

the Soc 0154
Université des Sciences et Technologies de Lille
Faculté des Sciences Economiques et Sociales

REGLES ET RATIONALITE ECONOMIQUE

Thèse de doctorat nouveau régime
Sciences Economiques

Présentée et soutenue publiquement par
Nicolas POSTEL
le 21 janvier 2000

Directeur de Thèse :

M. Henri PHILIPSON, Maître de Conférences à l'Université de Lille 1

Membres du Jury :

M. Eric BROUSSEAU, Professeur à l'Université de Nancy II, Rapporteur

M. Olivier FAVEREAU, Professeur à l'Université de Paris X-Nanterre, Rapporteur

M. Roger FRYDMAN, Professeur à l'Université de Paris X-Nanterre

M. Jean GADREY, Professeur à l'Université de Lille 1

à Laetitia

Remerciements

Je tiens tout particulièrement à remercier Henri Philipson, dont l'aide nous a été précieuse. Ses conseils et sa capacité d'écoute ont été déterminants pour l'élaboration de ce travail. Sa grande disponibilité et son soutien permanent ont été des ressorts essentiels de la dynamique de notre recherche.

J'ai bénéficié, à de nombreuses étapes, des conseils fructueux d'Arnaud Berthoud, Laurent Cordonnier, Olivier Favereau, Jean Gadrey, Benoît Lengaigne et Franck Van de Velde, et je leur en suis grandement reconnaissant. J'ai également pu présenter certains passages de ce travail dans le cadre des séminaires du *Clerse*, et plus particulièrement dans le cadre des groupes "*Histoire de la pensée*" et "*Service, Innovation, Evaluation*". Je remercie l'ensemble des participants à ces deux équipes de recherche pour leurs encouragements et leurs critiques.

Ce travail a par ailleurs été réalisé dans les meilleures conditions grâce au soutien logistique de l'*IFRESI* et de tous ses membres.

J'ai enfin une pensée toute particulière pour Zoé, Noémie et Luna.

Introduction Générale

“A la profondeur de l'essence correspond le besoin profond de convention.”
(Wittgenstein, 1937-1944, p.65)

La règle est partout présente dans la *vie* économique. Elle n'est presque nulle part dans la *théorie* économique. C'est à partir de ce paradoxe que s'est construit ce travail.

D'une certaine façon, notre étude traque une absence, elle cherche à retrouver un concept largement “refoulé” de l'analyse théorique. Pourtant même s'il est bien peu question de la règle en économie, le problème que résout la règle est au centre de la problématique de la théorie économique moderne. La règle est en effet un dispositif assurant la coordination des individus. Or, l'économiste cherche précisément à concevoir un dispositif permettant de coordonner les actions des individus le plus efficacement possible, de manière à utiliser au mieux une certaine quantité de ressources rares.

Au cœur de cette question se trouve la contradiction fondamentale entre la liberté des *individus* et la nécessité d'atteindre un état *collectivement* satisfaisant. Le but poursuivi par l'économie théorique est en effet de permettre à chaque individu d'atteindre le niveau de bien-être le plus satisfaisant possible. Mais cette question ne peut se poser qu'au niveau collectif : il y a une certaine quantité de ressources dont la société doit tirer un maximum de richesses et qu'elle doit ensuite répartir au mieux. Il faut donc parvenir à ce que la coordination de chacun des individus permette d'atteindre un état collectivement satisfaisant. Une fois cet état “optimal” atteint, la répartition des richesses devrait permettre à chacun des individus d'obtenir un maximum de richesses (compte tenu de la contrainte que représente le caractère limité des ressources). Cependant, lors de ce va-et-vient entre les niveaux individuel et collectif, une certaine forme de contradiction se noue entre ce qui préoccupe chacun des individus et ce qui préoccupe l'économiste. En effet, chaque individu souhaite être libre d'utiliser au mieux ses propres ressources et l'économiste se place lui du point de vue collectif en cherchant à comprendre comment utiliser au mieux l'ensemble des ressources disponibles. Entre la liberté de l'individu et la contrainte macroéconomique, une tension naît ; entre la rationalité individuelle et la rationalité collective, une disjonction se produit.

Cette opposition n'est cependant qu'apparente : l'individu a besoin des autres. Il ne peut atteindre seul ses objectifs propres : la liberté individuelle doit être conciliée avec une

certaine forme de contrainte collective. La question qui se pose est donc de savoir comment concilier la liberté et le désir de chaque individu avec la nécessité de s'organiser collectivement, et donc, fatalement, de se contraindre. Cette question de la coordination des activités économiques pose immédiatement le problème des repères collectifs sur lesquels les individus doivent s'entendre pour parvenir à se coordonner efficacement. Le choix de ces repères dénoue la tension entre individuel et collectif. La question qui se pose alors est la suivante : "quels repères collectifs doit-on choisir pour que l'activité de chacun permette à tous d'atteindre un niveau de bien-être maximum ?" Tel est, exprimé de manière concise, le "*problème de la règle*".

La modalité marchande et l'émergence du "problème de la règle"

Ce problème ne se poserait cependant pas si la théorie économique pouvait se passer de la notion de règle. Notre travail ne serait d'aucune utilité s'il était possible de montrer que les individus, lorsqu'ils agissent librement, atteignent spontanément un état collectif optimal. La science économique moderne a ainsi tenté de démontrer qu'une société composée uniquement d'individus rationnels, exclusivement préoccupés par la maximisation de leur propre bien-être, pouvait spontanément atteindre un état collectif stable répondant à des critères précis d'optimalité. L'opérateur de ce "miracle" est bien sûr le mécanisme marchand, qui permet d'articuler les niveaux de rationalité individuel et collectif sans les opposer.

Les résultats que la théorie économique centrée sur le marché est parvenue à obtenir dans son analyse de la coordination "spontanée" entre des individus égoïstes et rationnels constituent *a priori* un démenti à l'idée selon laquelle la prise en compte des règles serait nécessaire à une bonne compréhension de l'articulation de la rationalité des individus et de la rationalité collective. Pour justifier de l'importance ou de l'urgence d'une prise en compte des règles, il nous faut donc tout d'abord montrer en quoi les analyses théoriques visant à faire "l'économie" de la notion de règle ne sont pas entièrement satisfaisantes. Nous chercherons donc à montrer que, d'une façon ou d'une autre, le "problème du choix des règles" refait surface dans la théorie économique néoclassique qui prétendait l'exclure.

La théorie néoclassique du marché cherche à éliminer le problème de la règle en montrant qu'il est possible de décrire le comportement des individus sous forme de loi, à compter du moment où ils sont situés dans un espace marchand (considéré comme une donnée naturelle). Lorsque le comportement des individus suit une loi qui assure la cohérence collective des comportements individuels, la question du choix des règles

devient en effet sans objet. *L'énoncé d'une loi exclut nécessairement la question de la règle*, puisque la règle apparaît justement lorsqu'il n'existe pas de loi reliant mécaniquement les comportements individuels à un état collectif satisfaisant.

Cependant, le modèle de marché ne fonctionne qu'imparfaitement. On constate que le niveau collectif conserve une opacité minimale pour les individus, ce qu'exprime à la fois le théorème d'impossibilité de Arrow et ses réflexions sur l'incertitude. On cherchera donc à montrer qu'à travers l'étude des réflexions critiques d'Arrow, *le problème de la règle refait surface*.

La modalité organisationnelle et la question de la genèse des règles

Le marché n'est cependant pas l'unique moyen d'évacuer le problème de la règle. Il existe une seconde façon d'éliminer ce problème en énonçant une loi : c'est la modalité organisationnelle. L'étude de la modalité organisationnelle naît du constat selon lequel le marché ne suffit pas à coordonner les individus. Elle succède historiquement à la "modalité marchande". En effet, lorsque la loi marchande s'affaiblit, la question de la coordination collective se pose à nouveau. *La modalité organisationnelle consiste à imposer aux individus une certaine cohérence d'ensemble permettant de restaurer l'existence d'une loi de comportement*. C'est ce que nous observerons à travers la théorie néoinstitutionnaliste de Williamson.

Cette fois cependant, la loi découle explicitement d'une construction institutionnelle qui *contraint* les acteurs. Elle requiert donc une certaine forme de *coercition* initiale, sur le modèle du contrat social de Hobbes. Or, il est extrêmement difficile de faire de ce moment coercitif une résultante du choix libre d'agents rationnels. La modalité organisationnelle fait donc surgir un problème : celui de la genèse des structures institutionnelles qui imposent un comportement uniforme aux individus. C'est à ce problème que se confronte l'analyse néoinstitutionnaliste de l'émergence des institutions, en utilisant la théorie des jeux comme outil théorique. Cette analyse fait cependant apparaître qu'il existe une sorte d'opacité minimale du collectif : les règles institutionnelles ne peuvent être à la fois la *solution* aux problèmes de coordination qui naissent de l'incomplétude de la rationalité individuelle instrumentale et le *produit* spontané du comportement rationnel des agents.

Le moment de la genèse des organisations demeure donc inévitablement obscur puisque les raisons qui président à leur mise en place échappent à une représentation en termes de rationalité individuelle. Le basculement de la "liberté" marchande à la "contrainte" organisationnelle est un problème récurrent que la théorie ne parvient pas à traiter de manière satisfaisante.

En quelque sorte on ne parvient pas, avec le modèle de marché, à remonter des individus vers le collectif, et l'on ne parvient pas dans le mode organisationnel, à redescendre du collectif à l'individuel. Il manque une passerelle, un niveau intermédiaire, sans lequel l'articulation des niveaux individuel et collectif est problématique. Ce niveau, c'est justement la règle.

La prise en compte pratique des règles, et l'obstacle de l'hypothèse de rationalité individuelle

Il apparaît donc d'une part que le concept de règle est effectivement utile à la théorie économique et que d'autre part, sa prise en compte nécessite de considérer sous un autre angle le comportement individuel. Pour parvenir à comprendre comment les individus s'accordent sur des règles communes, il est en effet nécessaire de ne plus considérer que le comportement individuel est descriptible sous forme de loi. C'est en particulier la prise en compte de l'ancrage "*pratique*" de la rationalité des acteurs qui permet de concevoir le rôle et la genèse des règles. Une telle conception de l'individu implique de s'écarter d'une vision systémique de l'économie, que ce soit sous l'angle du système de marché ou du système organisationnel, et de s'appuyer sur une conception plus "*pragmatique*". Pour cette raison, il nous semble que l'analyse keynésienne, ou plus exactement l'univers théorique mis en place par Keynes, constitue un point d'appui à une théorie non-standard de la règle. Les difficultés que rencontre l'analyse néoinstitutionnaliste invitent à redécouvrir l'univers théorique de Keynes. Cette redécouverte est tout particulièrement visible en France, où elle se concrétise par l'émergence d'une "*économie des conventions*".

Toutefois cette "troisième" modalité de prise en compte du "problème de la règle" qui s'appuie explicitement, à la différence des modes marchand et organisationnel, sur le concept de règle, se heurte à une difficulté d'ordre analytique. La reconnaissance du concept de règle met sous tension l'hypothèse traditionnelle de rationalité économique. Or cette hypothèse est un des éléments fondamentaux de l'analyse économique. C'est pourquoi il est extrêmement difficile de proposer une théorie de la règle ayant une forme analytique rigoureuse : ce concept remet en cause une des hypothèses centrales permettant l'analyse. Ainsi cette troisième modalité reste, elle aussi, ambiguë dans ses rapports avec l'hypothèse traditionnelle de rationalité, et incomplète dans son traitement analytique du concept de règle.

Il se pourrait d'ailleurs que cette incomplétude soit, d'une certaine manière, inévitable. La règle apparaît comme un élément tout à la fois indispensable à la compréhension du fonctionnement de l'économie, et difficilement intégrable dans le cadre d'une

représentation scientifique de cette économie. C'est finalement sur ce paradoxe que débouche notre recherche. La règle résiste à l'analyse économique, parce qu'elle met en cause les fondements théoriques d'une représentation modélisée. En ce sens elle pose un problème épistémologique au paradigme dominant et représente plus généralement un obstacle à la constitution historique de l'économie comme science positive.

Ce travail se déroulera donc en trois temps. Dans un *premier temps*, nous chercherons à montrer de quelle manière le problème de la règle émerge à partir d'un modèle théorique qui cherchait à l'exclure. Ce premier temps justifie que l'on s'intéresse de près à la question du choix des règles. Dans un *second temps*, nous étudierons une première réponse qui consiste à montrer que l'accord collectif sur le choix d'une structure organisationnelle coercitive élimine *de facto* la question du choix des règles. Il apparaît cependant que, là encore, la question de la règle ne peut pas être éliminée une fois pour toutes sauf à faire dépendre la coordination de structures reposant sur une forme de rationalité inaccessible aux acteurs, ce qui est problématique dans une optique individualiste. C'est pourquoi il nous semble utile, dans un *troisième temps*, d'étudier une autre modalité qui "prend au sérieux" le problème de la règle en cherchant à éclairer le mécanisme pratique de la construction des repères collectifs nécessaires à la coordination.

Puisqu'au fil de cette étude nous nous référerons en permanence à un concept de règle, il semble donc utile d'en donner dès maintenant une définition précise.

Définition du concept de règle

Nous utilisons le concept de règle dans un sens proche de celui qu'a défendu Wittgenstein. C'est notamment à travers la lecture qu'en propose Bouveresse [1987] que nous sommes parvenus à nous forger une représentation de ce que pouvait signifier le concept de règle. La règle est, selon une première définition très générale, un repère collectif progressivement mis en place par les acteurs pour parvenir à se coordonner. La nature de ces repères collectifs repose cependant sur un paradoxe : ils sont indispensables, mais leur choix n'est dicté par aucune forme de nécessité. Pour cette raison, la règle est *contingente, signifiante, nécessaire et contraignante*.

Dire que la règle est *contingente* implique qu'elle n'apparaît que lorsque les individus sont libérés de toute forme de nécessité. On ne peut pas dire que le fait de manger pour vivre soit une règle : c'est une nécessité naturelle. Cette forme de nécessité

peut s'exprimer sous forme de loi. C'est la tâche même de la science que d'énoncer sous forme de lois les relations qui unissent nécessairement une même cause à un même effet. Ainsi une pomme qui se décroche d'un arbre tombe nécessairement à la verticale. Si elle pouvait parfois dévier de cette trajectoire selon ses fantaisies et remonter vers le ciel, l'ensemble de notre connaissance du monde en serait affectée. En ce sens, il existe des lois naturelles nécessaires qui s'imposent à nous. Dans ce domaine, il ne peut exister d'espace pour la règle. On ne décrète pas de règle afin de réguler les mouvements des planètes. Par contre, il est possible d'édicter des règles définissant la priorité routière, ou le sens de circulation des automobiles. On peut donc opposer la notion de règle à celle de loi naturelle. Une conséquence importante de cette conception de la règle est qu'il n'existe de règle que lorsque l'on a le "choix". C'est lorsque les individus sont *collectivement* libres de se coordonner d'une manière ou d'une autre qu'ils peuvent se mettre d'accord sur des règles communes. C'est uniquement dans la mesure où "l'organisation" choisie par les acteurs n'est en aucun cas imposée par une nécessité d'ordre naturel ou transcendant, que cette organisation peut être considérée comme une règle.

Le choix d'une règle est *signifiant*. Il nous semble important de ne pas confondre le caractère contingent d'une règle et son caractère "fortuit" ou "arbitraire". Dire que les règles choisies ne reposent sur aucune forme de nécessité ne signifie pas qu'elles ne revêtent pas un sens ou une importance particulière en elles-mêmes. En ce sens, le concept de règle se distingue de l'idée de convention. En effet, la caractéristique d'une convention est d'être "arbitraire", et si l'on considère que le choix de la règle est *fortuit ou arbitraire*, on perd de vue ce qui explique ce choix. Wittgenstein distingue ainsi nettement "contingence" et "arbitraire" : "*vous pourriez dire que le choix des unités de mesure est arbitraire. Mais en un sens important il ne l'est pas. Il y a une raison très importante, qui réside à la fois dans la taille, dans l'irrégularité de forme et dans l'usage que nous faisons d'une pièce, qui fait que nous ne mesurons pas ses dimensions en microns ou même en millimètres. La proposition nous indique le résultat de la mesure, mais la méthode et l'unité de mesure choisie, nous informe également sur le monde dans lequel cette mesure est effectuée.*" (Wittgenstein, cité par Bouveresse, 1987, p.109).

Ainsi le choix des règles, pour être contingent, c'est-à-dire sans nécessité formelle, n'en est pas moins signifiant. C'est ce sens qu'il convient de découvrir lorsque l'on s'intéresse aux règles économiques, et en particulier à leur dynamique. Les règles ont une dimension éthique et historique qui leur donne de "l'épaisseur". Une analyse satisfaisante des règles devrait précisément percer les raisons et le mécanisme qui président à leur choix, à leur "justification". Face à une situation dans laquelle plusieurs choix collectifs sont possibles, comment les acteurs économiques vont-ils arbitrer et pourquoi ? Quels

sont les mécanismes par lesquels les règles qui encadrent le processus de production et d'échange évoluent ? Comment les individus influent-ils sur ces règles, et en quel sens ? Prendre en compte la signification de la règle est un aspect essentiel lorsque l'on cherche à rendre compte du mécanisme ou de l'histoire de leur *émergence*.

Bien qu'elles soient contingentes, les règles sont absolument *nécessaires*. Sans règle commune, il serait en effet impossible de se comprendre. C'est même précisément de leur contingence que les règles tirent leur nécessité. En effet, lorsque le comportement des hommes est libre de toute loi, transcendante ou naturelle, il est aussi parfaitement indéterminé. De cette indétermination naît un problème de coordination. La mise au point de règles communes permet d'assurer une coordination des acteurs, et d'obtenir une certaine forme de "rationalité" collective. C'est le cas, en premier lieu, des règles de langage sans lesquelles tout accord, et donc toute nouvelle règle, ne saurait exister. C'est également le cas, à un niveau moins fondamental, des règles qui encadrent la circulation routière, l'éthique médicale, le code pénal, ou bien encore la production et l'échange. Lorsque l'individu est dégagé de toute nécessité, il lui est nécessaire de disposer des règles qui lui permettent d'agir. Plus précisément, ces règles sont nécessaires dans la mesure où l'individu désire agir avec les autres ou dans un espace public dans lequel son comportement influe sur, et est influencé par, les actions des autres¹. Si les individus ne souhaitaient pas agir avec les autres, et produisaient des richesses de façon solitaire, sans jamais entrer en contact les uns avec les autres, les règles seraient inutiles. Le besoin de règle naît de l'*incertitude* inhérente à l'interdépendance des acteurs. Dans le domaine économique, il *paraît* évident, à première vue, que les acteurs sont en contact les uns avec les autres et ont besoin de se coordonner. C'est de ce besoin que naît la *nécessité* de la règle.

Enfin la règle est *contraignante*. S'il existe une liberté attachée à l'idée de règle, c'est une liberté collective. La société peut choisir les règles sur lesquelles elle souhaite asseoir son organisation. Mais, une fois choisie, la règle s'impose à chacun. Les individus sont libres de se mettre collectivement d'accord. Ils sont ensuite contraints individuellement de respecter l'accord collectif.

La notion de règle-contingente repose donc essentiellement sur quatre caractéristiques :

- 1- Il est collectivement nécessaire de disposer de règles communes.
- 2- Le choix collectif d'une règle particulière est contingent. (C'est même précisément

¹ On notera que le concept d'action implique, en fait, l'existence d'un espace public dans lequel l'action fait sens. C'est en tout cas le fondement de l'action pour Arendt [1960].

parce que l'on a collectivement le choix que la règle est nécessaire.)

3- La règle est signifiante : elle s'insère dans une *histoire* collective qui lui donne du sens.

4- L'individu est cependant contraint par la règle. Une fois adoptée, il doit la respecter. Elle devient donc nécessaire à ces yeux.

Pour résumer ces quatre points : *il est nécessaire que la collectivité des acteurs s'accorde sur une contrainte commune, mais que ce choix ne repose en lui-même sur aucune forme de nécessité, car il découle d'une histoire qui rend certaines règles plus acceptables que d'autres.*

Si la règle est imposée aux acteurs, en vertu d'une nécessité transcendante ou naturelle, ce n'est pas une règle, c'est une loi. C'est précisément parce que les individus ont parfois le "choix" que les règles sont nécessaires. Du même coup surgit ce que nous qualifions de "*problème de la règle*" : pourquoi une règle particulière est-elle choisie ? Rendre compte des règles, c'est donc non seulement rendre compte du rôle des règles, mais également des mécanismes par lesquels les acteurs en sont venus à choisir ces règles en particulier.

Précisions méthodologiques

Avant de préciser le plan précis que nous suivrons, il nous semble utile de préciser la nature de notre travail, qui justifie la forme particulière que prend notre réflexion.

Il s'agit, tout d'abord, d'une réflexion de nature purement théorique. Nous ne mobilisons pas de "fait" particulier pour justifier notre démarche. C'est, précisons le immédiatement, un aspect que nous regrettons, dans la mesure où notre attention a été portée au concept de règle pour des raisons "empiriques". C'est en particulier la difficulté que l'on rencontre pour décrire et comprendre les relations de travail lorsque l'on ne dispose pas du concept de règle, qui a été l'élément déclencheur de cette recherche². Nous avons souhaité comprendre pourquoi la théorie évacuait le concept de règle, et saisir les difficultés que posait ce concept. Or il nous est rapidement apparu que cette question était une question exclusivement théorique, et que l'appui sur les faits empiriques n'éclairerait pas les raisons de ce blocage théorique. Il est assez aisé de désigner, dans n'importe quel

² Bien que cet aspect apparaisse très peu dans la suite de ce travail nous tenons à souligner que c'est en étant confronté à la question de la représentation théorique de la "relation de service" qu'il nous a semblé essentiel de réfléchir à la manière de "formaliser" ou plus simplement de représenter les mécanismes par lesquels une collectivité d'acteur en viennent à définir les règles du jeu définissant le "produit" échangé. Sur ce point, on peut se reporter aux nombreux travaux menés par l'équipe de Jean Gadrey (par exemple : Gadrey [1990], Gadrey [1992], Gadrey et De Bandt [1994], Gadrey [1994] Gadrey [1995], Gadrey [1996])

aspect de la vie économique, l'existence de règles particulières. Mais cela ne nous permet pas de comprendre pourquoi la théorie n'en parle pas, ni même pourquoi elle *devrait* en parler. Or nous cherchons précisément à montrer que nous avons besoin, *en théorie*, du concept de règle, et que l'intégration de ce concept pose des problèmes théoriques difficiles à surmonter si l'on ne renouvelle pas profondément l'appareil analytique.

La forme que prend cette recherche est donc d'abord épistémologique. Mais elle est aussi, de façon étroitement liée, historique. Il s'agit en effet d'une thèse d'*Histoire de la Pensée Economique*, dans la mesure où nous cherchons à retracer la manière dont, ces soixante dernières années, le "problème" que constitue le choix des règles est devenu absolument central dans la théorie économique (sans que celle-ci n'utilise forcément le mot de règle pour désigner ce problème). Il s'agit d'une thèse *épistémologique* dans la mesure où notre recherche historique prend une forme analytique visant à cerner la manière dont le cadre théorique évolue progressivement en se heurtant au problème de la règle.

Le choix des auteurs étudiés et les liens que nous cherchons à faire apparaître entre eux n'ont donc nullement été guidés par un souci d'exhaustivité. Nous ne cherchons pas à balayer un champ théorique, en donnant une image la plus exacte possible des forces en présence. Au contraire, nous cherchons à suivre au plus près un nombre très restreint d'auteurs, dont le discours et le parcours théorique nous semblent significatifs du problème épistémologique que pose l'analyse de la règle. Si notre thèse n'est pas érudite, au sens où elle ne propose qu'un choix très partiel parmi les auteurs qu'un tel sujet aurait pu convoquer, elle ne se présente pas non plus comme une analyse systématique de quelques auteurs particuliers. Nous ne cherchons à résumer ni la pensée d'Arrow, ni celle de Keynes, ni celle de Williamson, ni celle des écoles conventionnaliste ou néoinstitutionnaliste. La lecture et l'analyse que nous proposons de ces auteurs est entièrement guidée par notre sujet d'investigation, elle ne prétend pas atteindre un autre niveau de généralité. Tous ces auteurs sont convoqués dans la mesure où leurs analyses théoriques nous semblent contenir quelque chose d'essentiel concernant la manière d'évacuer, ou de prendre en compte, le problème du choix des règles. Notre analyse est donc historique, puisqu'elle retrace la manière dont la théorie évolue à propos le traitement d'un problème particulier, et épistémologique, dans la mesure où elle cherche à relier cette évolution au questionnement des fondements méthodologiques des différents discours étudiés.

Pour cette raison, le plan que nous suivrons sera dans une large mesure chronologique. En effet, le programme de recherche néoinstitutionnaliste prend tout son sens au regard des problèmes rencontrés par la théorie de l'équilibre général.

L'opposition que Williamson fait apparaître entre marché et organisation repose sur la formulation par Walras, puis par Arrow-Debreu, d'une théorie de l'équilibre général reposant sur une conception précise du marché. De même, la démarche Keynésienne prend sens dans son rapport critique avec la théorie "Klassique", selon l'expression de Favereau [1985]. Plus précisément, l'analyse keynésienne de la coordination et de l'incertitude s'oppose à la conception mathématique du comportement individuel, en terme de maximisation de l'utilité espérée, que l'on retrouve au fondement de la théorie de l'équilibre général. On cherchera donc à faire apparaître la séquence théorique suivante : partant du modèle de marché walrassien, porté à son degré d'achèvement par Arrow-Debreu [1954], nous montrerons en quoi ce modèle se désagrège face à l'incertitude et à la question de la correspondance entre rationalité individuelle et collective. L'essor, dans les années soixante-dix, des théories néoinstitutionnalistes répond directement aux failles du modèle de marché, repérées en particulier par Arrow lui-même. Mais la réponse néoinstitutionnaliste mobilise une conception standard, ou quasi-standard, de l'individu qui ne peut pas lui permettre de comprendre dans quelle mesure l'organisation collective est "agie" de l'intérieur par les individus qui la constituent. De ce fait elle est comme "prisonnière" d'une conception hobbenne de l'ordre social. On cherche alors à mettre à jour une tradition de recherche alternative qui trouve son origine en économie dans l'analyse keynésienne de la coordination, et qui se caractérise par une conception plus "pratique" de la rationalité et de la coordination. C'est cette tradition que mobilise selon nous le mouvement récent de l'économie des conventions.

Plan de la thèse

L'exposé de la problématique de cette étude a déjà largement dévoilé le chemin que nous prendrons pour exposer notre thèse. Nous exposerons successivement les modalités marchande, organisationnelle et pragmatique de traitement du problème de la règle. Nous montrerons dans quelle mesure la question du "choix" des règles particulières de coordination reçoit trois réponses différentes, incomplètes, qui se succèdent et se répondent en raison des problèmes spécifiques que chacune d'elle pose.

La **première partie** présente la *modalité marchande*. Le premier chapitre expose la tentative néoclassique visant à formuler une loi qui rendrait toute règle inutile. Le second chapitre expose la critique "interne" formulée par Arrow, qui le conduit à formuler un programme de recherche préfigurant le mouvement néoinstitutionnaliste.

Premier chapitre – Le moyen qu’a privilégié l’économie pour rendre compte de l’articulation entre rationalité individuelle et ordre collectif est le modèle de marché. La théorie néoclassique trouve son modèle théorique avec l’analyse walrassienne et son point d’orgue avec la formulation par Arrow et Debreu [1954] du premier théorème d’équilibre général. Le modèle de marché résout la tension entre le niveau collectif et le niveau individuel en se donnant une conception de l’individu qui fonctionne conformément à la loi marchande de l’offre et de la demande. Vus sous l’angle de leur comportement, tous les individus sont identiques et agissent conformément à cette loi. Ils ne sont jamais en contact les uns avec les autres. Leur seule différence, qui nécessite une coordination, est le caractère divergent de leurs désirs exprimés sous forme de paniers de bien. Cette différence est “gérée” par le système des prix incarné par le commissaire-priseur qui centralise les offres et demandes et “crie” un système de prix les rendant compatibles. Il y a une règle unique dans ce modèle : la règle qui fixe, institutionnellement, la “dotation initiale” de chaque individu. C’est une *contrainte* fixée de façon *exogène*, en dehors de l’économie. Elle n’intéresse pas l’économiste. Pour le reste, les individus ne sont pas en contact direct, et sont immédiatement “en phase” puisqu’ils se comportent tous selon la même loi. Cette loi est d’essence naturelle, c’est en tout cas ce que semble attester la référence permanente à la “main invisible” smithienne. Ce modèle décrit donc la coordination d’une infinité d’individus rationnels et égoïstes sans faire appel à la notion de “règle”, puisque le marché est régi par une “loi”. Cependant cette loi souffre de nombreuses limites. Elle ne fonctionne qu’à l’équilibre, et ne trouve pas d’application pratique. On ne sait pas comment atteindre cet équilibre, ni même s’il est stable une fois qu’on l’a atteint.

Second Chapitre – Arrow, à l’origine de la théorie moderne de l’équilibre général, formule deux lignes critiques de cette théorie qui visent à dépasser ces “blocages”. Il montre très tôt, par le biais de son théorème d’impossibilité, qu’il n’existe aucune correspondance entre la rationalité individuelle et la rationalité collective. Aucun mécanisme d’agrégation ne conserve la propriété de transitivité des ordres individuels rationnels. Le marché ne peut donc pas prétendre fixer lui-même ses propres fins, il ne peut fonctionner que comme un moyen visant à atteindre des fins déterminées de l’extérieur. Cette disjonction entre rationalité individuelle et collective, qui rejaillit sur la distinction fin/moyen, constitue une première voie de redécouverte du problème de la règle. Il faut en effet que les individus s’accordent sur certaines règles *a priori* pour que leurs ordres de préférences individuels soient rendus compatibles.

Sa seconde ligne critique porte sur l’incertitude. Arrow montre que le système de

marché est incomplet puisqu'il ne peut pas donner d'information sur les prix futurs. L'incertitude qui en découle provoque une hétérogénéité des anticipations des agents. Ceux-ci sont alors conduits à échanger entre eux leurs informations : ils entrent en contact et constatent l'hétérogénéité de leurs jugements sur l'avenir. Cette hétérogénéité pose un *nouveau* problème de coordination. Là encore, il est nécessaire que des règles communes permettent une diffusion de l'information et une harmonisation des anticipations.

Arrow indique alors une voie à suivre pour pallier ces deux problèmes : la socialisation des individus dans un espace leur assurant des *codes éthiques* communs (harmonisation des objectifs individuels et collectifs) et un *langage et des informations communes* (harmonisation des anticipations) permettrait de combler les failles du marché. C'est là le rôle des organisations. Cette critique d'Arrow réintroduit donc la question du rôle des règles communes dans la thématique néoclassique qui l'excluait. Toutefois, Arrow indique immédiatement une "solution" au problème de la règle en mettant en avant la solution organisationnelle. Il se contente cependant d'évoquer un programme de recherche qu'il ne développera pas lui-même, dans la mesure où il implique une critique plus nette du mode marchand.

La seconde partie présente la modalité organisationnelle de gestion du problème de la règle. Le troisième chapitre est consacré à l'analyse de Williamson. L'accent que met cet auteur sur le rôle de la hiérarchie l'empêche de saisir la dynamique individuelle des règles organisationnelles. Le quatrième chapitre est consacré à la théorie des jeux non coopératifs, dans sa branche néoinstitutionnaliste (Lewis, Schotter, Ullmann Margalit, Axelrod...). Ces développements sont complémentaires de ceux de Williamson, dans la mesure où ils visent à donner un *fondement individuel* aux règles collectives. Il apparaît cependant que la théorie des jeux ne parvient pas à percer l'opacité qu'ont les règles collectives du point de vue des individus "standard". De ce fait, cette théorie ne permet pas réellement de comprendre la dynamique des règles. L'analyse néoinstitutionnaliste reste donc fondamentalement statique.

Troisième chapitre – C'est au mouvement néoinstitutionnaliste qu'il revient d'avoir développé une théorie des règles économiques à partir du programme établi entre autres par Arrow.

On observe ce faisant un véritable renversement théorique. Cette fois en effet, l'analyse ne prétend plus déduire la rationalité collective de la rationalité individuelle, mais, au contraire, montre que la rationalité individuelle émane, via les "règles organisationnelles hiérarchiques", de la rationalité collective qu'incarne l'organisation.

Cependant les “règles” rationnelles dont il est question sont imposées aux acteurs et dominées par la *nécessité* de l’efficacité. Ce ne sont donc pas véritablement des règles, mais plutôt les éléments d’une “loi organisationnelle” imposée par le primat de l’efficacité collective.

On reconnaît donc, dans cette position théorique, la seconde modalité d’approche de l’ordre collectif visant à éliminer la question du choix des règles contingentes. C’est une modalité constructiviste héritée de Hobbes. Il s’agit cette fois d’imposer aux individus une série de règles rationnelles en vertu d’un principe supérieur non discutable. L’ordre collectif est donc artificiellement construit en dehors des individus qui s’y plient. Cependant cette modalité d’intégration du collectif pose un problème théorique essentiel : quelle est la rationalité individuelle qui fonde l’acceptation des règles assurant la rationalité collective ? C’est là le dilemme organisationnel³. Le renversement théorique qu’opère Williamson en faisant exclusivement dépendre la coordination de l’action d’un centre organisationnel omniscient s’oppose radicalement à l’individualisme méthodologique. Si on considère que les éléments collectifs sont toujours animés par des actions individuelles, l’analyse néoinstitutionnaliste est condamnée à être statique : on ne voit pas comment les individus pourraient fonder ou faire évoluer une structure qui leur donne les moyens d’être rationnels. C’est là un problème important pour cette “modalité organisationnelle” de traitement du problème de la règle.

Quatrième Chapitre - Les tentatives menées en théorie des jeux pour fonder l’existence des règles institutionnelles à partir de l’étude d’interactions récurrentes entre des individus rationnels visent précisément à résoudre ce problème. Elles mobilisent toutes, de façon plus ou moins explicite, le concept de convention formulé de manière rigoureuse par Lewis [1969]. Le concept humien de convention permet en effet d’intégrer le collectif sans la coercition que Hobbes et Williamson lui associent. Il dénoue donc, d’une certaine manière, le dilemme organisationnel. Toutefois cette “solution” est obtenue au prix d’un considérable assèchement du contenu des règles, qui dévient à peu près nul. La neutralité des règles conventionnelles est en effet liée à leur arbitraire. Cette tentative théorique ne permet donc pas d’éclairer les raisons pour lesquelles les individus choisissent les règles, puisqu’elle suppose que ce choix est fortuit. De ce fait, l’utilisation de ce concept en théorie des jeux ne répond pas véritablement à la question que pose la présence de règles collectives. Elle ne permet pas d’éclairer le mécanisme de l’accord des individus sur le choix d’une règle particulière, autrement qu’en supposant que cette règle était déjà là au départ, ou a été choisie par hasard...

Il semble donc que le maintien, dans le champ néoinstitutionnaliste, d’une conception

³ On reprend librement la formule de Ménard [1994].

classique de l'individu ne permette pas de comprendre la dynamique des règles.

La troisième partie vise à mettre en évidence l'existence d'une troisième modalité de compréhension de l'articulation de l'individuel et du collectif qui passe explicitement par l'utilisation du concept de règle. Le cinquième chapitre est consacré à l'analyse keynésienne de la coordination en situation d'incertitude. Il permet de défendre l'idée selon laquelle l'analyse de Keynes impulse une compréhension pratique de la coordination des acteurs économiques. Cependant, Keynes ne développe pas de théorie de la règle. C'est à travers le mouvement récent de l'économie des conventions qu'apparaît explicitement une théorie pragmatique de la règle. On cherchera, dans le sixième et dernier chapitre, à mesurer l'importance du projet et des réalisations conventionnalistes et, plus particulièrement, des problèmes que soulève cette stratégie théorique originale.

Cinquième Chapitre- Il existe une conception keynésienne de la rationalité, et elle est exposée dans le *Treatise on Probability*, qui précède la *Théorie Générale* d'une vingtaine d'année. Keynes fait dépendre la rationalité d'un comportement du sens pratique de l'acteur, c'est-à-dire de sa capacité à recueillir des informations pertinentes et à en déduire une vision raisonnable du futur. Cette définition de la rationalité, ou plus exactement de la "raison", implique que le choix raisonnable dépend étroitement de la personnalité de chacun des individus. Le comportement individuel ne peut donc plus être décrit sous forme de loi. Cette multiplicité des points de vue raisonnables qui peuvent être portés sur une même situation pose, en situation d'interdépendance, un problème de coordination. C'est ce qui apparaît dans la *Théorie Générale*.

