

Jeux évolutionnistes et normes sociales

André Orléan

(Economie Appliquée, n°3, tome L, 1997, 177-198
version du 7 juillet 1997)

Abstract

Evolutionary Games and Social Norms

Evolutionary Game Theory has lately developed important models of social norms and conventions. This paper aims at exposing this literature and some of its drawbacks. In particular, it is shown that such approaches to institutional forms, grounded on the sole notion of individual utility, are limited by the absence in the analysis of normative judgements. We then highlight the importance of sanctions and collective beliefs in the emergence of social norms.

Résumé

Au cours des dernières années, la théorie des jeux évolutionnistes a consacré d'importants développements à la modélisation des normes sociales et des conventions. Le présent article en propose une présentation critique. On cherche à montrer les limites d'une approche qui cherche à rendre compte du fonctionnement de ces formes institutionnelles en laissant totalement de côté le rôle des jugements normatifs. Seule la poursuite de l'utilité est prise en compte. Quelques pistes de réflexion sont alors présentées, en particulier à propos du rôle des sanctions et des croyances collectives.

Depuis les travaux fondateurs de John Maynard Smith (1982) et dans le prolongement de l'interprétation socio-économique qu'en ont proposé Robert Axelrod (1984) et Robert Sugden (1986), la théorie des jeux évolutionnistes (TJE) est devenu un outil de réflexion apprécié des économistes. Elle est à l'origine d'un important renouvellement dans l'approche évolutionniste de certaines formes institutionnelles comme les normes sociales, les conventions ou la tradition. A titre d'exemples, on peut citer les travaux de H. Peyton Young (1996) sur l'émergence et la stabilité des conventions ; de Sajiv Sethi (1996) sur les normes sociales ; de Robert Boyd et Peter J. Richerson (1993) sur le rôle de la tradition ; ou encore de Robert Boyer et André Orléan (1994) sur les changements de conventions. La force de la TJE est d'avoir fourni un cadre d'analyse permettant une étude formalisée de ces questions. Si cette modélisation a permis d'importants progrès dans la manière d'aborder le rôle économique des institutions, elle possède néanmoins de notables limites qui, à ma connaissance, n'ont guère été étudiées. Parmi ces limites, il en est une qui nous semble particulièrement problématique, à savoir le fait que, dans la TJE, les jugements individuels portent exclusivement sur l'utilité. Il n'est fait aucune place à l'évaluation normative des stratégies. Autrement dit, la question de savoir si une conduite est juste, normale ou conforme n'est jamais abordée : celle-ci est uniquement jugée en fonction des gains qu'elle engendre. Cela n'a certes rien de surprenant. Par nature, l'économie s'intéresse aux activités orientées vers l'utilité. Il n'en reste pas moins qu'on peut s'interroger sur la prétention d'une telle démarche à appréhender l'existence de normes sociales : peut-on penser les normes sociales en faisant l'impasse totale sur toute notion de normativité ? Comme on le voit, cette interrogation dépasse le strict cadre de la TJE. C'est la capacité de la théorie économique à proposer une analyse pertinente des institutions qui est alors posée.

Le présent texte s'intéressera pour l'essentiel aux outils de formalisation développés par la TJE. Nous en ferons une rapide présentation qui nous permettra d'en souligner

certaines insuffisances. Des voies de modélisation seront alors proposées qui devraient permettre une formalisation plus satisfaisante des normes sociales.

JEUX ÉVOLUTIONNISTES ET CONVENTIONS

Pour présenter la TJE, nous considérerons un exemple très simple, celui du code de la route. On y étudie une population composée de N individus où chaque agent a le choix entre deux stratégies pures : soit D (rouler à droite), soit G (rouler à gauche). L'interaction est formalisée par la matrice de gains suivante :

$$\begin{array}{cc}
 & \begin{array}{cc} D & G \end{array} \\
 \begin{array}{c} D \\ G \end{array} & \begin{array}{cc} 1, 1 & 0, 0 \\ 0, 0 & 1, 1 \end{array}
 \end{array} \tag{1}$$

La variable d'état qui décrit la situation de la population est le vecteur $\mathbf{n}(t) = [n_D(t), n_G(t)]$ qui donne le nombre d'individus ayant opté, à l'instant t , pour chacune des deux stratégies. On utilisera également le vecteur $\mathbf{p}(t) = [p_D(t), p_G(t)]$ des proportions¹. Le modèle de base de la TJE considère que les individus se rencontrent aléatoirement deux par deux. Il s'ensuit que l'utilité obtenue par un individu est une variable aléatoire qui vaut soit 1, soit 0, en fonction de la stratégie du partenaire qui a été tiré aléatoirement. On peut aisément calculer l'espérance mathématique de cette utilité, soit u_D pour les individus ayant choisi D et u_G pour ceux ayant choisi G . On obtient :

$$\begin{cases} u_D(\mathbf{p}) = p_D \\ u_G(\mathbf{p}) = p_G = 1 - p_D \end{cases} \tag{2}$$

L'utilité engendrée par une stratégie est d'autant plus grande que la proportion

¹On a bien évidemment $n_D(t) + n_G(t) = N$ et $p_D(t) + p_G(t) = 1$.

d'individus ayant choisi cette même stratégie est élevée. Que nous dit alors la TJE sur l'évolution de \mathbf{p} au cours du temps ? Vers quoi converge la population considérée ?

Dans sa formulation primitive issue de la biologie, où l'utilité correspond à la "fitness" darwinienne, à savoir un taux de reproduction, la TJE retient la dynamique suivante :

$$p_i(t+1) = p_i(t) \frac{u_i(t)}{u_m(t)} \quad (3)$$

où $p_i(t)$ est la proportion des individus ayant opté pour la stratégie i , $u_i(t)$ l'utilité moyenne que procure cette même stratégie et $u_m(t)$ l'utilité moyenne du groupe. Cette dynamique particulière est appelée "dynamique de réplication". Elle se déduit directement du fait que le nombre d'individus $n_i(t)$ ayant choisi la stratégie i croit proportionnellement à $u_i(t)$, soit :

$$n_i(t+1) = u_i(t) \cdot n_i(t) \quad (4)$$

C'est là une hypothèse qui n'a pas grand sens pour un économiste². Pour analyser la dynamique, les économistes font le plus souvent appel à une hypothèse d'apprentissage qui conduit les individus à comparer l'utilité qu'ils ont obtenue à celles qu'obtiennent les autres agents puis, sur la base de cette information, à abandonner leur stratégie pour une stratégie plus rentable. Cette hypothèse très générale, compatible avec une inertie plus ou moins grande des comportements, conduit à des dynamiques qui vérifient :

$$\frac{p_i(t+1) - p_i(t)}{p_i(t)} > \frac{p_j(t+1) - p_j(t)}{p_j(t)} \quad \text{ssi} \quad u_i(t) > u_j(t) \quad (5)$$

La dynamique de réplication en est un cas particulier. Un autre cas particulier important pour les économistes est la dynamique dite "de meilleure réponse". Elle est telle qu'à chaque instant, tous les individus choisissent la stratégie la plus performante.

