i jouera sa meilleure réponse a_i^* : $a_i^* = \operatorname{argmax}_{a_i} U_i(a_i, b_j, c_i)$. Dans la mesure où à l'équilibre, croyances et comportement effectif doivent correspondre, il s'ensuit que le couple (a_1^*, a_2^*) constituera un équilibre de bienveillance réciproque si :

$$a_i^* = \operatorname{argmax}_{a_i} U_i(a_i, b_j, c_i) \quad i=1,2.$$

La définition suivante résume notre propos.

Définition

Soit S_i l'ensemble des stratégies potentielles du joueur i , $i = 1, 2$. $(a_1^*, a_2^*) \in (S_1, S_2)$ est un équilibre de bienveillance réciproque si, pour $i = 1, 2$, $j \neq i$:

$$\text{Condition (1) : } a_i^* \in \operatorname{argmax}_{a_i \in S_i} U_i(a_i, b_j^*, c_i^*)$$

$$\text{Condition (2) : } c_i^* = b_j^* = a_j^*.$$

L'introduction dans les préférences individuelles d'éléments d'ordre psychologique détermine la manière dont les individus attribuent une cause à des événements et la façon dont cela affecte leur comportement. Ainsi sur la base du concept d'équilibre de bienveillance réciproque, il devient possible de rendre compte du comportement coopératif dans le DP : lorsqu'il est connaissance commune pour les joueurs qu'ils opteront pour l'issue (C,C), alors chacun d'eux sait que l'autre consent à sacrifier son propre bien-être matériel afin de lui rendre service. Chacun est ainsi incité à faire preuve de générosité à l'égard de son partenaire (en jouant C) tant que les gains matériels résultant de T ne paraissent pas significatifs.

Le problème du concept de Rabin c'est qu'il ne garantit pas, pour le DP, l'existence d'une unique solution. C'est même le contraire puisque l'issue (T,T) constitue également un équilibre de bienveillance réciproque. S'il est connaissance commune que les agents sont enclins à jouer (T,T), chacun d'eux sait alors que l'autre ne consent pas à sacrifier son gain de trahison et donc aucun d'eux n'obtiendra le gain coopératif. D'où un comportement hostile réciproque. La pertinence de la solution (T,T) s'explique par le fait que les joueurs se satisfont doublement : en réalisant leur désir de punir l'autre et par la défense de leur intérêt personnel.

Aboutir à une situation d'équilibres multiples constitue, dans une certaine mesure, une avancée car cela donne tout son sens au comportement coopératif, sans pour autant délaiss

la faisabilité de l'issue non-coopérative. Toutefois cela s'accompagne aussi d'un nouvel handicap : il devient impossible de faire la moindre prédiction. En effet, pour une situation donnée, quel comportement est-il susceptible de voir le jour ? Quelle solution les acteurs vont-ils privilégier au concret ? La réponse ne pourra être donnée qu'après que le jeu ait été joué. Avec ce type de conceptualisation, on ne peut pas faire mieux que fournir une justification ex-post. On pourrait éventuellement procéder à l'élimination (ou à la justification) de la solution la moins plausible, mais cela nécessite l'introduction de critères de sélection ad hoc, affaiblissant dans le même la portée de l'analyse.

Ce concept d'équilibre est activement utilisé en économie expérimentale, c'est à dire la branche de l'économie qui recourt à l'expérimentation en reproduisant en laboratoire une situation économique (stylisée) afin d'en tester le modèle théorique associé. Cependant, on s'aperçoit que pour faire émerger une solution, notamment dans des jeux séquentiels, il est toujours nécessaire de procéder à des raffinements du concept. Ceux-ci s'interprètent comme la définition de formes particulières de préférences, et donc une analyse en terme d'équilibre de bienveillance réciproque revient, pour beaucoup, à spécifier les préférences correspondant au comportement que l'on cherche à étudier¹⁸.

3. CONCLUSION

Aborder la question de la coopération sous l'angle des jeux non-coopératifs, que cette coopération soit guidée par des considérations morales, laisse en suspens une question importante. En effet, l'équilibre de Nash nous apprend que si $n-1$ individus respectent un comportement, il est dans l'intérêt du $n^{\text{ème}}$ d'en faire autant. Cependant, cela ne nous dit pas comment les $n-1$ premiers agents en sont venus d'eux-mêmes à jouer la stratégie d'équilibre ? Cette question se pose pour chacun des acteurs du jeu. Dans ce cadre, l'élaboration d'une théorie de l'émergence du comportement coopératif n'est possible que si on fait l'hypothèse que chacun sait déjà quel est le comportement adopté par les autres, et qu'il s'agit bien du même ; ici le comportement coopératif étant favorisé par des aspects moraux.