Dans son analyse des déterminants de l'incitation à investir, Keynes montre que l'organisation particulière du financement de l'investissement par les marchés financiers tend à faire triompher le point de vue le moins "raisonnable" : celui des individus "ignorants". Ces individus ne disposent pas des informations qui leur permettraient d'agir raisonnablement, ils tendent donc à s'épier et s'imiter et ce processus mimétique donne naissance à des conventions largement arbitraires. Or, ces conventions déterminent d'une part le taux d'intérêt, d'autre part le rendement escompté des actions, et bloquent l'émergence du seul point de vue vraiment raisonnable (c'est-à-dire fondé sur une compréhension pratique) : celui des entrepreneurs. C'est là une des causes de la faiblesse du niveau de l'investissement, et, par voie de conséquence, de l'absence de plein emploi.

L'analyse de Keynes revêt, pour ce qui nous concerne, une double importance. D'une part, Keynes donne une plus grande liberté, et donc une plus grande indétermination au

comportement des individus. Il lui est donc possible de concevoir le rôle des règles, puisque les individus ne sont plus mus par une loi naturelle. D'autre part, Keynes propose une conception du marché qui met en évidence l'importance de l'organisation "humaine" de ces marchés, et éclaire le caractère contingent des règles d'organisation de ces marchés. Ces deux éléments théoriques, novateurs, lui permettent de mettre en évidence le rôle des règles dans l'arbitrage entre les points de vue des différents acteurs. Cependant, Keynes arbitre lui-même en faveur d'un de ces points de vue en donnant une importance particulière au regard "réaliste" de l'entrepreneur. Il plaide donc pour la mise en place d'une règle particulière, mais ne développe pas vraiment de théorie permettant de comprendre comment, en pratique, les individus déterminent collectivement des règles économiques. Son attachement au "réalisme" ne lui permet pas entièrement d'accéder au point de vue de Wittgenstein sur le caractère contingent de toute règle.

Sixième Chapitre— La réflexion de Keynes sur les règles n'a pas véritablement été prolongée ou poursuivie jusqu'à la redécouverte récente de cet aspect de l'économie keynésienne notamment par Favereau [1985,1988] et Orléan [1986,1988]. Il est cependant possible d'établir un lien entre les conceptions de Keynes, et celle d'Herbert Simon, en termes de rationalité procédurale. C'est à partir de cette hypothèse de rationalité, ainsi que de l'analyse keynésienne de la coordination, que s'est constituée l'économie des conventions. Cette école de pensée a mis au point un programme de recherche important, permettant d'envisager une meilleure compréhension non seulement du rôle, mais aussi du choix et donc de la dynamique des règles.

Toutefois, la réalisation effective du projet de recherche de l'économie des conventions est entravée par le caractère ambigu de son hypothèse de rationalité. En s'appuyant sur Simon, cette école de pensée reste en effet proche d'une conception instrumentale de la rationalité, bien que par ailleurs elle enrichisse considérablement l'analyse des raisons d'action des individus. C'est pourquoi l'économie des conventions peut difficilement mener à bien son projet de recherche. La dimension cognitive de la coordination, sur laquelle elle tend à se replier, ne recouvre pas, en effet, tous les motifs, y compris éthiques, qui président au choix des règles.

A travers les avancées et les difficultés du projet conventionnaliste, on identifie un problème fondamental concernant l'intégration du concept de règle dans le corpus de la théorie économique. L'idée de règle remet directement en cause la possibilité d'un comportement qui ne serait qu'instrumentalement rationnel. Elle questionne donc un aspect fondateur de la science économique moderne, déjà dénoncé par Polanyi : le primat

de l'instrumentalité. De ce fait, la question se pose de savoir dans quelle mesure la théorie économique ne se nie pas elle-même lorsqu'elle intègre le concept de règle. En quelque sorte la règle représente un obstacle théorique, mais aussi un fait empirique résistant, sur laquelle l'analyse économique bute inévitablement, dans la mesure où il remet en cause un de ses postulats fondamentaux, à savoir : une conception exclusivement instrumentale de la rationalité. La question qui se pose est alors de savoir dans quelle mesure la théorie peut, ou doit, se reconstruire sur une conception différente de l'action individuelle.

Première partie

La modalité marchande

La règle n'appartient pas au champ de connaissance naturel de l'économiste. En effet l'économiste ne connaît que des lois, il décrit un univers ordonné dans lequel, précisément, les agents sont débarrassés de la question du choix des règles. Cependant, le monde économique n'est pas d'emblée ordonné. Il se présente à l'économiste sous une forme chaotique. Cela ne signifie sûrement pas, qu'il n'existe pas de lois économiques, simplement : il est nécessaire, pour les découvrir, de les exhumer de l'enchevêtrement de règles et conventions sous lesquelles les hommes les ont dissimulées. Pour cette raison, le travail de l'économiste comporte une étape supplémentaire par rapport à celui du scientifique étudiant la nature. Il ne peut se contenter de décrire une réalité naturelle ordonnée. Il doit ordonner, selon des lois naturelles qu'il lui faut exhumer, une réalité économique et sociale qui ne l'est pas. Pour cette raison l'économiste raisonne à partir de modèles idéaux représentant l'économie non pas telle qu'elle est, mais telle qu'elle devrait être. Le modèle de l'économiste représente l'essence de l'économie, au-delà du désordre apparent de la réalité. Il lui faut donc *définir* l'économie, en produisant ce que Walras qualifie justement de "théorie pure". La première étape de ce travail consiste à établir l'existence d'une loi de comportement pouvant servir de base à la mise en place d'un système économique ordonné. C'est à l'expression de cette loi, et à la description du système dans lequel elle s'exprime que vise la théorie néoclassique. Or, exposer une loi est une manière de supprimer le problème de la règle.

Le premier chapitre décrit le mouvement par lequel la théorie néoclassique cherche à exposer la loi marchande. L'étude du mécanisme de marché a précédé la théorie néoclassique. Mais, le mécanisme fondamental de la main invisible n'a jamais été fondé en logique : c'est ce manque que la théorie néoclassique cherche à combler. Cet objectif

consiste tout d'abord à montrer que le marché est un espace dans lequel le comportement des individus peut être décrit par une loi. La formulation de cette loi devrait ensuite permettre de décrire un processus automatique de convergence vers un état social optimal. Autrement dit, la théorie néoclassique vise à libérer les individus de la question du choix des règles économiques. Le mécanisme marchand assure à lui seul la coordination des agents et leur permet d'atteindre le meilleur état collectif possible.

Le modèle Arrow-Debreu représente, dans cette perspective, le modèle le plus achevé, ou, pour reprendre une expression de Hahn [1986, p.209] : “*le bout du chemin*”. Il est, de ce point de vue, la quintessence de la science économique moderne. Le moment où les lois régissant l'économie sont exprimées dans toute leur clarté. C'est donc aussi une représentation du système économique qui exclut absolument l'idée de règle-contingente. Les actes des acteurs peuvent en effet être entièrement décrits sous forme de loi.

Le bout du chemin, pourtant, nous fait découvrir un paysage que nous n'imaginions pas. Hahn[1986, p.209] écrit ainsi : “*maintenant que nous y sommes, nous le trouvons moins révélateur que nous l'escomptions*”. En effet, la loi ne fonctionne qu'à l'équilibre. Elle n'assure pas que cet équilibre puisse, en pratique, être atteint. La théorie nous dit donc peu de chose, presque rien, sur la manière dont il est possible d'organiser la société afin d'atteindre un équilibre. Elle a peu de contenu pratique.

La "critique" d'Arrow, un des pères de la théorie moderne de l'équilibre général, vise précisément à redonner un contenu *pratique* à la théorie *formelle* de l'équilibre général. Arrow développe deux critiques du modèle de marché, à travers lesquelles, le problème de la règle refait surface...

Il montre tout d'abord, avec son *théorème d'impossibilité* qu'il n'existe aucun mécanisme permettant d'agrégier les préférences individuelles en préférence collective. Autrement dit, “*le mécanisme de marché ne peut pas constituer un choix social rationnel*” (Arrow, 1963, p.59). Cette impossibilité souligne la nécessité d'assigner au marché des fins collectives, déterminées en dehors de lui. Or, comme le montre Arrow cet accord nécessite en amont une certaine mise en cohérence des profils de préférence de chacun des individus. C'est là, précisément, que le problème de la *règle* réapparaît : comment harmoniser les préférences ? Arrow donne cependant un sens particulier à cette question. Il suppose qu'il existe un consensus "moral" qui permettrait de fonder un accord collectif, si toutefois le marché pouvait supprimer toute incertitude en faisant correspondre à chaque allocation de départ un état social donné. Il suffirait alors de déterminer l'allocation de départ correspondant à l'état collectif que, pour des raisons morales, chacun des agents souhaite voir se réaliser. Pour cela, il faut cependant que le marché puisse éliminer toute incertitude quant à l'avenir.

C'est l'objet du second travail critique d'Arrow. Le marché devrait dans l'idéal être capable de "gérer" l'incertitude et permettre aux individus de choisir une fois pour toute l'état qu'il souhaiterait voir se réaliser. Mais le marché est un moyen imparfait. Il ne donne pas aux individus les repères suffisant pour se guider dans un univers rendu incertain par la nature imparfaite de notre connaissance du futur. Il faut donc mettre en place des repères collectifs, des règles, afin de suppléer aux insuffisances du mécanisme marchand.

Arrow pose donc le problème de la règle, mais pour l'éliminer immédiatement en indiquant une solution. Il identifie en effet, dans le même temps qu'il formule ses critiques, une solution aux problèmes que rencontre le marché : *l'organisation*. On ne fait donc qu'apercevoir le problème de la règle, il est immédiatement résolu par la mise en évidence de l'efficacité d'une règle particulière : la règle organisationnelle hiérarchique. Arrow ne fait cependant qu'évoquer cette solution. Il énonce un programme de recherche, que les néoinstitutionnalistes expliciteront et développeront.

Cette partie explore la "modalité marchande" d'analyse du problème de la règle. On observe que la modalité marchande vise à "éliminer" ce problème en montrant qu'on peut décrire les interactions économiques sous forme de loi. Mais, la loi marchande ne fonctionne qu'imparfaitement, et, à travers ses failles, le problème de la règle refait surface.

Premier chapitre

*Le modèle Arrow-Debreu :
la formulation d'une loi ?*

La démonstration par Arrow et Debreu [1954] du théorème d'existence d'un équilibre général concurrentiel constitue le résultat le plus achevé de la théorie économique néoclassique. Ce théorème d'existence, si on l'associe aux deux "théorèmes fondamentaux" du bien-être établis par Arrow [1951a], démontre en effet la supériorité du mécanisme marchand par rapport à n'importe quel autre mécanisme de coordination. Le marché tire son efficacité de l'existence d'une loi, la loi de l'offre et de la demande, qui conduit des individus égoïstes à se coordonner malgré eux, et à atteindre une efficacité collective qu'ils ne recherchaient pas. Cette loi de marché affranchit donc les individus de la question, centrale en économie, de la coordination. Autrement dit, elle supprime définitivement le problème du choix des règles collectives. Formuler l'existence de cette loi revient à rendre sans objet toute nouvelle réflexion sur les supports de la coordination des acteurs.

Arrow et Debreu n'ont cependant pas "découvert" les propriétés du mécanisme marchand. Smith avait déjà perçu l'efficacité du marché en matière de coordination. Mais il n'en avait pas fourni une démonstration rigoureuse. Or, la théorie de l'équilibre général cherche précisément, à partir de l'analyse de Walras, à démontrer l'efficacité du marché en donnant un fondement logique au concept d'équilibre. Nous suivons donc dans un premier temps l'histoire qui transforme progressivement l'étude "pratique" du phénomène marchand, à laquelle se sont livrés les classiques, en une analyse logique, qui vise non plus seulement à *comprendre* le fonctionnement du marché, mais à *démontrer* sa supériorité. Cette "logicismation" de l'économie se traduit mécaniquement par une disparition de la question pratique du choix des règles.

La démonstration d'Arrow et Debreu repose sur le principe de l'analyse logique. Or, l'analyse logique s'appuie sur le principe de la déduction. Elle ne fait rien apparaître de "nouveau" qui n'était déjà dans les prémisses. L'analyse néoclassique cherche à démontrer l'existence mathématique d'un équilibre général, à partir de l'hypothèse de rationalité individuelle. L'hypothèse de rationalité est donc conçue de façon à ce que le comportement d'un individu rationnel puisse être décrit sous forme d'une loi mathématique. L'énoncé d'une loi doit cependant être précédé de la définition de son domaine d'application, ici l'espace marchand. Cet espace est conçu de façon à éliminer d'une part tout contact direct entre les individus, d'autre part toute temporalité réelle. Il peut de ce fait être analysé comme un espace de nature mathématique. Dans un tel espace, le comportement individuel prend nécessairement la forme d'une loi permettant de déterminer la demande de chaque individu, étant donné sa fonction d'utilité, pour chaque nouveau mouvement de prix. Ainsi, il existe une forme de codétermination entre le marché et l'individu. Cette fusion de l'individu dans le système marchand élimine toute "trace" d'institution extérieure¹ ou de comportement non conforme à la loi marchande.

Cependant, la théorie de l'équilibre générale n'aboutit pas au résultat espéré. Elle tue le mythe smithien, plus qu'elle ne le fonde. En effet, la "loi" marchande n'existe pas, ou, plus exactement elle n'existe que de façon virtuelle au moment de l'équilibre. Mais cet équilibre est tout aussi mythique et "métaphorique" que la main invisible. En effet, la théorie ne parvient pas à établir l'existence d'une loi permettant d'assurer la convergence systématique des demandes individuelles vers l'équilibre. Certaines propriétés des fonctions de demande individuelle disparaissent lorsque ces fonctions sont agrégées. Il est donc nécessaire d'ajouter des hypothèses supplémentaires pour s'assurer que, compte tenu du fait que les individus sont rationnels, ils atteindront un équilibre de marché. Il s'agit là, dans le cadre d'une analyse logique, d'un constat d'échec : on ne peut pas déduire la "stabilité" de l'équilibre directement du mode de comportement des agents. Cet échec touche les fondements de la théorie. Il remet en question la pertinence de cette représentation du marché construit sur un principe de fusion de l'individu rationnel et du système des prix.

¹ Si l'on accepte de mettre pour le moment de côté la question de la détermination, hors du marché, de la "dotation de base".

1- De la main invisible au modèle d'équilibre général Arrow-Debreu

Arrow et Hahn [1971] exposent, dans la préface de leur bilan de la théorie de l'équilibre général, une justification historique de leur méthode d'analyse économique :

“ Il est naturel de se demander pourquoi il vaut la peine d'étudier une économie apparemment si éloignée du monde réel. Nous pourrions répondre, de façon classique, en faisant valoir que l'extraordinaire complexité de la matière qu'étudie l'économiste requiert une simplification urgente, et donc nécessite l'abstraction. Toutefois, cela laisse ouverte la question des raisons du choix de ces simplifications particulières, et de leur caractère approprié.

Notre réponse est d'une nature différente. Il existe aujourd'hui une longue et véritablement impressionnante lignée d'économistes, d'Adam Smith à nos jours qui ont cherché à montrer qu'une économie décentralisée, motivée par l'intérêt personnel et guidée par des signaux-prix, serait sans doute compatible avec une répartition cohérente des ressources économiques pouvant être considérée, en un sens bien défini, comme étant supérieure à une large classe d'autres répartitions possibles. Qui plus est, le mécanisme des signaux sous forme de prix devrait agir de façon à établir ce degré de cohérence. Il est important de comprendre à quel point un tel résultat peut apparaître surprenant à toute personne non initiée à cette tradition. En effet, la réponse de bon sens que l'on ferait immédiatement à la question : « A quoi ressemblerait une économie motivée par la cupidité individuelle et contrôlée par un très grand nombre d'individus ? » serait probablement : « ce serait le chaos ». Qu'une réponse tout à fait différente ait longtemps été défendue et ait effectivement influencé et structuré la pensée économique d'un grand nombre de personnes qui n'étaient pas du tout économistes, est en soi une raison suffisante de l'étudier sérieusement. La proposition ayant été avancée et très sérieusement considérée, il est important non seulement de savoir si elle est vraie, mais encore si elle peut être vraie. Une large part de ce qui suit porte sur cette question, qui nous semble réclamer très fortement l'attention des économistes.” (Arrow et Hahn, 1971, p.VI et VII)

Cette justification historique de la théorie de l'équilibre général souligne clairement l'objectif poursuivi par les auteurs. La “main invisible” est considérée comme un mythe “poétique”, selon l'expression d'Arrow [1968, p.107]. La théorie de l'équilibre général vise à fonder en logique ce mythe persistant. Les auteurs justifient ainsi leur objet, mais aussi leur méthode abstraite, la seule permettant d'énoncer des “vérités logiques”. Cependant la justification logique induit, en fait, un changement d'objet, et, plus encore

une modification du sens du mot "vérité". A bien y regarder, en effet, la logique ne nous dit pas ce qui est vrai "en pratique", mais uniquement ce qui est mathématiquement "sensé". En mettant l'accent sur la possibilité logique du modèle de marché, les auteurs perdent donc de vue la signification "pratique" du mythe smithien. L'histoire qui les conduit à procéder à une démonstration logique est donc aussi l'histoire de la disparition du concept de règle et de l'apparition de l'idée d'une loi logique nécessaire.

11- La référence à Smith.

"L'idée selon laquelle, par le biais du fonctionnement d'un système complet, les effets puissent être très différents, voire opposés, aux intentions est sûrement la contribution intellectuelle la plus importante que la pensée économique ait apporté à l'intelligence générale des processus sociaux." (Arrow, 1968, p.107-108).

Les théoriciens de l'équilibre général considèrent Smith comme l'initiateur d'un mode de représentation du processus économique de marché en lequel ils se reconnaissent. Blaug [1992, p.167] estime cependant qu'il s'agit là "*d'un travestissement de l'histoire*". En identifiant la démarche smithienne à la leur, les théoriciens de l'équilibre général semblent en effet réinterpréter les théories passées à la lumière de leurs propres analyses². L'analyse de Smith n'apparaît plus alors comme une théorie *différente* du marché, mais simplement comme une théorie du marché *formellement moins satisfaisante*, comme une intuition à laquelle il est nécessaire d'apporter une justification rigoureuse. Or cette vision "condensée" de l'histoire, conforme à la compréhension de la science économique comme science cumulative, dissimule en réalité un *changement* de représentation du rôle et de fonctionnement du marché.

La question que cherche à résoudre Smith est celle de la richesse des nations. Or, l'origine de cette richesse se situe d'abord dans le gain de temps que permet d'obtenir la division du travail. Le mécanisme marchand n'est vertueux que dans la mesure où il permet une diffusion des bénéfices tirés de la division du travail. Chez Smith, c'est donc l'organisation sociale de la production qui est, en quelque sorte, première. Le marché n'est qu'une courroie de transmission des gains liés à cette organisation sociale. Il n'est pas le lieu premier où se socialisent les individus : la société smithienne est plus large que le marché.

² Ce n'est d'ailleurs pas une attitude spécifique à la théorie de Smith. Samuelson [1957, 1959] a ainsi proposé une retranscription des théories de Marx et Ricardo dans les termes de l'analyse néoclassique.

Cette prégnance de la société sur l'individu se manifeste jusque dans l'acte marchand. A travers le concept de main invisible, Smith désigne en effet une véritable "manipulation" de l'individu par la société. L'acte marchand est un acte de reproduction de la société, à travers les individus : *"Chaque individu s'efforce continuellement de trouver l'emploi le plus avantageux pour tout le capital dont il peut disposer. Il est bien vrai que c'est son propre bénéfice qu'il a en vue et non celui de la société, mais les soins qu'il se donne pour trouver son avantage personnel le conduisent nécessairement à préférer précisément ce genre d'emploi qui se trouve être le plus avantageux à la société... en cela comme en de beaucoup d'autres cas il est conduit par une main invisible pour remplir une fin qui n'entre nullement dans ses intentions. Tout en ne cherchant que son intérêt personnel, il travaille souvent plus efficacement pour l'intérêt de la société que s'il avait réellement pour but d'y travailler."* (Adam Smith, 1776, p.43). Ainsi, pour Smith la main invisible agit véritablement l'acteur. Elle le conduit à servir la société. Le comportement marchand apparaît comme la manifestation d'une causalité transcendant l'individu et assurant le bien être de la société. L'action de l'individu lui échappe au profit d'une action d'ordre supérieur³. Dans la théorie de Smith, le processus marchand manifeste la prégnance de la société sur l'individu.

Ce n'est plus le cas dans le cadre de la théorie de l'équilibre général. D'une part, la théorie de l'équilibre général, à la suite des marginalistes, sépare le marché de l'organisation sociale qu'est la division du travail. Le marché n'est donc plus une "courroie" de transmission des bienfaits tirés d'un autre aspect de l'organisation sociale. Il devient le lieu unique où se détermine la valeur des biens et est conçu comme étant un mécanisme indépendant de tout autre institution sociale. D'autre part, l'individu marchand n'est plus "agi" par la société. Il est au contraire dénué de toute dimension sociale. Le marché n'est donc plus un outil permettant à la société de s'enrichir. Le système n'a plus de finalité autre que celle de rendre compatible des désirs individuels *a priori* divergents. Il a même une fonction inverse à celle de Smith dans la mesure où, loin de "réunir" les individus, il leur permet de se coordonner sans même se rencontrer. Or, c'est précisément le désir d'une rencontre directe avec l'autre qui est, chez Smith, à l'origine du penchant pour l'échange⁴. Au contraire, dans la théorie de l'équilibre général, le marché remplace la société. La "société" est donc réduite à une série de calculs

³. Sur l'analyse de la main invisible chez Smith, ainsi que sur les interprétations ultérieures de ce concept, on peut se reporter à Defalvard [1990a et b].

⁴ Pour Smith, l'individu échange par goût de la rencontre autant que par intérêt. La sympathie est donc pour Smith un moteur aussi puissant que l'intérêt (Mathiot [1990]). L'idée selon laquelle l'individu est en quelque sorte "poussé" à échanger nous semble bien accréditer la thèse d'une expression de la "société" à travers l'acte individuel. Le principe du "penchant pour l'échange", souligne à quel point l'acteur smithien est un animal social avant d'être un calculateur isolé.

individuels, elle n'est plus un concept distinct de la somme de l'ensemble des individus⁵. Or, en mobilisant une telle conception de la société, la théorie néoclassique désincarne l'idée de main invisible. Elle supprime le sens du mythe smithien. Le principe d'une causalité sociale est rabattu sur celui d'une simple mise en cohérence des choix individuels. Au concept totalisant de main invisible est substitué le concept d'un équilibre entre des forces concurrentes.

Ce glissement apparaît clairement dans la définition du concept de main invisible que proposent Arrow et Hahn (1971, p.1) : "*l'idée selon laquelle un système social mû par des actions indépendantes poursuivant des valeurs différentes est compatible avec un état final cohérent d'équilibre*". Le concept de main invisible devient ici strictement synonyme de celui d'équilibre général. Or, l'idée selon laquelle le marché permet de satisfaire des désirs égoïstes est importante. Elle est constitutive d'un lien réel entre les modèles d'équilibre général et l'intuition smithienne. Cependant, chez Smith, cette idée prend corps à travers un *processus* effectif d'échange, motivé en particulier par le "penchant à échanger" des acteurs. Comme le souligne à juste titre Blaug [1992, p.167-168], Smith décrit le mécanisme de l'échange marchand, mais s'intéresse assez peu au résultat général de ce processus, c'est-à-dire à l'équilibre général. A l'inverse, la théorie de l'équilibre général, dans sa version Arrow-Debreu, n'est véritablement parvenue qu'à la formulation d'un théorème d'existence. Les résultats concernant la stabilité, qui visent précisément à établir l'existence d'un processus automatique de convergence vers l'équilibre, sont beaucoup moins satisfaisants. Ainsi, tandis que Smith prête attention à l'acte effectif d'échange, au mécanisme pratique par lequel les "bienfaits" de la division du travail se trouvent répandus dans l'ensemble de la société, les théoriciens de l'équilibre général ne mettent l'accent que sur la possibilité d'un équilibre général, sans jamais chercher à décrire l'acte d'échange effectif. La théorie perd donc de vue le processus *pratique* de l'échange marchand. L'échange n'est plus que l'application mécanique d'un plan assurant l'équilibre. Il n'intervient qu'après que les multiples demandes aient atteint un point d'équilibre entre elles.

La main invisible ne désigne pas chez Smith un mécanisme de marché anonyme régi par un système de prix fondé sur l'utilité subjective des individus. Elle est l'expression d'une appartenance sociale des acteurs, qui disparaît dans la révolution néoclassique. De ce fait, la reprise par Arrow du thème de la main invisible dans un cadre néoclassique rehausse le caractère "miraculeux" du mécanisme de marché, puisque l'accord collectif survient entre individus non socialisés. Dans le même mouvement, la métaphore de

⁵ Arrow [1965, p.80] : "*Après tout, la société est juste un label commode pour désigner un ensemble d'individus*".

Smith perd son sens premier. Elle devient plus obscure, dès lors que l'idée de "société" disparaît. Dans la théorie de l'équilibre général, le système des prix devient en effet l'unique lien entre des individus absolument isolés.

12- L'héritage walrassien

La filiation qu'établissent Arrow et Hahn entre Smith et les modèles d'équilibre général est assez ténue bien que légitime. En réalité, le véritable initiateur des modèles d'équilibre général est bien sûr Léon Walras :

“On peut défendre l'idée selon laquelle Smith fut l'auteur d'une théorie de l'équilibre général, bien que la cohérence et la consistance de son œuvre puissent être questionnées. Cela est encore plus vrai de ceux qui, après lui, exposèrent de façon systématique le système classique, comme Ricardo, Mill, Marx, dans les travaux desquels quelques failles logiques du raisonnement smithien furent comblées, et qui, tous, peuvent être regardés comme les premiers théoriciens qui exposèrent une théorie d'équilibre général. (...) Pourtant, il y a une raison importante pour laquelle aucun des économistes classiques ne possède une véritable théorie d'équilibre général : aucun ne fait jouer de rôle explicite à la demande. (...) La véritable découverte du concept d'équilibre général peut sans erreur possible être attribuée à Walras, bien que plusieurs éléments du système néoclassique furent mis au point indépendamment par Jevons et par Menger.” (Arrow, 1968, p.109)

La rupture néoclassique vis-à-vis de la théorie classique porte principalement, comme le rappelle Arrow, sur le rôle qu'il convient de faire jouer à la demande, et donc au marché. Cette rupture se traduit par une remise en cause du lien étroit qui existe entre sphère de la production et sphère de la répartition dans la théorie classique, comme on vient de le voir chez Smith. Le marché devient en effet l'unique objet d'attention. La valeur des biens n'est plus déterminée en dehors de lui, dans la sphère de la production. On passe en effet d'une théorie de la valeur travail à la théorie de la valeur-utilité. La valeur des biens ne se situe plus dans le travail, socialement organisé, mais dans la subjectivité de l'individu émettant un jugement d'utilité. De ce fait, le prix de marché devient le seul véritable prix, il ne gravite plus autour d'un prix naturel déterminé par les conditions de la production. La loi de l'offre et de la demande devient alors la seule véritable loi régissant l'économie. C'est à cette condition que l'économie peut prendre une forme mathématique. L'économie peut se construire en science, sur le modèle des

sciences physiques-mécaniques, puisqu'elle prend pour objet d'étude un système autonome régi par une unique loi déterminant les variables quantitatives que sont les prix. C'est précisément parce qu'ils n'ont pas clairement séparé la loi de l'offre et de la demande des conditions sociales de la production que les classiques peuvent paraître, aux yeux d'Arrow, manquer de cohérence et de rigueur. Avec les néoclassiques, et en particulier avec Walras, la théorie économique devient une théorie pure, une science logique. Elle se débarrasse de la dimension pratique qui entachait la pureté de l'économie classique⁶.

Ces deux ruptures de la théorie néoclassique vis-à-vis de la théorie classique sont importantes pour le sujet qui nous préoccupe ici, à savoir la question du rôle des règles. Dans l'univers classique, la question de la coordination n'est pas uniquement traitée sur un plan analytique, elle reçoit également une réponse pratique, fondée sur l'observation de la réalité, en termes d'institutions adéquates. Ce n'est plus le cas dans le cadre néoclassique : la coordination y est traitée de manière purement mathématique, en dehors de toute observation pratique, et prend la forme d'une démonstration établissant l'efficacité de la loi de l'offre et de la demande. Sous le coup de cette loi, le comportement des individus les mène, *a priori*, automatiquement à l'équilibre. Ainsi, alors que l'univers classique permet d'aborder la question des règles économiques et de leur choix (parce qu'il est contingent et n'est pas mathématiquement ordonné) l'univers néoclassique évacue cette question, faisant redescendre la question du choix en amont, au moment où l'individu établit ses préférences en termes de biens.

La détermination subjective de la valeur, en termes d'utilité, et la séparation d'une sphère économique marchande des autres institutions sociales, ont donc pour effet tout à la fois de permettre une science des choix de l'acteur, en termes de préférences individuelles, et de rendre sans objet une théorie des choix institutionnels, puisque la loi de l'offre et de la demande impose un mode d'action unique. Cette conception walrassienne de l'économie récuse une conception "substantiviste" de l'économie, et défend une conception purement analytique qui rend possible la spéculation théorique pure. Walras est le premier auteur à proposer une définition rigoureuse, au niveau de la théorie pure, de ce qu'est l'objet et la nature de l'économie : l'étude de la manière dont des individus égoïstes et rationnels, procédant à des calculs d'utilité concernant des biens rares, liés entre eux par cette concurrence autour de ces biens, et coordonnés par le biais d'un système de prix s'imposant à eux, parviennent sans le chercher à atteindre une solution d'équilibre satisfaisante. Cette définition de l'économie consacre l'idée

⁶On sait que les efforts de Ricardo pour établir une détermination naturelle de la valeur visaient précisément à séparer l'économie de l'incertaine "pratique" des acteurs.

d'équilibre entre des forces concurrentes comme le concept fondamental. Walras pose ainsi les bases d'un paradigme économique qui reste aujourd'hui dominant.

Les théoriciens de l'équilibre général s'inscrivent directement dans la lignée de la démarche walrassienne. Certaines évolutions théoriques modifient cependant de façon non négligeable l'optique walrassienne. C'est en particulier le cas de la théorie de la stabilité fondée, chez Walras, sur la célèbre figure du commissaire-priseur.

Walras propose en effet une théorie "pratique" de la stabilité : le "tâtonnement", associé à l'existence d'un commissaire-priseur criant les prix. Le commissaire-priseur recense les demandes excédentaires pour un système de prix donné, et modifie en conséquence les prix relatifs, calculant un nouveau système de prix qu'il fait parvenir à chaque individu qui revoit son plan de consommation en fonction de ce changement des prix. Il répète l'opération jusqu'à ce que la fonction de demande excédentaire soit nulle : c'est l'équilibre. Arrow qualifie cette approche de "*célèbre mais assez maladroite*" (Arrow, 1968, p.110). L'analyse walrassienne du processus de convergence du système des prix vers les prix d'équilibre souffre de deux défauts rédhibitoires dans la perspective moderne de la théorie de l'équilibre général. En premier lieu, elle fait apparaître l'extrême centralisation du marché, ce qui entre en contradiction avec le principe de la main "invisible" à laquelle la théorie cherche à se référer. D'autre part, elle est de nature "pratique" : elle met en lumière le caractère institutionnalisé du marché, ce qui contredit l'idée selon laquelle la loi de marché serait naturelle.

Berthoud [1988, p.82] a souligné le rôle crucial que joue chez Walras l'existence du commissaire-priseur : "*on interprète le crieur comme la simple représentation pédagogique d'une causalité mécanique. On doit dire plutôt que l'unicité du crieur figure l'unicité de la causalité mise en œuvre en toute machine*". Selon l'analyse de Berthoud, le projet walrassien est précisément de rendre la liberté aux hommes par la machine. En construisant les conditions de réalisation d'un "échange naturel", au sens où il respecte les conditions de régularité des lois naturelles, Walras cherche à rendre chacun "quitte de tous les autres". Dans cette optique, le marché, symbolisé par le crieur, est une construction distincte des individus, une machine construite dans le but, précisément, de libérer les individus de leur interdépendance. La théorie de la stabilité n'est donc pas, dans l'analyse walrassienne, d'essence mathématique. De plus, elle vient avant la théorie de l'existence, comme condition institutionnelle permettant de penser l'équilibre. Walras, selon Berthoud, adresse donc d'abord un message au réformateur social. L'existence d'une institution sociale symbolisée par le crieur est la condition de l'existence de l'équilibre. Il existe donc, dans la théorie walrassienne, une condition institutionnelle, plus ou moins mythique, pour que l'espace économique devienne entièrement

appréhensible sous forme mathématique. Le savoir économique mathématique est encore conditionné par la mise en place d'une règle institutionnelle. L'analyse logique est distinguée, mais subordonnée, à une réflexion de nature "pratique" portant sur la mise en place des règles d'organisation du marché⁷. La temporalité du processus marchand est donc circonscrite, elle n'est pas supprimée.

Dans la théorie de l'équilibre général, au contraire, la stabilité de l'équilibre n'est envisagée qu'après la démonstration de son existence et de son unicité et selon le même procédé logique. L'hypothèse du crieur est considérée comme un obstacle à franchir, une commodité à dépasser, une représentation "faute de mieux" du mécanisme de marché. La conception walrassienne du marché, centralisée, n'est pas considérée comme une conception satisfaisante, puisqu'elle remet en cause l'idée smithienne d'une "main invisible". La théorie moderne de l'équilibre général cherche donc à éliminer l'institution centrale walrassienne en la considérant comme "l'image" d'un mécanisme de va et vient entre chacun des individus et le cœur du marché. Elle cherche à fournir une analyse purement logique et mécanique de ce phénomène de tâtonnement. La dimension institutionnelle du marché disparaît donc tout-à-fait⁸.

Ainsi, la constitution du paradigme de l'équilibre général se fait par réduction successive de l'idée d'économie, depuis une conception "substantive" chez les classiques, jusqu'à une conception purement analytique et axée non pas sur l'institution de marché, mais sur la rationalité individuelle. Dans l'idéal, toute la théorie de l'équilibre général devrait pouvoir être dérivée des hypothèses relatives à la rationalité individuelle. Cet ancrage au concept de rationalité constitue le trait principal de l'économie néoclassique, elle lui permet de développer une analyse purement mathématique du marché.

13- Une question mathématique

Si on laisse de côté l'émergence de l'économie du bien-être initié par Pareto, disciple de Walras à Lausanne, on peut considérer, avec Arrow [1968] que la théorie de l'équilibre général tombe dans l'oubli pendant une quarantaine d'année, puis resurgit

⁷On peut sur ce point se reporter à l'étude de Tricou [1994].

⁸Cela ne signifie pas que les théoriciens de l'équilibre général n'aient pas en fait besoin de supposer l'existence d'un cadre institutionnel précis, que l'on pourra repérer derrière leurs hypothèses. Mais à aucun moment, cette dimension de la théorie n'est assumée. Les hypothèses sont présentées comme des représentations plus ou moins naturelles, pas comme des réquisits institutionnels. Leur éloignement de la réalité est présenté comme l'effet nécessairement "simplificateur" de la théorie.

dans les années trente autour de deux courants : “l'un débutant dans la littérature de langue allemande et traitant principalement des questions d'existence et d'unicité de l'équilibre, l'autre, de langue anglaise, traitant de stabilité et de statique comparative.” (Arrow, 1968, p.111)

Le premier courant reprend la démarche walrassienne en constatant que Walras n'a pas formellement prouvé qu'il existait une solution à la série d'équations différentielles constituant le système des prix d'équilibre. Sa démonstration s'arrête avec le constat que le système contient autant d'inconnues que d'équations, ce dont Walras déduit, sans le démontrer, qu'il est soluble. Cette idée est remise en cause par trois auteurs au même moment⁹, et cette remise en cause conduit finalement à la première formulation, par un pur mathématicien, d'un théorème d'équilibre général dans les années trente¹⁰. Toutefois, c'est l'existence de progrès mathématiques, et en particulier le théorème du point fixe de Brouwer, qui permet à l'économie d'intégrer les résultats de Wald¹¹. La théorie des jeux se constitue en effet à cette période et utilise la “topologie” pour formuler des théorèmes généraux d'équilibre, qui vont être entièrement repris en économie. Von Neumann formule les premiers résultats, ils sont repris et simplifiés par Kakutani, toujours dans le champ de la mathématique pure. Ce n'est finalement qu'au début des années cinquante que sont formulés des théorèmes d'existence en économie : Mac Kenzie [1954] et Arrow-Debreu [1954], ce dernier s'imposant comme référence¹².