²En particulier si l'économiste raisonne à population constante.

Autrement dit, à partir de la connaissance de $\mathbf{p}(t)$, ils choisissent la stratégie qui maximise $u_i(\mathbf{p})$; ou l'une quelconque d'entre elles s'il y en a plusieurs.

On peut montrer que toutes les dynamiques évolutionnistes vérifiant la condition (5) convergent vers un même état d'unanimité. Si $p_D(0) > \frac{1}{2}$, l'unanimité se fait sur la conduite à droite, à savoir l'état $\{p_D = 1\}$. Si $p_D(0) < \frac{1}{2}$, c'est la conduite à gauche qui est sélectionnée, à savoir l'état $\{p_D = 0\}$. Ce résultat est censé rendre compte de l'émergence des conventions "rouler à droite" ou "rouler à gauche". Comme on le voit, on retrouve strictement les équilibres de Nash du jeu initial. Examinons ce résultat.

Comme nous l'indiquions dans l'introduction, cette dynamique a un contenu cognitif très pauvre. Les agents ne s'y interrogent pas sur les raisons qui motivent le choix des autres ; seules importent les valeurs constatées des $u_i(\mathbf{p})$. C'est là une conséquence du contexte biologique qui a vu naître ce type de modélisation. Le but du modèle biologique est d'étudier les dynamiques de sélection au sein de populations animales. La dimension cognitive y est alors totalement absente. En effet, la diffusion d'une mutation s'y fait mécaniquement sans qu'à aucun moment les individus porteurs de cette mutation n'aient à se représenter, de quelque manière que ce soit, le processus global qui les entraîne : un avantage en termes d'adaptation ou de "fitness" conduit à un taux de fécondité plus grand, ce qui mène à la domination progressive de cette mutation. Dans sa version économique, ce contenu cognitif n'est guère plus élaboré. Il implique cependant une comparaison des utilités obtenues par les diverses stratégies en prenant comme référence l'état des choix de la population à la date précédente, soit $\mathbf{p}(t)$. Les agents sont myopes dans la mesure où ils ne prennent pas compte le fait qu'en $(t + 1)$, les autres joueurs auront à leur tour également modifié leur choix stratégique, ce qui influera sur l'utilité qu'engendrera alors la stratégie qu'ils auront choisie. Seules importent les valeurs $u_i(\mathbf{p}(t))$.

Autrement dit, contrairement au modèle biologique, l'effet de l'utilité ne se traduit

pas mécaniquement par une progression de la sous-population des stratégies qui ont obtenu les meilleurs scores (équation 5). Cette progression suppose une activité réfléchie des protagonistes, à savoir comparer les résultats des diverses conduites et préférer celles ayant fait un bon score. Cette réflexion des agents peut être plus ou moins complexe selon les versions retenues du modèle de base. On en trouve la forme la plus élaborée dans le cas de la dynamique de “meilleure réponse” où les agents choisissent la meilleure stratégie compte tenu de \mathbf{p} . Cependant, même dans le cadre de cette hypothèse, il importe de souligner que toute dimension normative est absente des jugements individuels. La seule question pertinente est de savoir quelle stratégie produit les gains les plus élevés sans que jamais n’intervienne une évaluation portant sur la pertinence, la justesse ou la conformité des conduites. La question est alors de savoir si une telle approche permet une description pertinente des institutions, qu’il s’agisse de conventions ou de normes sociales³. Celles-ci, comme nous le verrons, donnent lieu à un fort investissement affectif qui conduit l’individu à les respecter et à désapprouver profondément les personnes qui en violent les interdits. Une analyse satisfaisante de la formation d’une convention ou d’une norme sociale doit rendre compte de l’apparition de jugements qui ne portent pas uniquement sur les utilités respectives des stratégies, mais qui en évaluent également la “normalité”. Il n’est pas nécessaire dans le cadre de cet article de définir avec précision les notions de normalité, de justesse ou de conformité. Il nous suffira de considérer que les évaluations normatives sont des évaluations qui visent, à partir des éléments propres à la culture du groupe considéré, à juger de l’adéquation d’une conduite aux habitudes, traditions et valeurs de cette communauté.

On est certainement ici à la limite de ce que peuvent faire les modèles économiques, car intégrer ce type de jugement à l’analyse suppose que soient pris en compte des

³La distinction entre conventions et normes sociales sera présentée et discutée dans une prochaine section.

ordres de justification fort éloignés du calcul d'utilité⁴. Il convient néanmoins d'en souligner l'importance pour qui prétend à une théorie pertinente des formes institutionnelles. Ces jugements de valeur interviennent à la fois dans la détermination de la dynamique, dans le choix de l'équilibre et dans la stabilité de la convention. Considérons ces différents aspects.

JUGEMENTS DE VALEUR ET LÉGITIMITÉ DES INSTITUTIONS

Depuis T. Schelling (1960), on sait que le seul exercice de la rationalité ne permet pas de résoudre les problèmes de coordination du type (1). Pour se coordonner, les individus sont amenés à mobiliser des éléments d'appréciation disparates dans le but de construire des repères pour l'action collective à partir de saillances cognitives. Pour ce faire les individus se livrent à ce qu'on pourrait appeler une "surinterprétation" des données contextuelles de façon à faire apparaître des points focaux. Schelling souligne tout particulièrement l'importance des précédents. C'est ce principe qui justifie l'emploi de l'état passé de la population, $\mathbf{p}(t)$, pour déterminer le choix en $(t + 1)$. Dans un même ordre d'idées, on comprend le rôle que peuvent jouer certaines valeurs partagées dans la construction des points focaux. Les jugements sur le caractère adéquat de telle ou telle conduite vont faire apparaître certaines solutions comme plus plausibles que d'autres. Il existe ainsi une étroite imbrication entre ces jugements de type normatif et la manière dont les individus apprennent. On peut conjecturer que les élaborations intellectuelles visant à évaluer le caractère conforme des choix en présence ont une influence sur le processus d'apprentissage, comparable à celle qu'exercent les évaluations qui portent sur les utilités comparées des stratégies. Lorsqu'est en cause l'élaboration d'une convention, c'est à dire l'émergence d'un code de conduite partagé par tous les membres du groupe, des appréciations quant au caractère normal des comportements en compétition sont nécessairement mobilisées.