Le point à admettre, partant de cette remarque, est que le comportement coopératif est, dans une certaine mesure, à l'opposé de ce que présente la théorie des jeux : inconditionné. Si l'on veut ménager un point de passage entre l'absence du comportement et sa généralisation au terme d'un processus d'extension, on est conduit à admettre que les individus se fixent des règles d'action indépendamment du fait de savoir si tous les autres vont s'y tenir. En d'autres termes, le comportement coopératif serait une affaire plus personnelle qu'il n'y paraît au départ. L'effet foule n'étant qu'une conséquence, et non une condition, même auto-réalisatrice ex-post, de l'adoption de ce comportement au niveau collectif.

On serait tenté d'objecter, à cet endroit, que le respect de la coopération mutuelle est loin d'être une affaire personnelle dans la mesure où toutes les règles sociales prescrivant un

¹⁸ Pour des exemples, voir la thèse de Nadège Marchand (1999).

comportement particulier sont assorties d'un système de sanctions garantissant son respect. Mais ceci revient à admettre que la coopération existe déjà lorsqu'elle commence à être adoptée. Le système de sanctions ne peut, en toute logique, précéder le comportement qu'il est censé renforcer, si l'on reste dans la perspective d'une émergence spontanée de la coopération. Dans une certaine mesure, cela prend à contre-pied le regard que l'économiste porte sur le comportement coopératif. L'attitude coopérative consiste en fait en un comportement qui permet, sinon qui a pour but, d'affranchir l'individu des contraintes liées à la variabilité de son environnement. Ceci se concrétise de la façon suivante. Lorsque les autres adoptent un comportement coopératif, cela a pour conséquence d'introduire de l'ordre et de la fiabilité dans l'environnement de l'agent. A ce niveau, la coopération permet de faire des prévisions, et non seulement des conjectures, qui rendent possibles la prise de décision. Elle a donc pour caractéristique majeure de réduire l'incertitude lié au manque de confiance, ce trait étant d'ailleurs conforme aux différentes représentations s'appuyant sur l'équilibre de Nash.

BIBLIOGRAPHIE

- Ballet J. (2000), "Altruisme et Biens Collectifs, une Présentation de la Littérature", *Revue Economique*, 4, p. 789-811.
- Becker G.S. (1974), "A Theory of Social Interactions", *Journal of Political Economy*, 82, p. 1063-1093.
- Becker G.S. (1981), *A Treatise on the Family*, Cambridge Mass., Harvard University Press.
- Bénard J. (1985), *Economie Publique*, Paris, Economica.
- Bordignon M. (1990), "Was Kant Right? Voluntary Provision of Public Goods under the Principle of Unconditional Commitment", *Economic Notes*, 3, p. 342-372.
- Collard D.A. (1978), *Altruism and Economy*, Oxford, Martin Robinson.
- Collard D.A. (1983), "Economics and Philanthropy: A Comment", *Economic Journal*, 93, p. 637-638.
- Comte A. (1852/1966), *Catéchisme Positiviste*, Paris, Garnier Flammarion.
- Cornes R., Sandler T. (1996), *The Theory of Externalities, Public Goods and Club Goods*, Second Edition, Cambridge, Cambridge University Press.
- Darwin C. (1859/1973), *L'Origine des Espèces*, Verviers, Marabout-Université. Darwin C. (1871/1999), *La Filiation de l'Homme*, Paris, Syllepse.
- De Alessi L. (1969), "Implications of Property Rights for Government Investment Choices", *American Economic Review*, 59, p. 13-24.
- Etzioni A. (1987), "Toward a Kantian Socio-Economics", *Review of social Economy*, 45, p. 37-47.
- Etzioni A. (1990), "Pour une Science Sociale Déontologique", *Revue du MAUSS*, n° 9, p. 14-32.
- Gauthier, D.P. (1986), *Morals by Agreement*, Oxford, Oxford University Press.
- Geanakoplos J., Pearce D., Stacchetti E. (1989), "Psychological Games and Sequential Rationality", *Games and Economic Behavior*, 1, p. 60-79.
- Harsanyi J.C. (1980), "Rules Utilitarianism, Rights, Obligations and the Theory of Rational Behaviour", *Theory and Decision*, 12, p. 115-133. Hobbes T. (1651/1971), *Léviathan*, Editions Sirey.
- Hochman H.M., Rodgers J.D. (1969), "Pareto Optimal Redistribution", *American Economic Review*, 59, p. 997-1002.
- Kant (1798/1992), *Fondements de la Métaphysique des Mœurs*, Paris, Vrin.
- Kolm S.-C. (1966), "La Production Optimale de Justice Sociale", Conférence de l'A.I.S.E, dans *Economie Publique*, Paris, C.N.R.S.
- Koulibaly M. (1998), "Solidarité Internationale, Altruisme et Sous-Développement Africain", dans F.R. Mahieu et H. Rapoport (éds.), *Altruisme, Analyses Economiques*, Paris, Economica, p. 263-294.
- Kropotkine A. (1891/1989), *La Morale Anarchiste*, Fresnes, éds Volonté Anarchiste.
- Kropotkine A. (1902/1979), *L'Entraide, un Facteur d'Evolution*, Paris, Editions de l'Entraide.
- Laffont J.-J. (1975), "Macroeconomics Constraints, Economic Efficiency and Ethics: An Introduction to Kantian Economics", *Economica*, 42, p. 430-437.
- Locke J. (1690/1985), *Deuxième Traité du Gouvernement Civil*, Paris, Vrin.