Le second courant s'intéresse, en Angleterre et indépendamment du précédent, à des questions de stabilité et de statique comparative. Hicks [1939] et Samuelson [1941-1942] en sont les deux principaux représentants. Leur approche est différente de la précédente dans la mesure où elle étudie d'emblée l'évolution de la fonction de demande nette d'une valeur positive à une valeur nulle. Il s'agit donc de décrire la marche vers l'équilibre, sans s'être assuré de son existence ou de son unicité. Ces auteurs mettent l'accent sur la nécessité de fonder une théorie moderne de la stabilité, qui donnera naissance en particulier à la statique comparative. En fait, de par les outils mathématiques utilisés (le calcul différentiel) autant que par leur façon de poursuivre le programme de recherche walrassien, ces deux auteurs se distinguent de la perspective moderne de l'équilibre général. Ils privilégient l'étude d'équilibres locaux, ou s'intéressent au rôle des chocs modifiant l'équilibre atteint. Ce faisant, ils empruntent un chemin différent de la théorie d'équilibre général d'Arrow et Debreu, qui cherche d'abord à s'assurer de l'existence et

⁹:Zeuthen [1932], Neisser [1932], Von Stackelberg [1932].

¹⁰:Wald [1933-35], voir Arrow [1968] pour une présentation plus détaillée.

¹¹ L'outil mathématique cesse alors d'être principalement le calcul différentiel. L'émergence de la théorie des ensembles (topologie) simplifie la présentation et augmente la généralité des résultats.

¹² Debreu [1959] reprend la démonstration sur des bases différentes, sans utiliser le théorème de Nash [1950]. Cette contribution est sans doute la référence la plus souvent citée, à côté de son article avec Arrow.

de l'unicité de l'équilibre. Cette dernière optique semble s'imposer pour des raisons logiques : l'étude de la stabilité, ou de la statique comparative, est conditionnée par la démonstration de l'unicité. Or, les résultats en termes d'unicité et de stabilité seront, comme nous le verrons, assez négatifs. Cette façon d'aborder la théorie de l'équilibre général aura donc tendance à s'effacer derrière les résultats du premier courant. Le noyau dur de la théorie d'équilibre général se constitue à partir de la démarche et des résultats d'Arrow-Debreu [1954], plutôt qu'autour de ceux de Hicks-Samuelson.

Cette dernière étape historique de la constitution d'une théorie moderne de l'équilibre général est singulière pour plusieurs raisons, qui tiennent toutes à l'accélération du mouvement de la théorie économique vers une forme exclusivement analytique. Durant cette période en effet, la théorie de l'équilibre général évolue en fonction des progrès opérés dans une autre discipline : les mathématiques. La nouveauté ne tient pas tant ici à l'étroitesse des liens entre ces deux disciplines, qui date au moins de Walras, mais dans la nature purement mathématique des réflexions alimentant cette branche de la théorie économique. Le fait même que l'évolution de la science économique soit conditionnée par les progrès réalisés en mathématique est un signe du changement de nature du discours économique, ou, plus précisément, de la radicalisation de son analytisme. Avec le passage au second plan d'une idée substantive du marché et des préférences de l'acteur, et une organisation de la théorie autour de la notion tout à fait neutre de rationalité, la théorie de l'équilibre général a gommé ce qui la distinguait encore des mathématiques.

Reste, bien sûr, la question de l'interprétation des résultats obtenus, qui ancre la théorie de l'équilibre général dans l'économie. Or, cette interprétation s'avère très délicate, dans la mesure où ces résultats ne parviennent que très imparfaitement à s'appliquer à la situation économique réelle. De ce point de vue le parti pris méthodologique qui se dessine dans cette période et qui consiste à donner la priorité à la progression analytique de la connaissance et à la rigueur formelle des résultats obtenus grâce aux outils mathématiques récents, plutôt qu'à leur sens économique, constitue un pari dont le résultat peine aujourd'hui encore à être tout à fait évalué. L'essentiel pour notre propos n'est de toute façon pas lié à la pertinence de ce pari, mais au pari lui-même, dont on comprend bien qu'il conduit automatiquement à rejeter la question du choix des règles en dehors de la théorie économique, sinon en dehors de l'économie.

Franck Hahn [1980] qualifie la théorie de l'équilibre général de "*théorie de la main invisible*". Cette appellation mérite d'être nuancée. Il est certain qu'un lien théorique unit la métaphore smithienne et l'analyse néoclassique de l'équilibre général concurrentiel,

mais ce lien est ténu et s'accompagne d'un certain nombre de profondes ruptures théoriques. Pour s'exprimer de façon lapidaire, on peut dire que la théorie de l'équilibre général est un palimpseste de la théorie de la main invisible : elle n'en garde que l'idée selon laquelle le mécanisme de marché a pour effet de permettre aux individus d'atteindre collectivement un état qu'aucun d'eux n'a cherché à atteindre.

Mais, tandis que chez Smith, cette idée n'était que la conséquence d'une socialisation des acteurs (la sympathie et traduisait la manifestation d'une force naturelle d'ordre supérieur garantissant la conservation de la société, elle n'est dans le cadre de l'équilibre général qu'un produit direct de la rationalité purement instrumentale des acteurs non socialisés. Autrement dit, ce qui cause l'équilibre chez Smith disparaît chez Arrow-Debreu. Entre temps, Walras parvient à matérialiser ces "forces sociales" abstraites en intégrant à sa théorie une institution centralisatrice et chargée de la réalisation de l'équilibre. Il incombe alors au législateur de créer pour permettre l'existence d'un équilibre de marché. Mais ce réquisit institutionnel disparaît chez Arrow-Debreu, qui ouvrent une tradition dans laquelle la question de la stabilité est seconde par rapport à celle de l'existence.

Ce double "nettoyage" de la théorie économique dessine un paradigme articulé autour de l'hypothèse de rationalité, bien plus qu'autour de l'idée de marché. La science économique se constitue en dépouillant l'acteur de ses attributs sociaux pour le réduire à sa rationalité. Ce faisant, la loi du marché est d'emblée intégrée à la définition de l'individu. L'analyse abstraite du marché, à partir de prémisses portant exclusivement sur l'hypothèse de rationalité individuelle, nécessite en effet de situer dans ces prémisses l'origine de la loi. La déduction logique ne fait en effet rien apparaître de "nouveau" : elle met simplement à jour ce qui était déjà là.

L'histoire qui relie le mythe de la main invisible à la théorie moderne de l'équilibre général, via l'analyse walrassienne, est l'histoire d'une redéfinition de l'économie autour d'un individu mu par une loi de comportement de type mathématique. Cette histoire est donc aussi celle de la disparition du concept de règle économique.

2- La théorie de l'équilibre général : le cadre d'analyse

“Le modèle néoclassique est fondé sur deux concepts de nature très différente. L'un est la notion d'agent individuel économique, dont le comportement est régi par l'optimisation sous contraintes. (...) L'autre est le marché. (...) Nous tendons à considérer comme une vertu du marché le fait que l'agent individuel ait besoin de savoir si peu de choses. (...) Cette simplification de la prise de décision individuelle n'est rendue possible que parce que le marché nous fournit de l'information de manière économique, sous forme de prix.” (Arrow, 1974b, p.158)

La théorie de l'équilibre général est d'une grande cohérence. Cela tient au fait qu'elle est entièrement abstraite. En effet, pour répondre rigoureusement à la question de l'efficacité du marché comme mode de coordination d'individus égoïstes et concurrents, la théorie de l'équilibre général construit tout autant un type de marché particulier, et purement abstrait, qu'un type d'individu, lui aussi absolument abstrait. Cette double construction permet d'accorder la figure de l'individu et celle du marché dans lequel il évolue. Une question tentante est de chercher à cerner laquelle de ces deux constructions, de l'individu ou du marché, est première, et en quelque sorte détermine l'autre. Ce n'est sans doute pas une bonne question, tant il semble que l'individu et le marché, dans leur acception néoclassique, se contiennent l'un l'autre, et sont dans un rapport de détermination mutuelle qui ne permet pas de faire apparaître une relation causale à sens unique. L'individu ou le marché ? L'œuf ou la poule ? L'un ne va pas sans l'autre et c'est cet indéfectible lien qui est le meilleur gage de la cohérence du cadre d'analyse néoclassique.

Nous allons tenter d'exposer la loi de comportement par laquelle il est possible de décrire le comportement d'un individu rationnel sur un marché de type “Arrow-Debreu”. Décrire cette loi nécessite tout d'abord de définir l'espace dans lequel l'individu se déplace, et la temporalité qui caractérise cet espace. Le comportement de l'individu répond aux caractéristiques de cet espace qui, en le situant hors du temps et de la société, lui impose un mode de comportement instrumentalement rationnel. C'est la raison pour laquelle son comportement peut être décrit sous forme de loi.

21-L'espace

Deux hypothèses essentielles permettent de définir le cadre marchand du modèle Arrow-Debreu : l'existence d'une définition préalable des biens et l'existence d'un système complet de marché permettant d'associer un prix à chacun des biens. Ces deux hypothèses déterminent un univers dont la temporalité est absente, ou, plus précisément, dont la temporalité est purement logique et formelle. Elles font de l'univers Arrow-Debreu un univers de nature mathématique.

211- Les biens Arrow-Debreu

La définition des biens constitue la première pierre de la théorie de l'équilibre général. En effet, la théorie se donne pour objet d'analyser le partage (ou la répartition) des biens disponibles et désirés. La première question à résoudre est donc de savoir ce que l'on se partage. Quelle est la nature de cet ensemble, objet des convoitises de tous les individus ?

Les biens que s'échangent les agents sont définis de façon rigoureuse, à la suite de Hicks [1939], comme étant des biens situés dans l'espace et dans le temps de façon précise. Ainsi des biens matériellement identiques seront néanmoins différents s'ils ne sont pas délivrés au même endroit, ou à la même date. Cette hypothèse traduit le fait que, par exemple, l'achat d'un parapluie livrable en Espagne dans la période d'été de l'année suivant l'achat, n'est pas identique à l'achat, en Écosse et durant le mois de novembre, d'un parapluie immédiatement disponible.

Cette définition de la liste des biens disponibles inclut une description censée épuiser entièrement les questions que peuvent se poser les acheteurs à leur propos. Leur qualité ne souffre donc d'aucune ambiguïté, les biens sont supposés être objectivement et exhaustivement décrits. Le degré de raffinement de la liste des biens "Arrow-Debreu" est maximal : "Les descriptions sont si précises qu'on ne peut imaginer aucun raffinement supplémentaire susceptible de donner lieu à des allocations nouvelles améliorant la satisfaction des agents." (Geanakoplos, 1985, p.116). Cette liste rend donc compte de manière infiniment précise de l'éventail des biens désirés par les agents. Les individus choisissent un panier de biens parmi cette liste de biens disponibles.

Cette représentation de l'espace des biens permet de définir l'espace des biens comme un espace mathématique : "Nous supposons qu'il n'y a qu'un nombre fini de marchandises distinctes, qui seront repérées par un indice allant de 1 à l . Nous

supposons aussi que la quantité de chacune peut être n'importe quel nombre réel." (Debreu, 1959, p.35)¹³. L'espace des marchandises ainsi défini correspond à \mathbf{R}^l . Cette définition de l'espace des biens constitue, comme on l'a souligné, le premier mot de la théorie, préalable à la question de savoir comment les répartir. Ce premier mot est mathématique, et cela conditionne nécessairement l'ensemble de la théorie.

C'est dans cet espace mathématique que les individus vont évoluer pour prendre possession de ces biens. La seconde étape de la description de cet espace consiste à comprendre, précisément, comment ces biens sont rendus accessibles aux individus.

212- Un système complet de marché

La mise en rapport des biens et des individus est assurée par le marché. Mais pour que le degré de précision avec lequel les biens sont supposés être définis ne soit pas altéré par la mise sur le marché, il faut supposer qu'il existe un marché par bien. Il y a donc l marchés pour les l types de biens.

Un tel système suppose une infinité de marchés, dont l'ouverture et l'organisation seraient coûteuses et, parfois, impossibles. Mais, là encore, cette construction joue un rôle qui n'est pas de permettre une description de la réalité¹⁴, mais plutôt de proposer une théorie pure du marché qui en souligne l'essence mathématique. L'hypothèse d'un système complet de marché permet en effet de définir le système des marchés comme une correspondance de \mathbf{R}^l dans \mathbf{R}^l , associant un prix à chaque quantité de chacun des biens de l'ensemble "L" des types de biens. Les individus expriment leur désir sous forme d'un "panier de biens" représentable sous la forme d'un vecteur quantité (q_1, q_2, \dots, q_l) comportant une certaine quantité q de chacun des biens disponibles dans l'ensemble "L". Le marché fait correspondre à chaque vecteur quantité ainsi déterminé un vecteur prix $(p_1 \dots p_l)$ qui fait correspondre un prix à chaque quantité de chacun des biens demandés. Le système des prix représente alors un point de \mathbf{R}^l .

L'association de l'hypothèse de nomenclature¹⁵ et de l'hypothèse de système complet de marché permet d'ordonner entièrement le domaine des interactions. Le contact direct des individus aux biens est facilité par l'existence d'une définition objective des biens et d'un système de marché permettant d'associer un prix à chaque quantité de biens. La

¹³ Les biens sont donc indéfiniment divisibles.

¹⁴ Geanakoplos [1985, p.116] note : "il est très rare de trouver un marché pour un bien «pur» Arrow-Debreu. (...) [mais] le modèle Arrow-Debreu, en proposant une conception idéale d'un marché séparé pour chaque bien, se rencontrant tous simultanément, constitue un étalon à partir duquel l'économie réelle peut-être mesurée". Cette conception de l'utilité de la représentation Arrow-Debreu semble pour le moins optimiste : l'étalon semble si loin de la réalité qu'il en devient tout à fait mythique.

¹⁵ C'est ainsi que Benetti et Cartelier [1980] qualifient la représentation de l'espace des biens. Ils entendent par là souligner qu'une telle représentation suppose déjà un accord collectif institutionnel.

mathématisation de l'espace économique est rendue possible par la conjonction de ces deux hypothèses établissant une possibilité d'interprétation mathématique de la répartition marchande des biens. Il faut souligner que ce rapport transparent des individus aux biens, et à leur valeur, n'est pas altéré par l'existence d'une monnaie. Les prix qui émergent de chaque marché sont des prix relatifs, déterminés par tous les autres prix de tous les autres biens. Le système de prix mesure donc l'exacte valeur des biens relativement à tous les autres, leur exacte place dans l'échelle de l'utilité collective. La comparabilité du prix des biens est assurée par le choix d'une unité de compte, non désirable en tant que telle. Un des biens est arbitrairement choisi comme base permettant d'exprimer la valeur relative de tous les autres biens : c'est le numéraire¹⁶. Les prix reflètent donc l'exacte comptabilité des désirs multiples se portant sur les biens, ils ne sont pas biaisés par l'existence d'un voile monétaire¹⁷. Le système de marché associe à chaque bien un prix représentant les rapports d'échange réels entre les biens (sur le modèle du troc). Ce mode d'organisation des échanges permet d'attribuer, à chaque bien particulier, une valeur d'échange "collective" sous forme de nombre réel, représentant sa valeur dans certaines conditions d'offre et de demande, sans pour autant remettre en cause la relation directe de chaque individu aux biens matériels.

Ce système complet de marché comporte une vertu supplémentaire : il assure une information complète de chacun des individus. On a souligné que l'individu peut, puisqu'il est situé dans l'espace des biens, exprimer chacun de ses désirs sous forme de bien. L'hypothèse d'un système de marché complet lui donne de plus la possibilité de connaître avec certitude combien il lui coûtera d'acquérir ces biens. De ce fait, l'individu dispose d'une parfaite correspondance entre ses désirs et le moyen de les assouvir. C'est ce que souligne l'expression "information parfaite" : chacun des individus sait tout ce qu'il lui est utile de savoir pour satisfaire ses désirs. Le système de marché rend la réalité transparente pour l'individu, il lui permet d'établir une correspondance parfaite entre ses désirs et cette réalité¹⁸.

¹⁶Dans le système de Walras, le système d'équation finale comporte n équations, pour $n+1$ biens. Le choix d'un des biens comme étant le numéraire permet d'égaliser le nombre d'inconnues et le nombre d'équations, et donc, pour Walras, de prétendre ce système "solvable".

¹⁷Arrow [1981, p.235] le rappelle : "*La théorie néoclassique de l'équilibre général soutient que les prix relatifs guident l'allocation des ressources. On peut déterminer tous les rapports d'échange entre des biens pris deux à deux si l'on distingue un bien spécial et si l'on connaît les rapports d'échange de l'ensemble des autres biens à celui-ci. Ces rapports peuvent être considérés comme le prix des biens. Walras nomma «numéraire» ce bien ainsi distingué. On considère que le prix du numéraire est égal à 1. En théorie pure le numéraire n'a pas de signification spéciale. Passer d'un numéraire à un autre n'affecterait pas l'allocation des ressources (...). Le niveau absolu des prix n'a aucune signification*". Il insiste cependant sur le fait que l'existence d'un chômage persistant l'incite à modifier cette perspective : "*Ce sur quoi je voudrais insister ici, est que le mécanisme d'ajustement des prix n'est pas indépendant du choix du numéraire.*" (ibid., p.237).

¹⁸Ce qui bien sûr ne signifie pas que l'individu peut assouvir tous ses désirs. Le marché ne permet pas de dépasser la contrainte que constitue la "dotation initiale" dont chacun dispose avant d'entrer sur le marché.

Deux obstacles séparent, *a priori*, l'individu des biens objets de ses désirs : la connaissance des biens disponibles (localisation, date, caractéristiques physiques...) et le moyen de se porter acquéreur de ces biens. L'existence d'une nomenclature d'une précision parfaite et d'un système établissant un marché spécifique pour chaque bien, permet de représenter le rapport de l'individu aux biens comme une relation directe ne nécessitant aucune socialisation, aucune recherche d'information, aucune médiation. Le cadre marchand attribue une "valeur" à chaque bien, et assure une compatibilité générale des désirs individuels, sans que cette médiatisation altère le rapport exclusif de chaque individu aux biens désirés.

L'univers Arrow-Debreu permet donc une relation directe des individus aux biens. Cette relation directe est rendue possible par l'existence d'un système de marché qui place cette relation sous un angle quantitatif. Pour cette raison, l'univers marchand est de nature mathématique. Il reste, pour achever de le caractériser, à rendre compte de la forme de temporalité qu'il mobilise.

213- La temporalité de l'univers Arrow-Debreu.

On peut, pour clarifier la forme de temporalité qui caractérise l'univers marchand, distinguer deux grandes formes de temporalité : une forme irréversible et marquée par l'imprévisibilité ; une forme réversible et parfaitement prévisible. Ce sont, respectivement, le temps *pratique* de l'action et le temps *mécanique* qu'utilise la physique. Or, le modèle Arrow-Debreu exclut ces deux formes de temporalité et en mobilise une troisième, qui lui est spécifique, et qui correspond à une forme extrêmement abstraite et formelle de temporalité, que l'on qualifiera de temps *logique*. Ce mode de temporalité formelle se caractérise par l'absence de mouvement, et donc, par l'absence de temps "réel".

Le premier des deux modes temporels classiques correspond au temps pratique de l'action. Lorsqu'un acteur agit, dans le monde humain, cette action a pour effet de modifier le contexte initial dans lequel elle apparaît¹⁹. Pour cette raison, la temporalité de l'action est irréversible. De plus, l'action se déroule dans un environnement social. La multitude des acteurs et la simultanéité et l'interdépendance des actions de chacun rendent *imprévisibles* les conséquences de l'action. Il n'est pas possible de démêler par avance l'écheveau constitué par l'enchevêtrement de multiples actions ayant toutes pour

¹⁹ Ricœur [1986, p.172] rappelle : "*agir signifie avant tout opérer un changement dans le monde*". Cette définition de l'action correspond à l'interprétation par Arendt [1960] du concept d'action chez Aristote. Le propre de l'action est d'être imprévisible, pour cette raison l'action nécessite une forme particulière de sagesse, qui n'est pas la rationalité économique classique.

caractéristique d'apporter de la nouveauté dans l'environnement. Cette temporalité de l'action, *imprévisible* et *irréversible*, met donc l'acteur aux prises avec la *contingence* : il doit agir et décider sans pouvoir avec certitude connaître les conséquences des diverses actions qu'il peut entreprendre. L'action fait apparaître de la nouveauté, et cette nouveauté n'est pas *déductible* des prémisses de l'action. Il n'existe pas, dans cette temporalité, de structure logique unissant *ex ante* une action à ses conséquences.

Le second mode temporel est le temps "expérimental" de la physique. Le temps de l'expérience physique est réversible et prévisible, puisque les mécanismes étudiés s'enchaînent logiquement selon une série de causalités que l'on peut parcourir dans un sens ou dans un autre. Ainsi l'application au temps t d'un vecteur force d'intensité f sur un objet "x" provoquera, en fonction de la nature de l'espace dans lequel x est situé, un déplacement de x qui est prévisible. De plus, si l'on souhaite faire revenir "x" à sa position initiale, il suffira de lui appliquer un vecteur force d'une intensité telle que, compte tenu de la nouvelle situation atteinte par "x", il reprenne sa position initiale. Le temps est donc cette fois homogène, mesurable, prévisible, sur le modèle des mouvements permanents d'un balancier qui ne connaîtrait aucun frottement. Ce temps est celui qui permet la connaissance scientifique.

L'espace marchand Arrow-Debreu ne correspond ni au temps de l'action, ni même au temps mécanique de la physique, mais à une troisième forme encore plus abstraite : la temporalité logique. La temporalité pratique représente un processus au cours duquel les actions s'enchaînent de façon *imprévisible* et *irréversible*. La temporalité de la physique correspond à une relation mécanique qui se déroule dans le temps. Une cause met un certain temps pour produire un effet, même si cette relation de cause à effet est réversible et prévisible. Dans le cadre logique, on peut dire qu'il existe une forme de temps, puisqu'il existe une relation d'implication logique, dans laquelle une cause produit nécessairement un effet. Mais la relation d'implication logique, qui relie par exemple la demande de l'individu aux mouvements de prix, est "hors du temps". La vérité logique n'est pas située dans le temps, elle se situe à un niveau d'abstraction, de transcendance, qui la situe en dehors de toute "histoire". Il existe donc une certaine forme de temporalité, qui correspond à l'enchaînement logique des causalités. Mais ce temps est virtuel : il n'existe pas. Il ne correspond pas au temps de l'action puisqu'il ne fait pas apparaître de nouveauté. Il ne correspond pas non plus au temps physique, durant lequel l'objet étudié se déplace réellement. Ici le temps ne décrit aucun *mouvement* réel. Il est entièrement contenu dans la logique du raisonnement. Bien que les choix des agents portent sur une période potentiellement infinie (quoique nécessairement bornée), ces choix sont entièrement compressibles dans le moment de la déduction logique. L'espace Arrow-Debreu se caractérise donc par une absence de temps réel.

Cette affirmation peut paraître cavalière. En effet, les biens sont datés, et à chaque moment correspond un bien, unique compte tenu de son type. A chacun de ces biens correspond un marché unique. Chacun des marchés constituant le système porte donc sur un moment précis. Les paniers de biens demandés par les individus comportent des biens livrables à des dates futures, et les offres et les demandes portant sur ces biens sont confrontées sur des marchés portant sur des transactions qui ne seront effectives que dans le futur, même si elles sont décidées de manière irrévocable à la date présente. Pour cette raison, on désigne par l'expression "marchés futurs" ces marchés portant sur des "biens futurs". Mais il est essentiel de remarquer que ces marchés futurs sont en réalité tout-à-fait actuels. Les paiements des biens futurs sont en effet immédiats. La réflexion logique des agents leur permet donc de décider à un instant précis de l'ensemble de leurs actes à venir. Leur déduction logique détermine donc entièrement ce que sera le futur, c'est-à-dire quels échanges auront lieu et à quelle date. La temporalité formelle de l'univers Arrow-Debreu est entièrement contenue dans l'analyse logique.

L'univers Arrow-Debreu est ainsi marqué par une forme de temporalité très particulière et très limitée. Cette forme logique de temporalité est purement formelle. Elle n'est même pas, comme c'est le cas en physique, une dimension de l'espace. Dans un tel espace, la rationalité de l'individu peut être exprimée sous forme de loi mathématique.

22- La loi de comportement de L'individu rationnel

L'univers Arrow Debreu dicte à l'individu son comportement. L'individu est en effet installé dans un rapport aux biens qui ne lui laisse aucune latitude quant au choix du comportement approprié. C'est la raison pour laquelle son comportement est descriptible sous forme de loi.

La caractéristique fondamentale de la rationalité individuelle est son instrumentalité. Elle découle directement de l'isolement de l'individu. L'individu est en effet situé dans un contexte dont l'incertitude et les questions éthiques sont exclues, précisément parce qu'il est placé dans un rapport direct avec les biens, qui exclut toute relation sociale. Les raisons qui déterminent cette attitude instrumentale imposent également à l'individu une attitude paramétrique. L'individu ne peut pas peser sur la décision des "autres" qu'il ne croise jamais. Il doit donc se contenter de réagir à des mouvements de prix qui s'imposent à lui. Enfin le caractère mathématique de l'univers Arrow-Debreu impose une

attitude calculatoire. L'individu exprime ses désirs sous forme mathématique. Il les met en équation, et cette équation comporte une solution unique. L'unicité de la solution lui impose donc un comportement unique. C'est le fondement de la loi de la demande.

221- Une rationalité instrumentale

L'individu au centre du modèle d'équilibre général a pour unique horizon l'ensemble des biens disponibles, et pour unique souci celui du calcul des moyens les plus efficaces pour accaparer un maximum de biens. Le choix d'un comportement approprié est exclusivement guidé par ce souci d'efficacité. Les conséquences finales de l'action, en termes de dotation en biens, sont donc son unique critère de choix. La fin de l'action détermine entièrement le choix des moyens appropriés, au détriment de toute réflexion sur la valeur intrinsèque de ces moyens. Autrement dit, l'individu rationnel de la théorie de l'équilibre général a une attitude *conséquentialiste*. Cette attitude conséquentialiste, détermine une forme de rationalité *instrumentale*²⁰.

Le mode de rationalité instrumentale nécessite que la fin et les moyens de l'action puissent être parfaitement séparés. En effet, dans une perspective instrumentale, la fin est donnée, et la rationalité de l'individu ne s'applique qu'au choix des moyens relativement à leur efficacité. C'est le cadre marchand, en isolant l'individu, qui impose ce mode de rationalité. Si l'acteur était en permanence confronté aux autres, en tant que détenteurs de biens, son rapport aux biens serait nécessairement pollué ou biaisé par la médiation que constituerait alors le rapport aux autres. Mais, dans l'univers Arrow-Debreu, l'individu n'est jamais en rapport *direct* avec les autres. L'interdépendance des acteurs est "gérée" par un système de marché qui permet à chacun de l'oublier. L'individu est donc situé dans un univers qui est, de son point de vue, purement matériel et quantitatif. Cet univers "technicise" la question de l'obtention de la meilleure dotation en biens possible. C'est parce que le problème à résoudre devient technique que l'individu peut être instrumentalement rationnel. Plus précisément, cette technicisation, qui permet de séparer clairement fin et moyens de l'action et de ne traiter que des moyens, repose sur deux conditions.

La première est que l'individu puisse s'affranchir du problème de l'effet des moyens utilisés sur les autres acteurs. Dans un univers "humain" l'acteur est confronté à la question de l'effet de ses actes sur autrui. Il doit respecter certaines règles éthiques, ne pas tuer, ne pas spolier, ne pas nuire à autrui. Il n'a donc pas toute latitude dans le choix

²⁰ Par rationalité instrumentale on entend donc désigner un mode de rationalité ne concernant que l'efficacité des moyens choisis relativement aux fins poursuivies. Ce mode de rationalité ne se prononce pas sur la validité des moyens eux-mêmes, relativement par exemple à leur dimension éthique ou politique.

des moyens. L'univers Arrow-Debreu, en isolant l'individu, l'affranchit de cette question. L'individu n'a que son propre bien-être comme souci. C'est la raison pour laquelle l'individu économique est souvent présumé être "égoïste". Mais, pour être égoïste, il faudrait encore que cette attitude résulte d'un choix de valeur. Il faudrait que l'individu connaisse l'altérité et le rapport à l'autre. Il n'existe rien de tel dans l'univers marchand, l'individu est d'emblée mis en face des biens, et cette situation impose à son raisonnement une fin unique : comment obtenir le plus de biens possible. L'individu est tout-à-fait incapable de réfléchir sur les fins de l'action elles-mêmes ou de se poser la question de la valeur intrinsèque des moyens qu'il met en œuvre. Ces questions éthiques ne peuvent pas apparaître dans l'univers Arrow-Debreu. Elles sont "hors sujet". Voilà pourquoi l'univers Arrow-Debreu impose à l'individu une réflexion purement instrumentale. Le seul choix qui lui soit donné de faire est purement technique, il n'a aucune dimension éthique.

L'autre condition nécessaire pour qu'une rationalité de type instrumentale puisse s'appliquer, est de pouvoir parfaitement évaluer les conséquences de ses actes. Évaluer une action en fonction de ses conséquences nécessite en effet une parfaite connaissance de l'avenir. La connaissance du futur est indispensable dans une perspective instrumentale. Or, là encore, l'interdépendance des acteurs et l'enchevêtrement des actions constitue dans un univers social un obstacle insurmontable. C'est parce qu'il se déroule dans un tel univers que le temps de l'action est marqué par la contingence et l'incertitude. Mais, le système de marché "ordonne" cette interdépendance et élimine cette contingence. L'univers dans lequel l'individu évolue lui assure une parfaite prévisibilité de l'avenir. L'existence d'un système de marché complet est un point clé de cette maîtrise de l'avenir. Il assure une information parfaite sous forme de prix qui permet à l'individu de déterminer entièrement et avec certitude toutes les conséquences de ses choix pour l'ensemble des périodes à venir. Cette relation causale logique qui unit nécessairement les choix d'un comportement à ses conséquences est au fondement d'une attitude instrumentale. Elle permet de poser techniquement la question du choix des moyens, compte tenu de la fin donnée de l'individu. Là encore, c'est le cadre marchand qui, en inscrivant le comportement de l'individu dans un mode temporel qui n'est pas le temps obscur de l'action mais le temps limpide de la logique, rend possible cette *technicisation*.

L'individu ne peut exercer sa rationalité qu'à un unique problème : parmi les paniers de biens qui me sont accessibles, compte tenu de mes revenus, lequel assurerait une satisfaction *maximale* de mes désirs ? C'est une question technique, elle appelle un raisonnement instrumental. L'interdépendance des acteurs, source d'une part d'un conflit entre éthique et efficacité, d'autre part d'une imprévisibilité de l'avenir, est ici dénouée, pour chaque individu par l'existence d'un système rendant chacun "quitte de tous les

autres". Dans l'univers Arrow-Debreu, l'individu est séparé de tous ses semblables, ce qui lui permet de mettre en œuvre une rationalité de type purement instrumental²¹.

222- Une rationalité paramétrique

L'acteur est "quitte de tous les autres". Cela lui permet de s'affranchir de la question du respect des autres, mais cela lui impose également de renoncer à influencer leur comportement. L'individu ne peut donc pas agir stratégiquement. Sa situation est telle qu'il ne peut pas modifier unilatéralement le prix d'un bien, ou jouer d'une manière ou d'une autre sur les paiements des autres acteurs. Il ne négocie pas, il prend acte de ce que sa situation est déterminée en partie par tous les autres qui lui apparaissent comme une sorte d'entité fictive communiquant par le biais du système des prix. C'est ce que traduit l'expression "*price taker*". L'individu ne peut pas agir stratégiquement et unilatéralement sur les prix pour améliorer sa position. Il reçoit les prix comme une donnée définitive définissant sa situation. Son choix se limite à l'expression de ses désirs sous forme de panier de bien. Il ne peut pas agir sur les données de son problème, ce qui caractérise une rationalité de type paramétrique.

Les raisons qui imposent à l'individu une rationalité de type "instrumental", le contraignent également à avoir une attitude "paramétrique". En effet, l'univers qui lui permet d'être libéré de son rapport aux autres, lui impose également une attitude passive. L'individu n'a pas les moyens d'agir stratégiquement en influençant les autres en un sens qui lui serait favorable. Il est informé du "résultat" de l'ensemble des comportements des agents par le biais des mouvements de prix, mais ne peut en aucun cas influencer sur ces mouvements. L'individu est contraint à une attitude paramétrique, il ne peut agir sur les données de son problème²². Il réagit au mouvement des prix, il ne les maîtrise pas.

223- Une rationalité calculatoire

L'environnement de l'individu technicise les données de son problème, et le contraint à une attitude passive. De plus, cet environnement dans lequel l'acteur est plongé, et à partir duquel il doit tenter de satisfaire ses désirs matériels est de nature mathématique. La

²¹ Ce qui bien sûr ne règle pas la question du caractère juste ou injuste du système de marché, et ne solde donc pas la question des rapports entre éthique et économie en général. C'est uniquement le dilemme individuel qui se trouve ici dénoué. C'est précisément la thèse de Berthoud [1988] concernant le projet Walrassien. Selon lui, Walras, comme Descartes, sont des tenants de la "liberté par la machine". La machine marchande libère l'individu de son interdépendance avec les autres.

²² Le concept de rationalité paramétrique s'oppose dans son principe à celui de rationalité stratégique, qui est notamment utilisé en théorie des jeux, et dans le cadre de laquelle l'acteur utilise ses rapports directs avec les autres individus pour modifier l'environnement à son avantage.

réflexion mathématique est l'unique technique efficace dans ce type d'univers. L'individu doit donc raisonner sous forme mathématique : il *calcule* le meilleur moyen de satisfaire ses désirs. L'univers Arrow-Debreu impose à l'individu une rationalité instrumentale de type calculatoire.

La mise en équation de ses désirs passe nécessairement par une expression de ses préférences en termes de biens, les biens étant le seul moyen de satisfaire ses désirs... Sa rationalité consiste alors d'une part à exprimer des préférences *cohérentes*, de manière à classer les paniers de biens selon une échelle croissante correspondant à ses désirs, et, d'autre part, à choisir le panier de biens accessible situé le plus haut dans cette échelle, c'est-à-dire celui qui *maximise* sa fonction d'utilité.

La première opération consiste donc à retranscrire ses désirs en termes de préférences portant sur des paniers de biens. L'individu au fondement des modèles d'équilibre général va donc exprimer sa rationalité dans son rapport aux biens. Chaque bien lui apporte une certaine utilité, et il doit être capable, à partir de la liste des biens, d'opérer un classement rationnel des différents paniers de biens. Ici la rationalité désigne donc une capacité au classement d'états matériels.

Un individu rationnel est capable de produire un classement d'ordre total, transitif et réflexif. (*p.e.* : Arrow [1969, p31-52], Arrow et Hahn [1971, chap.4], Quirk et Saposnik [1968,p.9-19]). Ainsi, soit S : "l'ensemble des états possible de l'économie"²³ et R une relation d'ordre signifiant "au moins aussi désiré que". On peut alors définir un individu comme rationnel s'il parvient à classer S selon une relation d'ordre R vérifiant les propriétés suivantes :

1- Propriété de totalité : si $Z' \in S$ et $Z'' \in S$, alors $Z'RZ''$ ou $Z''RZ'$

2- Propriété de transitivité : si $Z' \in S$, $Z'' \in S$, $Z''' \in S$ et si $Z' \in Z''$ et $Z''RZ'''$ alors $Z'RZ'''$

3-Propriété de réflexivité : si $Z' \in S$, alors $Z'RZ'$.

Ces caractéristiques du choix du consommateur sont complétées en particulier par deux axiomes :

* *L'axiome de non-satiété* signifiant qu'il n'existe aucun panier de bien préféré à n'importe quel autre intégrant une unité de bien supplémentaire.

²³ Dans le vocabulaire de la théorie, une certaine allocation de biens pour un individu est interprétée comme un état de l'économie. L'individu classe donc des états de l'économie. Un état complet de l'économie se traduit pour l'individu par un panier de biens particulier. Lorsque l'on introduit la production, le temps de travail est interprété comme un bien ayant une utilité négative.

** *L'axiome de convexité des préférences* signifiant qu'un consommateur préfère le panier de biens le plus mixte parmi les paniers accessibles²⁴.

*** *L'axiome de continuité* : les préférences du consommateur constituent un ensemble continu : si un ensemble de biens est préféré à un second, tout complexe de biens suffisamment proche lui sera également préféré. Autrement dit, l'individu n'a pas de discontinuité dans ses préférences, de "saute de goût", en ce qui concerne les biens²⁵.

Si les préférences d'un individu vérifient toutes ces conditions, il est possible de définir une fonction attribuant une valeur numérique à chaque panier de biens en fonction de sa place dans le classement. La fonction d'utilité ainsi définie ne mesure pas l'utilité absolue que chaque panier de biens représente pour l'individu, mais traduit simplement la place que ce panier occupe dans l'échelle de ses préférences. Le fait qu'un panier de biens A ait une valeur inférieure à un panier B selon la fonction d'utilité d'un individu I , traduit simplement que A apporte plus d'utilité que B à l'individu I . Elle ne nous informe donc pas sur l'intensité avec laquelle tel ou tel panier de biens est préféré à tel ou tel autre. Elle est de nature purement lexicographique, et reste du côté de l'utilité ordinale. Cette fonction d'utilité est croissante et quasi-concave.