⁴C'est la voie choisie par l'Economie des conventions. Voir par exemple Thévenot (1989).

Ces appréciations peuvent se révéler un puissant facteur de cohésion et permettre une coordination rapide et aisée.

Détermination de la dynamique

Pour toutes ces raisons, les individus n'ont pas le comportement passif et mécanique que leur attribue la TJE. L'observation de $\mathbf{p}(t)$ engendre au sein de la population une activité cognitive plus complexe que le seul calcul des $\{u_i(\mathbf{p})\}$. Des hypothèses sont inductivement formées à partir de $\mathbf{p}(t)$ quant à la nature et aux penchants du groupe. Cette activité d'interprétation peut avoir des effets importants sur la dynamique d'apprentissage. Par exemple, dans une situation où n'existent guère de précédents de telle sorte que toutes les stratégies apparaissent comme des prétendants également légitimes, on peut imaginer une grande diversité initiale de choix. L'émergence purement aléatoire d'une stratégie A recueillant un nombre de voix significativement supérieur aux autres, sans être majoritaire, peut alors être interprétée par nos individus en quête de sens, comme l'expression d'une normalité cachée, comme révélant une identité souterraine, et cela d'autant plus facilement que ses caractéristiques objectives se prêtent à un tel jugement. Si cette interprétation est retenue, elle peut modifier radicalement la dynamique et conduire à l'émergence rapide d'une unanimité sur A. On voit ici l'importance des jugements normatifs. Selon que les agents adhéreront à la proposition, "la stratégie A est celle qui recueille le plus grand nombre de suffrages", ou à la proposition alternative, "la stratégie A est la conduite normale du groupe", la dynamique connaîtra des profils différents. Dans les deux cas la variable d'état est la même, mais l'interprétation qui en est faite diffère.

Sélection des équilibres

Cette logique d'interprétation dépend fortement de la manière dont les individus se représentent l'identité du groupe. En effet, les jugements de normalité, de justesse ou de conformité n'ont de sens que relativement à une culture. C'est conditionnellement à celle-ci que sont tentées des hypothèses interprétatives. Pour en voir l'importance, on peut considérer l'exemple suivant. Soit un jeu de coordination du type (1) où les deux protagonistes doivent séparément choisir deux objets dans l'ensemble suivant : {Aix-en-Provence, Béthune, Lille, Marseille}. Si leurs choix se recoupent, ils obtiennent 0. Dans le cas contraire, ils obtiennent 1. Les équilibres de Nash correspondent alors aux trois partitions suivantes, {Aix-en-Provence, Béthune / Lille, Marseille}, {Aix-en-Provence, Marseille / Béthune, Lille} et {Aix-en-Provence, Lille / Béthune, Marseille}. Deux principes saillants émergent : l'ordre alphabétique qui conduit à la première réponse et l'ordre géographique qui conduit à la seconde réponse. Très clairement, le choix de l'un ou l'autre des équilibres va dépendre de la manière dont les participants se représentent le groupe. Si le groupe est formé de français, on optera vraisemblablement pour la seconde partition. S'il est composé de non français connaissant l'alphabet latin, on retiendra la première partition. Théoriquement la situation est plus complexe dans la mesure où ce qui compte pour un participant n'est pas sa propre opinion sur l'identité du groupe, mais bien l'opinion des autres, lesquels, comme le dit Keynes, "examinent tous le problème sous le même angle". Il s'agit donc de comprendre comment le groupe se représente lui-même. Nous examinerons cette question de la modélisation de telles représentations collectives dans une prochaine sous-section. Mais auparavant, considérons la manière dont les jugements normatifs affectent la stabilité des conventions.

Stabilité des institutions

L'analyse des conventions que propose la TJE peut être dite instrumentaliste ou fonctionnaliste dans la mesure où la convention n'y est appréhendée que comme un instrument permettant aux individus d'accroître leur utilité. La conception instrumentale est la conception qui prévaut au sein d'un discours théorique qui ne retient que l'utilité comme source de légitimité. Autrement dit, l'institution S y est légitimée pour autant que l'utilité de tous les agents est telle que :

$$u^i(e, S) \geq u^i(e) \quad \forall i \quad \forall e \in E \quad (6)$$

où E désigne l'ensemble des états du monde possibles, et où $u^i(e, S)$ désigne l'utilité de l'individu i quand l'état du monde est e et quand est présente l'institution S . On comprend la force de la vision instrumentaliste : la relation des individus à l'institution y est totalement dégagée de tous les forts liens émotifs par lesquels s'expriment, dans les ordres traditionnels, les rapports aux formes collectives. Certains théoriciens critiquent cette approche. Ils font valoir que les règles sociales ne peuvent jouer leur rôle normatif sans être l'objet d'un intense investissement affectif qui va bien au-delà de ce que peut produire la conception instrumentaliste et sa distance critique aux institutions. La légitimité d'une institution suppose des formes de justification beaucoup plus radicales que le seul calcul des performances. Au contraire d'une évaluation qui ramène sans cesse la norme sociale à ses résultats, l'acte de légitimation évalue la norme par delà ses manifestations pratiques, dans sa conformité supposée à l'ordre universel. Dans la légitimité, c'est la qualité intrinsèque de l'institution qui est directement reconnue. Comme l'écrit Mary Douglas (1989), "une convention est légitimée quant, à la question de savoir pourquoi on agit ainsi ... l'on peut répondre in fine en se référant au mouvement des planètes dans le ciel ou au comportement naturel des plantes, des animaux ou des hommes" (page 42). Autrement dit, c'est la conformité de l'institution à une certaine conception de l'ordre universel qui est