- Machiavel N. (1532/1983), *Le Prince*, Librairie générale française (Préface de R. Aron).
- Maître P. (2000), "L'Hypothèse de Sympathie : Une Perspective Historique", *Economies et Sociétés*, série *Æconomia*, n° 30, p. 39-67.
- Marchand N. (1999), *Ultimatum et Logique Economique de la Négociation : L'Apport de l'Economie Expérimentale*, Thèse de Doctorat en Sciences Economiques, Université Lumière Lyon II.
- Marglin S.A. (1963), "The Social Rate of Discount and the Optimal Rate of investment", *Quarterly Journal of Economics*, 77, p. 95-111.
- Margolis H. (1982), *Selfishness, Altruism and Rationality*, Cambridge, Cambridge University Press.
- Marshall A. (1890), *Principles of Economics*, New York, MacMillan.
- Meidinger C., Robin S., Ruffieux B. (1999), "Confiance, Réciprocité et Cheap Talk", *Revue Economique*, 50, p. 5-44.
- Monroe K.R. (1996), *The Heart of Altruism. Perceptions of a Common Humanity*, Princeton, Princeton University Press.
- Moulin H. (1981), *Théorie des Jeux pour l'Economie et la Politique*, Paris, Hermann. Olson M. (1965/1987), *Logique de l'Action Collective*, Paris, PUF.
- Quattrone G.A., Tversky A. (1986), "Self-Deception and the Voter's Illusion", in J. Elster (ed.), *The Multiple Self*, Cambridge, Cambridge University Press, p. 35-58.
- Rabin M. (1993), "Incorporating Fairness into Game Theory and Economics", *American Economic Review*, 83, p. 1281-1302.
- Rabin M. (1998), "Psychology and Economics", *Journal of Economic Literature*, 36, p. 11-46.
- Raiffa H. (1992), "Game Theory at the University of Michigan, 1948-1952", in E.R. Weintraub (ed.), *Toward a History of Game Theory*, supplement to vol. 24, Duke University Press, p. 165-175.
- Roemer J.E. (1996), *Theories of Justice*, Cambridge Mass., Harvard University Press.
- Rosanvallon P. (1989), *Le Libéralisme Economique. Histoire de l'Idée de Marché*, Paris, Seuil.
- Rousseau J.-J. (1762/1992), *Du Contrat Social*, Paris, GF-Flammarion.
- Schotter A. (1981), *The Economic Theory of Social Institutions*, Cambridge University Press.
- Selye H. (1974), *Stress sans Détresse*, Montréal, Editions La Presse.
- Smith A. (1759/1976), *Theory of Moral Sentiments*, Oxford, Clarendon Press.
- Smith A. (1776/1976), *Recherches sur la Nature et les Causes de la Richesse des Nations*, Paris, Gallimard.
- Steedman I. (1995), "Rationality, Economic Man and Altruism in P.H. Wicksteed's Common Sense of Political Economy", in S. Zamagni (ed.), *The Economics of Altruism*, Aldershot, Edward Elgar.
- Straffin P. (1980), "The Prisoner's Dilemma", *UMAP Journal*, 1, p. 101-103.
- Sugden R. (1984), "Reciprocity: The Supply of Public Goods Through Voluntary Contributions", *Economic Journal*, 94, p. 772-787.
- Sugden R. (1991), "Rational Choice: A Survey of Contributions from Economics and Philosophy", *Economic Journal*, 101, p. 751-758.
- Ulmann-Margalit E. (1977), *The Emergence of Norms*, Oxford, Oxford University Press.