Il reste alors à l'individu, parvenu à exprimer ses préférences par le biais de cette fonction d'utilité, à déterminer quel panier de biens accessible est le meilleur. Cette opération de "maximisation sous contrainte" est l'aboutissement du processus rationnel par lequel l'individu calcule le meilleur moyen d'atteindre le bien-être maximum. On notera pourtant qu'arrivé à ce stade, l'essentiel du travail rationnel est déjà fait. Il reste à l'acteur à établir à quel niveau se situent ses revenus, et à calculer quel est le panier de biens ayant l'utilité la plus élevée pour un coût total équivalent à ce revenu. C'est un problème comptable. Il faut donc à ce stade introduire la notion de prix. On sait qu'ils sont donnés à l'acteur. L'individu est informé des prix, calcule sur cette base le coût des différents paniers de biens et la hauteur de ses ressources qui forment sa "contrainte budgétaire". L'opération de maximisation sous contrainte consiste donc à déterminer le panier de biens apportant la plus grande utilité à l'acteur, tout en étant accessible, c'est-à-dire ayant un coût inférieur ou égal à la contrainte budgétaire de l'agent.

²⁴ L'axiome de convexité des préférences s'énonce comme suit : X_i et X_j , deux biens appartenant à l'ensemble R^1 des biens. I étant une relation d'ordre signifiant "est indifférent à" et P signifiant "est strictement préféré à" : Si $X_i I X_j$, alors $[tX_i + (1-t)X_j] P X_j$. Cet axiome est souvent traduit par l'expression : "les consommateurs aiment les mélanges" et dénoncé comme irréaliste. Certains biens ne gagnent rien à être mélangés ou scindés en petite quantité : Churchill préférerait sans doute un grand verre de whisky qu'un petit verre de whisky et un grand verre de lait. On peut souvent se passer de cet axiome quand le nombre de consommateurs est suffisamment élevé.

²⁵ Mathématiquement, la continuité s'exprime comme suit : "Pour tout x et y appartenant à R^1 , les ensembles $(x / x \geq y)$ et $(x/y \geq x)$ sont fermés".

Cette contrainte budgétaire est calculée en fonction des dotations de base de l'individu. On estime en effet qu'au début du processus, chaque individu possède une dotation de base $Q^* = \sum(q_i \dots q_n)^{26}$. Cette dotation peut être écoulee aux prix de marché associés à ces biens pour déterminer le revenu de l'individu : $r = \sum(q_i p_i \dots q_n r_n)^{27}$.

L'individu détermine donc son panier de biens Q souhaité tel qu'il soit celui qui lui apporte l'utilité maximale (valeur donnée par la fonction d'utilité) tout en respectant $Q \leq r$. Sa demande nette est alors donnée par $(Q - Q^*)^{28}$. L'opération de maximisation est donc une opération purement comptable. Le véritable choix de l'acteur se traduit dans le classement cohérent des paniers de biens qu'il opère au préalable, et qui traduit sa capacité à établir une relation de préordre sur l'ensemble des biens.

Le problème posé à l'individu marchand est donc d'essence mathématique. Sa résolution prend donc nécessairement la forme d'un calcul. Dans ce calcul, la solution est unique. Pour cette raison sa rationalité est non seulement calculatoire, mais aussi *optimisatrice*. Le principe de l'optimisation permet en effet de souligner ce qui n'est qu'une conséquence du mode de choix calculatoire et de la simplicité de l'environnement de ce choix : l'individu n'atteint pas seulement un panier de biens qui le satisfait, il obtient celui qui le satisfait le plus. Tout écart par rapport à cet optimum est considéré comme une erreur et caractérise donc un comportement qui n'est pas rationnel. Le comportement ne peut donc être partiellement rationnel. Il n'existe pas de processus de convergence progressive vers la solution. A information identique, ce qu'assure l'unicité de la source d'information, tout individu doit, en fonction de ses préférences propres, atteindre un résultat unique. L'énoncé des conditions de choix de l'individu permet d'en déduire un mode de rationalité instrumental et calculatoire déterminant un comportement unique. Compte tenu de ses préférences et de sa contrainte de revenu, l'individu n'a pas le choix. L'univers Arrow-Debreu lui prescrit un itinéraire unique. L'unicité du comportement rationnel rend possible une description extérieure. Compte tenu des préférences d'un individu et de sa contrainte de budget, la théorie est capable de décrire avec précision son comportement, puisque le lien entre les prémisses de son

²⁶ On peut intégrer à cette dotation de base une certaine quantité de temps disponible, que l'individu écoule à certain prix, sous forme de travail. Ainsi cette représentation est "adaptée" à un monde avec production.

²⁷ Il est important de souligner que cette dotation est déterminée en dehors du marché. C'est donc un élément de nature institutionnelle, qui intervient avant que le processus marchand ne se déroule. "*Cette loi institutionnelle détermine la façon dont sont réparties les ressources initiales, ou simplement la façon dont elle ne sont pas distribuées.*" (Guesnerie, 1973, p.84). C'est à partir de cette entrée, étroite, que peuvent se développer les réflexions sur la nature juste ou injuste de telle ou telle "société" marchande. Nous y reviendrons dans le chapitre suivant.

²⁸ La fonction de demande nette est homogène et de degré zéro, puisque l'individu ne "souffre" d'aucune "illusion monétaire". Son rapport au bien n'est pas biaisé par la monnaie.

comportement (ses préférences) et le comportement lui-même est de nature purement logique. Le comportement de l'individu peut donc être décrit sous forme de loi.

224- La loi de la demande

Le comportement d'un individu dont la rationalité respecte l'ensemble des critères énoncés est descriptible sous forme d'une loi. L'augmentation du prix relatif d'un bien, diminue la demande de l'individu pour ce bien et augmente relativement sa demande pour les autres biens. Cette loi de comportement est synthétisée par l'axiome des préférences révélées.

Cet axiome stipule que si un consommateur choisit un panier de biens X_a plutôt qu'un panier de biens X_b , également accessible dans le cas où le système des prix est P_a , alors ce même consommateur ne préférera pas le panier X_b au panier X_a dans une autre situation où ils sont tous les deux accessibles. Autrement dit, le choix de X_a révèle la préférence du consommateur, et, dès lors, le choix de X_b ne peut s'expliquer que par l'existence d'un système de prix rendant X_a inaccessible à l'acteur. Le respect de cet axiome est une condition nécessaire et suffisante pour qu'une fonction puisse décrire la demande nette d'un consommateur rationnel²⁹. Or, il se déduit entièrement des hypothèses définissant la rationalité individuelle. Ce sont ces hypothèses qui fondent donc l'existence d'une loi de demande établissant la décroissance relative de la demande d'un bien lorsque son prix augmente³⁰.

L'ensemble des hypothèses définissant la rationalité individuelle permet donc finalement de décrire le comportement individuel sous forme de loi.

Frydman [1992, p.10] souligne que dans l'univers Arrow-Debreu, "*la loi de l'agent est celle du système*". De fait, la rationalité de l'individu, comme le remarque Tricou [1994b, p.39], n'est pas indépendante du mode d'organisation des marchés. L'individu n'est pas autonome vis-à-vis du système marchand. L'univers dans lequel il est situé définit l'unique problème qu'il a à résoudre, et la manière de le résoudre. L'agent est, en fait, un élément du système marchand, et ce système détermine son comportement. Il n'a

²⁹ Il existe deux versions de cet axiome. La version faible, qui est présentée ici parce qu'elle est plus intuitive, ne suffit pas à définir une fonction d'utilité lorsque il y a plus de deux biens. Il faut alors que l'axiome fort des préférences révélées soit vérifié. Autrement dit, il faut que l'axiome faible puisse être généralisé à n biens : aucun panier appartenant à l'ensemble des paniers de biens non préférés au panier X_a , ne peut jamais être choisi, sauf si X_a est inaccessible. L'axiome fort des préférences révélées constitue en quelque sorte une itération de l'axiome faible, et permet de définir rigoureusement une fonction de demande nette reflétant la rationalité du consommateur. (Quirk et Saposnik [1968, p.63-64], Tallon [1997, p.28-29].

³⁰ Tallon [1997, p.37,38] précise que cet axiome montre qu'une loi de demande existe si l'on suppose que le revenu est adéquatement compensé lorsque le prix du bien augmente.

pas plus de libre arbitre que la pomme lorsqu'elle subit l'influence de l'attraction terrestre. Compte tenu de ses caractéristiques (ses préférences) l'agent se déplace nécessairement de manière conforme à la loi marchande, puisque l'univers dans lequel il est placé détermine entièrement sa façon de se comporter.

Cette conception de l'individu articule des compétences logiques, calculatoires, très importantes, et des compétences pratiques nulles. En effet, l'univers Arrow Debreu élimine tout rapport aux autres et toute temporalité réelle. Il supprime ce faisant les deux éléments qui distinguent un raisonnement logique d'un comportement pratique. L'individu marchand n'agit pas. Il calcule. Il est comme poussé, déplacé, animé par un mouvement entièrement déterminé d'une part par ses caractéristiques, d'autre part par son environnement. Compte tenu de ce qu'il préfère (ce qui ne fait l'objet d'aucune décision rationnelle), le comportement de l'individu dépend entièrement des caractéristiques de son environnement. Il n'a donc rien à décider. Il se conforme à la loi marchande qui caractérise l'univers dans lequel il est situé. Il est un élément de cet univers, un rouage de cette machine. Debreu résume de manière lapidaire cette situation en soulignant : "*Une action d'un agent a est un point de R^l , l'espace des marchandises. Un système de prix p est un point de R^l . La valeur d'une action a par rapport au système de prix p est le produit interne pa .*" (Debreu, 1959, p.39).

La description du cadre d'analyse nous a amené à l'expression de la loi qui décrit le comportement d'un individu³¹. A partir de cette loi de comportement individuel, la théorie de l'équilibre général cherche à exprimer sous forme de loi mathématique le fonctionnement de l'ensemble du système marchand, qui n'est que la réunion d'un ensemble, potentiellement infini, d'individus répondant à la loi individuelle. L'exposé de cette tentative fait l'objet de la section suivante.

³¹ Cette présentation du cadre de la théorie de l'équilibre général vise à mettre l'accent sur son projet : formuler l'existence d'une loi assurant l'existence d'un équilibre de marché. Dans cette optique nous n'avons pas jugé nécessaire de présenter les hypothèses portant sur le producteur. En effet, aucune des difficultés spécifiques liées à la production ne sont prises en compte dans la théorie de l'équilibre général, qui construit un individu producteur sur le modèle du consommateur. Elle n'intègre ni firmes, ni entrepreneur, et n'envisage pas la création d'entreprise. Tallon souligne d'ailleurs : "*l'introduction de la production ne pose pas de problème, ce qui tendrait à suggérer que la théorie de l'équilibre général est principalement une théorie de l'échange, qui s'accommode bien avec l'introduction du secteur productif*" (Tallon, 1995, p.173). En fait l'exclusion de toute temporalité "réelle" dans le cadre d'analyse Arrow-Debreu vide l'activité de production de toute sa substance et sa spécificité. (Voir sur ce point : Quirk et Saposnik [1968,p.84] et Arrow (1971b)).

3– Une loi qui n'existe qu'à l'équilibre...

La théorie de l'équilibre général vise à démontrer que le marché, en situation de concurrence parfaite, permet de rendre compatibles des plans d'agents égoïstes ne se préoccupant pas des conséquences collectives de leur action individuelle, et d'obtenir une allocation optimale des ressources, au sens de Pareto. Les théoriciens de l'équilibre général affirment qu'ils cherchent par ce biais à fournir des fondements logiques au mythe smithien de la main invisible. Ce projet est décliné, dans l'optique Arrow-Debreu de la théorie de l'équilibre général, en trois étapes distinctes : il faut tout d'abord démontrer qu'un équilibre de marché existe, puis qu'il est unique, puis, enfin, qu'il est stable³². Existence, unicité, stabilité sont les trois étapes nécessaires, et sans doute suffisantes, pour la formulation d'une loi assurant la convergence des interactions concurrentielles d'une multitude d'agents vers un équilibre de marché optimal.

De ces trois étapes, seule la première a pu être franchie. Les deux suivantes sont progressivement apparues comme inaccessibles. En effet, le cadre d'analyse d'Arrow et Debreu permet d'obtenir un résultat d'existence formel : il existe un système de prix annulant chacune des fonctions de demande nette. Mais il ne permet pas de dire quoi que ce soit du phénomène pratique, ou au moins mécanique, de convergence vers l'équilibre. Une tel processus sort en effet du cadre strictement logique qui est celui de la démonstration du théorème d'existence. Or, la démonstration formelle du théorème d'existence ne prouve rien quant à la possibilité d'atteindre *effectivement* l'équilibre.

31- Existence et optimalité de l'équilibre concurrentiel

La théorie de l'équilibre générale moderne s'est construite historiquement, après Walras, autour de la question de la résolution du système d'équations différentielles que Walras pensait être trivial et automatiquement soluble. Pour cette raison la première étape du projet Arrow-Debreu est la formulation d'un théorème d'existence. Toutefois, cinquante ans après Walras, la question de l'existence d'un équilibre paraît encore plus cruciale. En effet, pendant cette période, Pareto a proposé une définition formellement

³² C'est précisément sur ce programme que le projet Arrow-Debreu se distingue de la démarche Hicks-Samuelson. Ces derniers cherchent en effet à traiter directement de la stabilité, à partir des formes des courbes de demande et d'offre, sans s'assurer ni de l'unicité, ni de l'existence de l'équilibre.

satisfaisante de l'optimalité et, sur cette base, Arrow [1951a] a démontré l'optimalité du système marchand. Il n'est pas anodin, que la démonstration par Arrow [1951a] de l'optimalité d'un équilibre concurrentiel ait précédé la démonstration de référence de l'existence d'un équilibre, par Arrow et Debreu [1954]. Comme le rappellent Quirk et Saposnik [1968, p.110] : *“une fois portés les jugements de valeur suffisants pour définir un état désirable de l'économie, l'économie positive permet de déterminer si un tel état est réalisable avec une certaine organisation de l'économie”*. Ainsi, la justification de la théorie moderne de l'équilibre général est la démonstration par Arrow [1951a] des deux théorèmes “fondamentaux” du bien-être.

311- L'optimalité de l'équilibre concurrentiel

La définition de l'optimalité au sens de Pareto revient à considérer qu'un état de l'économie est optimal à compter du moment où il n'existe aucune situation qui soit meilleure pour au moins une personne, sans être plus mauvaise pour qui que ce soit. Cette définition de l'optimalité n'exclut donc pas les situations injustes caractérisées par une inégale répartition des richesses. Elle stipule simplement que la répartition des biens est la meilleure possible, en cela qu'il n'existe aucun moyen d'améliorer la situation de *chacun* des individus.

Arrow [1951a] a établi l'optimalité de la répartition marchande sous forme d'équilibre concurrentiel. Ce résultat est contenu dans les deux “théorèmes fondamentaux” de l'économie du bien être :

1- Toute position d'équilibre concurrentiel est optimale au sens de Pareto.

2- Quel que soit l'état optimal de l'économie, il existe un choix approprié des paramètres tel que cet état apparaisse comme une position d'équilibre du système des marchés concurrentiels.

Ainsi la démonstration fournit une première base rigoureuse à la métaphore de la main invisible : le mécanisme de marché permet de réaliser une allocation des ressources optimale, en un sens précis. Plus encore, toute situation optimale, en ce sens, peut-être atteinte par le biais d'un mécanisme de marché concurrentiel, pourvu que chacun des individus reçoive la dotation appropriée en ressources. Le marché est donc non seulement un des moyens d'atteindre l'optimum, mais, plus encore, un moyen qui permet d'atteindre *tous* les optimums possibles. Ceci n'implique pas qu'il n'existe pas d'autres moyens d'atteindre un optimum, mais établit, sinon la supériorité absolue du mécanisme de marché, du moins la certitude qu'il n'existe pas de mécanisme qui soit à la fois aussi efficace, au sens de Pareto, et plus général.

Ces deux théorèmes fondamentaux ont été complétés, postérieurement, par les théorèmes qui établissent qu'un groupe d'agents organisés pour défendre leurs intérêts, une coalition, n'obtiendrait pas individuellement davantage que le paiement obtenu par chacun dans la situation d'équilibre concurrentiel. Cette idée est formellement définie à partir de la théorie du "cœur" ou du "noyau" (*core*). Le cœur contient l'ensemble des allocations possibles et non bloquées : c'est-à-dire des allocations individuellement plus intéressantes, ou équivalentes, à celles qu'auraient individuellement pu obtenir des individus organisés en groupe afin de défendre leurs intérêts communs. Tout équilibre concurrentiel, optimal au sens de Pareto, appartient au cœur. On démontre également que dans le cas où le nombre d'agent est suffisamment grand, c'est à dire quasi infini, tout état appartenant au noyau de l'économie est une position d'équilibre concurrentiel³³. Ainsi, même si un groupe d'agents ayant des intérêts convergents cherchait à obtenir davantage que l'allocation concurrentiel d'équilibre, il ne pourrait y parvenir.

Cet ensemble de résultats est de la plus haute importance puisqu'il établit l'absolue inutilité d'une organisation des acteurs visant à atteindre un bien-être maximum pour chacun. Le marché, dans lequel personne n'a le souci de l'efficacité collective, fait au moins aussi bien que le ferait n'importe quelle organisation. Réfléchir sur d'autres "règles" possibles n'a dès lors aucun intérêt économique³⁴. Du moins, c'est le cas, à condition qu'un tel équilibre existe.

312- Le théorème d'existence

Arrow et Debreu [1954] établissent l'existence d'un équilibre de marché concurrentiel³⁵. Leur démonstration, qui n'est pas la première, s'est imposée comme la référence, en raison de sa relative simplicité au regard des précédentes, et de sa

³³ Ce résultat est dû à Aumann [1964].

³⁴ Ce qui ne veut pas dire qu'il épuise les questions ayant trait à la justice. Le concept d'optimum de Pareto ne dit rien du caractère juste ou injuste de la répartition des ressources. Sen, comme Arrow, ont à de nombreuses reprises rappelé qu'une situation dans laquelle un petit nombre d'agents avait tout, et une grande majorité rien du tout (ce qui n'est pas si éloigné de la situation actuelle au niveau mondial) pouvait néanmoins être tout à fait optimale au sens de Pareto. En effet, l'optimum signifie simplement que dans un tel cas, l'amélioration du sort des pauvres se ferait au détriment des riches. La question de l'équilibre ne règle donc pas la question des corrections à apporter à la répartition, après que l'équilibre a été réalisé, de manière à la rendre plus juste. L'intervention étatique est envisageable, pour établir *ex-ante* une allocation cohérente des ressources, ou corriger *ex-post* une répartition injuste. Cela ne remet cependant pas en cause, le fait que le marché soit un mécanisme de répartition dont l'efficacité est optimale. Or, l'efficacité constitue, dans l'optique néoclassique, une condition nécessaire et préalable à la justice.

³⁵ Il ne nous a pas semblé nécessaire de présenter ici la démonstration du théorème Arrow-Debreu. De même nous nous contenterons de présenter les résultats obtenus en matière de stabilité et d'unicité sans chercher à entrer dans le détail de leur démonstration. L'important dans notre optique est de comprendre les problèmes posés par ces résultats, pas de reprendre leur démonstration. Pour la démonstration de l'équilibre on peut se reporter, outre les contributions fondatrices d'Arrow-Debreu [1954] et Debreu [1959], à Quirk et Saposnik [1968, p.84-102].

généralité. Les deux auteurs s'appuient sur le concept d'équilibre de Nash pour un jeu concurrentiel à "n" personnes. Debreu [1959] reprend cette démonstration à partir d'une méthode complètement différente (directement à partir de la théorie du point fixe de Kakutani que Nash utilise également).

Le concept d'existence prend, dans cette démonstration, un sens précis, différent du sens commun. Il s'agit en effet d'un concept d'existence au sens mathématique, qui ne nous apprend rien sur la possibilité, même théorique, que cet équilibre existe effectivement dans une économie réelle. Autrement dit, le théorème montre seulement que le système d'équations différentielles posé par Walras a effectivement une solution, en employant des méthodes mathématiques nouvelles fondées sur la topologie. On peut interpréter ce résultat de la façon suivante : "il existe un système de prix d'équilibre annulant l'ensemble des demandes nettes, c'est-à-dire rendant compatibles l'ensemble des plans de production et de consommation des agents"³⁶. Après avoir montré que l'équilibre est optimal, la théorie démontre donc qu'un état d'équilibre est logiquement concevable.

Hahn [1980,p.17] salue ce résultat dans les termes suivants : "il faut être devenu bien philistin pour ne pas apprécier la nature tout-à-fait surprenante de ce résultat, ou bien être insensible à l'élégance des moyens utilisés pour le prouver. Cette théorie aboutit en effet à la conclusion surprenante qu'il est logiquement possible de décrire une économie dans laquelle des millions d'agents, n'ayant d'autre souci que leur intérêt propre et répondant uniquement à l'information sous forme de prix, peuvent néanmoins réaliser une allocation cohérente des ressources." Cette interprétation enthousiaste est sans doute à nuancer. A proprement parler, en effet, le théorème d'existence ne décrit pas une économie où des acteurs réalisent une allocation d'équilibre. Il montre uniquement que les hypothèses du modèle sont compatibles avec l'existence d'un système de prix d'équilibre. Il ne dit rien sur la manière dont il peut être réalisé, ni même si cela est possible, ne serait-ce que logiquement. Le théorème d'existence nécessite donc d'être complété par un théorème montrant d'une part que l'équilibre peut être atteint, et d'autre part qu'une fois atteint, il est stable.

³⁶Une fonction de demande nette correspond à la différence entre l'offre et la demande totale se portant sur un bien. Pour le système de prix d'équilibre, l'offre et la demande sont identiques pour tous les biens. Le système de prix d'équilibre annule donc l'ensemble des demandes nettes.

32- Comment atteindre l'équilibre ?

Dans une économie "théorique" ayant les caractéristiques de l'univers Arrow-Debreu, la question de l'existence d'une procédure assurant la convergence vers un équilibre unique et optimal est qualifiée de question de la "stabilité". Cette question reçoit chez Walras une réponse d'une nature tout-à-fait différente de la "démonstration" logique de l'existence. Comme on l'a déjà souligné, Walras apporte une réponse *pragmatique* à cette question en décrivant une procédure de va-et-vient entre chaque individu et un "centre" du marché assurant, par tâtonnement, le calcul d'un système de prix d'équilibre. Ce centre est qualifié de "commissaire-priseur", de "crieur" ou encore de "secrétaire de marché". L'objectif poursuivi par la théorie moderne de l'équilibre général vise au contraire à apporter à cette question une réponse qui se situe au même niveau abstrait et logique que la question de l'existence. Autrement dit, la théorie vise à formuler des conditions satisfaisantes, c'est-à-dire déductibles de la loi de la demande établie au niveau individuel, qui permettent d'assurer que les fonctions de demandes nettes convergent systématiquement vers un point d'équilibre où chacune d'entre elles est nulle. Mais cet objectif ne sera pas atteint, et le seul résultat obtenu consiste précisément à montrer qu'il ne peut pas l'être. Le théorème de Sonnenschein établit en effet l'absence de lien logique entre la forme des fonctions de demande nette et les fonctions d'utilité des individus. En quelque sorte, la théorie échoue à donner un sens pratique à la démonstration d'équilibre. Bien qu'elle parvienne à décrire le comportement des individus sous forme de loi, cette loi ne "s'agrège" pas. Elle ne nous dit rien sur le moyen d'atteindre l'équilibre. Le problème pratique de la coordination des individus reste donc entier.

321- Le commissaire-priseur

La présentation des forces en présence et du cadre de la bataille n'épuise pas l'histoire de la défaite de Waterloo. La démonstration qu'il est logiquement possible de *décrire* un état du marché, pour lequel l'ensemble des plans d'agents concurrents sont compatibles, ne nous dit rien sur la manière dont ces agents, dans ce système, peuvent *atteindre* un tel état. Dans un cas comme dans l'autre, il manque un élément essentiel pour que la description soit complète : la manière dont les forces en présence se coordonnent.

Comme on l'a déjà dit, l'objectif de la théorie de l'équilibre général est précisément de montrer qu'il existe un système dans lequel les agents n'ont pas à se soucier de leur coordination. Le système de marché doit être conçu comme un mécanisme qui assure la cohérence des actions malgré les acteurs. Dans l'idéal la description du système et des

individus devrait suffire à établir l'existence d'une convergence automatique vers l'équilibre. Les particules de métal, en raison de leurs caractéristiques physiques, convergent nécessairement vers un aimant. De la même manière les plans de consommation des agents devraient nécessairement converger vers l'équilibre. En quelque sorte, l'équilibre devrait déjà être en eux.

Mais Walras n'a pas été jusqu'à faire une telle hypothèse. La description de la détermination logique de la fonction d'utilité individuelle n'épuise pas, dans sa théorie, la description de la machine marchande. Il reste le moteur : le crieur. Le crieur walrassien, qui n'est pas un individu mais une entité mécanique, représente l'organe central de coordination. Ses fonctions sont multiples. Il doit en effet crier le système de prix, calculer le système de prix d'équilibre, et s'assurer qu'il n'existe aucun échange hors des prix d'équilibre. Cette entité centrale assure donc que toutes les transactions portant sur un même bien se réalisent toutes au même prix³⁷, et, qu'elles ne deviennent effectives que lorsque l'on s'est assuré que ces prix correspondaient au système de prix d'équilibre, de manière à ce que les marchés puissent être définitivement fermés après les échanges.

Le mécanisme du tâtonnement assuré par le commissaire-priseur se traduit par une succession de séquences identiques :

- (a) Les individus calculent leur demande pour le système de prix en vigueur.
- (b) Le crieur calcule l'ensemble des demandes nettes excédentaires, puis en déduit un nouveau système de prix qui les annule.
- (c) Il indique ce nouveau système de prix à chaque agent

Cette séquence est amenée à se répéter un certain nombre de fois. En effet, la modification du système de prix modifie également les préférences des agents. Plus exactement : le panier de biens préféré, parmi ceux qui sont accessibles, n'est plus le même qu'aux prix précédents. Les individus indiquent donc leur nouveau choix, et la procédure se répète à l'identique. Elle se répète jusqu'à ce qu'aucune nouvelle demande insatisfaite ne parvienne au crieur. L'ensemble des demandes nettes sont alors nulles : c'est l'équilibre.

L'existence d'un crieur chargé d'assurer le tâtonnement constitue une aporie pour un modèle cherchant à déduire l'équilibre du libre choix d'individus rationnels. La théorie moderne de l'équilibre général a donc tenté, dans la mesure du possible, de rendre compte du processus de tâtonnement, sans s'appuyer sur un acteur central qui en serait responsable. Ainsi Arrow et Hahn [1971, p.264] indiquent simplement : "*on peut imaginer, pour donner quelque incarnation concrète à l'abstraction que nous nous proposons d'étudier* [id: le processus de tâtonnement], *l'existence d'un super-secrétaire*

³⁷ Gadrey [1981] qualifie cette hypothèse de "postulat du prix unique".

qui crierait un ensemble donné de prix et recevrait les transactions de l'ensemble des agents dans l'économie." Mais cette incarnation doit bizarrement perdurer, alors même que le procédé qu'elle incarne est abandonné... Ainsi, dans leur introduction à l'étude des processus dits de "non-tâtonnements", Arrow et Hahn indiquent : "Nous devons continuer à supposer qu'il existe un commissaire qui crie les prix. (...) La question : "comment changent les prix en l'absence de secrétaire de marché ?" est bien plus difficile à traiter" (p.325). On est donc amené à penser que le commissaire-priseur n'est pas seulement une fiction commode, une illustration pédagogique... Le problème n'est pas tant que l'on fasse l'hypothèse que les prix évoluent "comme si" il y avait un commissaire-priseur, mais bien qu'on ne sache pas du tout comment ils évolueraient s'il n'y en avait pas.

Là où Walras indiquait clairement l'existence d'une centralisation des échanges, et d'un mécanisme institutionnel assurant une série de tâches que les individus ne parviendraient pas à réaliser d'eux-mêmes, spontanément, la théorie moderne préfère entretenir un certain flou derrière l'expression "fiction". On sait qu'il n'y a en fait pas de crieur, mais que l'on doit faire comme s'il y en avait un, puisque le marché fonctionne comme s'il y en avait un. Autrement dit, la théorie de l'équilibre général construit une représentation du marché qui rend nécessaire l'existence d'un crieur, mais cette existence est justifiée par l'argument surprenant que le marché fonctionne réellement "comme si" il existait. Or, cette réalité n'est que celle du modèle lui-même. La fiction du crieur n'est nécessaire que dans le cadre du modèle théorique, où les prix évoluent comme s'il existait, mais pas dans la réalité, où ils évoluent de façon plus chaotique. Ici l'argument classique par lequel l'économiste justifie ses hypothèses les moins réalistes, en arguant du fait qu'elles représentent une bonne abstraction de la réalité, est tout-à-fait circulaire, puisque la "réalité" à laquelle il se réfère n'est autre que le modèle qu'il a construit... Le commissaire walrassien n'est pas seulement une image. Il est nécessaire pour assurer la convergence vers un équilibre dans un univers Arrow-Debreu.

Ce qui apparaît à travers cet embarras de la théorie néoclassique à propos de la figure du commissaire-priseur, que Walras assumait parfaitement, c'est une différence de projet théorique. Walras sait qu'une construction institutionnelle est un préalable nécessaire à l'existence d'un équilibre de marché. Il situe donc la question de la stabilité sur le plan pratique, ou au moins mécanique, mais en aucun cas formel. La temporalité du processus de tâtonnement n'est pas la temporalité logique du raisonnement déductif. C'est une forme de temporalité qui implique de pouvoir penser le changement d'état, le mouvement. C'est donc, au minimum, une temporalité "physique". De fait l'analyse de Walras prend comme modèle la mécanique et cherche à faire apparaître le marché non pas comme une construction logique abstraite, mais comme un mécanisme concret. C'est ce

mécanicisme que rejette la théorie moderne, qui entend donner à la question de la stabilité, une réponse située sur le même plan formel que la question de l'équilibre. Cela s'avérera impossible.

322- Le théorème de Sonnenschein.

Dans les termes de l'analyse moderne, la question de la convergence vers un point d'équilibre se décompose logiquement en deux questions successives : à quelles conditions l'équilibre est-il unique, puis, à quelles conditions cet équilibre est-il atteignable ?

La question de l'unicité est donc évoquée avant celle de la stabilité. Arrow et Hahn rappellent ainsi, que cette question est en particulier nécessaire pour utiliser le modèle à des fins de prospective³⁸. Or, précisément, les résultats en termes d'unicité sont assez négatifs : *“Presque tous les économistes, qu'ils soient ou non théoriciens, si on leur demande d'évaluer les conséquences probables de long terme d'un changement dans l'un des paramètres de l'économie, tenteront tout d'abord de comparer l'équilibre avant et après le changement. Si l'équilibre n'est pas unique cette procédure peut être erronée. Malheureusement, les conditions nécessaires (pour l'unicité) ont très peu de chances d'être remplies”* (Arrow et Hahn, 1971, p.242).

En fait, le problème principal n'est pas tant que ces conditions aient peu de chances d'être remplies. La théorie de l'équilibre général ne prétend pas être une description satisfaisante, mais une démonstration logique rigoureuse. Elle fait donc abstraction d'un grand nombre de “réalités” économiques, et on voit mal pourquoi l'adjonction d'une hypothèse non-réaliste supplémentaire ferait problème. La difficulté se situe ailleurs. La véritable ambition de la théorie est de fournir une démonstration cohérente entièrement déduite de la loi de comportement établie au niveau individuel. Or, c'est cette ambition qui est remise en cause par les “résultats” des études sur l'unicité. Le principal d'entre eux consiste en effet à faire dépendre l'unicité d'une condition portant sur les biens, et non sur les individus : l'hypothèse de substituabilité brute³⁹. Cette hypothèse suppose en effet que tous les biens pris deux à deux sont des “substituts” bruts. Elle porte sur la forme que prennent les courbes de demandes nettes, et ne peut être déduite des postulats de base concernant la rationalité, les goûts, les dotations des agents. Cette solution n'est donc pas seulement irréaliste, elle est surtout logiquement non satisfaisante, puisqu'elle

³⁸ Ainsi, la “statique comparative” qui a pour objet de comparer les positions d'un système avant et après un choc exogène, c'est-à-dire une crise, repose sur cette hypothèse d'unicité de l'équilibre. A défaut, il est impossible de prévoir vers quel équilibre le système converge. Or, comme le rappellent Quirk et Saposnik [1968, p. 207], la statique comparative constitue “la plus grande partie de ce qu'on peut appeler le corps de la théorie économique traditionnelle”.

³⁹ Ce résultat est dû à Arrow Block et Hurwicz [1959].

revient à faire dépendre la solution d'une hypothèse supplémentaire, non comprise dans les prémisses de la démonstration.

Ce problème se pose de façon identique en ce qui concerne la question de la stabilité de l'équilibre. Arrow-Block et Urwicz [1959] ont ainsi établi que l'hypothèse de "substituabilité brute" est également une condition suffisante, mais nécessaire, pour que la stabilité de l'équilibre soit assurée. La formulation de cette condition s'apparente à un aveu d'échec : le processus de stabilité ne se déduit pas logiquement des caractéristiques des individus marchand.

L'espoir d'établir une relation logique entre la forme que doivent nécessairement prendre les fonctions de demande nettes pour que l'unicité et la stabilité de l'équilibre soient assurées, et la rationalité des agents, s'est en effet rapidement évanoui. Dans une série d'articles indépendants, Sonnenschein, Mantel, et Debreu⁴⁰ ont montré, dans les années soixante-dix, que : *"toute fonction de demande définie de R^+ vers R , respectant la loi de Walras et qui est continue et homogène de degré zéro, peut-être considérée comme une fonction de demande nette d'une économie d'échange dont les n individus ont des préférences continues et strictement convexes"* (Guerrien [1985, p.172]). Rien ne permet donc de supposer que la fonction de demande nette agrégée prenne une forme assurant l'unicité et la stabilité de l'équilibre. Les axiomes définissant la rationalité individuelle sont compatibles avec n'importe quelle fonction de demande nette. Ils ne suffisent donc pas à établir l'existence d'une quelconque loi assurant qu'un seul équilibre sera atteint, et qu'il sera stable.

L'ajout d'une hypothèse *ad-hoc* assurant la forme de la fonction agrégée de demande nette est donc nécessaire pour assurer la stabilité de l'équilibre. Hahn conclut ainsi les "résultats" de la théorie néoclassique en matière de stabilité : *"Dans le cadre de l'équilibre général, dans lequel beaucoup de prix changent simultanément, l'analyse de la "la loi de l'offre et de la demande" devient plus complexe et la convergence vers un équilibre plus problématique. Nous devons en tirer la conclusion qu'il y a une grande partie des économies pour lesquelles l'équilibre est instable dans le cadre de la forme la plus populaire du mécanisme des prix"* (Hahn, 1980, p.745). Ainsi, l'espoir de déduire de la loi individuelle une loi collective assurant, en pratique, l'efficacité du marché tourne court. La question de la stabilité ne reçoit pas, dans le cadre d'analyse Arrow-Debreu, de fondations logiques satisfaisantes. La main invisible semble donc condamnée à demeurer un mythe. L'analyse logique ne parvient pas à lui donner une consistance.

Le programme de la théorie de l'équilibre général, dans sa version Arrow-Debreu tout au moins, a trouvé son point d'orgue, mais aussi d'achèvement, avec la démonstration

⁴⁰ Sonnenschein [1972] et [1973]. Mantel [1974]. Debreu [1974].

du théorème d'existence. Or, ce théorème d'existence perd considérablement de sa portée, s'il n'est pas complété par une théorie présentant les moyens d'en assurer l'effectivité.

33- Une économie circonscrite à l'équilibre

“La théorie économique moderne se présente jusqu'à l'excès comme une théorie de l'équilibre. Elle étudie des positions dont rien n'incite à s'éloigner parce que les plans et les désirs des agents économiques y sont mutuellement compatibles, mais elle est presque silencieuse sur la question de savoir comment de telles positions sont atteintes, et comment les grandeurs économiques se forment lorsqu'elles ne s'établissent pas d'emblée à leurs valeurs d'équilibre”. (Fisher, 1990, p.21).