évaluée. Ces réflexions rejoignent les analyses que Max Weber a consacrées à la notion d'ordre légitime : “La validité d'un ordre signifie quelque chose de plus qu'une simple régularité, conditionnée par une coutume ou par une situation commandée par l'intérêt”. C'est le cas lorsqu'un individu suit une conduite et qu'il “ne le fait pas par habitude prise, ni parce que ses propres intérêts le lui commandent, mais il le fait parce qu'il respecte par devoir la validité de la conduite qui s'il la violait, non seulement lui causerait certains préjudices, mais encore lui inspirerait du point de vue de la rationalité en valeur, un remords pour n'avoir pas accompli son devoir” (p. 64-65). Les implications de cette conception sur notre compréhension de la stabilité des conventions sont clairement énoncées par Weber : “L'ordre que l'on respecte uniquement pour des motifs rationnels en finalité est incomparablement moins stable que celui qui s'affirme grâce au prestige de l'exemplarité et de l'obligation” (p. 65). C'est sur ce point que l'analyse de la TJE est peut-être le plus limitée. La conception instrumentale ne comprend qu'une faible part des processus qui concourent à la reproduction des formes institutionnelles. Introduire la notion de légitimité permet d'en penser l'autonomie par rapport à la seule dynamique des utilités : une convention n'est pas immédiatement détruite lorsqu'un choc exogène modifie à la baisse $u^i(e, s)$. Avancer dans cette voie est indispensable pour les économistes s'ils veulent prétendre à une analyse pertinente des institutions. Très clairement, les difficultés sont nombreuses, non seulement concernant la compréhension théorique fondamentale des ressorts de la légitimité, mais également à propos des outils en permettant une modélisation satisfaisante. C'est plus spécifiquement la dimension collective des croyances qu'il est délicat d'analyser. Nous lui consacrerons la sous-section qui suit.

Autonomie des représentations collectives

Les analyses développées jusqu'à maintenant ont, à plusieurs reprises, insisté sur le rôle que jouent les représentations collectives dans la coordination. On a vu que

la manière dont le groupe se représentait lui-même était un élément central pour analyser l'émergence de telle ou telle convention. Par ailleurs, ces représentations semblent posséder une grande inertie et une grande autonomie par rapport aux croyances individuelles. Penser cette autonomie est capital pour évaluer correctement l'importance des dispositifs cognitifs collectifs. Pour mener cette analyse, il est utile d'introduire quelques notations simples, soit :

$$C_j P \tag{7}$$

qui signifie "j croit que la proposition P est vraie". On pourrait dire, de manière équivalente, que P est la croyance de j . On peut également définir une croyance collective par le fait qu'elle est la croyance de tout le groupe. C'est précisément cette croyance collective qui importe à j : c'est elle qui déterminera l'action à choisir. Autrement dit, j agira en fonction de la proposition Q définie par :

$$C_j C_h Q \quad \forall h \neq j \tag{8}$$

En effet, l'équation (8) signifie que j croit que tous les individus h ont Q pour croyance personnelle. Q est donc ce que j croit être l'opinion collective du groupe. C'est l'élément-clef pour j . Pour comprendre la complexité de la détermination des croyances collectives et leur relative autonomie par rapport aux croyances individuelles, considérons la situation décrite par la structure de croyances suivante :

$$\left\{ \begin{array}{l} C_j P \quad \forall j \\ C_j C_h Q \quad \forall h \neq j \quad \text{et ceci } \forall j \end{array} \right. \tag{9}$$

avec P différent de Q . Autrement dit, tout le monde croit la proposition P . P est donc la croyance individuelle de chacun des membres du groupe. Mais, simultanément, tout le monde croit que l'opinion collective est Q .

L'intérêt de cette structure est de mettre en avant les différences existant entre des croyances de niveaux différents. Un exemple d'une telle situation nous est donné par ce que j'ai appelé dans un article antérieur "l'effet Reagan", car une telle configuration a été observée à propos d'une déclaration du président Reagan rapportée par le *New York Times* du 12 novembre 1987. Dans cette déclaration, Reagan affirmait que le dollar avait suffisamment baissé. Cette intervention a été suivie d'une remontée du dollar. Cependant, ayant interrogé divers cambistes, le journaliste du *New York Times* s'est rendu compte que personne n'accordait de crédit aux déclarations du président Reagan dont l'expertise économique était fortement mise en doute par les opérateurs financiers. Le journaliste surpris demande alors à ces interlocuteurs pour quelles raisons ils avaient néanmoins acheté du dollar. La réponse qu'il obtint fut que, si chacun d'eux étaient bien persuadé de l'inanité des propos de Reagan, ils tenaient cependant pour fort probable que les autres cambistes lui apporteraient un certain crédit. Pour cette raison, il était judicieux d'acheter du dollar en prévision de sa hausse, mécanisme qui autoréalise la hausse anticipée. La structure des croyances est donc bien conforme à celle décrite par l'équation (9) : chacun croit qu'aucune information pertinente n'est contenue dans la déclaration de Reagan (proposition P) mais, simultanément, chacun croit que les autres peuvent croire à la pertinence de cette déclaration (proposition Q).

Notons que dans ce type de structure autoréférentielle, c'est la croyance de plus haut niveau qui compte de telle sorte que la structure suivante aurait conduit au même résultat :

$$\left\{ \begin{array}{l} C_j P \quad \forall j \\ C_j C_h P \quad \forall h \neq j \quad \text{et ceci } \forall j \\ C_j C_h C_k Q \quad \forall k \neq h \quad \text{et ceci } \forall h \neq j, \quad \forall j \end{array} \right. \quad (10)$$

Il suffit qu'à un certain niveau, Q apparaisse comme une croyance collective. Mais limitons nous au cas décrit par l'équation (9).

Une telle structure est très importante en matière financière. Elle permet de rendre intelligibles certaines dynamiques financières qui resteraient sinon mystérieuses. Par exemple, durant la crise du peso mexicain qui a débuté en décembre 1994, de nombreux analystes ont fait remarquer que la dépréciation du peso avait atteint une ampleur disproportionnée par rapport à ce que nécessitait la situation économique. Comment expliquer un tel schéma ? Les opérateurs sont-ils devenus fous ? Pas nécessairement. La structure précédente des croyances (9) fournit une explication plus plausible. Tous les participants au marché peuvent être intimement persuadés que le peso est tombé trop bas et qu'en conséquence, on doit éviter une dépréciation supplémentaire et, simultanément continuer à jouer à la baisse parce qu'ils pensent que le marché, quant à lui, continue de croire à la baisse. C'est alors la croyance attribuée au groupe, en l'occurrence le marché, qui s'impose.

L'étrangeté de cette structure cognitive permet de bien saisir la particularité des croyances collectives. Elles ne répondent pas à la même logique que les croyances individuelles. On ne saurait non plus les interpréter comme des moyennes d'opinions personnelles. Il s'agit de *croyances collectives formées intersubjectivement*. L'analyse précédente montre que ces croyances intersubjectives possèdent une certaine forme d'objectivité en ce qu'elles s'imposent partiellement aux individus, indépendamment de leurs croyances personnelles. Le groupe y apparaît de ce fait comme un être autonome. Cette autonomie trouve une illustration exemplaire dans l'opinion des marchés financiers. Cette opinion s'impose non seulement aux dirigeants politiques, mais également aux opérateurs qui interviennent sur les marchés. Cette extériorité joue évidemment un grand rôle dans les processus de coordination, dans l'émergence et dans la stabilité des conventions. Les croyances de type Q intègrent tous les stéréotypes propres au groupe considéré. Elles définissent partiellement ce qui doit être

considéré comme normal. Approfondir l'analyse de ces croyances collectives intersubjectives permettrait d'avancer grandement dans la compréhension des équilibres de points focaux.

Jusqu'à maintenant nous avons étudié principalement la manière dont la TJE pensait les conventions. Ce faisant, il nous est arrivé de parler de normes sociales, mais dans une acception voisine. On va maintenant s'intéresser plus spécifiquement aux normes sociales. Pour ce faire, on s'appuiera sur l'analyse qu'en propose Jon Elster (1989).

NORMES ET SANCTIONS

Elster définit les normes en les opposant à la rationalité. Le comportement rationnel est un comportement tourné vers des résultats futurs hypothétiques. Les impératifs que dicte aux agents la théorie du choix rationnel sont de la forme : "Si tu veux obtenir Y, fais X". Ces impératifs sont hypothétiques ou conditionnels plutôt que catégoriques ou inconditionnels. Les normes sont d'une nature différente. Elles ne sont pas orientées vers un résultat futur et prennent, le plus souvent, une forme impérative : "Fais X" ou "Ne fais pas X", comme, par exemple "Il est interdit de manger de la chair humaine". Les normes qui ont une forme hypothétique ne font pas dépendre l'action d'un résultat futur à réaliser, mais d'états passés, contemporains ou contrefactuels. "Rationality is essentially conditional and future-oriented. Social norms are either unconditional or, if conditional, are not future-oriented" (1989, p.99).

Par ailleurs les normes sont dites sociales dans la mesure où (i) elles sont partagées avec d'autres membres de la société (ou d'un groupe social plus restreint, par exemple un groupe professionnel) et où (ii) elles sont maintenues en partie par des sanctions. Ces sanctions sont l'expression spontanée de la désapprobation du corps social. L'intensité des sanctions est, de ce fait, une mesure de l'adhésion des so-

ciétaires à la norme qui a été transgressée. Cette forme de sanction se distingue fortement des sanctions pénales auxquelles donnent lieu une violation des normes légales. Dans le cas des normes légales, c'est un groupe spécialisé de sanctionneurs qui intervient, les juges et les policiers. Dans le cas des normes sociales, la sanction prend sa source dans l'ensemble du corps social. Elle est pour cette raison beaucoup plus efficace. L'ostracisme est une forme extrême de la sanction sociale : chez certains peuples, l'individu responsable de la transgression est socialement mort. Personne ne lui adresse plus la parole, ni ne communique d'aucune manière avec lui.

Dans les sanctions, c'est la cohésion de la société autour de certaines valeurs qui s'exprime. Par exemple, aux USA, le conformisme social et le respect des normes est très fort dans certaines communautés. C'est ainsi qu'un ami français ayant emmenagé dans une banlieue cossue de Washington s'est aperçu que l'entretien des gazons individuels y fait l'objet d'une attention sociale particulière. Peu attiré par le jardinage, cet ami avait négligé à son arrivée de soigner son gazon. Très rapidement des voisins vinrent lui rappeler ces obligations : ne pas soigner le gazon est une faute qui fait perdre de la valeur au quartier et entraîne la réprobation du voisinage. Sans qu'il y ait eu une quelconque concertation, ces voisins étaient les représentants informels de la communauté. Cette pression sociale diffuse qui trouve sa source dans l'ensemble de la communauté joue un rôle important dans le maintien des normes. Plus la communauté se sent concernée, plus la perte de réputation pour un individu qui ne se conformerait pas aux normes est coûteuse et dissuasive.

Dernière caractéristique des normes sociales chez Elster : elles sont l'objet d'un fort investissement affectif. Cette dimension affective s'exprime au travers d'intenses émotions comme l'anxiété, la honte ou la culpabilité, ressenties par celui qui viole une norme ou par celui qui est le témoin d'une telle violation. "Social norms have a *grip on the mind* that is due to the strong emotion their violations can trigger" (page 100).

Cette analyse conduit Elster à distinguer fortement normes et conventions. Pour Elster, conformément à la définition proposée par Sugden (1986), les conventions sont des équilibres. On retrouve ici l'approche développée par la TJE, approche que nous avons présentée dans la première partie. Dire que les conventions sont des équilibres signifie que la recherche de l'intérêt individuel par les agents suffit à assurer la stabilité de la convention. En effet l'utilité qu'obtient un individu qui dévie de la convention est inférieure à celle qu'il aurait obtenue en s'y conformant. La déviation est une stratégie pénalisante. Alors que dans le cas d'une convention, il est de l'intérêt des parties d'en perpétuer l'existence, cela n'est pas vrai pour une norme sociale qui peut aller à l'encontre des intérêts individuels. Pour cette raison, le respect de la norme sociale a une autre source que l'utilité : c'est la crainte des sanctions ou la force émotionnelle propre à la norme qui en assurent le maintien.

Pour illustrer cette différence, Elster prend l'exemple du code de la route. Le code de la route, écrit-il, est une convention parce que violer le code de la route fait encourir deux sortes de désagrément, avoir un accident ou être blâmé par des spectateurs parce que le fait de mal conduire peut leur causer des accidents. Autrement dit, les deux sortes de désagrément sont liées, directement ou indirectement, à une mauvaise performance en terme d'utilité. La seule référence à l'utilité peut permettre de comprendre le maintien de la convention. Par opposition, continue Elster, si je choisis la mauvaise fourchette lors d'un repas, ou si je mets un mauvais costume pour un enterrement, le seul désagrément qui peut m'arriver provient de la désapprobation des autres pour mes mauvaises manières (page 102).