La théorie de l'équilibre général tire sa force de sa cohérence logique. Cette cohérence naît d'une conception intégrée de l'individu au système de marché. L'individu est un des éléments du système marchand. La loi à laquelle son comportement se conforme est une loi qui émane de l'univers marchand. Or, les raisons qui font que cette conception de l'individu permet d'établir logiquement l'existence d'un équilibre concurrentiel empêchent également de comprendre comment les individus s'approchent de cet équilibre. Cette question pose le problème de l'existence *pratique*, et non plus logique, de l'équilibre. La question de la stabilité est donc une question de nature pratique. Dès lors tout est dit, cette question n'a aucun sens dans l'univers Arrow-Debreu⁴¹. Cet univers repose sur une temporalité purement logique. Dans un tel univers, il ne peut donc pas y avoir de “processus”, de “convergence”, de “sentier d'équilibre” parce qu'il ne peut y avoir de mouvement.

On peut s'en convaincre en observant les développements concernant la stabilité qui ont été menés en dehors de la perspective Arrow-Debreu. L'analyse “orthodoxe” d'Arrow, Block et Hurwicz [1959] ne constitue pas, en effet, l'unique mode d'étude du processus de stabilité. L'un des spécialistes de cette question, Franck Fisher, indique ainsi : *“l'échec du tâtonnement n'implique cependant pas celui de la stabilité, et le début des années soixante vit l'ébauche de développements plus fructueux”* (Fischer, 1990, p.32). Hahn [1962a] puis Hahn et Negishi [1962b] ont notamment contribué à remettre en cause le principe du tâtonnement, le commissaire-priseur, l'absence de monnaie

⁴¹ Les doutes qu'entretient Debreu quant à la pertinence de la question de la stabilité nous semblent cohérents avec cette interprétation.

(Arrow et Hahn [1971]). Mais, ces tentatives demeurent dominées par l'idée d'équilibre. Les individus agissent comme s'ils savaient qu'ils convergent vers l'équilibre sans chercher à profiter directement des situations de déséquilibre en cherchant à arbitrer⁴². La véritable échappée hors du "camp de base", selon l'expression de Hahn, est réalisée par la théorie du "search". Cette perspective vise, précisément, à prendre en compte le fait que les individus, non convaincus de l'existence d'un équilibre à long terme, cherchent à profiter des opportunités à court terme. Les individus agissent donc dans le temps, et retrouvent une certaine autonomie par rapport au système. Or, à compter du moment où l'individu n'est plus entièrement incorporé à la machine marchande, l'équilibre n'est plus assuré, et le concept même d'équilibre change : *"Même si le processus d'échange est globalement stable, l'équilibre finalement atteint ne sera généralement pas un de ceux qui correspondaient aux dotations initiales, au sens statique de la correspondance walrassienne. L'équilibre en question dépendra plutôt du sentier suivi, donc de la dynamique du processus engendré par le déséquilibre"*.(Fischer, 1983, p.31, c'est moi qui souligne).

Ce que nous donne à voir l'étude d'une économie marchande en dehors de l'équilibre, c'est que l'idée de processus marchand implique deux choses : une temporalité "réelle", ne serait-ce que mécanique, et une certaine autonomie de l'individu par rapport à l'environnement marchand. Pour comprendre comment l'équilibre peut se réaliser, il faut donc s'écarter de la conception standard de l'individu. Dans le cadre de cette conception logique, l'idée d'autonomie de l'individu par rapport au système des prix n'a pas plus de sens que celle de l'autonomie de la pomme par rapport à la loi de l'attraction terrestre. Le modèle résout en effet les deux problèmes classiques de l'interdépendance et de l'incertitude en éliminant tout contact direct entre les individus et en supprimant toute dimension temporelle à leur comportement (qui est réduit à un calcul formel). Cette simplification vise à toucher du doigt le cœur du comportement marchand, à n'en conserver que son essence profonde. Mais, si elle permet de forger une conception précise de ce que peut abstraitement être un équilibre de marché, elle échoue à rendre compte de ce que peut être l'économie de marché en dehors de l'équilibre⁴³. Comme le

⁴² Les tentatives menées par Hahn visent à prendre en compte une partie des échanges hors équilibre : seulement ceux qui s'inscrivent dans la perspective d'une allocation d'équilibre (Fischer, 1990, p.32-37). Sur cette question de la stabilité, on peut consulter le survey par Negishi [1962] comprenant l'analyse des premiers problèmes de non-tâtonnement. Arrow et Hahn [1971] proposent un bilan d'étape très complet. Hahn [1980, p.745-793] dresse un bilan de la théorie de la stabilité qui fait aujourd'hui office de référence. Fischer [1990] réactualise ce bilan, en se montrant plus critique, et en introduisant l'analyse du "search".

⁴³ Même à l'équilibre, Ostroy et Starr [1974] ont démontré qu'il était impossible de décrire une procédure décentralisée d'effectuation des échanges, sous forme de troc comme le veut la théorie : *"Les réquisits informationnels du marchandage impliquent la nécessité d'une coordination centralisée des échanges ; la fonction de la monnaie comme médium commun est de permettre une décentralisation de ces échanges"* (Ostroy et Starr [1974, p.1094). Toutes les tentatives d'intégration de la monnaie dans un modèle

note Cartelier [1995, p.232], “la définition de l'individu et les hypothèses relatives à son comportement ne fournissent pas une base adéquate pour développer une théorie des relations individuelles”. L'individu standard ne croise, en effet, jamais les autres. Il n'en a pas besoin pour relier son propre désir aux biens objectivement définis. Il n'en a pas besoin pour déterminer à quels prix ces biens sont accessibles, puisque le système des prix l'informe parfaitement. Il n'en a pas besoin pour adapter ses propres choix aux possibilités qu'ouvrent l'économie concurrentielle, puisque le crieur l'informe parfaitement des “désirs des autres” en les répercutant sur les prix relatifs de l'ensemble des biens. Donc, l'individu n'est pas programmé pour une situation dans laquelle il serait amené à rencontrer les autres. Son mode de rationalité ne peut fonctionner que dans l'isolement⁴⁴.

De ce fait, dès lors que le système est en dehors de l'équilibre, l'individu rationnel est défaillant. Hors de l'équilibre, l'économiste n'a plus rien à dire de l'économie. La théorie prétend se fonder sur l'étude des choix individuels. Mais le comportement de l'individu n'est finalement rien d'autre que l'action du système lui-même : “L'individu de la théorie économique moderne ne peut agir librement qu'en respectant strictement un schéma (pré)-déterminé au niveau de l'ensemble de la société ”(Cartelier [1995, p.238]). Or, lorsque le système est en déséquilibre, il est nécessaire de disposer d'un individu capable d'agir afin de le ramener à l'équilibre. Un individu dont les compétences ne se résument donc pas à recevoir l'information sous forme de prix donnés, mais qui au contraire cherche de l'information et arbitre en fonction des opportunités qu'il rencontre⁴⁵. L'individu au fondement du modèle Arrow-Debreu [1954] est loin de pouvoir accéder à une telle autonomie.

Le modèle Arrow-Debreu paraît donc condamné à rester une démonstration d'existence sans son complément indispensable qui serait une démonstration de la stabilité, et de l'unicité, de l'équilibre. Cette incomplétude de la théorie de l'équilibre générale est une conséquence directe de l'articulation particulière qu'elle construit entre l'individu et du système. L'individu contient le système, et, dans le même temps est contenu par ce système. Le modèle repose à la fois sur trop d'individualisme et pas assez

d'équilibre général ont cependant inévitablement échoué. Ainsi Helwig [1993, p.215] affirme : “Je pense que nous ne disposons pas encore d'un cadre théorique adapté pour étudier le fonctionnement d'un système monétaire. Le principal obstacle pour le développement d'un tel cadre est notre habitude de penser les problèmes dans un cadre organisé, sans friction, c'est-à-dire celui des marchés walrassiens.”

⁴⁴ C'est un des points d'insatisfaction de Morgenstern qui souligne que l'analyse du marché concurrentiel met si peu en rapport les individus que l'on peut en déduire que c'est “Hamlet sans Hamlet” (Morgenstern, 1974, p. 291)

⁴⁵ Outre les références déjà citées (Fischer [1990], Cartelier [1995], Benetti et Cartelier [1994]) on peut se reporter aux articles de Kirman [1985,1991], qui cherche précisément à étudier ce que peut être une économie dans laquelle les individus communiquent.

d'individualité, pour saisir l'effectivité de la relation marchande dont la bonne compréhension est un préalable à l'analyse de la stabilité.

Conclusion

La théorie néoclassique de l'équilibre général vise à formuler une loi. Elle cherche par là à donner un fondement logique et rigoureux au concept de main invisible, en montrant que le mécanisme marchand permet une répartition Pareto-optimale. C'est donc une démarche analytique, visant à mettre à jour l'existence d'une relation logique entre les mouvements de prix au niveau global, et les offres et demandes des individus. Elle construit pour ce faire un univers permettant d'exprimer ces relations dans toute leur pureté. Cet univers se caractérise par une conception logique du temps, c'est-à-dire en fait par une absence de temporalité réelle, et par un profond isolement de l'individu, qui ne connaît pas l'altérité. De ce fait, le modèle est dénué de toute dimension "pratique". Il repose en effet sur la suppression du problème de la gestion du temps, et donc de l'incertitude, et de celui de l'accord collectif, c'est-à-dire de la coordination. Le problème économique de la répartition est donc traduit sous une forme purement logique et peut recevoir une réponse mathématique.

Mais, ce projet échoue. Les démonstrations de l'existence et de l'optimalité de l'équilibre constituent sans aucun doute un résultat remarquable et essentiel. Mais il ne prend sens que s'il est associé à une démonstration de l'unicité et de la stabilité de l'équilibre. Or, il apparaît qu'il n'est pas possible de déduire logiquement l'existence d'un processus de stabilité à partir des caractéristiques de la loi de comportement de l'individu, sur le modèle de la démonstration de l'existence de l'équilibre. Le "théorème de Sonnenschein" fait en effet apparaître qu'il n'existe aucune relation logique permettant de déduire la forme des fonctions de demande nette à partir des fonctions d'utilité individuelle.

Ce que fait apparaître cet échec de la théorie de la stabilité c'est, d'une certaine manière, le "prix" d'une analyse purement mathématique et l'étroitesse de son champ d'application. L'analyse logique rend nécessaire l'exclusion du temps et des interactions directes entre individus. Or, cette exclusion nous empêche de penser le processus pratique de la détermination des prix d'équilibre. Pour comprendre la stabilité il apparaît nécessaire de remettre en question la conception de l'univers marchand. Cet univers manque d'une dimension pratique.

C'est ce manque que les critiques de Kenneth Arrow tentent de pallier. D'une part en cherchant à reposer la question de la compatibilité des niveaux de rationalité individuelle

et collective, ce qui est, au fond, la question de la coordination. D'autre part en cherchant à introduire une certaine dimension temporelle dans le modèle d'équilibre général. En remettant en jeu les problèmes de la coordination et de l'incertitude, ses critiques vont faire resurgir le "problème de la règle" que l'énoncé d'une loi de comportement avait fait disparaître.

Chapitre 2

La critique d'Arrow

Kenneth John Arrow a été à l'origine de la plupart des résultats théoriques de la théorie de l'équilibre général, il est le premier à avoir formulé les deux théorèmes fondamentaux du bien-être (Arrow [1951a]), il est associé au plus célèbre théorème d'existence (Arrow-Debreu [1954]), et au théorème de stabilité le plus général (Arrow, Block et Hurwicz [1959]). Cette contribution décisive à la théorie de l'équilibre général s'est cependant accompagnée, très tôt, d'une lucidité sur les limites de cette entreprise théorique, ou, plus exactement, sur la nécessité d'étayer le modèle de marché d'autres hypothèses institutionnelles. Arrow réalise donc le tour de force d'être tout à la fois un des pères de la théorie moderne de l'équilibre général, et l'un de ceux qui a le plus contribué à la prise de conscience du caractère limité des résultats de ce paradigme.

L'analyse "critique" d'Arrow répond aux problèmes rencontrés par la théorie de l'équilibre général, en cherchant à réintroduire un type de questionnement *pratique* sur le fonctionnement du mode marchand. Or, ce questionnement pratique, en mettant en lumière non plus seulement la *cohérence* des choix individuels mais leur *fondement*, fait apparaître les limites d'une forme de rationalité exclusivement instrumentale. Elle présente en effet deux types de limites.

La première porte sur la question des fins. Le système marchand formalisé par Arrow et Debreu repose sur un mode de comportement exclusivement instrumental. La rationalité, dans ce modèle, ne désigne qu'une réflexion concernant les moyens d'une fin donnée : la maximisation d'une dotation individuelle en biens et services. Or, le résultat d'un système fondé sur une telle forme de rationalité peut être jugé collectivement non souhaitable. Arrow [1951b] dans sa célèbre étude sur la détermination du choix collectif montre ainsi que le système de marché ne peut être considéré comme un mécanisme

satisfaisant d'agrégation des préférences individuelles, lorsqu'elles portent sur des états sociaux. Son théorème d'impossibilité démontre que le marché ne peut pas déterminer à lui seul l'état social que la collectivité souhaite majoritairement. Le mécanisme marchand nécessite, en amont, une forme de consensus sur les fins "collectives" à atteindre. Or, ce consensus repose sur une forme de rationalité qui n'est pas instrumentale. Il est réalisé, *en pratique*, parce que l'individu est socialisé et respecte certaines valeurs collectives. Autrement dit, l'existence d'un accord éthique constitue un préalable à l'analyse logique du phénomène marchand.

La seconde ligne critique d'Arrow porte sur le thème de l'incertitude. La distinction entre fin et moyen, au fondement de l'attitude instrumentale, nécessite une parfaite prévisibilité des conséquences des actions possibles. Or, Arrow argue du fait que le système marchand ne peut, *en pratique*, assurer cette parfaite prévisibilité. Il est incomplet, et cette incomplétude produit de l'incertitude. De plus, lorsque les individus sont incertains, ils cherchent à s'informer au-delà du système des prix. Rien n'assure qu'ils accèdent finalement aux mêmes informations : l'incertitude crée de l'hétérogénéité. Il faut alors mettre en place des "règles communes" pour harmoniser les anticipations des agents, et "sauver" ainsi le processus marchand. L'introduction d'une forme de temporalité "réelle" met donc en évidence la nécessité des repères collectifs.

Ces deux lignes critiques se rejoignent : l'existence d'un consensus social sur les fins collectives ne peut-être efficace que si le marché représente un mécanisme capable d'éliminer l'opacité du futur. En effet, si le marché ne permet pas à coups sûrs d'atteindre une fin déterminée en dehors de lui-même, parce qu'il gère mal l'incertitude quant au futur, il est nécessaire que le choix collectif soit en permanence rappelé, imposé à chacun des individus dans le cours de son action. Il faut que l'accord collectif à atteindre s'incarne dans des repères collectifs qui, au cours du processus marchand, puissent guider les individus dans une direction satisfaisante. Les deux critiques d'Arrow montrent donc qu'il est nécessaire d'établir une *passerelle* entre l'individu et le collectif. C'est précisément le problème de la règle qui se retrouve ainsi posée. Il ne pouvait pas apparaître dans le cadre logique de l'univers Arrow-Debreu. Le questionnement critique d'Arrow, son souci de donner un contenu pratique plus important à la théorie, le fait réapparaître.

Pourtant Arrow ne se pose pas la question du *choix* des règles. Les repères collectifs qu'il identifie sont des *instruments* que la société utilise pour atteindre ses fins. Pour répondre à cette exigence d'efficacité, un type particulier de règle s'impose : la règle *organisationnelle*. Les règles organisationnelles sont les relais des normes collectives. Elles assurent une conformité des choix individuels et imposent une cohérence entre ces choix et l'objectif d'efficacité collective.

Arrow met donc finalement en évidence les *conditions* “holistes” de l’expression de l’individualisme marchand. Il éclaire la face cachée du modèle de marché. Il démontre l’incomplétude de la *modalité marchande* du traitement du problème de la règle. Le marché n’est autonome que parce que les organisations prennent en charge ce problème, en imposant des règles efficaces.

1- Le théorème d'impossibilité

“Dans une dictature idéale il y a une volonté unique impliquée dans le choix social, de même, dans une société idéalement régie par une convention, il existe une seule volonté qui peut être d'essence divine, ou encore, par hypothèse, être une volonté commune à tous les individus. Dans ces deux cas, il n'existe aucun conflit entre les individus concernant le choix social. A l'inverse, les méthodes du vote et du marché construisent le choix social en agrégeant [amalgamating] les goûts de plusieurs individus. Or, les méthodes de la dictature et de la convention sont, ou peuvent être, rationnelles, dans le sens où n'importe quel acteur isolé peut être rationnel dans ses choix. Mais, peut-on attribuer la même consistance logique aux modes de choix collectifs dans lesquels la volonté de plusieurs personnes est impliquée ?” (Arrow, 1951b, p.2)

On connaît la réponse d'Arrow à cette question : elle est négative. Le théorème d'impossibilité qui synthétise ce résultat est formulé par Arrow dès 1950¹. Il ne le modifiera plus par la suite². Quelques pages lui suffisent pour démontrer qu'il n'existe aucune fonction de choix social permettant de déduire le choix collectif à partir de l'énoncé des choix individuels, si toutefois les individus sont supposés être rationnels et absolument libres de déterminer leurs choix comme ils l'entendent. Ces quelques pages se concluent sur un constat sans appel : “le marché ne peut pas déterminer un choix social rationnel ” (Arrow, 1950, p.23 ; 1951b, p.59).

Il est utile de revenir aux premières lignes de l'ouvrage d'Arrow pour comprendre les liens précis qu'entretient le théorème d'impossibilité avec le mécanisme marchand. Arrow, s'appuyant sur Knight, y distingue trois modes de prise de décision collective : la dictature, la coutume ou la convention, et le “consensus”. Les deux premiers modes impliquent l'existence d'une *unité* de choix entre des différents individus. Cette unité rend triviale la détermination du choix collectif. Le troisième de ces modes de décision repose

¹Arrow formule le théorème dans un court article dès 1950 (Arrow, 1950). Mais on fait plus volontiers référence à l'ouvrage dans lequel il expose et commente ce théorème. Cet ouvrage paraît en 1951 (Arrow, 1951b). Il est republié intégralement en 1963, et augmenté d'une préface dans laquelle Arrow répond à certaines critiques, notamment celles de Bergson et de Little, concernant le sens de ce théorème pour la nouvelle économie du bien-être (Arrow, 1963a). Il est traduit en français en 1974 (Arrow, 1974). Nos indications de pagination renvoient toutes à l'ouvrage de 1963.

²Suite à une remarque critique de Blau [1957], qui a noté l'existence d'une erreur dans la formulation de la condition U dans la version initiale du théorème, Arrow [1963], dans la réédition de son ouvrage, a reformulé son théorème en corrigeant cette erreur. Il a également proposé, dans sa réédition, ainsi que dans un article en français (Arrow, 1952) une démonstration différente montrant la proximité entre certaines de ces conditions (U, M,I) et le principe de Pareto. Mais, fondamentalement, Arrow n'a rien concédé sur le plan théorique.

au contraire sur le respect de la *pluralité* des individus, conformément à l'idéal démocratique moderne. Il ne nie donc pas l'existence de conflits entre acteurs. Il cherche au contraire à les résoudre par le biais d'un mécanisme de choix social respectueux de la diversité des opinions. Il s'agit alors d'agréger des décisions individuelles, pour en faire une décision collective. Arrow (1951b, p.1) note : "*il existe essentiellement deux méthodes [de ce type] : le vote (...) et le mécanisme marchand (...)*". C'est à ce type de mécanisme qu'Arrow tente de donner un fondement logique.

Sa démonstration se situe donc sur le plan de la logique formelle. Or, à ce niveau d'abstraction les multiples différences pratiques qui existent entre ces deux modes d'agrégation des préférences individuelles disparaissent : "*dans ce qui va suivre, la distinction entre le vote et le mécanisme de marché va être mis de côté, les deux étant considérés comme des cas spéciaux d'une catégorie plus générale : la fonction de choix collectif*" (Arrow, 1951b, p.5). Arrow teste la cohérence logique du *principe* de l'agrégation des choix individuels, que ce soit le marché ou le vote. La démarche d'Arrow aboutit finalement à une démonstration "par l'absurde". Ce qu'elle fait apparaître, c'est l'existence d'une discontinuité logique entre les niveaux de rationalité individuelle et collective. Il est impossible de passer directement de l'un à l'autre. La cohérence interne de chacun des ordres de préférence individuels n'assure pas la cohérence des différents ordres entre eux. Cette discontinuité brise l'espoir de disposer d'un mécanisme d'agrégation qui soit logiquement satisfaisant.

C'est un résultat purement logique, mais il est lourd d'implications pratiques. Il montre en effet que pour que le problème posé par Arrow admette une solution, il ne suffit pas que les préférences individuelles soient cohérentes. Arrow prolonge donc sa démonstration par une réflexion sur leur *fondement*. C'est en effet, selon lui, le fondement moral des préférences sociales qui rend possible la formulation d'un choix collectif. Or, le mécanisme marchand exclut cette dimension morale. C'est la raison pour laquelle le choix collectif ne peut être déterminé qu'en *amont* du marché. Le marché n'est qu'un *instrument* permettant de l'atteindre.

Ainsi, Arrow, en montrant que la solution au problème du choix collectif ne se situe pas sur un plan logique mais sur un plan pratique, fait émerger un problème que l'univers formel Arrow-Debreu excluait : celui du choix des règles collectives.

11- Enoncé du problème

“La question fondamentale de la théorie du choix social est la suivante : soit une série de décisions sociales possibles dont l'une doit être choisie. Etant donné le critère de décision associé aux individus dans la société, trouvez une méthode permettant de déterminer celle qui doit être choisie.” (Arrow, 1985, p.124).

111- Platon versus Bentham

Avant d'étudier les difficultés liées au mécanisme de choix social, Arrow (1951b, p. 22-23) établit clairement l'alternative qui va dominer l'ensemble de sa réflexion. Il oppose en effet une conception “classique”, de tradition platonicienne, à la conception moderne, mise en place par Bentham. La conception platonicienne définit le bien commun comme une entité transcendante, qui ne serait perceptible que par quelques philosophes éclairés. A cette conception, Arrow associe d'emblée, même s'il souligne que ce n'est pas une conséquence obligatoire, l'idée d'un gouvernement des élites, d'une dictature. Il lui oppose la conception utilitariste, héritée de Bentham, qui définit le choix social comme l'agrégation des choix individuels. Ainsi, tandis que dans une optique platonicienne le bien commun est immuable et ne dépend pas des préférences individuelles, dans l'optique benthamienne, il s'agit de rassembler, d'agréger, les préférences individuelles pour atteindre une expression du bien collectif. Arrow s'inscrit donc dans la modernité en cherchant à établir une procédure cohérente d'agrégation des choix individuels en un choix collectif.

Cette démarche est très proche de la théorie de l'équilibre général. En effet, il s'agit de déduire le choix collectif à partir du choix d'une collection d'individus sans aucun rapport entre eux. Le collectif doit en quelque sorte être rendu absolument transparent aux individus. La solution proposée par Bentham reposait sur la sommation des utilités associées aux divers états sociaux à comparer. Une telle solution suppose qu'il existe une unité commune à tous, de manière à ce que l'on puisse mesurer objectivement l'utilité associée à chacun de ces états sociaux soumis au choix des individus. Cette idée de mesure, qui suppose que l'on puisse opérer des comparaisons interpersonnelles d'utilité, est rejetée par la théorie moderne qui, on l'a vu, substitue à une conception “cardinale” de l'utilité une conception purement ordinale. C'est donc à partir d'une conception strictement ordinale que s'opère la démonstration d'Arrow³.

³ En cela il s'inscrit dans le champ de la nouvelle économie du bien être, qui, précisément oppose à la conception “ordinale” benthamienne une conception “ordinale” héritée de Pareto.

112- Un classement portant sur des "états sociaux"

La question du choix social porte sur la possibilité d'agréger des choix individuels. Ces choix individuels portent donc eux-mêmes sur des "états sociaux". Est-ce que la nature de cet objet modifie la nature du choix rationnel lui-même ?

Répondre à cette question relève de la gageure, comprendre précisément comment Arrow lui-même y répond relève de l'impossible⁴. Car, en fait, Arrow n'y répond pas : il estime que cette question, tant que l'on se situe au niveau logique du théorème, n'est pas légitime.

Il souligne cependant que le choix social est spécifique. Dans le cadre des choix individualistes tels qu'ils s'expriment par exemple sur le marché, le "choix" individuel est unidimensionnel : seules des questions d'intérêt personnel et matériel sont posées à l'acteur. Dans le cadre de la théorie du choix collectif, on suppose que les acteurs classent des "états sociaux" dont "*la définition la plus précise serait une description complète des quantités de biens aux mains de chacun des individus, la quantité de travail que chacun doit offrir, la quantité de chaque ressource productive, l'importance des diverses activités publiques comme les services municipaux, la diplomatie, ou encore l'érection de statuts en l'honneur de certains hommes célèbres*" (Arrow, 1951b, p.17). Le choix porte donc sur une organisation sociale, intégrant une répartition des richesses, et une série de choix en matière de développement social et de justice. La délibération de l'individu porte donc, *potentiellement*, sur des critères beaucoup plus larges que la maximisation de sa propre dotation en biens. Des motifs non égoïstes *peuvent* venir "polluer" la pureté du choix marchand. Des préférences "fondamentales", ayant trait aux valeurs, peuvent venir se substituer, lorsque le choix individuel porte sur des états sociaux, aux préférences "individualistes" ou "égoïstes" qui déterminent les choix marchands. C'est un point qu'Arrow ne conteste pas, même s'il doute du caractère logiquement "valable" de cette distinction⁵.

Cet aspect du choix social complique singulièrement l'interprétation du choix individuel, lorsque l'on cherche à en éclairer les fondements. Mais tel n'est pas le cas ici,

⁴ "Quant à Arrow lui-même ses textes présentent sur ce point essentiel un merveilleux mélange de lucidité, de confusion et de contradictions" (Kolm, 1986, p.319).

⁵ "On pourrait réserver le terme de "valeurs" pour un ensemble de choix particulièrement nobles ou supérieurs. Les choix en général pourraient être désignés par le terme de "goûts". Habituellement, on ne considère pas que la préférence d'une quantité de pain supplémentaire par rapport à une quantité de bière supplémentaire relève d'une enquête philosophique profonde. Je crois cependant que **cette distinction ne peut être faite logiquement**, et certainement pas en ce qui concerne les choix d'un individu isolé. S'il existe une distinction entre goût et valeur, elle appartient au domaine des relations interpersonnelles." (Arrow, 1967, p.60, c'est moi qui souligne). Cette distinction est due initialement à Friedman (Arrow, 1951b, p.18, note 14).

puisque Arrow s'intéresse, au moins dans un premier temps, non pas au fondement, mais à la *cohérence* des préférences entre elles. Ces préférences doivent respecter des principes logiques portant sur leur ordonnancement, mais aucune exigence ne pèse sur leur fondement. Ce n'est qu'après avoir formulé le théorème d'impossibilité, attestant de l'existence d'un problème logique, qu'Arrow reviendra sur la nature particulière du choix social. C'est en effet là un problème pratique, qui ne se pose qu'une fois constatée l'impossibilité logique.

Dans un premier temps, on ne cherche donc pas à éclairer les raisons des choix de chaque individu. La nature du critère de choix des individus, qu'ils soient matérialistes ou idéalistes, égoïstes ou généreux, n'altère pas la forme logique de leur classement des états sociaux. Sur le plan formel, donc, le critère du choix n'est pas une question pertinente. L'objectif d'Arrow est simplement de comprendre comment de tels choix peuvent être agrégés, comme le prévoit la théorie utilitariste du choix collectif. L'individu au fondement de sa démonstration est donc en tout point identique à l'individu standard, et fait preuve d'une rationalité du même type. L'hypothèse de rationalité individuelle renvoie uniquement, dans un premier temps, à des propriétés logiques d'ordonnancement. Elle ne comporte aucune dimension pratique. "Rationnel" se dit d'un individu sachant organiser les états sociaux selon un ordre complet transitif et réflexif.

113- La rationalité du choix collectif

Le choix rationnel est donc défini comme étant une relation d'ordre complète. Cette définition se décline en deux axiomes. Ainsi, un individu *i* est rationnel, si et seulement si, ses préférences respectent les axiomes suivants :

Axiome I (complétude) : quelques soient *x* et *y* : «*x* R_i *y*» ou «*y* R_i *x*»

Axiome II (transitivité) : quelques soient *x*, *y* et *z* : si «*x* R_i *y*» et «*y* R_i *z*», alors, «*x* R_i *z*»

("x" et "y" étant des alternatives et "R" étant une relation de préférence individuelle purement ordinale signifiant "est préféré à" ou "est indifférent à")⁶.

Cette définition standard de la rationalité individuelle n'est pas problématique. Le problème réside dans son extension au niveau de la collectivité : "Tout au long de cet ouvrage, on supposera que les individus sont rationnels, c'est-à-dire que les relations R_i

⁶ A notre connaissance Arrow [1951b] n'expose pas la propriété supplémentaire de "réflexivité" (id : quelque soit "x", xR_x). C'est une propriété évidente de la relation R (qui correspond au symbole mathématique : " \geq "). Réflexivité et transitivité font de R un "préordre". La complétude fait de "R" un "préordre total".

satisfont les axiomes I et II. Le problème sera de construire une relation d'ordre, pour la société dans son ensemble, qui reflète un choix rationnel en admettant que R respecte également les axiomes I et II." (Arrow, 1951b, p. 19).

Avant toute chose, le choix collectif doit donc être rationnel, c'est-à-dire respecter les conditions classiques de complétude et de transitivité. Avec cette première exigence, le problème du choix collectif se précise. Arrow cherche en effet à articuler le principe du respect de toutes les préférences individuelles, à travers le mécanisme de l'agrégation, avec le principe de rationalité des décisions collectives. La volonté de respecter le principe d'agrégation est liée à l'éthique utilitariste. La volonté d'obtenir un choix collectif rationnel est liée à l'objet du choix collectif : "*le problème n'est pas de conseiller les citoyens en général, mais les décideurs publics*", ou encore, "*l'économie du bien-être, quelle que soit sa définition, est liée à l'adoption, par les pouvoirs publics, d'une politique économique*" (Arrow, 1963a, p.107,106).

Le choix collectif doit permettre l'action de la puissance publique dans une direction précise⁷. Cette visée du choix collectif implique que le classement des états sociaux soit transitif. Dans le cas contraire, en effet, on est confronté à une situation "indécidable" (du type $A < B < C < A$). Pour que le décideur puisse agir, il faut qu'une préférence claire se détache du choix collectif, et, donc, que le classement des états soit transitif. Dans le cas contraire, c'est le principe du *statu quo* qui prime, bien que ce choix ne corresponde à aucun ordre de préférence individuel. Un tel choix négatif constitue donc un échec dans l'optique de celui qui cherche à obtenir un choix collectif fidèle aux préférences des individus. Il revient en effet à choisir une alternative qui n'était pas proposée⁸. Pour Arrow, la méthode d'agrégation doit donc conserver la propriété de rationalité qui caractérise les ordres individuels. Le classement collectif des états sociaux doit être cohérent afin de rendre possible le choix d'une des issues.

Arrow s'attache donc à établir une méthode permettant d'agréger des préférences individuelles portant sur des états sociaux, afin de déterminer une préférence collective sur ces mêmes états, et selon une méthode préservant la rationalité des choix individuels. Le résultat de son étude est négatif : il n'en existe pas.

⁷ Il est par exemple au fondement du problème de la détermination des dotations initiales, c'est-à-dire des redistributions avant et après le processus marchand. Le mécanisme de marché fait correspondre, dans l'idéal, un équilibre à chacune des configurations dessinées par l'ensemble des dotations initiales de chacun des agents. La détermination de ces dotations devrait permettre, au cas où le mécanisme de marché serait infallible, de faire en sorte que l'équilibre atteint soit non seulement Pareto optimal mais aussi "juste".

⁸ Arrow (p.217) affirme ainsi : "*la rationalité collective dans le mécanisme de choix public est une attribution importante d'un système authentiquement démocratique.*"

12- L'axiomatique

121- Exposé de l'axiomatique

Arrow établit une série de cinq conditions que toute procédure de choix social devrait respecter pour être satisfaisante : “ces conditions sont naturellement des jugements de valeur et pourraient être mises en question ; pris ensembles, elles constituent de façon très générale les principes de souveraineté et de rationalité du citoyen.”(Arrow, 1950, p.19). Elles sont donc conçues comme des impératifs minimaux, raisonnables, visant à s’assurer que le choix collectif est bien *déduit* des choix individuels, selon un mode logique.

Ces conditions sont de deux ordres. Les trois premières visent principalement à assurer une cohérence logique au mécanisme de choix collectif. Les deux dernières portent plus explicitement sur des impératifs éthiques de liberté de choix. Elles ont pour objet de s’assurer que le choix collectif est bien obtenu par agrégation, et non pas en raison d’un pré-ordonnement qui rendrait cette agrégation triviale (Arrow, 1952, p.52).

Cette axiomatique s’appuie sur quelques termes techniques. Un *ordre de préférence individuel* est un classement individuel des états sociaux respectant les deux axiomes de rationalité. Un *profil* de préférences est la description de l’ensemble des ordres individuels. Une *constitution* est une fonction qui assigne à chaque profil un *ordre* collectif (ou social). L’axiomatique se décline alors selon les cinq conditions suivantes⁹ :

1- Condition U (universalité) : La constitution est définie pour tous les ordres de préférence logiquement admissibles sur l’ensemble des alternatives.

2- Condition M (monotonie) : Supposons un profil dans lequel x est socialement préféré à y. Supposons maintenant un nouveau profil, dans lequel x monte dans l’échelle de préférence d’un individu et ne diminue dans l’échelle d’aucun autre. Alors x est toujours préféré à y dans cette nouvelle situation.

3- Condition I (Indépendance vis-à-vis des alternatives hors de propos) : Soit S un ensemble d’alternatives. Soient deux profils dans lesquels l’ordonnement des alternatives contenues dans S est identique pour tous les individus, alors le choix social, concernant S, est identique pour ces deux profils.

4- Condition N (Non Imposition) : La constitution n’est pas imposée.

5- Condition D (non dictature) : La constitution n’est pas dictatoriale.

⁹Je suis littéralement la présentation corrigée d’Arrow [1985, p.125].

122- Commentaire de l'axiomatique

La première condition assure la généralité de la constitution. Elle assure le respect de la pluralité des choix individuels. Elle implique donc qu'aucun réquisit ne pèse sur les choix individuels, et que tous les profils soient possibles, ce qui est une conséquence logique de la liberté de choix. La seule restriction qui existe ici, comme en ce qui concerne les choix marchands, est l'impératif de rationalité dans la forme du choix. Conformément à la posture ordinale et individualiste, la constitution n'impose rien quant au(x) critère(s) déterminant les préférences des individus.

La seconde condition est proche du principe de Pareto¹⁰. Elle assure que les choix individuels et collectifs évoluent dans le même sens. C'est là un principe logique et éthique assez généralement admis. Le but étant de définir une constitution respectant les choix individuels, cette condition semble aller de soi.

La troisième condition est une condition classique du choix social. Pour l'explicitier, Arrow [1950, p.28] prend l'exemple de deux individus se partageant du vin et de la bière. L'un préfère avoir plus de bière, l'autre plus de vin. Si dans une autre situation les deux individus n'ont que de la bière à se partager, la condition I stipule que le fait qu'un des deux individus soit davantage brimé que l'autre par l'étroitesse du choix possible ne doit pas être pris en compte dans le partage. Autrement dit, le désir de vin étant "hors de propos", l'individu qui en est friand ne peut pas se servir de cette préférence brimée pour exiger plus de bière que l'autre. Cet exemple suggère bien que la levée de cette condition pose d'importants problèmes d'information. Comment vérifier en effet que l'amateur de vin dit vrai ? D'une façon générale, si le choix dans une situation donnée dépend d'alternatives "hors de propos", il faut alors qu'il y ait une information égale sur ces alternatives. Dans ces conditions, le choix collectif requiert une information parfaite sur toutes les situations possibles. Cet impératif ne pouvant être respecté, Arrow estime que la condition I est nécessaire¹¹.

Nous porterons davantage notre attention sur les deux dernières conditions. D'une part, elles signent la volonté d'échapper aux solutions "de type platonicien" du choix social, et sont donc plus profondément liées à l'éthique utilitariste et démocratique. D'autre part chacune de ces deux conditions rejette un des deux mécanismes de décision

¹⁰ Arrow [1963, 1985] a reformulé son théorème à partir du principe de Pareto en montrant que les conditions M, I et D, lorsqu'elles sont réunies, lui sont équivalentes.

¹¹ C'est sans doute la condition de l'axiomatique qui est la plus controversée, parce qu'elle ne découle pas directement de la liberté du choix, et que son rapport à la rationalité est ambigu. Arrow (1963, p.110) la justifie par le fait que le marché et le vote la respectent naturellement tous les deux. Elle a donc "*un grand avantage pratique*". Elle découle de plus de la transitivité : si, dans une société, on a $xRyRz$, alors on a aussi xRz . Le choix entre z et x est indépendant de l'alternative y .

collective écartés par Arrow au début de son analyse : la dictature et la convention. Ainsi, ces deux conditions N et D récuse respectivement les solutions que Hume et Hobbes apportent à la question de l'ordre social. Dans les deux cas, c'est le principe de "l'unité" des choix qui est rejeté, dans la mesure où il apparaît, d'une part comme étant le fruit d'une pression sur l'individu, d'autre part comme un mode "trivial" d'agrégation des préférences.