Dans la définition qui est mise en avant par Elster, la convention ne serait maintenue que par la force des intérêts rationnels. C'est pourquoi une convention pourrait s'autoentretenir sans action extérieure, alors que la norme dépendrait, pour sa survie, des sanctions. Les raisonnements que nous avons consacrés, dans la deuxième partie, à l'analyse de la légitimité, nous conduisent à trouver cette différence discutable. On

peut certes se cantonner à la définition formelle de la convention qui est ici proposée, mais il nous semble que, dans la pratique, la plupart des conventions économiques qui nous intéressent sont en fait des *conventions légitimées*. Comme l’écrit Douglas : “Pour qu’une convention devienne une institution sociale légitime, il faut une convention parallèle de type cognitif qui la soutienne” (page 42). Dans ces conditions, le rapport à la convention et les processus qui en assurent la stabilité ne participent plus de la seule utilité. Ils reposent également sur des jugements qui établissent que la convention *doit* être respectée. Il s’opère, ce faisant, un investissement affectif qui transforme le respect de la convention légitimée en norme sociale et qui établit que la violation du code conventionnel de conduite peut entraîner une réprobation sociale. C’est ainsi que même dans le cas du code de la route, la violation des règles de conduite donne lieu à un mouvement de désapprobation collective qui va au-delà des seuls désagréments d’utilité. Ceci montre que la convention originelle est devenu une convention légitimée qui possède, pour cette raison, certaines des caractéristiques de la norme. Il s’ensuit que l’analyse des sanctions qui va suivre n’a pas comme champ d’application les seules normes sociales ; elle concerne également les conventions légitimées. Avant de passer à cette analyse, signalons que la distinction entre normes et conventions défendue par Elster n’est pas nécessairement retenue par la TJE. Celle-ci emploie le plus souvent indifféremment l’un ou l’autre terme. Par exemple, Young (1996) utilise exactement le même cadre d’analyse pour étudier le code de la route ou les normes de distribution (“norms in distributive bargaining”). Dans ce cas, la norme est considérée à la manière d’un équilibre.

Le modèle d’Akerlof

L’introduction des sanctions constitue un pas important vers une modélisation satisfaisante des institutions, pour plusieurs raisons. D’une part, les sanctions, tout particulièrement les sanctions dont l’effet transite par la réputation, permettent de

saisir une dimension essentielle des processus institutionnels, comme l'a déjà souligné Elster. D'autre part, avec les sanctions, il devient possible de modéliser des normes qui vont à l'encontre de l'intérêt immédiat des individus considérés, ce que la modélisation précédemment présentée ne permettait pas de faire. Enfin, une bonne théorie des normes doit pouvoir rendre compte de l'affaiblissement de certaines normes, voire de leur disparition. C'est au travers de sanctions d'intensité variable qu'il nous sera possible d'appréhender ce phénomène. Les réflexions qui vont suivre ont été élaborées à partir de l'analyse des normes proposée par Akerlof en 1990⁵.

Akerlof s'intéresse à la norme qui impose aux individus un salaire équitable, soit $\{w = \bar{w}\}$. Il est supposé que ce salaire équitable ("fair wage") est plus élevé que le salaire walrassien. L'existence de cette norme n'a rien d'évident puisque il est dans l'intérêt de certaines personnes d'y désobéir, à savoir ceux qui n'ont pas d'emploi et qui préféreraient travailler pour un salaire inférieur à \bar{w} que demeurer chomeur ; et les employeurs qui seraient prêts à embaucher ces chomeurs à un salaire inférieur au salaire équitable. Par ailleurs, si l'on admet le fait que, lorsque certains ne respectent pas la norme, cela entraîne un amoindrissement de son rayonnement, cela conduit à nous interroger sur la plausibilité du maintien de cette norme. Examinons la réponse qu'Akerlof apporte à cette question.

On note x la proportion des individus qui obéissent à la norme et μ la proportion de ceux qui croient à la justesse de la norme. Notons qu'on peut respecter la norme même si l'on ne croit pas à sa justesse, comme on peut la transgresser même si l'on la juge juste. L'idée centrale est alors de considérer que la fonction d'utilité se définit comme la somme de deux termes : d'une part V , l'utilité classiquement prise en compte par les économistes ; d'autre part, une utilité d'une nature spéciale qui modélise les effets

⁵En fait Akerlof emploie peu le terme de norme. Il parle le plus souvent de "social custom" ou de "code of behavior". Mais l'accent sur le rôle des effets réputationnels montre bien qu'il s'agit d'une norme au sens qu'Elster donne à ce concept.

sociaux et psychologiques particuliers provoqués par la norme, à savoir les sanctions sociales et la force émotionnelle intrinsèque.

Akerlof propose, pour les travailleurs, la fonction d'utilité suivante :

$$U = V + cR - d^R d^C \bar{C} \quad (11)$$

avec d^R , une dummy égale à 0 si le travailleur considéré obéit au code de conduite de la communauté, et égal à 1 s'il désobéit ; d^C , une dummy égale à 1 si l'individu croit au bien fondé de la norme et égale à 0 sinon ; R , la réputation ; c et \bar{C} sont des paramètres positifs. On note que l'action de la norme passe par deux effets de nature différente. Le terme \bar{C} mesure la force émotionnelle propre à la norme. Elle est indépendante du comportement des autres : une personne qui croit à la norme ($d^C = 1$) et qui lui désobéit ($d^R = 1$) ressent une perte de \bar{C} en termes d'utilité. Cet effet ne joue pas pour les individus qui ne croient pas à la pertinence de la norme ($d^C = 0$) puisque, dans ces conditions, ne pas la respecter ne crée pas de remords ou de honte personnels. L'amplitude de \bar{C} dépend cependant de la norme considérée. Pour certaines d'entre elles, intimement liées à notre culture, le coût intérieur de passer outre peut être très élevé (ne pas voler, ne pas tuer, etc.). Pour d'autres, moins essentielles, \bar{C} peut être faible (aller voter, ...). Le terme R exprime l'effet de la désapprobation sociale vis-à-vis d'un individu qui ne respecterait pas le code de conduite sociale $\{w = \bar{w}\}$. Ce terme ne dépend que du respect ou non de la norme ; le fait que l'agent considéré croit ou non dans la norme n'a ici aucune importance. Le modèle d'Akerlof analyse cette désapprobation sociale comme un effet réputationnel. Autrement dit, un individu qui viole la règle subit une perte d'utilité au travers des effets réputationnels négatifs qu'engendre sur sa personne le non respect du code social. Par exemple, beaucoup de personnes choquées par sa conduite n'accepteront plus de transacter avec lui. Pour ces raisons l'activité économique de l'individu en question sera rendue plus difficile, ce qui réduira d'autant ses gains. C'est l'ensemble

de ces effets que mesure R .