La définition du "choix imposé" est la suivante : " *Une fonction de choix social sera dite imposée si pour un couple quelconque d'alternatives distinctes x, y , on a xRy , quelques soient les préférences individuelles $R_1...R_n$ (où R est l'ordre social correspondant au profil $R_1...R_n$).* " (Arrow, 1951b, p.29).

Autrement dit, le choix est imposé si, quelques soient les préférences individuelles, " y " ne peut être collectivement préféré à " x ". La condition de non-imposition se traduit donc par le rejet des constitutions qui excluent certains ordres de préférence. Dans une telle constitution, il existe en effet des préférences collectivement taboues, qui ne peuvent donc exister qu'à l'échelon individuel. Le rejet, a priori, de certains profils de préférence entre en contradiction avec la généralité qu'institue la condition U, et s'oppose au principe même d'une agrégation " neutre " qui ne modifierait pas le profil qu'elle tente de synthétiser.

Cette condition conduit à rejeter les théories conventionnalistes de l'ordre social, fondées sur le respect de la coutume ou des conventions sociales. Le principe de la coutume est en effet de rejeter certaines préférences comme mauvaises, et de prôner le bien fondé de certaines autres, de manière à préserver une certaine forme d'ordre ou de cohésion sociale. En rejetant tout choix "imposé", Arrow passe sous silence le phénomène de la conformation sociale. Dans bien des cas, l'apprentissage du conformisme conduit les individus à aligner leurs désirs sur les conventions en usage. A un certain niveau il devient donc impossible de distinguer la préférence individuelle de la préférence imposée par la coutume. Le principe de monogamie, ou de polygamie, est un exemple de préférence où individuel et social sont étroitement mêlés. Mais c'est là un problème "pratique" qui n'a aucune validité logique. Si Arrow rejette toute solution qui ne fonctionne qu'à la condition que les préférences des individus portent déjà en elles la marque de la société, c'est parce qu'il cherche à isoler la question *logique* de l'agrégation de celle, *pratique*, de l'origine d'une éventuelle "harmonie" des préférences. Pour tester le mécanisme d'agrégation, et seulement lui, il faut s'assurer que les préférences individuelles sont bien, au départ, divergentes.

Le postulat individualiste est donc pris au pied de la lettre. Les choix doivent être absolument libres, et absolument individuels. Cela conduit à rejeter toute solution fondée

sur le conditionnement social des individus. Il faut considérer des individus tout-à-fait libres de toute pression sociale, c'est-à-dire, en fait, hors de toute société : comme dans l'univers marchand Arrow-Debreu.

La définition de la dictature est la suivante : *“Une fonction de choix social est dite dictatoriale s'il existe un individu i tel que pour tout couple d'alternative (x,y) , “ $x P_i y$ ” implique “ xPy ”, quelques soient les autres ordres individuels $R_1...R_n$. (P étant la relation de préférence collective, P_i étant la relation de préférence de l'individu-dictateur)”* (Arrow, 1951b, p.30).

La condition de non-dictature s'oppose directement, quoique non explicitement, à une solution de type hobbiennne. En effet, cette condition stipule que toute constitution reposant sur une équivalence directe entre le choix d'un individu particulier et le choix collectif doit être rejetée. Cette condition se comprend aisément dans le cas où un individu proclame ses préférences individuelles comme étant celles de tous, alors que le profil des préférences montre en réalité une grande divergence entre ses préférences et celles des autres individus. Toutefois, la définition formelle de la dictature inclut le cas où cette dictature est souhaitée par les individus. Lorsque les préférences entre individus ne sont pas très différentes, un individu particulier peut s'avérer être l'interprète honnête de la volonté générale. Là encore, la frontière entre préférences individuelles et collectives peut être étroite, en particulier si les individus intègrent les préférences du souverain comme étant les leurs. Pour cette raison, le principe de non-dictature se traduit, comme le précédent, par le rejet de toute solution unanimiste. Dans le cas limite d'une unanimité parfaite, en effet, le choix d'un individu résume le choix de tous les autres : sa préférence est donc identique à la préférence collective. C'est au fond l'idée même de Hobbes, selon laquelle le *Léviathan* naît de la volonté générale

Ces deux dernières conditions sont donc construites sur un même modèle. Elles visent à rejeter des modalités d'expression de l'ordre social qui ne seraient pas entièrement fondées sur l'agrégation des préférences exprimées par des individus “souverains”, c'est-à-dire, ici, libres de toute pression sociale. Arrow [1952, p52] souligne ainsi : *“Il existe certainement des méthodes triviales d'agrégation qui satisfont ces trois conditions [les trois premières].(...) Aucune de ces méthodes ne règle le problème de l'agrégation dans une société démocratique”*. En rejetant les méthodes reposant sur une forme de consensus préexistant, que celui-ci repose sur la contrainte violente ou sur l'adhésion volontaire, Arrow écarte donc des méthodes d'agrégation “triviale”. Il cherche précisément à isoler le processus logique d'agrégation lui-même, à le distinguer d'autres modalités d'expression du choix collectif. Il s'assure ainsi que la réussite de la procédure d'agrégation soit bien liée aux qualités intrinsèques de cette procédure, et non pas à des

phénomènes extérieurs de nature pratique, comme par exemple l'existence d'une habitude ou d'un tabou social. Pour que le principe d'agrégation soit réellement testé, il importe que le choix collectif ne soit pas, d'une façon ou d'une autre acquis à l'avance, en étant d'emblée inscrit dans chaque ordre de préférence individuel...

Ces cinq conditions forment en quelque sorte le pendant nécessaire au principe de rationalité du choix individuel, et assurent que l'individu soit affranchi de toute contrainte sociale. Comme dans le cadre des choix marchands, l'individu tire sa souveraineté de sa solitude. Sa liberté absolue s'entend comme un affranchissement de toute influence des autres sur son propre choix ou, tout au moins, de toute influence de la norme sociale sur les préférences individuelles. Comme dans le cadre de l'équilibre général, le but est de faire émerger un résultat collectif à partir d'individus n'ayant a priori aucun lien entre eux. On cherche à déduire le collectif d'ordres "absolument" individuels. Or, le théorème d'impossibilité semble bien mettre un terme à cet espoir d'une parfaite transparence du collectif.

13- Le théorème d'impossibilité

131- Démonstration.

Arrow démontre qu'il n'existe pas de procédure de choix social qui respecte les cinq conditions et produise néanmoins un choix collectif rationnel. Sa démonstration repose sur la notion de "groupe décisif"¹². On démontre que lorsque, à propos d'une alternative quelconque, on peut isoler un groupe dont la préférence est "décisive" (*id* : détermine la préférence sociale), alors ce groupe est décisif pour n'importe quelle autre alternative. On détermine ensuite le plus petit groupe décisif possible, celui qui contient le moins d'individus. Ce plus petit groupe comporte nécessairement au moins deux membres, sinon la condition "D" n'est plus respectée. On scinde donc ce groupe en deux sous-ensembles : (v1) et (v2). Un troisième ensemble (v3) rassemble tous les individus n'appartenant ni à l'un ni à l'autre de ces deux groupes.

A partir de là, on suppose qu'il existe trois états sociaux possibles : x, y, z. Les membres de (v1) établissent l'ordre suivant : $x \geq y \geq z$. Les membres de (v2) établissent l'ordre suivant : $y \geq z \geq x$. Comme (v1) et (v2) sont des sous-ensembles du plus petit groupe décisif, aucun n'est lui-même "décisif". Comme, hormis dans (v2), tout le monde

¹² Je suis la présentation qu'Arrow (1985, p.125-126) fait de sa démonstration.

préfère (ou est indifférent à) x par rapport à y on a forcément : xRy . Par ailleurs, l'union de (v1) et de (v2) constitue un groupe décisif ce qui permet de déterminer : yRz . La transitivité impose donc le choix collectif suivant : $xRyRz$. Or, ce choix est aussi celui de (v1). (v1) est donc un groupe décisif. Or, ce résultat contredit les hypothèses de départ, puisque (v1) est, par définition, une sous partie du plus petit groupe décisif. On a donc démontré que le choix collectif est soit intransitif, soit dictatorial ou imposé.

132- Enoncé

“Il n'existe pas de constitution vérifiant les conditions U, M, I, N et D” (Arrow, 1985, p.125).

Le théorème d'impossibilité d'Arrow atteste donc de la contradiction qui existe entre les cinq conditions définissant une constitution raisonnable, et le fait que cette constitution doive permettre d'exprimer un ordre social rationnel. Il y a en particulier contradiction entre l'impératif de transitivité du choix collectif et la liberté de choix individuels telle que les conditions N et D l'expriment. Arrow propose ainsi une autre formulation du théorème : *“S'il existe au moins trois situations possibles que les individus peuvent classer comme ils l'entendent [condition 1], toute fonction de bien-être collectif satisfaisant les conditions 2 et 3 et définissant un ordre social satisfaisant aux axiomes I et II (complétude et transitivité) est soit imposée soit dictatoriale”* (Arrow, 1951b, p.59). Arrow pointe donc une incompatibilité entre liberté absolue des opinions, au sens bien précis que l'idée de liberté implique dans la philosophie utilitariste et individualiste, et rationalité du choix collectif.

Un exemple classique attestant cette impossibilité nous est fourni par le paradoxe du vote, du initialement à Condorcet (Arrow, 1951b, p. 3) :

Trois électeurs (1,2,3) ont à choisir entre trois situations (A,B,C). On peut lire les résultats dans le tableau suivant (la transitivité du choix individuel est vérifiable de haut en bas, la dernière colonne donne le choix collectif en fonction de la majorité) :

| <i>Electeur 1</i> | <i>Electeur 2</i> | <i>Electeur 3</i> | <i>Choix Collectif</i> |
|-------------------|-------------------|-------------------|------------------------|
| <u>A > B</u> | <i>B > C</i> | <i>C > A</i> | <u>A > B</u> |
| <i>B > C</i> | <i>C > A</i> | <u>A > B</u> | <i>B > C</i> |
| <i>A > C</i> | <u>B > A</u> | <i>C > B</i> | <i>C > A</i> |

Dans cet exemple, chaque électeur est rationnel puisque ses choix sont transitifs. Mais le choix collectif ne l'est pas, il est intransitif. Il devrait normalement être $A > C$: du point

de vue logique en effet, $A > B$ et $B > C$ donnent $A > C$. Toutefois, si l'on respecte le choix démocratique à la majorité, il donne $C > A$ (en gras, sont indiqués les choix des trois électeurs concernant l'alternative A / C). Dans une telle situation, le décideur est confronté à une "indécidabilité" paralysante : il y a en effet circularité du choix collectif. Le décideur doit choisir A plutôt que B , B plutôt que C , et... C plutôt que A !

Le théorème d'impossibilité démontre qu'il n'existe pas d'implication logique permettant de relier rationalité individuelle et collective. De ce fait, Arrow est amené à reformuler le problème du choix collectif. Quittant la question formelle de l'agrégation des préférences, il cherche à établir les conditions sociales dans lesquelles un choix collectif est possible. Le constat de l'incompatibilité d'une "liberté" de choix absolue et d'une rationalité collective se traduit donc par la recherche d'une conception différente de la liberté de choix et de l'indépendance de l'individu par rapport à la société.

14- Une reformulation du problème du choix collectif

"Si on adopte comme attitude la recherche du consensus fondé sur l'impératif moral, étant entendu que ce consensus existe, le problème du choix d'un mécanisme électoral, ou d'un autre mécanisme de choix, ou, encore, plus généralement, celui du choix d'une structure sociale, prend une forme complètement différente de celle qui a été discutée pendant la majeure partie de notre étude." (Arrow, 1951b, p.85).

Dans le dernier chapitre de son ouvrage, Arrow développe la solution qui lui semble la plus pertinente au problème qu'il a constaté. Le titre de ce chapitre résume à lui seul la philosophie de cette solution : "*La conformité, fondement des jugements sur le bien-être collectif*" (Arrow, 1974, p.139). Arrow cherche donc à montrer que le paradoxe du choix collectif ne peut-être résolu qu'en supposant l'existence d'une forme de volonté générale se manifestant à l'intérieur des préférences de chacun des individus. Cette volonté générale résout le problème du choix collectif, puisqu'elle assure l'unanimité des individus. L'agrégation logique est donc rendue possible, voire triviale, par l'existence d'un consensus pratique limitant la pluralité des ordres individuels. Le problème du choix collectif change donc de nature : ce n'est plus une question de cohérence logique, mais une question pratique visant à déterminer comment le consensus peut être réalisé. Ce changement d'attitude implique de modifier l'axiomatique, mais aussi la conception de la rationalité individuelle. Pour rendre compte de l'existence d'un consensus collectif, il faut en effet éclairer les *fondements* du choix individuel, et non plus seulement sa cohérence.

Or, l'éclairage des fondements du choix collectif écarte le marché comme mécanisme susceptible de le déterminer.

141- Un changement dans l'axiomatique

Arrow, dans la dernière partie de son ouvrage, indique l'ultime condition permettant l'existence d'une fonction de choix collectif : *"Il faut qu'il existe une forme de consensus sur les fins poursuivies par la société sinon aucune fonction de bien-être collectif ne peut être formulée"* (Arrow, 1951b, p.83).

Or, cette "condition" entre en contradiction avec les deux conditions "N" et "D", qui visent à préserver la pluralité des choix individuels. Arrow s'affranchit allègrement de l'hypothèse de non-dictature : *"Avec les hypothèses posées à cette section, n'importe qui peut être considéré comme étant le dictateur, la condition de non-dictature est donc privée de sens"* (Arrow, 1951b, p.74). Il ne revient pas en revanche sur la condition de non-imposition. Toutefois, lors de la présentation de cette condition, il anticipe sur sa remise en cause en soulignant : *"Il est possible que l'on souhaite que la fonction de bien-être collectif ne soit définie que pour les ensembles d'ordres individuels compatibles avec les normes socio-éthiques de la collectivité. Ceci nécessite un affaiblissement de la première condition (U). Ce point sera discuté plus en détail dans le chapitre VII."* (Arrow, 1951b, p.29). Ainsi, les conditions N et D peuvent être contournées en restreignant les ordres individuels admissibles, c'est-à-dire en modifiant la condition U. C'est précisément la solution qu'étudie Arrow dans le dernier chapitre de son ouvrage.

Dans cette partie de l'ouvrage consacrée à l'étude d'une voie de dépassement du théorème d'impossibilité, Arrow abandonne donc, sans le souligner clairement, les conditions auxquelles il avait soumis sa démonstration. Cet abandon s'explique par le fait qu'il change radicalement de point de vue sur le choix collectif, ne cherchant plus à comprendre comment il peut être déterminé par une agrégation des préférences individuelles, mais comment il peut, en quelque sorte, être mis à jour, exhumé. Ce changement d'attitude est motivé par la conviction selon laquelle il existe un consensus latent dans la société, même si celui-ci n'est pas toujours exprimé. Mais il lui faut justifier cette conviction, et, plus encore, justifier la restriction de la condition U, assurant la pluralité des préférences individuelles, restriction qui vide de sens les conditions N et D. C'est en modifiant sa conception de la rationalité individuelle qu'Arrow parvient à justifier cette modification. Il montre, en effet, que le problème du choix collectif nécessite, dès lors que l'on a constaté l'échec de tout mécanisme d'agrégation formel, de s'intéresser au fondement des préférences individuelles. L'existence d'un "consensus" dépend en effet de l'existence d'un "impératif moral" déterminant les choix des individus.

142- Un individu complètement rationnel ?

Arrow va donc chercher à justifier la notion de consensus en éclairant la spécificité des préférences individuelles lorsqu'elles portent sur des états sociaux. Pour ce faire, Arrow fait référence à la philosophie idéaliste et s'appuie en particulier sur Kant¹³. Il emprunte à ce dernier la distinction entre les impératifs contingents (*technique* et *pragmatique*) et l'impératif catégorique (impératif *moral*). Les impératifs contingents désignent le choix du moyen le plus efficace compte tenu d'une fin donnée (que cette fin soit la résolution d'un problème *technique* extérieur à l'individu ou bien consiste à améliorer son propre "bien-être" selon un objectif *pragmatique*). Il n'a donc pas de dimension éthique. L'impératif catégorique règle à l'inverse les choix ayant trait à l'éthique et aux valeurs collectives transcendant l'intérêt individuel. Un individu qui agit sous impératif catégorique considère "l'autre" non plus comme un moyen d'améliorer son propre bien-être mais comme une fin en soi. Cette distinction kantienne recouvre, selon Arrow, la distinction entre volonté individuelle et volonté générale. Arrow souligne ainsi :

*"En un sens l'impératif moral correspond à notre concept d'ordre social, mais il constitue aussi un ordre individuel pour chacun des individus : il représente la volonté que chaque individu manifesterait s'il était **complètement (fully) rationnel**.(...) La doctrine idéaliste peut donc se résumer en disant que tout individu a deux relations d'ordre. L'une commande ses actes quotidiens, l'autre interviendrait sous certaines conditions idéales et, en un sens, est plus authentique (truer) que la première. C'est cette dernière que l'on considère comme pertinente en ce qui concerne le choix social, et on suppose qu'il existe une unanimité complète en ce qui concerne cet authentique (truer) classement individuel."* (Arrow, 1951b, p.82,83, c'est moi qui souligne).

Le cœur de la "solution" esquissée par Arrow pour régler le problème que pose la détermination du choix collectif se situe dans l'expression "*fully rational*" par laquelle il qualifie le choix sous impératif catégorique, par opposition au choix individualiste réalisé sous impératif pragmatique.

A bien y réfléchir cette expression n'a strictement aucun sens si l'on s'en tient à la définition standard de la rationalité qu'Arrow revendique et qu'il traduit par les axiomes de complétude et de transitivité. Il ne remet jamais cette hypothèse en cause, soulignant même : "*l'abandon de cette hypothèse [de rationalité] est si coûteuse qu'il semble préférable d'étudier les conséquences qui découlent de son maintien*" (Arrow, 1951b, p.60). Pourtant, que peut signifier une rationalité "complète" lorsque la rationalité désigne

¹³ Arrow s'appuie sur Rousseau, Green et Kant, qu'il rassemble sous le label "idéalistes" sans préciser les différences entre la philosophie contractualiste et la philosophie kantienne de l'éthique de responsabilité.

la capacité purement logique d'ordonnement ? On ne peut être partiellement rationnel lorsqu'il s'agit de logique : on l'est ou on ne l'est pas. Il n'est évidemment pas possible de produire un ordre incomplètement complet, ou à peu près complet, ou encore un ordre complet qui soit rationnel mais seulement partiellement transitif. Soit l'individu rationnel produit un ordre complet et transitif, soit il n'est pas rationnel.

C'est donc le *sens* même du mot rationnel qui change. Il intègre brusquement une dimension morale qu'il ne comportait pas jusqu'alors. Les conditions formelles du choix ne sont pas remises en cause, mais de nouvelles exigences viennent peser sur le critère du choix. En ce sens, être complètement rationnel peut vouloir dire : "émettre un choix en accord non seulement avec notre intérêt et nos goûts, mais aussi avec nos valeurs morales". Cette fois la rationalité touche alors au fondement du choix, et non plus seulement à sa forme logique. C'est donc une notion de la rationalité tout-à-fait différente de la rationalité standard...

Ce changement de concept de rationalité demeure implicite, et ne fait l'objet d'aucun commentaire de la part de Kenneth Arrow. On est donc contraint d'interpréter le sens de cette modification conceptuelle, qui entraîne, comme on vient de le voir, un changement de l'axiomatique elle-même. Tout se passe comme si Arrow tirait les conséquences du théorème d'impossibilité, et abordait donc de façon différente le problème du choix collectif. L'échec d'une stratégie générale visant à établir un choix collectif à partir d'un ensemble très vaste de préférences individuelles le conduit à poser le problème du choix collectif non plus en termes d'agrégation, mais en termes de consensus à réaliser. Dans le dernier chapitre de son ouvrage, Arrow n'étudie plus le problème de l'agrégation des préférences. Il cherche à comprendre comment et pourquoi elles pourraient *a priori* converger. Dans cette nouvelle optique, il devient donc nécessaire d'éclairer le *fondement* du choix rationnel, ce qui implique de mobiliser un nouveau concept de rationalité.

143- Le rejet du marché

"Tout point de vue qui dépend du consensus, considéré comme fondement de l'action collective, exclut le mécanisme de marché de l'ensemble des fonctions de bien être collectif, puisque ce mécanisme ne peut tenir compte des motivations altruistes nécessaires à l'établissement du consensus." (Arrow, 1951b, p.86).

L'étude des solutions permettant de dépasser le théorème d'impossibilité amène donc Arrow à éclairer les limites de la figure de l'individu économique. L'existence de choix sociaux, la nécessité pour les individus de s'entendre sur des objectifs communs, impliquent de redécouvrir une dimension de l'individu que la démarche analytique de

l'économie néoclassique avait fait disparaître. En quelque sorte, Arrow semble nous inviter à une resocialisation et à une moralisation de l'individu néoclassique. L'expression d'un choix collectif nécessite d'une façon ou d'une autre que l'individu s'intègre dans une communauté partageant des valeurs similaires, terreau d'un consensus nécessaire à l'action collective.

Cette nécessité explique pourquoi le marché ne peut pas être un mode d'agrégation satisfaisant. En effet, sur le marché, l'impératif pragmatique dicte seul les choix individuels¹⁴. Le marché ne peut donc être le lieu d'expression d'un jugement moral. Il ne peut être qu'un *moyen* efficace permettant d'atteindre des *fins* collectives. Celles-ci sont déterminées en dehors du marché, par référence à l'impératif moral. Toutefois : "*cela n'exclut pas la possibilité d'un recours limité au marché comme moyen de réalisation de l'optimum social.*" (Arrow, 1951b, p. 86). La réflexion d'Arrow ne remet donc absolument pas en cause les résultats de la théorie de l'équilibre général. Au contraire : elle défend leur utilité et leur pertinence. Mais elle établit les conditions de possibilité d'une autonomie du marché et de l'analyse logique qui lui est associée. Pour que le marché soit un mécanisme économique efficace, il faut qu'un accord collectif de nature morale détermine en amont les fins qu'il doit permettre à la collectivité d'atteindre.

Arrow souligne que le théorème d'impossibilité met en évidence l'incomplétude d'une rationalité strictement instrumentale. Cette rationalité, répondant exclusivement à des critères de cohérence formelle, ne permet pas de passer de l'individuel au collectif. Elle ne contient pas en elle même ce qui rend cohérents entre eux les niveaux individuel et collectif, c'est-à-dire : la réflexion éthique. Des préférences rationnelles purement individualistes, de nature instrumentales, ne peuvent servir de fondement au choix collectif lorsqu'il porte sur des états sociaux, parce qu'elles enferment le choix dans une optique purement instrumentale. Constatant que la logique ne suffit pas à régler le problème du choix collectif, Arrow va chercher dans la dimension morale des préférences portant sur des états sociaux, l'appui qui lui manque pour parvenir à surmonter ce dilemme. Il va chercher à faire apparaître le rôle des valeurs communes qui, en assurant un accord parfait sur les fins poursuivies, permettent de dégager un espace purement instrumental, pouvant être étudié de façon logique : l'univers marchand.

¹⁴ "*Le marché ne peut exprimer que des impératifs pragmatiques.*" (Arrow, 1951b, p.86).

15- La "solution" d'Arrow

"Le fait que des hypothèses purement individualistes ne nous permettent pas de formuler une fonction de choix social satisfaisante souligne qu'il existe sans doute une divergence entre les avantages sociaux et privés qui apparaissent lorsque l'on évalue un optimum social. Une part de chaque système de valeur individuel doit être constitué par des normes socio-éthiques, dont l'expression ne peut pas être atteinte par le biais d'un comportement marchand individualiste. Ces normes doivent être suffisamment similaires pour tous les membres de la société afin d'éliminer la difficulté soulignée ci-dessus." (Arrow, 1950, p.25).

Pour Arrow, la contradiction entre rationalité individuelle et collective, mise en évidence par le théorème d'impossibilité, se dénoue à compter du moment où l'on prend en compte la dimension éthique du jugement individuel. En quelque sorte, le conflit irréductible qui existe entre les ordres de préférences purement "individualistes" se transforme en parfaite harmonie dès lors qu'on prend en compte la dimension morale des préférences des individus. La solution vient donc du fait qu'il existe un consensus des acteurs, en ce qui concerne les préférences fondamentales. D'où vient ce consensus ?

151- Un consensus d'origine "biologique et culturelle"

Arrow ne justifie pas véritablement sa conception "unanimiste" du choix fondé sur les valeurs. Il s'appuie, dans son développement, sur la philosophie du contrat social de Rousseau et de Green, ainsi que sur la philosophie morale de Kant. Le rapprochement qu'il opère entre ces deux philosophies repose sur une identification, là aussi implicite, entre les concepts de "volonté générale" de Rousseau et "d'impératif catégorique" chez Kant. Ainsi, à suivre Arrow, pour Rousseau comme pour Kant, il existe un niveau de choix rationnel à partir duquel le choix individuel doit pouvoir être étendu à toute la société¹⁵. Ce principe d'universalisation du choix personnel définit un choix "complètement rationnel"¹⁶. A ce niveau de "rationalité", selon la définition kantienne, le choix individuel doit pouvoir être universalisable. L'individu s'exprime comme s'il se prononçait au nom de l'ensemble de la société.

¹⁵ Arrow identifie donc le mécanisme par lequel, chez Rousseau, chaque individu surmonte son intérêt personnel et parvient à saisir l'intérêt et la volonté générale, et le mécanisme par lequel, pour Kant, un acteur parvient à agir en fonction d'un impératif moral reposant sur la prise en compte des autres acteurs comme fin et non comme moyen de son action.

¹⁶ Le terme de raison serait plus adéquat pour désigner un jugement sur des états sociaux intégrant des impératifs éthiques. Mais Arrow emploie uniquement le terme de rationalité.

Toutefois, chez Kant l'impératif catégorique est transcendantal. Il définit le "royaume des fins" qui est un idéal métaphysique. A l'inverse, Arrow donne un sens pratique et matériel à cet idéal kantien : "le "royaume des fins" correspond, dans notre terminologie, à une société disposant d'une fonction de choix social satisfaisante" (Arrow, 1951b, p.82). Il ne s'agit pas, en fait, uniquement d'un problème de terminologie, mais d'une différence d'optique. Arrow utilise les distinctions kantienne dans un sens beaucoup moins abstrait que Kant lui-même. Il cherche, en pratique, à déterminer les moyens de rendre "effectif" un consensus correspondant à l'idéal kantien. Il lui faut donc donner un fondement *non métaphysique* au consensus moral conceptualisé par Kant : "le consensus moral ne repose pas nécessairement sur un absolu métaphysique, mais peut être construit à partir de normes socio-éthiques relatives à une culture particulière." (Arrow, 1951b, p.83). Arrow précise encore sa conception du consensus kantien en soulignant, quelques lignes plus loin, que les valeurs constituant ce consensus ont une origine "en partie biologique et en partie spécifique à des standards culturels" (Arrow, 1951b, p.86).

Ce faisant, Arrow transforme la "délibération interne", qui chez Kant définit une forme de raison pratique, de sagesse individuelle, en une forme de "contrainte" biologique et culturelle. L'impératif moral ne naît plus d'une délibération éthique, fondée sur le respect des autres, mais d'une appartenance sociale et d'une donnée naturelle. De ce fait, cet impératif n'est plus l'expression d'une individualité fondant l'autonomie de l'acteur, mais au contraire le signe de l'influence de la collectivité dans laquelle il s'insère. L'éthique échappe à l'individu, elle s'impose à lui selon un processus social quasi-durkheimien. L'individu, souverain dans ses préférences individualistes marchandes, est doublement aliéné par la "nature" et la "société" en ce qui concerne sa réflexion morale. C'est à ce prix qu'Arrow peut postuler que l'accord sur les fins est toujours déjà résolu.

C'est pourquoi, la question du choix collectif ne consiste pas en la construction d'un espace de discussion, de délibération collective permettant de "forger" un consensus moral. Celui-ci est toujours déjà là. Le véritable problème à résoudre est qu'il n'est pas forcément "actualisé" : il peut être enfoui dans l'inconscient des individus, et ne jamais se manifester en pratique. La question n'est donc pas de construire le consensus, mais de mettre en place un système obligeant les individus à choisir conformément aux valeurs morales qui le constituent. Il faut révéler aux individus leurs préférences "réelles"¹⁷. Le problème du choix collectif se résume donc à faire en sorte que chacun puisse accéder à ce niveau de rationalité à partir duquel les valeurs d'origine "socio-biologique" déterminent le choix. Il faut parvenir à ce qu'en *pratique* chaque individu exprime ses choix comme

¹⁷ Arrow (1951b, p.83) dans le passage déjà cité page 93, souligne que les préférences issues de l'impératif catégorique "sont, en un sens, plus authentiques" et que "l'on suppose qu'il existe une parfaite unanimité concernant cet authentique classement individuel".

s'il agissait sous impératif catégorique. Il faut faire coïncider le comportement quotidien des individus (sous impératif pragmatique) avec leurs préférences "réelles", bien que leur perception soit rendue difficile par le caractère *incertain* de l'univers d'action.

152- L'obstacle de l'incertitude.

L'obstacle auquel doit désormais s'attaquer la théorie du choix collectif est donc, en définitive, l'incertitude. En effet, le consensus potentiel sur les valeurs ne peut devenir effectif que dans la mesure où chacun parvient à connaître parfaitement ses préférences fondamentales concernant les états sociaux alternatifs. Or cette connaissance parfaite est un idéal, dans la mesure où les individus sont confrontés à un déficit d'information et une incertitude qui opacifie la réalité. L'individu peut donc se tromper et émettre des préférences qui s'avèrent contredire son impératif moral, parce qu'il a mal évalué les conséquences de son choix. Le caractère opaque des choix, en particulier lorsqu'ils engagent l'avenir, peuvent l'amener en pratique à se "tromper" de préférences. L'absence de consensus est donc, pour Arrow, largement explicable par des problèmes d'information : *"les différences d'opinion sur les questions collectives proviennent d'un manque d'information et peuvent être éliminées en découvrant la vérité et en la diffusant largement. (...) Les liens entre incertitude et organisation optimale de l'économie ont été insuffisamment explorés dans les récents travaux formels d'économie du bien-être, sans doute parce que nous ne disposons pas d'une théorie suffisamment aboutie du comportement dans l'incertain."*(Arrow, 1951b p.88).

Les préférences fondamentales portent sur des fins sociales et ces fins sociales sont souvent incarnées dans des décisions sociales visant à les atteindre. C'est sur les décisions elles-mêmes que les individus sont amenés à se prononcer. L'écart entre la décision et la fin qu'elle est censée atteindre est fonction de l'incertitude quant aux conséquences de la décision. C'est à la réduction de cet écart que concourent, selon Arrow, les statisticiens, et les théoriciens de l'incertitude, permettant par là même de relier chaque décision et ses conséquences finales précises. Le problème du choix collectif devient donc celui de la gestion de l'incertitude¹⁸.

Cette citation permet de mettre en perspective le théorème d'impossibilité avec l'ensemble du projet de recherche de la théorie de l'équilibre général d'une part, et des développements critiques d'Arrow d'autre part.

¹⁸ Pour prendre une image "provocatrice" : on pourrait dire qu'il faut faire sortir les individus de la "caverne". Comme chez Platon (contre qui, pourtant, Arrow construit sa réflexion) le conflit vient en dernier lieu de la contingence et du manque de clairvoyance des individus pris dans cette contingence.

La théorie "logique" du marché cherche à concevoir un système permettant de faire correspondre toute allocation générale de départ à un équilibre de marché prévisible. Par ailleurs, les deux théorèmes fondamentaux du bien-être, formulés par Arrow la même année que le théorème d'impossibilité, établissent une forme d'équivalence entre équilibre de marché et optimum social au sens de Pareto. L'espoir qui anime alors la théorie de l'équilibre général est de parvenir à associer à chaque allocation "de départ" un optimum social atteignable par le marché. Resterait alors à déterminer quelle est l'allocation de départ souhaitée par une majorité d'individus, compte tenu de l'optimum social auquel elle mènera avec certitude par la grâce du mécanisme marchand. Or, les individus, en dehors du marché, devraient être capables, sous impératif catégorique, de s'accorder sur une allocation juste. Une fois ce choix initial réalisé, le marché fonctionne comme un mécanisme assurant *de facto*, quelques soient les choix marchands des individus, une étroite correspondance entre ce choix initial sous impératif catégorique et les choix "individualistes" réalisés lors du processus effectif de marché. Le marché est donc utile, "surtout si on l'utilise délibérément pour faire coïncider les comportements obéissants à des impératifs pragmatiques avec ceux qui découlent d'impératifs moraux." (Arrow, 1951b, p.86).

Mais Arrow souligne aussi l'obstacle qui lui apparaît déjà comme constituant le problème principal : l'incertitude. Son appel pour une meilleure compréhension du comportement dans l'incertain annonce précisément son propre programme de recherche, en dehors de la perspective de la démonstration de l'équilibre général. On voit ainsi apparaître, dès 1951, un lien étroit entre les trois thèmes "critiques" étudiés par Arrow : la théorie du *choix collectif* nécessite une réflexion sur l'*incertitude*, cette réflexion sur l'incertitude doit permettre de mettre en place une "*organisation*" optimale de la société. La question de la détermination du choix collectif nécessite de régler la question de la maîtrise organisationnelle de l'incertitude. Pour que les individus puissent s'accorder une fois pour toutes sur des décisions collectives, il est nécessaire, dans l'optique "conséquentialiste" d'Arrow, que chacune des alternatives possibles soit reliée à des conséquences certaines.

Ainsi, Arrow met l'accent sur un point qui va s'avérer crucial dans le reste de son œuvre : le choix collectif repose sur un consensus quant aux *fins* poursuivies, l'expression d'un tel choix nécessite de maîtriser l'incertitude. La compréhension du comportement dans l'incertain, et la mise au point de mécanismes organisationnels permettant de réduire cette incertitude, constituent précisément les deux autres thèmes critiques d'Arrow.

153- Un rapprochement entre Arrow et Hobbes.

Arrow met en évidence la contrepartie nécessaire à l'existence d'un ordre marchand, régulé de façon autonome, et débarrassé de toute règle contingente. Il montre en effet, comme Hobbes avant lui, que l'existence d'un accord général sur les valeurs, d'un consensus sur les fins collectives, est un préalable à l'existence d'un espace autonome descriptible par les mathématiques.

Il existe une certaine proximité entre le projet de la théorie de l'équilibre général et le projet hobbesien¹⁹. Ces deux théories visent en effet à décrire un univers organisé de façon logique, permettant de maintenir le conflit entre des individus égoïstes dans des limites compatibles avec un certain bien-être collectif. Dans les deux cas, l'ordre se caractérise par l'élimination des conflits directs entre individus et par la mise en place de relations interindividuelles de nature logico-mathématique.

Hobbes construit sa défense de l'absolutisme en opposant deux mondes : l'état de nature et l'état civil²⁰. L'état de nature est régi par un droit naturel : le droit à la vie qui se prolonge en un droit à "tuer plutôt que d'être tué". Cet état de nature est marqué par l'omniprésence de la mort violente et de la lutte de tous contre tous. A cet enfer naturel, Hobbes oppose l'état civil régi par deux lois naturelles : les individus se déchargent de tous leurs droits au profit du souverain, le souverain leur assure en échange paix et vie sauve. Ce qui permet au législateur de proposer ce contrat social, c'est sa maîtrise du langage mathématique qui lui permet de construire un artifice logiquement organisé. La première pierre de cet artifice est la mise au point d'un langage non ambigu, attribuant une signification unique à chaque chose : un langage logique. Ainsi, cet univers artificiel est un univers de paix, parce que c'est un univers dont la structure logique exclut la possibilité de l'incompréhension mutuelle. Les relations entre les individus deviennent donc limpides, prévisibles, mécaniques, tout comme dans l'univers Arrow-Debreu.

Cependant, Hobbes met en évidence la contrepartie d'un tel univers. Pour que le législateur puisse construire un artifice logique assurant la paix des individus, il est nécessaire que tous les individus, lors d'un contrat social "mythique", se soumettent à un seul d'entre eux : le *Léviathan*. Le *Léviathan*, investi de la puissance collective, devient alors suffisamment puissant pour construire l'ordre : il peut imposer une organisation de type mathématique. L'artifice hobbesien naît d'une renonciation de tous les individus à leur libre arbitre. La solution hobbesienne est donc l'archétype de la "dictature" dans

¹⁹ Berthoud souligne cette proximité : "la pensée de Hobbes définit de manière frappante l'ambition, la méthode et l'axiome anthropologique qui donnent à l'orthodoxie néoclassique son principe d'unité." (Berthoud, 1994, p.134).

²⁰ Cette analyse de la pensée hobbesienne est directement empruntée à Henri Philipson [1997, chap. 1].

l'axiomatique d'Arrow. Cette dictature naît cependant d'un consensus collectif permettant de passer de la nature chaotique à l'artifice bien organisé. La mise en place de cet artifice règle une fois pour toute la question de la coordination. Le Léviathan, en la prenant seul en charge, libère tous les individus de cette question.

Dans son ouvrage, Arrow parcourt un chemin inverse à celui de Hobbes. Hobbes, pour justifier l'existence d'un ordre politique autoritaire fondé sur le Léviathan démontre d'abord que l'état de nature est caractérisé par le chaos, la lutte de tous contre tous, la mort violente. L'état de nature est une construction "mythique" permettant de faire apparaître les vertus de l'état civil. Arrow applique une *méthode* similaire, mais partant de l'univers *logique* pour faire apparaître sa nécessaire contrepartie. Il démontre d'abord que la simple mise en correspondance des individus par le biais d'une procédure d'agrégation respectant leur pluralité ne permet pas d'obtenir un choix social. Dans cette première partie, il s'arc-boute sur des hypothèses très limitatives dans le seul but de tester le mécanisme de l'agrégation et lui seul, et de démontrer ainsi sa déficience. En posant ainsi des conditions restrictives, qu'il relâche une fois le théorème démontré, Arrow cherche précisément à pointer les raisons qui permettent à un mécanisme de choix collectif de fonctionner. Il montre ainsi que le principe de l'agrégation ne peut être à lui seul à l'origine des choix collectifs, mais qu'il nécessite au préalable que les individus partagent certaines valeurs morales. De ce fait, on bascule, chez Arrow comme chez Hobbes, de la contradiction entre les multiples désirs individuels au consensus parfait qui caractérise le moment du contrat social. En rappelant que la possibilité d'un univers régi de façon mathématique nécessite, en amont, l'existence d'un "consensus", d'une forme d'unanimité sur les "valeurs", Arrow retrouve le fondement de la théorie hobbesienne.

Il existe donc une communauté de pensée entre ces deux auteurs en ce qui concerne l'articulation de la liberté et de la contrainte. L'individu ne peut être absolument libre et déchargé du poids de la coordination dans l'univers marchand que s'il existe en amont un accord collectif permettant, malgré l'individu, de faire en sorte qu'il agisse dans une direction conforme à l'intérêt collectif. Au fondement de la liberté marchande, existe non seulement la possibilité d'un clivage absolu entre la question des fins de l'action et celle des moyens, mais aussi l'existence de normes collectives, imposant à l'individu des fins communes à celles des autres. L'individualité marchande se déploie sur fond de respect des normes sociales et morales.

Il existe cependant une différence importante entre les réflexions d'Arrow et Hobbes. Tandis que Hobbes met l'accent sur la procédure coercitive qui permet de réaliser l'accord des volontés, Arrow considère que cet accord est d'emblée présent. La conversion de

l'individu, lorsqu'il passe du "royaume des fins" à celui des moyens, c'est-à-dire le marché, ne se traduit pas par une perte de souveraineté individuelle²¹. Arrow n'éclaire pas, à la différence de Hobbes, la manière dont se construit et s'instrumente le consensus social. Il ne propose pas de réflexion sur l'institution proprement dite. Il n'a donc pas besoin du concept de règle : l'existence de normes communes suffit à assurer l'accord collectif nécessaire à l'utilisation du mécanisme marchand comme instrument de bien-être collectif.

Arrow parvient, dans son étude, à établir les limites dans lesquelles le marché peut constituer un mécanisme socialement efficace. Il doit être soumis à des objectifs décidés collectivement, hors de lui-même. Ce faisant, il nous informe sur les conditions qui sont nécessaires pour pouvoir se passer de règles communes, et, du même coup, rejeter ce concept hors de l'économie : il faut pouvoir clairement distinguer fin et moyen, et s'assurer qu'il existe un consensus solide sur les fins communes.

Le dernier chapitre de l'ouvrage d'Arrow met clairement en évidence la manière dont les concepts de norme et d'autonomie individuelle peuvent être conciliés. Pour construire un espace marchand déchargé du problème des fins et des règles collectives, il est commode de supposer que celles-ci font l'objet d'un consensus moral, social, normatif. En quelque sorte, le marché s'appuie sur l'existence de structures collectives normatives, comme l'a déjà montré Hobbes. On bascule de la pure individualité marchande, dans le domaine des moyens, à la fusion de tous les individus en un seul, lorsque l'on se situe dans le domaine des fins. C'est précisément ce clivage constructiviste, ce basculement, qu'interroge et que remet en cause le concept de règle.

Cependant cette construction binaire, reposant sur un clivage de l'individu entre respect absolu des normes sociales et recherche de son intérêt personnel, ne fonctionne qu'à condition que l'incertitude soit jugulée. C'est parce que le consensus peut être scellé "une fois pour toutes" que la question des règles peut être éliminée, et c'est parce que l'incertitude est jugulée que ce compromis peut être scellé "une fois pour toutes". Dans le

²¹ La solution hobbesienne n'est cependant pas si loin qu'il y paraît : *"Le désir d'un système dictatorial, ou d'un dictateur particulier, peut s'avérer être dominant dans certaines situations. Le problème de la fonction de bien-être collectif peut alors être considérée comme étant résolue, puisque l'unanimité quant au processus de décision résout également les conflits concernant les décisions elles-mêmes."* (Arrow, 1951b, p.90). La dictature acceptée, fruit d'un consensus, ou d'un contrat social, ne constitue donc pas une entorse au principe de la décision collective, mais une solution. Or, c'est précisément le cas de figure dans lequel se situe Hobbes. La "solution" hobbesienne n'est donc pas incompatible avec la conception d'Arrow. Elle en est un cas particulier, une modalité de réalisation d'un choix collectif consensuel.

cas inverse, l'accord collectif est toujours à reconstruire et à rediscuter. La question de la règle va donc refaire surface lors de l'étude des problèmes que pose la gestion de l'incertitude.

2- L'analyse de l'incertitude

“Cette présentation sera l'occasion d'exprimer un mécontentement et une attente. Comme je vais essayer de le montrer, nos incertitudes quant à l'économie trouvent leur origine dans notre manque de compréhension de l'économie de l'incertain ; nos lacunes dans les connaissances économiques tiennent pour une bonne part aux difficultés que nous rencontrons pour modéliser l'ignorance des agents.” (Arrow, 1974b, p.154)²².

C'est probablement à travers sa conception du temps que la théorie de l'équilibre général s'est le plus clairement éloignée de la réalité. L'absence de temporalité réelle, ou plus exactement la conception du temps purement logique qui sous-tend le modèle Arrow-Debreu, tranche en effet avec la dimension profondément temporelle du processus économique. Arrow cherche précisément à réintroduire une temporalité réelle dans le modèle, et, immédiatement, constate que celui-ci ne permet pas cette ouverture. En effet, le temps pratique comprend nécessairement de l'incertitude : or le modèle est entièrement construit autour du principe d'une prise de décision ponctuelle portant sur l'ensemble des périodes à venir. C'est précisément ce que l'incertitude ne permet pas. Lorsque l'avenir est incertain, il est impossible de s'engager de manière définitive dans un contrat impliquant un paiement immédiat pour un bien qui ne sera livré que dans le futur. Arrow a proposé de prendre en compte cette incertitude à travers le concept de *marchés contingents*. Il estime cependant que ce concept ne résout pas le problème : la “non certitude” provoque inmanquablement un “non engagement” des individus et interdit tout échange.

Cette “dissolution” du marché Arrow-Debreu en situation d'incertitude est directement liée au *fondement* des choix individuels en univers incertain. Il ne s'agit cependant plus de leur fondement moral, comme dans la section précédente, mais de leur fondement informationnel. Pour percevoir cet aspect du choix en univers incertain, il faut s'écarter d'une conception strictement standard de l'incertitude telle que Savage [1954] la modélise. Savage propose en effet une conception du comportement dans l'incertain qui impose exclusivement une exigence de *cohérence* entre les différents jugements d'un individu concernant l'avenir. Cette conception s'intègre de ce fait parfaitement dans l'univers Arrow-Debreu. Mais, une fois encore, Arrow montre que l'exigence de cohérence logique n'est pas suffisante. Pour être pertinent en pratique, il faut que le

²² Extrait de “l'allocation du président” (presidential address) d'Arrow, alors président de l'AEA.

jugement concernant l'avenir repose sur des informations fiables. Il faut donc éclairer les "fondements bayésiens" du jugement dans l'incertain.

Cet éclairage met en évidence l'importance de l'information. Or, l'information est liée à l'expérience des individus, et cette expérience est difficilement transmissible. C'est la raison pour laquelle l'incertitude met en évidence l'*altérité* des individus. C'est cette altérité qui ruine la logique marchande de l'univers Arrow-Debreu. Elle pose en effet "*un nouveau problème : celui de la coordination*" (Arrow, 1973c, p.146). Ce problème est aussi celui du choix des règles permettant d'harmoniser les croyances des individus.

21- *L'incomplétude du système de marché*

Arrow entreprend de donner un contenu pratique au modèle formel de marché. La première étape consiste à introduire dans le modèle une forme de temporalité réelle. Or, l'introduction d'une temporalité réelle brise la cohérence de l'univers Arrow-Debreu. En effet, le système des prix ne peut juguler l'incertitude liée à l'existence d'un futur imparfaitement connu. Même l'introduction de biens contingents, visant à circonscrire cette incertitude à certains marchés, ne suffit pas à régler ce problème. Le modèle formel du marché ne tolère pas l'existence d'une forme de temps non entièrement réductible à l'enchaînement logique de la déduction.

211- *L'absence de marchés futurs*

"L'information à propos des biens futurs doit inclure leur prix. Ces prix doivent être déterminés sur un marché adapté, sur lequel s'échangent des offres futures et des demandes futures. Malheureusement, de tels marchés n'existent pas" (Arrow, 1974b, p.159).

Le modèle d'équilibre général repose sur l'harmonie qui existe entre le mode de rationalité de l'individu et le fonctionnement du système marchand. Le système marchand fournit en effet sous forme de prix toute l'information dont l'individu a besoin pour déterminer son plan d'optimisation. Or Arrow constate que le système des prix est défaillant, dans la mesure où il n'indique pas aux individus les prix des biens futurs.

Cette absence de marchés futurs est d'abord un constat empirique. Cependant, Arrow dépasse ce constat pour chercher à comprendre les raisons qui font qu'il n'existe pas de marchés futurs, ou, plus précisément, qu'il *ne peut pas* en exister. Il précise ainsi : "*Cette faille du marché doit être regardée comme un problème analytique autant que comme un*

présupposé” (Arrow, 1974b, p.159). Ce qu'Arrow cherche à mettre en lumière, c'est que la représentation standard du marché ne permet pas de rendre compte du problème pratique fondamental : celui de la méconnaissance du futur. Le système des prix ne parvient pas, et ne peut pas parvenir, à éliminer cette incertitude. C'est la raison pour laquelle il ne peut exister de marché futur.

En effet, l'incertitude des acteurs concernant le futur les empêche de s'engager. La signature d'un contrat d'achat/vente d'un bien suppose que chacun des contractants soit sûr que l'autre tiendra ses engagements le moment venu. Mais, dès lors que ce contrat est “à venir” et porte sur une période future, personne ne peut avoir aucune certitude. Dès lors, personne ne souhaite s'engager. Or, lorsqu'il n'y a plus de proposition ni d'offre ni de vente, il n'y a plus de marché. De plus, la faille que représente l'absence de marché futur peut rapidement se transformer en gouffre, dans la mesure où l'incertitude est un phénomène contagieux. Le fait qu'on ne connaisse pas le prix de certains biens futurs nous empêche d'acheter ou de vendre aujourd'hui, de peur de manquer une opportunité. L'interdépendance de l'ensemble des marchés provoque donc une généralisation de l'incertitude, même lorsqu'elle est initialement locale. Avec l'introduction d'une temporalité pratique, c'est donc l'ensemble du système de marché qui est menacé.

La réflexion d'Arrow sur l'absence de marchés futurs le conduit donc à mettre en évidence un problème lié à la représentation “standard” du marché : elle ne fonctionne que “hors du temps”. Or, cette limite théorique empêche la perception d'un problème économique fondamental : l'incertitude. Il faut donc, pour que la représentation théorique du marché puisse avoir un contenu non seulement formel mais aussi pratique, parvenir à prendre en compte cette incertitude. C'est l'objectif de la notion de marché contingent.

214- *Le principe des marchés contingents.*

“Il existe une approche ultranéoclassique du traitement marchand de l'incertitude dont je suis assez fier. C'est le concept de marché contingent. Plutôt que de laisser l'incertitude ruiner les marchés existant, nous pouvons la prendre explicitement en compte en achetant et en vendant des engagements qui ne seront effectifs que si un événement incertain survient.” (Arrow, 1974b, p.165).

Le concept de marché contingent est initialement proposé par Arrow [1953]. Il est intégré au modèle d'équilibre général par Debreu [1959]. L'intérêt du principe des marchés contingents est qu'il permet d'intégrer une forme d'incertitude, sans pour autant modifier la structure formelle du modèle.

Les marchés contingents sont des marchés sur lesquels s'échangent, à la date présente, des “biens contingents”. Un bien contingent est un bien dont la livraison

effective est conditionnée à la réalisation d'un état du monde donné. Le bien est acheté aujourd'hui, mais il ne sera livré que si un événement, spécifié lors de la transaction, se produit. Ainsi, l'achat "*d'une portière de voiture neuve au cas où l'actuelle portière de ma voiture serait endommagée*" constitue un bien contingent. On aura bien sûr reconnu ici le principe des polices d'assurance qui constituent l'archétype des biens contingents²³. Sur les marchés contingents, on échange donc des biens dont la livraison est incertaine. Le prix de ces biens, outre leur qualité intrinsèque, intègre le risque de non livraison, que chacun évalue subjectivement. La valeur du bien est donc en quelque sorte pondérée négativement par une prime correspondant au risque de non livraison du bien, que l'acheteur supporte.

L'incertitude attachée à la non livraison des biens contingents est donc supposée pouvoir prendre une forme quantifiable. Cette méthode d'analyse de l'incertitude repose sur une conception "probabilisable" de l'avenir en terme "d'états du monde possible". Un état du monde est un état dont les caractéristiques sont connues avec certitude, et dans lequel l'ensemble des conséquences présentes de nos actes peuvent être anticipées sans erreur. On suppose de plus que la liste des états du monde est *complète*. L'avenir continue donc d'être entièrement réductible au présent : on sait quelles formes possibles il peut revêtir, même si on ne sait pas exactement quelle forme il revêtira effectivement. L'incertitude concernant l'avenir porte seulement sur les chances de réalisation des états du monde possibles. Il s'agit donc d'une conception extrêmement restreinte de l'incertitude. Elle implique en effet que tous les changements potentiels à venir, toutes les innovations futures possibles, sont parfaitement connus, mais que seul leur *degré* d'occurrence reste à déterminer. Cette conception probabiliste du futur permet de ne pas modifier la *nature* mathématique de l'univers Arrow-Debreu. Elle ajoute simplement une dimension supplémentaire à la définition des biens. Ceux-ci sont maintenant non seulement situés dans le temps, mais également, dans un des univers appartenant à l'ensemble des univers concevables.

L'hypothèse des marchés contingents permet d'introduire dans le modèle une caractéristique propre au temps de l'action : l'incertitude. Cette intégration se fait cependant sans modifier la nature formelle de la temporalité du modèle. Le temps futur est toujours entièrement compressible dans le moment présent de la déduction. Simplement, cette temporalité logique n'est plus circonscrite à un seul univers, elle est en quelque sorte élevée à une puissance correspondant au nombre d'états du monde possibles. S'il existe deux mondes possibles, la déduction logique se déroule

²³ C'est en travaillant dans l'assurance, dans les années quarante, qu'Arrow a pris conscience de la particularité des biens assurantiels. L'assurance est donc sans doute à l'origine du concept de bien contingent qu'il a lui-même proposé (Arrow [1953]).

simultanément dans ces deux mondes. A compter du moment où la liste des états du monde possibles est close, la temporalité du modèle reste purement logique, tout en intégrant, d'une certaine manière, le principe de l'incertitude.

215- L'échec des marchés contingents

Arrow [1974c] se félicite donc de l'idée des "marchés contingents", mais doit immédiatement reconnaître que, bien qu'ingénieuse, elle n'élimine nullement le problème de l'incertitude. En effet, les individus évaluent eux-mêmes le degré d'occurrence des biens contingents. Or, pour réaliser cette tâche, tous les individus ne sont pas égaux. La réalisation de l'état du monde peut en effet être liée à leur propre comportement ou à des caractéristiques du bien connues seulement de certains individus. Enfin, tous les états du monde ne sont pas observables par tous les individus. Il peut donc exister des inégalités entre les individus en ce qui concerne leur capacité à déterminer correctement l'occurrence des états du monde. La théorie qualifie d'*asymétries d'information* ces inégalités. Les individus, soucieux d'exploiter au maximum leur surplus d'information, sont amenés à profiter de ces asymétries. Ainsi, le principe des biens contingents introduit un phénomène tout à fait nouveau par rapport à l'univers Arrow-Debreu : un individu peut ne pas "tout savoir". Le doute qui découle de cette possibilité d'une connaissance imparfaite freine le développement des transactions portant sur des biens contingents. Le problème de l'incertitude n'est donc pas réglé par cet "artifice" théorique.

Arrow [1963c, 1965, 1968b, 1973b, 1974b, 1978] s'appuie sur un vocabulaire issu de la politique assurantielle pour identifier précisément deux types de problèmes ruinant la possibilité des marchés contingents : la *sélection adverse* et le *hasard moral*.

Le problème de la *sélection adverse* naît du fait qu'un agent peut choisir de *dissimuler une information* qui serait nécessaire pour évaluer les chances de réalisation de l'événement auquel est conditionnée la livraison du bien. Il est en effet possible que les acheteurs aient connaissance d'informations que le vendeur ignore, en particulier si la réalisation de l'état du monde est directement liée à leur propre comportement (comme c'est le cas dans l'exemple des assurances). Un assuré peut ainsi chercher à dissimuler une grave maladie à un assureur au moment de souscrire une assurance vie. Il n'y a aucune raison de supposer que les individus rejettent *a priori* tout comportement malhonnête. La théorie de l'équilibre général étudie en effet le comportement d'individus instrumentalement rationnels et dénués de toute dimension morale. Le vendeur doit donc se protéger contre cet opportunisme latent, en majorant le prix de vente du bien contingent. Il fait acquitter à l'acheteur une prime qui lui permet de se prémunir contre le

risque d'une tricherie au moment du contrat. Or, une telle prime a un effet pervers : seuls les individus qui savent que les chances de réalisation de l'état du monde qui leur est favorable sont plus importantes que la prime de risque sont susceptibles de souscrire un tel contrat. Le niveau du prix *sélectionne* inévitablement tous les individus qui pensent qu'il est fondé sur une sous-évaluation du risque encouru. Ainsi, l'assureur proposant des contrats avec des primes de risque élevées risque fort de n'attirer en fait que des clients ayant un risque d'accident encore plus élevé. Il est difficile, voire impossible d'échapper au risque de *dissimulation d'information* : tous ceux qui achèteront le bien auront de bonnes raisons de penser que l'acheteur sous-estime les chances de réalisation de l'état du monde qui leur est le plus favorable. Le mécanisme de sélection adverse condamne donc, en quelque sorte, le vendeur de biens contingents à la faillite.

L'autre difficulté spécifique des contrats contingents, le problème du *hasard moral*, est liée à la *dissimulation d'un comportement* opportuniste, une fois le contrat passé. Le comportement des contractants peut en effet influencer sur la réalisation des états du monde auxquels le contrat est conditionné. Les contractants vont alors chercher à faire advenir l'état du monde qui leur est le plus favorable. Ainsi, le salarié va faire en sorte de travailler le moins possible, dans la limite où le contrat l'autorise. Dans le cas où le contrat définit un comportement requis, il faut encore que le comportement de chacun des contractants soit observable par l'autre. Il faut enfin que l'état du monde auquel la livraison du bien est conditionnée puisse être observé par les deux contractants. Le contrat contingent ne peut donc avoir lieu que dans le cas où la réalisation de l'état du monde est indépendante de la conduite d'un des deux contractants. Si tel n'est pas le cas, il est nécessaire d'encadrer le contrat par des dispositifs incitatifs et contraignants²⁴.

Ces deux phénomènes limitent donc considérablement le nombre de marchés contingents. Ils démontrent que ces marchés ne constituent pas une réponse appropriée au problème de l'incertitude, parce qu'ils reposent sur une recherche "individuelle" de l'information. En effet, avec la notion de bien contingent, Arrow modifie considérablement l'esprit du modèle de base : il introduit des biens imparfaitement définis. Les états du monde auxquels la livraison des biens est conditionnée sont certes parfaitement décrits, mais le système n'indique pas quelles sont leurs *chances* de réalisation. Il est donc nécessaire que les individus évaluent *eux-mêmes* leur degré

²⁴ Ces dispositifs feront l'objet de la très vaste littérature "principal-agent". Pour une synthèse des fondements théoriques de ce type d'analyse on peut se reporter à Pratt et Zechauser [1985]. Arrow [1985a] constate que cette voie de recherche privilégie l'étude de dispositifs très sophistiqués requérant d'une part une très forte rationalité des agents, et, d'autre part, un coût très important. Il constate qu'en réalité la relation principal-agent est plus simple, et que c'est cette simplicité que la théorie doit parvenir à modéliser. Sa propre voie de recherche tient plus dans l'étude de l'influence du groupe sur l'individu, et sur l'influence des normes sociales et des codes éthiques qui évitent les comportements opportunistes.

d'occurrence. Or, il est tout à fait étrange d'attendre un tel travail d'évaluation de la part de l'individu rationnel, qui n'a aucune compétence pratique.

Les deux problèmes spécifiques pointés par Arrow sont donc emblématiques d'un problème plus crucial et fondamental : dans l'incertain, le système des prix n'assure plus aux individus une information parfaite. Dès lors l'hypothèse classique de rationalité ne suffit plus à décrire leur comportement. Il ne leur suffit plus en effet d'être *cohérents*, il leur faut également avoir un comportement pertinent du point de vue pratique. Le constat des failles du marché provoque immédiatement une remise en cause de la conception de l'individu marchand, en raison de la codétermination de l'individu et du système dans l'univers Arrow-Debreu. Ce constat nécessite une théorie du comportement dans l'incertain décrivant l'activité cognitive, informationnelle, des agents. Cependant, pour traiter de ce problème, il faut d'abord prendre des distances avec la théorie "standard" mise au point par Savage [1954], et comprendre pourquoi, bien qu'elle parvienne à étendre le comportement rationnel dans l'incertain sans changer sa nature, elle ne permet pas, de l'aveu d'Arrow, de bien percevoir les problèmes les plus cruciaux liés à l'existence de l'incertitude.

22- La représentation standard du comportement rationnel dans l'incertain.

“Pour formaliser la théorie du choix en univers incertain, il est commode d'introduire le concept “d'état du monde”, qui représente une description du monde si complète, que , si elle est vraie et connue, les conséquences de chaque action doivent elles aussi l'être. La signification de l'incertitude est que chaque agent ne peut pas connaître l'état du monde. (...) Il est clair que toute fonction décrivant le choix d'une action doit être construite à partir de deux paramètres, la valeur des conséquences et la force relative de la croyance en la réalisation des différents états du monde. Tout comme dans la théorie du consommateur, le choix est subjectif. Dans ce qui détermine le choix d'une action, la valeur associée à cette action ainsi que les croyances qui guident ce choix sont subjectives : elles sont entièrement déterminées par les valeurs et les croyances de l'individu. La manière dont celles-ci peuvent différer des valeurs et croyances “objectivement” déterminées, dans un sens ou un autre, est sans importance” (Arrow, 1971a, p.173).

La représentation du comportement rationnel dans l'incertain a trouvé son point d'achèvement avec l'axiomatique de Savage [1954]. Cette représentation repose d'une part sur le principe de la maximisation de l'utilité espérée, d'autre part sur la notion de probabilité subjective. Cette théorie est purement logique. Elle impose exclusivement un impératif de *cohérence* à l'ensemble des croyances d'un individu rationnel. Elle ne requiert aucune exigence quant au *fondement* informationnel de ces croyances. De ce fait elle s'intègre parfaitement dans l'univers logique Arrow-Debreu, et en particulier dans le cadre des marchés contingents.

Cependant, ce qui la rend compatible avec le modèle théorique la vide, comme l'univers Arrow-Debreu, de tout contenu pratique. Ainsi, cette définition ne permet pas de distinguer un ensemble de croyances ne s'appuyant sur aucune base informationnelle sérieuse, d'un ensemble de croyances raisonnables. Or c'est précisément le *fondement* pratique, informationnel, du jugement rationnel qu'Arrow cherche à retrouver. Il lui faut donc aller plus loin que la théorie "standard".

221- La maximisation de l'utilité espérée

La rationalité dans l'incertain consiste, comme en univers certain, à maximiser une fonction d'utilité²⁵. Cependant l'individu, pour connaître l'utilité totale de son choix, doit faire la somme de l'utilité qui lui est associée dans chacun des états du monde possibles. Il calcule ainsi l'utilité *espérée* de son action : il pondère chacune des conséquences possibles de son action par la probabilité représentant les chances de réalisation à l'état du monde à laquelle cette conséquence possible est associée, puis fait la somme de l'ensemble des conséquences possibles ainsi pondérées. C'est le principe de base de la représentation standard de la prise de décision dans l'incertain : la *maximisation de l'utilité espérée* (désormais S.E.U.).

Le chapitre précédent nous a permis de décrire de quelle manière le jugement subjectif en termes d'utilité pouvait donner lieu à une fonction d'utilité. L'acteur classe les paniers de biens accessibles sous forme de préordre. On définit ensuite une fonction lexicographique associée à ce préordre. La difficulté concernant les choix incertains est que la livraison des biens n'est que probable. Il faut donc intégrer une probabilité associée à chacun des biens de ce préordre, de manière à ce que la fonction d'utilité prenne en compte le caractère incertain de la livraison des biens. On détermine ainsi une fonction d'utilité espérée. Le concept d'espérance mathématique est ainsi au cœur de

²⁵ Arrow [1951c, 1958] retrace l'historique de la constitution de la théorie standard de la maximisation de l'utilité espérée. Nous suivrons également sa présentation synthétique de la théorie du risque (Arrow [1971a]), ainsi que celle de Laffont [1985, p.1-18] et de De Montbrial [1988, p.289-322].

l'extension du jugement de l'individu à un univers contingent. Le calcul de l'espérance mathématique d'une variable s'obtient en faisant la somme de l'ensemble des valeurs que peut prendre cette variable. Le calcul de l'utilité espérée nécessite donc de *quantifier* l'incertitude. On suppose donc, comme on vient de le voir, qu'il existe une liste *finie* d'états du monde possibles, tous parfaitement connus. L'individu doit déterminer les chances de réalisation de ces différents états possibles.

Comme le souligne Arrow, cette interprétation est commode, parce qu'elle permet d'associer à chaque état du monde une probabilité numérique, qui sert de fondement à la maximisation de l'utilité. Sur cette base la maximisation de l'utilité espérée se construit comme la maximisation de l'utilité en univers certain, à ceci près que l'acteur ne maximise pas une fonction d'utilité associée à des paniers de *biens* mais une fonction d'utilité associée à une combinaison de "*loteries*"²⁶ : c'est à dire à un panier constitués de biens dont la livraison est, pour chacun d'entre eux, conditionnée par la réalisation d'un état du monde spécifique. L'individu doit classer de façon cohérente les différentes "*loteries*" possibles en fonction, d'une part de l'utilité qu'il associe à chacun des biens, d'autre part de la probabilité qu'il associe à chacun des états du monde possibles.

C'est à Bernoulli [1738], que revient la paternité de l'approche en termes d'espérance d'utilité²⁷. Bernoulli s'appuie cependant sur une conception cardinale de l'utilité, interprétable comme une véritable *mesure* de l'utilité associée aux états du monde, et sur une conception *objective* des probabilités associées à ces états du monde. Cette double hypothèse de *mesure*, de l'utilité et des probabilités, va être progressivement abandonnée par la théorie standard moderne. Ainsi, Von Neumann et Morgenstern [1944], deux cents ans après Bernoulli, établissent une axiomatique permettant de définir une fonction d'utilité, continue et bornée, permettant d'associer directement un chiffre réel représentant l'utilité espérée, en terme *ordinal*, à partir d'une relation d'ordre, établie sur l'ensemble des loteries. Ils postulent l'existence, pour chaque acteur, d'un classement cohérent ordonnant l'ensemble des "*loteries*" ("*valeurs*" des états du monde pondérées de leur "*probabilité*"). "*Tri cohérent*" signifie, comme on l'a vu pour les préférences certaines,

²⁶ Par loterie on entend ici le choix entre différentes alternatives possibles pondérées des probabilités associées. Par exemple, le choix entre les alternatives "investir dans un parapluie avec un risque de 2/3 qu'il fasse beau" ou "partir à pied avec un risque de 1/3 qu'il pleuve" constitue un choix entre deux loteries possibles.

²⁷ Il identifie de plus un premier obstacle, connu sous le nom du "*paradoxe de Saint Petersburg*". Imaginons une loterie proposant : [un gain de 2^n si "face" sort pour la première fois au $n^{\text{ième}}$ tirage d'une pièce à "pile ou face"]. L'utilité espérée associée à cette loterie est : $[\lim(n \Rightarrow +\infty) [\sum 1/2^n \cdot 2^n] = +\infty]$. Autrement dit, cette loterie peut être proposée à un prix infini. Pour répondre à ce paradoxe, il est nécessaire de postuler l'existence d'une borne supérieure et inférieure à la fonction d'utilité. C'est à Menger [1934] que revient la paternité de cette précision sur la forme nécessaire des fonctions d'utilité.

“existence d'un *préordre total* (transitivité, réflexivité, totalité) et continu²⁸ ”. A ces conditions, communes avec celle de la maximisation de l'utilité certaine, Von Neumann et Morgenstern ajoutent un *axiome d'indépendance* impliquant que les préférences entre les différentes loteries possibles ne sont pas sensibles à la manière dont elles sont combinées²⁹. Von Neumann et Morgenstern établissent alors que si les préférences respectent ces axiomes, il existe une fonction d'utilité bornée, continue, de type lexicographique, associant une variable réelle à chaque loterie. Autrement dit, Von Neumann et Morgenstern démontrent que, pour peu que l'on dispose d'une série de probabilités associées aux différents états du monde possibles, la maximisation de l'utilité peut être faite directement à partir des préférences sur les événements incertains comme elle l'est à partir des événements certains.

222- Des probabilités subjectives

Toutefois, Von Neumann et Morgenstern ne se posent pas la question de la *détermination* des probabilités permettant d'opérer un tri. Ils se situent en effet, comme leurs prédécesseurs, dans un cadre théorique supposant *a priori* la possibilité de *mesurer* objectivement ces probabilités associées à chaque état du monde possible. Or, dans la plupart des cas, ceci est impossible, ne serait-ce qu'en raison de l'interdépendance des acteurs. Il faut donc renoncer à mesurer objectivement les chances de réalisation des différents états du monde possibles. Néanmoins, la maximisation requiert l'existence de probabilités numériques. Il faut donc, d'une manière ou d'une autre, expliquer comment ces probabilités sont déterminées. La théorie moderne résout ce problème en se refusant à dire quoi que ce soit du fondement des croyances. Elle s'affranchit ainsi du problème pratique que représente la complexité du réel. Elle situe l'origine des probabilités numériques dans l'obscur subjectivité de l'acteur, qu'elle se refuse à éclairer. Cette conception *subjectiviste* des probabilités est élaborée successivement par Ramsey [1926], De Finetti [1937] et Savage [1954].

²⁸ La continuité s'exprime comme une possibilité d'exprimer toutes les préférences comme étant suffisamment proches les unes des autres pour que l'une d'entre elles puisse être exprimée comme équivalente à une fonction linéaire de deux autres préférences l'encadrant [quelques soient a_1, a_2, a_3 , trois loteries possibles telles que $a_1 \geq a_2 \geq a_3$, il existe $\delta \in [0,1]$ tel que $\delta a_1 + (1-\delta) a_3 \approx a_2$]. Cette hypothèse suppose que l'on puisse comparer, par exemple, la mort avec le gain d'une somme d'argent. Comme le remarque Arrow ceci n'est qu'une affaire de multiples petits écarts : “*tout voyage, quelque soit sa durée, commence par un simple pas...*” (Arrow, 1971, p.176).

²⁹ Formellement : “Quelques soient, p et q avec $P \succ I \succ Q$, alors, $[ap+(1-a)r] \sim [aq+(1-a)r]$ ” (avec : $0 \leq a \leq 1$; et I signifiant “est indifférent à”) et “Quelques soient p et q avec $P \succ Q$, alors $[ap+(1-a)r] \geq [aq+(1-a)r]$ ” (avec : $0 \leq a \leq 1$). Cet axiome sera critiqué par Allais [1953], ainsi que par les psychologues, en particulier Kahneman et Tversky [1976] qui ont apporté plusieurs expériences invalidant le caractère réaliste de cet axiome. Nous y reviendrons dans le chapitre suivant.

Ramsey [1926] inaugure l'approche subjectiviste en proposant de représenter le degré de croyance d'un individu en une proposition P , en l'identifiant à la somme monétaire qu'il serait prêt à parier sur sa véracité. La probabilité numérique qui représente les anticipations de l'agent n'est pas censée être le fruit d'une mesure effectivement réalisée par l'individu. On suppose que l'acteur rationnel agit *comme si* il pariait. Sur cette base, Ramsey impose des conditions qui ne portent que sur la cohérence interne de l'ensemble des croyances, comme le feront Von Neumann et Morgenstern. Cette description reste donc résolument à la surface du comportement de l'individu sans chercher à représenter ses motivations profondes. La rationalité s'attache à la cohérence des choix, pas au jugement concernant le futur.

Savage [1954] reprend la théorie de Ramsey et propose une axiomatique complète, cohérente avec l'axiomatique Von Neumann-Morgenstern, mais qui permet de construire la fonction d'utilité espérée directement à partir des probabilités subjectives³⁰. Savage part du principe que la préférence d'un acteur concernant un acte est uniquement liée aux conséquences que cet acte peut avoir dans les états du monde non exclus³¹. Le point de départ de Savage consiste donc à considérer qu'il existe des *préférences conditionnelles*, qu'Arrow [1971, p176] présente ainsi : "*le principe des préférences conditionnelles certifie que l'ordonnancement ne dépend que des conséquences des actes dans les états du monde non exclus par l'information disponible*". Cette conception des préférences implique que l'acteur puisse tout-à-fait séparer d'une part, sa préférence pour les conséquences d'un acte dans un état du monde donné et, d'autre part, son évaluation des chances de réalisation de cet état du monde. Pour qu'un acteur puisse avoir des préférences conditionnelles, il faut qu'il puisse déjà estimer les chances de réalisation de l'ensemble des états du monde. Savage souligne donc que sa théorie suppose que l'individu dispose d'une série de probabilités subjectives portant sur l'ensemble des états du monde possibles. Ses préférences conditionnelles révèlent en quelque sorte ses "*a priori*" sur l'avenir. L'hypothèse "forte" de Savage est que cette probabilité subjective sera la même quelque soit l'objet du calcul³². Cette hypothèse implique que les

³⁰ L'axiomatique de Savage est exposée dans ses grandes lignes par Arrow [1971a], De Montbrial [1985, p.319-321] et par Laffont [1985, p.10-12]. On peut se reporter aussi à l'exposé de la démonstration par Guesnerie et Jaffray [1971].

³¹ Très simplement cela signifie que l'on ne préfère pas "en général" un parapluie à un chapeau de soleil, mais, que l'on préfère le parapluie uniquement s'il l'on suppose qu'il va pleuvoir, et le chapeau de soleil uniquement si l'on suppose qu'il va faire beau. Notre préférence pour le parapluie dans le cas où il va pleuvoir est donc tout à fait indépendante de notre préférence pour le parapluie dans le cas où il va faire beau.

³² Ainsi, le choix du parapluie doit entraîner avec lui le choix d'un parcours abrité, le choix d'un vêtement de pluie, voire de chaussures étanches. Un individu qui parierait sur le mauvais temps en ce qui concerne le choix du parapluie, et sur le beau temps en ce qui concerne le choix du manteau ou des chaussures serait irrationnel, même dans le cas, fréquent, où l'on ne peut en fait pas savoir quel temps il va réellement faire.

préférences en termes d'utilité et celles en termes d'espérance soient parfaitement indépendantes. L'individu rationnel doit donc produire deux relations d'ordre cohérentes, respectivement pour l'ensemble des actes possibles et pour l'ensemble des états du monde possibles. L'indépendance et la cohérence de ces deux ordres de préférence permettent de définir une fonction d'utilité qui agrège ces deux préordres, et exprime directement un jugement en termes d'utilité espérée.