Une hypothèse importante est celle qui détermine la forme de la fonction de réputation. Elle dépend bien évidemment du fait que l'individu suit ou ne suit pas la norme, mais surtout elle dépend du nombre de personnes μ qui croient ou non à la pertinence de cette norme. Nous y reviendrons. Akerlof écrit :

$$R = R(d^R, \mu) \tag{12}$$

avec R , fonction croissante de μ . Il retiendra plus spécifiquement la forme suivante :

$$R = -\mu d^R \bar{R} = \begin{cases} 0 & \text{si l'agent obéit à la norme} \\ -\mu \bar{R} & \text{si l'agent désobéit} \end{cases} \tag{13}$$

Comment fonctionne alors le modèle ? Au début de la période, μ est donné. Puis les individus maximisent leur utilité (11) et choisissent soit d'obéir à la norme, soit de lui désobéir, en fonction du meilleur gain. On choisira de violer le code social si les gains attendus en termes d'utilité V dépassent la perte de réputation R et de \bar{C} , pour ceux qui croit à la justesse de la norme. Très clairement la désobéissance sera plus rare chez ces derniers ($d^R = 1$), car ils doivent alors faire face à une double perte d'utilité. Une fois ces choix faits, on calcule x la proportion des individus qui respecteront la norme. Cela permet de déterminer l'évolution future de l'adhésion à la norme selon la formule :

$$\dot{\mu} = g(\mu, x) = \beta(x - \mu) \tag{14}$$

L'idée est que μ augmente si $x > \mu$. Akerlof montre alors que, sur la base de cette ensemble d'hypothèses, la norme d'équité peut se maintenir malgré la présence de chômeurs sous certaines valeurs des paramètres.

Ce modèle est très intéressant même si beaucoup en ont critiqué le caractère trop orthodoxe. Pourtant, il nous semble qu'utiliser la fonction d'utilité pour intégrer à

l'analyse les effets des normes conduit à des résultats satisfaisants. Pour le voir, nous considérerons successivement les trois mécanismes essentiels correspondant aux trois équations : l'obéissance aux normes (11) ; la force des normes (12) ; et l'adhésion aux normes (14).

(a) L'obéissance aux normes. L'existence de normes ne transforme pas les hommes en robot. La vie quotidienne montre au contraire que nous faisons un usage raisonné des normes. La force de l'équation (11) est de modéliser un tel comportement. Lorsque les normes imposent des choix trop en désaccord avec nos intérêts, nous passons outre les impératifs normatifs. C'est ainsi que je n'hésiterai pas à partir en vacances, même si une échéance électorale tombe au milieu de ces vacances. Cela est vrai également des normes les plus ancrées, les plus sacrées : les passagers d'un avion perdu dans les Andes décidèrent de manger de la chair humaine. Notons, dans cet ordre d'idées, que toutes les religions autorisent à contourner certains interdits religieux quand l'intérêt individuel ou collectif est en jeu. Autrement dit, il existe une norme supérieure qui autorise, et quelquefois contraint, à violer les règles sociales dans certaines situations exceptionnelles. Ainsi, dans les faits, il existe des marges de choix entre intérêts individuels et respect des normes. La fonction d'utilité utilisée intègre cet usage intelligent des normes. Notons que, dans le cadre du modèle d'Akerlof, le choix auquel les individus sont confrontés est binaire : agir selon la norme ou violer la norme. Dans certains contextes, l'interaction entre norme et rationalité peut être plus complexe, comme le montre le jeu suivant.

On demande à deux individus A et B de se partager une somme de 10 francs, que leur donne le meneur de jeu, en respectant la procédure suivante. C'est A qui commence par proposer un partage. Il revient alors à B d'accepter ou de refuser. Si B accepte, la somme est partagée conformément à la proposition de A ; si B refuse le partage, les deux agents ne reçoivent rien. Les études expérimentales démontrent que le plus souvent A propose une division égale, 5 francs chacun. Quant aux individus

B, ils rejettent tout partage qui leur donneraient moins de 2 francs. On constate ici le résultat d'une interaction complexe entre la norme d'égalité et la rationalité. En effet, si les individus étaient totalement rationnels, B accepterait tout partage lui donnant une somme positive puisqu'un centime est mieux que rien. Cela n'est pas le cas puisqu'il refuse tout montant inférieur à 2 francs. Les individus ne sont pas non plus totalement mus par la norme d'égalité. Si tel était le cas, B refuserait tout partage inférieur à 5 francs. En fait il accepte le partage inégale 8 francs pour A et deux francs pour B. Autrement dit, le comportement observé est un mixte des deux comportements.

(b) La force des normes. Le point ici essentiel est que toutes les normes ne sont pas équivalentes ; leur intensité est variable, y compris dans le temps. Il existe deux sources de variabilité : Akerlof distingue un effet réputationnel, mesuré par \bar{R} , d'un effet propre à la force intrinsèque de la norme, mesuré par \bar{C} . Il est en effet clair que certaines normes sont plus fondamentales que d'autres et qu'en conséquence, leur violation provoque d'importants troubles psychologiques intérieurs. C'est le cas, par exemple, de l'interdit du cannibalisme ou du meurtre. Cette première composante variable est mesurée par \bar{C} .

D'autre part, l'équation (12) considère que l'effet réputationnel est lui-même variable. Il dépend du nombre de personnes qui adhèrent à la norme. Autrement dit, la force d'une norme, sa capacité à produire des sanctions négatives, est liée à la force des convictions sociales qui la soutiennent. Ce qui est mis ici en avant est le rôle de l'exemplarité. La nature sociale des normes impose une adhésion réelle des sociétaires. C'est là un point important qui permet de comprendre l'obsolescence de certaines normes. Cela s'observe, par exemple, dans le cas de la corruption : le fait de savoir qu'autour de nous beaucoup de personnes acceptent la corruption produit une dynamique cumulative qui débouche sur une destruction de la norme, en tant que norme sociale. Elle peut demeurer en tant que norme légale, mais son action est

alors moins forte. Elle reste à la surface du corps social. On voit ici que la normativité, contrairement à une vision naïve, n'est pas un phénomène de tout ou rien. Les gradations quantitatives ont des effets qualitatifs.

(c) L'adhésion aux normes. Il s'agit d'un processus endogène où l'exemplarité joue à nouveau un rôle déterminant. A long terme, une norme qui n'est pas obéie finit par perdre son caractère normatif. C'est ce qu'exprime l'équation (14).

Le dilemme du prisonnier

L'intégration des sanctions à la manière d'Akerlof permet d'étendre les outils de formalisation à des situations qui ne sont plus du type coordination pure, comme dans le cas des conventions, mais à des situations du type dilemme du prisonnier. Cela permet de comprendre pourquoi certaines conduites pénalisantes du point de vue du calcul de l'utilité sont néanmoins suivies. Akerlof a traité du cas du salaire équitable. Dans les lignes qui suivent, nous nous proposons d'étendre cette réflexion au dilemme du prisonnier, tout en restant dans un cadre de jeux évolutionnistes.

	<i>C</i>	<i>D</i>	
<i>C</i>	<i>R, R</i>	<i>S, T</i>	(15)
<i>D</i>	<i>T, S</i>	<i>P, P</i>	

tel que $T > R > P > S$. Utilisant le modèle évolutionniste, on obtient :

$$\begin{cases} u_C(p) = p(R - S) + S \\ u_D(p) = p(T - P) + P \end{cases} \quad (16)$$

avec $p = p_C$, la proportion d'individus ayant choisi la stratégie de coopération, notée C . On vérifie immédiatement qu'on a toujours u_C inférieure à u_D ; ce qui n'est

qu'une autre manière d'établir que la défection est une stratégie dominante. Toutes les dynamiques d'apprentissage vérifiant (5) convergent alors vers $\{p = 0\}$, l'unanimité sur la défection. Que se passe-t-il si des effets réputationnels sont pris en compte ? Le fait qu'une stratégie acquiert une certaine audience transforme la perception qu'on a d'elle. Elle n'est plus uniquement perçue par les gens au travers de l'utilité qu'elle permet d'obtenir. Elle acquiert de facto un caractère de normalité qui peut donner naissance à des effets réputationnels. On peut alors réécrire les équations (16). On obtient :

$$\begin{cases} u_C(p) = p(R - S) + S + R(p) \\ u_D(p) = p(T - P) + P + R(1 - p) \end{cases} \quad (17)$$

avec $R(z)$ une fonction croissante en z . Comme on le voit, cette formalisation ne dit pas que la stratégie de coopération est, d'un point de vue normatif, supérieure à la défection, ce qui reviendrait à présupposer un certain type de normes. Le modélisateur ne prend pas parti pour telle ou telle stratégie. On tient compte simplement d'un effet réputationnel qui, au travers de l'adhésion à une certaine stratégie, conduit à une entraide entre les individus qui suivent la même conduite. Cela n'a rien d'un effet automatique. Justifier son introduction nécessite une investigation plus poussée. En effet toute conduite n'a pas prétention à engendrer de tels effets réputationnels. Cependant si cette vision normative des stratégies est supposée plausible dans le cas étudié, on peut alors aisément montrer que si :

$$R(1) - R(0) > T - R$$

alors la coopération est un équilibre de la dynamique évolutionniste. La condition précédente s'interprète aisément. Elle exprime que le gain que permet la défection, à savoir $(T - R)$, est inférieur à la perte de réputation qu'engendre le fait d'avoir recours à la défection dans un monde dominé par les conduites coopératives.

L'analyse que Harvey Leibenstein (1982) propose de la convention d'effort est une illustration d'un tel résultat, même si l'auteur n'en propose pas véritablement une modélisation. Leibenstein considère la relation employeur/employé. L'employeur détermine le niveau de salaire, w ; et l'employé détermine son niveau d'effort, π . Il montre qu'il s'agit là d'un dilemme du prisonnier. Un comportement individualiste de maximisation conduit à l'équilibre de Nash pareto-dominé (w_{\min}, π_{\min}) . L'observation des firmes montre que les entreprises réussissent à éviter cette situation inefficace. On y observe la création d'une convention d'effort stable (w^*, π^*) . C'est là la conséquence du mécanisme qui vient d'être présenté. Plus précisément, Leibenstein décrit un processus social, interne à l'entreprise, qui conduit à la formation de cette norme d'effort. Elle est alors l'objet d'une adhésion volontaire des employés. De ce fait, le groupe en contrôle le respect et l'apprend aux nouveaux salariés. C'est la pression du groupe ("peer group pressure") qui stabilise la convention d'effort au travers des effets réputationnels qu'il engendre.

CONCLUSION

Les réflexions du présent article trouvent leur origine dans un programme de recherches qui milite pour une approche formalisée des formes institutionnelles. Elles expriment le désarroi d'un chercheur face à des modélisations trop naïves et trop caricaturales qui ne permettent pas d'appréhender certains des mécanismes les plus fondamentaux à la base des institutions. Deux voies d'approfondissement sont alors apparues particulièrement prometteuses, d'une part autour des croyances collectives, d'autre part autour des sanctions. Il s'en faut de beaucoup que ces pistes de recherche est fait l'objet, dans cet article, d'un traitement satisfaisant. Il s'agissait essentiellement pour nous d'en faire sentir l'intérêt. Très clairement, leur développement suppose un renouvellement assez important des approches traditionnelles par l'accent qui y est mis sur le rôle social de la cognition.

REFERENCES

- [1] AKERLOF, George A. (1990), “A Theory of social custom, of which unemployment may be one consequence”, *Quarterly Journal of Economics*, 94 (juin), 599-617.
- [2] AXELROD, Robert (1984), *The evolution of cooperation*, Basic Books.
- [3] BOYER, Robert et ORLEAN, André (1994), “Persistence et changement des conventions”, in *Analyse économique des conventions* dirigé par ORLEAN A., Presses Universitaires de France.
- [4] BOYD, Robert et RICHERSON, Peter J. (1993), “Rationality, imitation, and tradition”, in *Nonlinear dynamics and evolutionary economics*, édité par Richard H. DAY et Ping CHEN, Oxford University Press.
- [5] DOUGLAS, Mary (1989), *Ainsi pensent les institutions*, Usher.
- [6] ELSTER, Jon (1989), *The cement of society*, Cambridge University Press.
- [7] LEIBENSTEIN, Harvey (1982), “The Prisoners’ dilemma in the invisible hand: an analysis of intrafirm productivity”, *The American Economic Review*, 72 (mai), 92-97.
- [8] MAYNARD SMITH, John (1982), *Evolution and the theory of games*, Cambridge University Press.
- [9] SCHELLING, Thomas (1960), *The strategy of conflict*, Harvard University Press.
- [10] SETHI, Rajiv (1996), “Evolutionary stability and social norms”, *Journal of Economic Behavior and Organization*, 29 (janvier), 113-140.
- [11] SUGDEN, Robert (1986), *The evolution of rights, cooperation, and welfare*, Basil Blackwell.

- [12] THEVENOT, Laurent (1989), "Equilibre et rationalité dans un univers complexe", *Revue Economique*, 40 (2), 147-197.
- [13] WEBER, Max (1995), *Economie et société*, collection Agora, Plon.
- [14] YOUNG, Peyton H (1996), "The economics of convention", *Journal of Economic Perspectives*, 10 (2), spring, 105-122.