Cette démarche permet donc d'obtenir, en se fondant exclusivement sur la subjectivité de l'individu, une distribution des préférences, portant sur des complexes de biens incertains, représentable par une fonction d'utilité lexicographique. La démarche de Savage est parfaitement décrite par Arrow [1971] dans la citation qui ouvre cette section. Il élargit le choix de l'agent à des biens incertains, sans modifier la nature de ce choix, subjectif mais cohérent. Il boucle ainsi le projet néoclassique en élargissant la rationalité de l'acteur à des situations incertaines, sans en modifier la nature, paramétrique et instrumentale. Elle est donc d'une grande commodité dans l'optique de la théorie de l'équilibre général. Elle parvient en effet à constituer une représentation des choix incertains en tout point identique à celle des choix certains, pour peu que l'on accepte l'idée d'une probabilité subjective. Ce tour de force se fait sur la base d'une représentation très complexe et peu réaliste de la prise de décision dans l'incertain, mais a le grand avantage de permettre une intégration de l'incertitude dans un modèle qui semblait nécessairement l'exclure³³.

23- La conception "bayésienne" d'Arrow

La théorie de Savage étend donc la représentation néoclassique d'un comportement rationnel à des situations incertaines, sans en modifier la nature, formelle et instrumentale. Elle gomme cependant les différences entre le désir d'un individu concernant des biens définis, et son jugement concernant un futur imparfaitement défini. Ce faisant, elle ne permet pas de cerner les raisons pour lesquelles l'incertitude entrave la régulation marchande. C'est pour mettre à jour ces raisons qu'Arrow va chercher à éclairer, en amont de la théorie de Savage, le *fondement* des jugements dans l'incertain. Il mobilise

³³ Laffont [1985, p.9] souligne ainsi : "L'économiste a besoin d'une théorie du comportement d'agents familiers et de leurs problèmes de décision qui soit assez simple pour être maniable. Tel est le véritable enjeu des efforts actuels pour dépasser le paradigme de l'espérance d'utilité, qui a fait la preuve de son efficacité, mais contre lequel suffisamment d'expériences ont été accumulées pour justifier, peut-être, la nécessité d'une généralisation, petite révolution scientifique au sens de Kuhn." (Laffont, 1985, p.9). Les critiques d'Allais [1953] et des psychologues cognitivistes (Kahneman et Tversky [1982]), n'ont pas conduit à une remise en cause de la théorie formelle de Savage, faute de candidats pour la remplacer.

pour se faire une conception bayésienne du jugement de probabilité qui, sans remettre en cause l'axiomatique de Savage, permet d'éclairer ce qu'elle laisse dans l'ombre : le rôle de l'information.

231- le jugement subjectif est un jugement de probabilité.

Debreu indique dans son chapitre consacré à l'introduction de l'incertitude dans le modèle d'équilibre général : *“Cette nouvelle définition d'une marchandise [sous forme de bien contingent] nous permet d'obtenir une théorie du cas incertain, libre de tout concept de probabilité et formellement identique à la théorie du cas certain développée dans les chapitres précédents.”* [Debreu, 1966, p.106, c'est moi qui souligne]. A le suivre, donc, l'intégration de l'incertitude ne change pas l'esprit du modèle d'équilibre général. C'est vrai en apparence : on suppose toujours que l'individu peut décider aujourd'hui de l'ensemble de ses actes futurs, par le biais du mécanisme des marchés contingents. Mais, pour pouvoir supposer que l'individu calcule de la sorte, il faut nécessairement introduire un jugement de probabilité par lequel l'individu classe les états du monde possibles en fonction de leur degré de vraisemblance.

Ainsi, la théorie de Savage, bien qu'elle dissimule la probabilité sous un jugement global en termes d'utilité espérée n'en requiert pas moins, implicitement, un *jugement* de probabilité de la part de l'acteur. Certes, la rationalité de l'individu ne tient qu'à la cohérence de ses choix. De plus, les deux classements concernant les alternatives possibles, celui portant sur l'utilité et celui portant sur le degré d'occurrence des “états du monde possibles”, se confondent, au moment de la prise de décision, en un jugement unique. Ce faisant, la détermination des probabilités associées aux différents états du monde possibles est un problème qui, en apparence, disparaît. Pourtant, en réalité, pour décider, il faut bien que l'individu évalue les chances de réalisation des différents états du monde. La théorie du choix rationnel dans l'incertain n'est donc pas, en fait, *“libre de tout concept de probabilité”*.

222- Le fondement informationnel des probabilités

Savage construit une représentation du choix dans l'incertain qui dissimule la spécificité de ce choix par rapport au choix dans un univers certain. En effet le jugement de probabilité perd ce qui fait sa spécificité par rapport au jugement en termes d'utilité : le rapport à un monde qui existe objectivement. L'individu rationnel défini par Savage n'a pas plus de sens pratique que l'individu rationnel évoluant dans l'univers certain Arrow-Debreu. Il est exclusivement “logique”. Or, cette conception de la rationalité fait problème

dans un environnement qui n'est pas absolument certain, même s'il est "probabilisable" (*id* : si la liste des états du monde possibles est finie).

En effet, selon cette représentation du choix rationnel dans l'incertain, on peut être rationnel même si le fondement de notre choix repose sur une négation de la réalité pratique. Or, si l'idée d'un choix purement subjectif s'applique bien à une situation où l'objet du choix, des paniers de biens certains et parfaitement définis, est parfaitement assuré, il s'applique avec plus de difficulté dans l'incertain. La détermination de la probabilité d'un événement n'est pas uniquement une affaire de goût. On ne choisit pas le degré d'occurrence d'un état du monde comme on choisit une tarte à la crème plutôt qu'une flamiche aux poireaux... Or, c'est bien la différence entre ces deux types de choix qui disparaît dans la théorie de Savage : elle introduit avec force une parfaite indépendance entre ces deux types de choix, même regroupés en un jugement unique, tout en les inscrivant tous deux dans l'étroite sphère de la subjectivité. Le jugement qui détermine la série des probabilités est maintenu dans le domaine des désirs et des *a priori* : il ne peut pas prétendre être rationnel en tant que tel. Seule compte la cohérence interne du classement qu'il détermine. Suppes note ainsi, dans un bilan critique de la théorie de Savage : "*La difficulté la plus profonde du modèle normatif est qu'il peut être satisfait par des personnes déficientes mentalement. En d'autres termes, la cohérence des calculs exigée par le modèle d'utilité attendue ne garantit pas l'exercice du jugement et de la sagesse au sens traditionnel.*" (Suppes, 1981, p.18).

En effet, l'absence de fondement pratique du choix rationnel dans l'incertain conduit à une hypothèse de rationalité ne permettant pas de différencier un fou d'un homme raisonnable. Un individu qui préfère prendre son parapluie et son imperméable lorsque le ciel est bleu azur au mois de juillet et qui s'enduit de crème solaire lorsque le ciel est gris au mois de janvier vit probablement dans un univers qui n'est pas le monde réel, mais il peut néanmoins être rationnel au sens de Savage. En effet, l'individu rationnel, selon la représentation de la S.E.U., peut s'affranchir du réel. Cette théorie ne nous dit rien de la pertinence pratique du comportement rationnel.

233- La dimension bayésienne de la prise de décision.

"Il est de la plus haute importance de remarquer que le caractère subjectif des probabilités ne doit pas exclure le fait qu'elles puissent être influencées par l'expérience. (...) Quand une croyance est représentée par des probabilités, alors la connaissance d'un événement conduit l'acteur à agir conformément aux probabilités conditionnelles consécutives à cet événement, plutôt que conformément aux probabilités qui avaient cours avant cet événement" (Arrow, 1971a, p.173, c'est moi qui souligne).

C'est précisément contre cet aspect de la S.E.U qu'Arrow développe sa propre conception du jugement dans l'incertain. Il ne rejette pas la théorie de Savage, mais, conscient des problèmes qu'elle pose, cherche à l'adapter de manière à mettre en lumière la question du fondement informationnel des probabilités. Il poursuit par là son objectif visant à donner un contenu pratique à la représentation standard du comportement marchand, et du mécanisme de marché. Cet objectif n'implique pas, selon lui, de refuser l'axiomatique de Savage. Il faut au contraire élargir les exigences qui pèsent sur la détermination des probabilités subjectives. Le jugement rationnel doit non seulement être cohérent, mais aussi pertinent en pratique.

Pour ce faire, il mobilise la conception bayésienne des probabilités. Il estime en effet que cette approche permet de réconcilier le principe des probabilités subjectives avec celui de l'utilisation de l'information pratique. Il ajoute donc à l'axiomatique de Savage une exigence supplémentaire : la règle bayésienne de révision des croyances subjectives³⁴. En mettant l'accent sur cette règle, Arrow entend redonner de l'importance à l'expérience des individus. Il vise ainsi à ajouter une dimension pratique à la représentation formelle, lui permettant de forger une conception du fonctionnement du marché moins éloigné de la réalité.

La prise en compte de cette dimension du choix rationnel dans l'incertain ouvre de nouvelles perspectives. Elle permet à Arrow d'envisager le principe d'un choix rationnel "contextualisé" et fondé sur une découverte progressive des informations pertinentes³⁵. Arrow progresse ainsi vers l'idée selon laquelle la rationalité du comportement d'un individu tient aussi à une compréhension "pratique" liée à la découverte puis la correction de ses erreurs³⁶. Il se donne ainsi les moyens d'intégrer certaines des critiques empiriques

³⁴ Bayes, en 1763, a proposé une formule permettant d'établir une méthode de calcul des probabilités d'événements généraux à partir d'une connaissance partielle de la réalité. La méthode de calcul des probabilités *a posteriori* fondée sur les probabilités *a priori* (règle de Bayes ou méthode des probabilités inverses) s'énonce ainsi : soit *a* un événement certain, et les (*b_i*) des événements mutuellement exclusifs formant une liste complète des états du monde possibles, alors on sait que $P(b_i/a) = [p(a/b_i)p(b_i)] / [\sum p(a/b_i)p(b_i)]$. Cette formule permet donc de relier les probabilités *a priori* aux probabilités fondées sur des informations connues. Elle peut donc être utilisée pour réviser des croyances *a priori* lorsqu'un événement nouveau se réalise : elle établit un lien, essentiel, entre information et probabilité *a priori*. Elle suppose cependant que l'on dispose déjà d'une série de probabilités *a priori*, qui peuvent être considérées comme étant équiprobables dans le cas où rien ne permet de les départager.

³⁵ Arrow attache une importance cruciale au phénomène d'apprentissage : "*l'apprentissage est certainement une des formes de comportement les plus importantes dans l'incertain. Les études menées sur ce thème devraient avoir des répercussions très importantes sur la pensée économique*". (Arrow, 1958, p.68).

³⁶ Arrow[1962] est le premier à avoir exposé l'efficacité du "Learning by Doing". Marchak et Radner [1972], fréquemment cités par Arrow, ont précisément construit leur théorie des équipes sur le principe de la gestion de l'information et de l'apprentissage.

de la conception standard qui mettent en avant l'importance du contexte de la décision ou l'existence de "biais heuristiques"³⁷.

Cette attitude d'Arrow est directement liée à sa volonté d'expliquer mieux le fonctionnement du système de marché. En effet, la conception de Savage ne permet pas d'analyser les raisons pour lesquelles l'incertitude fissure le système de marché. Ces raisons sont de nature pratique. Elles tiennent à la manière dont les individus cherchent, en s'informant, à se faire une idée du futur. C'est autour de la recherche et de la manipulation de l'information que se jouent les comportements qui, en pratique, mettent en péril la bonne marche du mécanisme marchand. Pour les comprendre et les analyser, il faut donc élargir la conception strictement formelle de Savage, et mettre l'accent sur ce que cette théorie laisse dans l'ombre : l'information, qui fonde, en pratique, les jugements de probabilité.

24- L'information

"Au minimum, la prise en compte du concept d'information et des changements possibles qu'elle peut subir avec le temps implique de réviser considérablement la théorie de l'équilibre général".(Arrow, 1973b, p.139).

Dans le modèle standard, l'information n'est pas un "bien" économique ayant une importance, puisque le système des prix délivre gratuitement aux individus toutes les informations nécessaires au calcul de leur plan optimal. Il en va tout autrement lorsque l'on introduit de l'incertitude dans le modèle : le système de prix n'informe plus assez les individus, et ceux-ci doivent donc chercher *en dehors* du marché, les informations qui leur sont nécessaires. Or, à l'extérieur du marché, l'individu rationnel rencontre les autres individus, dont il était isolé par le système. Ainsi, avec la reconnaissance de l'importance de l'information, c'est l'importance des rapports directs entre individus qui est mise en lumière.

³⁷ Cette notion a été développée durant les années soixante-dix par une équipe de psychologues cognitivistes dirigée par Kahneman, Slovic et Tversky [1982]. Arrow note à l'occasion d'une revue de leurs travaux : "j'espère avoir démontré que, sur une part importante des marchés intertemporels, des conduites systématiquement déviantes vis-à-vis du comportement individuellement rationnel se manifestent, et que ces déviances concordent avec des preuves réunies à partir de sources très variées par les psychologues." (Arrow, 1982, p.269-270).

241- *L'information n'est pas une marchandise.*

Arrow [1978, p.199] propose de définir l'information à partir de son effet sur le système de probabilités subjectives de l'individu. Ainsi, "*toute observation qui modifie effectivement les probabilités conditionnelles selon la règle de Bayes*" est considérée comme une information. Cette conception de l'information est donc "qualitative" et permet de mettre en évidence la nature profondément "endogène" du processus d'acquisition des connaissances. L'information ne constitue pas une donnée purement extérieure à l'individu, qui n'aurait qu'à l'utiliser sur le mode des biens et services classiques. Elle naît de la rencontre entre un signal extérieur, et, d'une certaine manière, un équipement individuel de l'acteur capable d'utiliser ce signal et de le *transformer* en information. Chaque individu crée de l'information à partir de signaux extérieurs qui n'en sont pas encore. De ce fait, l'information n'est pas un bien classique. Elle représente un enjeu essentiel, fait l'objet d'une demande forte, mais ne peut pas être distribuée par le marché. En effet, elle est trop intimement liée à l'individu qui la détient pour pouvoir circuler comme une marchandise. Arrow (1973b, p.142) signale les deux aspects qui la distinguent des autres biens :

- 1- Elle n'est pas divisible dans son usage, puisque son coût n'est pas lié à l'usage que l'on en fait.
- 2- Il est très difficile de se l'approprier.

Ces deux aspects expliquent qu'une économie fondée sur la concurrence ne produise pas d'information de façon optimale. En effet, l'indivisibilité de l'information pose des problèmes similaires à ceux des biens publics : personne ne souhaite produire individuellement une information qui profitera ensuite à tous. Si, à l'inverse, chacun protège secrètement les informations qu'il détient, la production d'information risque d'être redondante, chacun produisant des informations identiques, et donc collectivement sous-optimales. Le caractère inappropriable de l'information tient au fait qu'un individu qui détient une information à un moment donné ne la perd pas en la transmettant. Le coût de l'information est lié à la difficulté qu'il y a à la traiter, la comprendre, l'utiliser. Il n'est donc pas lié à son coût de production ni à la valeur qu'elle représente pour le vendeur qui, de toute façon, la conservera. Cette particularité de l'information en fait un bien économique qu'il n'est pas intéressant de produire à des fins commerciales, puisque son échange est très problématique. Le système de marché n'est donc pas un mode optimal de répartition et de diffusion de l'information. Il est nécessaire d'envisager le rôle, dans la diffusion de l'information, d'institutions non marchandes.

242- L'information fait resurgir l'altérité

“Deux individus partageant une même croyance probabiliste peuvent néanmoins avoir des probabilités différentes pour un même événement au moment où ils entrent sur le marché, parce qu'ils ont observé des événements différents.” (Arrow, 1978, p.199).

La prise en compte de l'information pose un autre problème au modèle théorique standard. On a souligné que le lien exclusif qu'entretenait l'individu rationnel avec le marché était une des conditions de fonctionnement du modèle. Cette condition n'est plus respectée dès lors que le système des prix ne donne plus à chacun toute l'information pertinente. L'information est liée, comme le souligne Arrow (1971a, p.173), à l'*expérience* des individus. Or, cette expérience est une donnée difficilement transmissible. Elle induit des facultés différentes quant à la compréhension pratique et au traitement des données, elle détermine des sensibilités différentes. De fait, chaque individu est, selon sa personnalité et son rôle économique, sensible à certaines informations plus qu'à d'autres. Avec la prise en compte de l'information, ces différences entre individus deviennent un phénomène économique essentiel. En effet, n'ayant pas la même sensibilité, les individus n'accèdent pas aux mêmes informations. Ces différentiels d'information provoquent des anticipations divergentes : les acteurs, diversement informés, n'ont en effet aucune raison d'affecter des probabilités similaires à chacun des états du monde possibles. Ainsi la prise en compte de l'information amène nécessairement la reconnaissance de l'*altérité* : ce que l'univers Arrow-Debreu excluait nécessairement.

Or, la prise en compte de l'altérité, de la diversité des acteurs, “*crée un nouveau problème : celui de la coordination*” (Arrow, 1973b, p.146). La remise en cause du caractère exclusif de la relation de l'individu au système de marché redonne en effet de l'autonomie aux individus. Chacun va chercher à tirer parti des informations dont il dispose. Le comportement des individus ne respecte plus tout-à-fait la loi du marché, puisque chacun peut jouer sur des données inaccessibles aux autres. En définitive, ce que la prise en compte de l'information fait resurgir, c'est le donc problème de la *pluralité* des individus. On sait que l'univers Arrow-Debreu l'excluait en "compartimentant" chacun des individus et en leur imposant un comportement de forme unique et universelle. On sait aussi qu'Arrow, à travers le théorème d'impossibilité, met en évidence une condition nécessaire à l'existence de cet univers ordonné : l'accord collectif sur les valeurs déterminant les fins collectives. Mais voilà que ce problème refait surface au cœur du processus de marché : l'incertitude obscurcit l'horizon final, et remet en cause la séparation nette entre fin et moyen. Les individus s'écartent de la loi de comportement qui assurait l'efficacité du mécanisme de marché. De nouveau l'optimalité collective est mise en cause par l'irréductible pluralité des objectifs individualistes des agents. Comme

l'annonçait Arrow [1951b] dans son ouvrage, l'incertitude remet en cause le clivage entre "valeurs" et "objectifs individuels". Or "l'autonomie" de l'individu marchand, qui est au fondement du mécanisme de marché, repose sur ce clivage. Pour lui redonner cette autonomie il faut donc parvenir à gérer l'incertitude et réintroduire une forme d'unité dans le comportement des individus marchands.

En introduisant dans le modèle de base une conception pratique du temps, comprenant de l'incertitude, Arrow cerne une autre limite du modèle de marché. L'incertitude rompt en effet le lien exclusif qui unit l'individu rationnel au système des prix. Ce faisant elle fait apparaître le problème de *l'altérité*. Cette altérité n'est pas intégrable dans le cadre logique Arrow-Debreu, fondé sur l'identité de tous les individus au système. Elle met donc en péril cette représentation du marché.

En cherchant à comprendre le comportement rationnel dans l'incertain, Arrow est donc à nouveau confronté à la question de la diversité des agents. Comme lorsqu'il entreprend de cerner les mécanismes du choix collectif, Arrow est amené à quitter la représentation logique et formelle du comportement des individus pour éclairer les fondements de ce comportement. Or, l'étude des fondements pratiques du comportement dans l'incertain met en évidence la diversité des regards portés sur l'avenir. Cette diversité entre en contradiction avec l'unité du choix collectif que les individus, au fond, poursuivent. Livrés à eux-mêmes sur le marché, les individus font donc des choix divergents. Cette divergence entrave le fonctionnement efficace du marché, et remet en cause la possibilité d'atteindre l'état social correspondant au choix collectif. Il est donc nécessaire de recréer de l'unité, de guider les individus. Il faut que l'accord fondamental sur les "valeurs" se manifeste au cœur même de l'échange marchand. Il faut que la société guide l'individu. Il faut donc des *règles*. Les règles qu'envisage Arrow sont cependant d'un type particulier : ce sont les règles organisationnelles.

3- Vers l'organisation...

“La logique et la limite du comportement concurrentiel idéal en univers incertain nous obligent à reconnaître le caractère incomplet d'une description de la réalité s'appuyant exclusivement sur le système de prix impersonnel.” (Arrow, 1963c, p.43).

L'introduction de l'incertitude dans l'univers marchand entrave le fonctionnement de l'ensemble du système. L'incertitude fait réapparaître l'altérité, ce que le mécanisme marchand ne peut supporter. En effet, l'altérité pose non seulement des problèmes de compréhension mutuelle, mais également de confiance : les individus ne peuvent donc plus échanger, et le marché est menacé de disparaître. Il est donc nécessaire de réintroduire une forme d'unité collective. C'est pourquoi la théorie doit faire place, à côté du marché, à un mécanisme assurant la gestion de l'incertitude restaurant l'unité collective. Ce mécanisme est, en quelque sorte, conçu comme le prolongement du choix social établi en amont du marché. Il contraint les individus à agir en un sens conforme à l'intérêt collectif. C'est en suivant ce raisonnement qu'Arrow intègre dans l'analyse économique une forme d'institution non marchande : l'organisation³⁸.

Cette ouverture est réalisée en deux temps. Arrow montre d'abord que l'échange marchand requiert l'existence d'une forme d'ordre social instrumentée par des codes éthiques et un langage commun. Il dessine ensuite les contours du mécanisme qui rend possible et institue cet ordre. C'est à ce stade qu'apparaît le concept d'organisation. Ce qui rend l'organisation efficace, c'est l'existence d'une relation d'autorité hiérarchique.

En définitive l'incertitude est jugulée parce que les individus se “soumettent” à une autorité supérieure chargée de gérer au mieux cette incertitude. L'individu rationnel se décharge du poids de la coordination et de la gestion des relations interpersonnelles au profit d'une structure autoritaire et transcendante. On repère donc, derrière la “solution organisationnelle” au problème de l'incertitude, une forme hobbenne de résolution du problème de la coordination et de construction de l'ordre social.

³⁸ La distinction entre institution et organisation n'est pas souvent clairement faite en économie, et ne l'est pas du tout par Arrow. Dans l'univers économique néoclassique, le marché est une forme de nature, et la seule véritable institution est donc l'organisation. Mais, il est également possible de mettre l'accent sur les structures institutionnelles du marché (commissaire priseur, définition de l'espace des biens, etc...). On peut donc considérer que l'institution est un concept général, dont le marché et l'organisation sont deux catégories particulières. Cette distinction sera reprise dans le chapitre suivant consacré aux néoinstitutionnalistes.

31- Un besoin d'unité

“Des contrôles non marchands, qu'ils soient internalisés, comme les principes moraux, ou imposés de l'extérieur, sont, jusqu'à un certain point, nécessaires pour assurer l'efficacité économique.” (Arrow, 1968, p.105, c'est moi qui souligne).

L'incertitude fait apparaître le problème de l'altérité. Juguler l'incertitude nécessite donc de limiter cette altérité en donnant aux individus des repères communs. Ces repères sont de deux ordres : des codes éthiques, assurant l'existence d'un “objectif commun”, et un langage commun permettant de répartir et de filtrer l'information. Ces deux éléments collectifs guident étroitement l'individu, ils le contraignent à agir dans un sens cohérent avec “l'intérêt collectif”.

311- Le rôle des codes éthiques

“Quand il y a une forte disparité des niveaux de connaissance entre les deux “côtés” du marché, l'existence de codes éthiques partagés peut être d'une grande efficacité économique.” (Arrow, 1973b, p.139)

Pour que les individus puissent surmonter leur incertitude, il est nécessaire qu'ils échangent leurs informations. Il faut donc que les individus ayant certaines informations en fassent profiter les autres, moyennant paiement, de façon à ce que l'information se répande dans l'économie. Mais, comme on l'a vu à propos des marchés contingents, le marché décourage de tels échanges. Il est en effet plus rentable pour un individu, de profiter des asymétries d'information jouant en sa faveur pour réaliser un bénéfice, même malhonnête, que de chercher à “jouer le jeu” en révélant à ses interlocuteurs échangistes les informations dont il dispose. C'est la raison pour laquelle, à ce stade, la morale doit nécessairement venir compenser l'effet collectivement dévastateur de l'application, par l'individu, d'un mode de rationalité exclusivement instrumentale³⁹.

Ainsi, on voit apparaître, à travers ce problème des asymétries d'information provoquées par la différenciation des individus, la question de la confiance. Pour que le marché fonctionne, dans un univers incertain, il faut que les individus puissent établir des relations de confiance. L'échange repose en effet sur une règle de réciprocité. Cette règle fonde la conviction selon laquelle l'individu avec lequel on fait échange va respecter sa part d'engagement, bien que sa rationalité lui dicte de prendre le bien qu'on lui propose,

³⁹ Arrow utilise sans distinction la notion d'éthique et la notion de morale. Cependant, l'éthique désigne l'activité rationnelle de mise en question critique des valeurs. Il n'existe pas, chez Arrow, de trace de ce travail critique : les normes s'imposent aux individus qui les respectent. Le terme de “morale” est donc ici plus approprié.

sans nous donner le bien ou la somme d'argent que l'on attend en retour⁴⁰. Sans confiance, il ne peut donc y avoir d'échange : le marché disparaît⁴¹. L'introduction de l'incertitude et de l'information met ainsi en lumière un problème crucial que l'univers logique Arrow-Debreu excluait.

Or, la confiance n'est pas un bien échangeable sur un marché : *“Si vous deviez l'acheter vous auriez déjà quelques doutes sur ce que vous venez d'acheter”* (Arrow, 1974c, p.23). Il faut donc l'introduire par le biais d'un mécanisme non marchand. C'est à la collectivité de “construire” la confiance en mettant en place des règles du jeu, des codes éthiques, des valeurs morales. Ces valeurs incitent chacun des contractants à respecter ses engagements, même lorsque ce n'est pas dans son intérêt⁴². La question du respect des engagements se trouve donc mise à distance de l'interaction marchande elle-même. Pour que celle-ci puisse se dérouler, il est nécessaire que des valeurs définissent un espace commun marqué par la confiance mutuelle. C'est donc de la contrainte collective, qui peut être intériorisée, que naît la possibilité de l'échange marchand. Cette contrainte restaure la confiance mutuelle, en introduisant dans le comportement individuel un frein collectif à l'avidité et l'opportunisme. Elle est nécessaire à l'existence du marché : *“des éléments éthiques entrent en jeu dans toute transaction, sans eux aucun marché ne pourrait fonctionner”* (Arrow, 1973b, p.150).

La prise en compte d'une forme de temps pratique ruine donc l'espoir d'une régulation reposant exclusivement sur des rapports strictement interindividuels, elle fait apparaître la nécessité des contraintes sociales collectives⁴³. Les “codes éthiques” représentent en effet une condition *extérieure* à l'univers marchand, nécessaire au fonctionnement du marché⁴⁴. Ainsi, l'existence d'une loi de comportement individuel n'est plus entièrement déductible de l'hypothèse de rationalité. On doit supposer que l'individu, en plus d'être

⁴⁰ La contradiction entre l'opportunisme et la nécessité de coopérer est en particulier formalisée par le dilemme du prisonnier. Sur l'étendue du sens de ce dilemme, et sur la question de la réciprocité qui en est le centre, on peut consulter l'ouvrage de Laurent Cordonnier [1997].

⁴¹ Akerlof [1970] parvient au même résultat à partir de l'analyse d'une incertitude sur la qualité des biens.

⁴² *“Une des caractéristiques d'un système économique efficace est l'existence de relations de confiance suffisamment fortes pour que les agents ne fassent pas défection, alors même qu'il serait économiquement rationnel d'agir ainsi”* (Arrow, 1968, p.104).

⁴³ Cette conception du rôle de la collectivité vis-à-vis de l'action strictement individuelle est très nette chez Arrow. Ainsi il estime : *“Même le plus simple, en apparence, des actes individuels nécessite la participation de la société dans son ensemble”* (Arrow, 1967, p.63).

⁴⁴ Ce rôle fondamental des “codes éthiques” est apparu pour la première fois à Arrow lors de sa célèbre étude du système de santé américain (Arrow, 1963c). L'asymétrie d'information atteint en effet des niveaux particulièrement importants entre le patient et le médecin. Pour que l'échange puisse malgré tout avoir lieu, il est nécessaire que les médecins assurent les patients de leur honnêteté. C'est le rôle que jouent les nombreuses prescriptions éthiques qui encadrent l'exercice des professions de santé. Arrow conclut : *“l'incapacité du marché à prémunir les individus contre l'incertitude a conduit à la création de nombreuses institutions sociales dans lesquelles les caractéristiques habituelles du marché sont, dans une certaine mesure, contredites. La profession médicale n'est qu'un exemple, même extrême, parmi d'autres.”* (Arrow, 1963c, p.41).

rationnel, respecte certaines règles morales. Elles doivent, en quelque sorte, être intégrées à la définition de l'individu marchand, pour que le comportement de celui-ci puisse ensuite être décrit sous la forme d'une loi. L'individu marchand est donc tributaire de règles "extérieures" au monde Arrow-Debreu. Cet univers n'est plus conçu comme étant "autonome". L'efficacité du marché requiert l'existence de la morale. La morale est utile, et même nécessaire au fonctionnement de l'économie de marché. Dans l'esprit d'Arrow, elle est même "*conçue*" pour cela.

Cette réflexion sur le rôle de la morale nous ramène au commentaire du théorème d'impossibilité. A la suite de son théorème, Arrow justifie l'utilisation du mécanisme de marché parce qu'il contraint l'individu à agir en un sens conforme à l'intérêt général⁴⁵. Le consensus est déterminé en amont du marché, grâce aux valeurs morales communes, mais ces valeurs peuvent ensuite être oubliées : le marché permet à la collectivité d'atteindre les objectifs qu'elle s'est fixée en s'appuyant uniquement sur les préférences "égoïstes" des individus. Toutefois, cette "solution" est idéale. En pratique, l'incertitude perturbe les choix des individus sur le marché. Leur comportement s'écarte donc de ce que prévoit la "loi de marché". De ce fait, le marché ne remplit plus parfaitement sa tâche. Il faut alors que la collectivité parvienne à concevoir un mécanisme collectif suppléant les défaillances du marché. L'objectif de ce mécanisme doit être de guider les individus afin qu'ils agissent en un sens conforme au choix social. Pour que cet objectif collectif puisse être atteint, il faut donc que la morale se manifeste jusque dans l'échange marchand.

Arrow montre donc que, lorsque l'on introduit de l'incertitude dans l'univers marchand, l'existence de "valeurs communes" n'est plus une condition nécessaire et suffisante au fonctionnement du marché. Il faut que ces valeurs soient instrumentées de manière à s'imposer concrètement aux individus au moment de leur choix. C'est le rôle que jouent les "règles", les "codes" moraux.

312- La gestion collective de l'information

L'incertitude rend nécessaire la mise en place des "relations de confiance" entre individus. Cependant, il ne suffit pas que les individus s'accordent sur le respect d'un code éthique commun pour que leur incertitude quant à l'avenir soit jugulée. Il faut également qu'ils parviennent collectivement à "gérer" au mieux les informations dont ils disposent. Or, l'information n'est pas une "marchandise". Elle ne peut donc être gérée, de façon décentralisée, par le mécanisme de marché. Elle requiert l'existence d'un "langage commun" permettant une spécialisation de chacun des membres de la société.

⁴⁵ Le marché peut être utilisé pour faire correspondre les impératifs moraux et pragmatiques.

L'information n'est pas une marchandise parce qu'elle est "endogène" : c'est l'individu qui, en s'accaparant un signal, le transforme, ou ne la transforme pas, en information. L'information n'est donc pas extérieure à l'acteur, et en est difficilement séparable, ce qui induit quelques caractéristiques spécifiques⁴⁶ :

- 1- Son coût d'acquisition est croissant, puisque les capacités cognitives de l'individu ne sont pas infinies.
- 2- Elle nécessite d'investir dans une série de techniques et connaissances préalables permettant ensuite de la traiter. Cette forme de dépense est irréversible, puisqu'elle constitue un savoir acquis non revendable.
- 3- L'individu n'est pas sensible de la même manière à toutes les informations. Selon le champ de l'information, l'acquisition de celle-ci sera donc plus ou moins aisée, et plus ou moins coûteuse.

Ces trois caractéristiques expliquent l'échec d'une gestion "décentralisée" de l'information. Du point de vue de l'efficacité collective, il est en effet nécessaire que l'information fasse l'objet d'une part d'une mise en commun, d'autre part d'une spécialisation. La spécialisation des individus permet en effet de rentabiliser les investissements irréversibles et de minimiser les coûts d'acquisition (puisque chaque individu peut se consacrer à un seul type d'information). La mise en commun de l'information au sein des organisations permet d'économiser sur le "stockage" : chacun des individus n'a à connaître qu'une petite partie de ce que toute la collectivité détient. Un traitement *efficace* de l'information passe donc par une répartition du travail de collecte, permettant d'éviter des redondances et d'augmenter l'efficacité de chaque agent.

Pour ce faire la société doit disposer d'un instrument que le marché ne produit pas : le langage. Dans l'univers Arrow-Debreu, dans lequel les individus sont tous identiques, le langage est inutile. Mais, lorsqu'on introduit l'altérité, l'existence d'un langage commun devient une condition indispensable pour parvenir à se coordonner efficacement. De plus, dans l'optique d'Arrow, le langage ne doit pas seulement servir à se coordonner. Il doit également permettre de traiter *efficacement* les données informationnelles. Cet impératif d'efficacité se décline en deux "objectifs" que le langage devrait permettre d'atteindre : il faut *trier* ce qui est utile parmi les données empiriques, et *diffuser* rapidement les informations que l'on tire de ces données. Le langage "efficace" constitue donc un instrument permettant de construire une représentation sociale de la réalité cohérente avec l'objectif collectif que l'ensemble des individus se sont fixé.

Cette conception du langage s'articule parfaitement avec la notion de "code éthique". Le langage permet une bonne compréhension des acteurs, facilitant donc leur coordination et leur coopération. Mais il sert aussi, comme les "codes", à guider les individus dans une

⁴⁶ Arrow les présente successivement dans le second chapitre de son ouvrage (Arrow, 1974c, p.44-51).

direction précise conforme à l'intérêt général. Les règles langagières sont donc le pendant des codes éthiques. Ces deux éléments collectifs structurent le comportement individuel. Ils le guident. Ils le contraignent. Ils imposent une conformité à l'ensemble des individus.

L'objectif poursuivi par Arrow est de disposer d'une compréhension pratique du mécanisme marchand. L'introduction d'une dose d'incertitude dans le modèle marchand, lui permet de mettre le doigt sur le problème de l'altérité des individus. Il montre que le marché ne peut pas faire face à ce problème, que la démonstration formelle Arrow-Debreu excluait. Il faut donc recréer de l'unité et ajouter, à la définition rationnelle de l'individu standard, deux types de règles collectives : un langage commun et des codes éthiques. Ces règles remplissent, dans l'esprit d'Arrow, une fonction précise. Elles sont conçues, par la société, pour atteindre l'optimum. Elles doivent donc, avant tout être efficaces.

Or, l'impératif d'efficacité est un impératif social, collectif. La conception des règles ne peut donc être laissée à la liberté de chacun des individus. Il est nécessaire qu'une structure les bâtit en ayant en vue l'impératif de l'efficacité collective. La mise en évidence du rôle des règles n'est donc, en quelque sorte, qu'un artifice théorique de présentation. Derrière ces règles, c'est en fait le rôle de cette structure collective régulatrice qu'Arrow va mettre en évidence. En observant la présence, sur le marché, d'éléments collectifs ne répondant pas au "modèle" théorique Arrow-Debreu, et en expliquant le rôle de ces éléments collectifs à partir d'une prise en compte de l'incertitude et de l'altérité, Arrow progresse, en fait vers une mise en évidence du rôle des *organisations*. Les codes éthiques et le langage n'existent qu'à travers *l'organisation* économique. Ces "règles" collectives sont l'émanation d'une structure collective plus large qui regroupe les individus et leur impose un comportement précis lorsque le marché ne parvient pas à assurer efficacement un comportement conforme à "l'intérêt collectif".

32- Le rôle des organisations

"Les organisations sont un moyen de tirer parti de l'action collective lorsque le système des prix est en défaut."(Arrow, 1974c, p.35).

Arrow justifie le rôle des organisations par l'existence d'une défaillance du marché. Elles apparaissent parce que le marché ne permet pas aux individus d'être, en pratique, "quittes de tous les autres". Dans l'esprit d'Arrow, les organisations permettent d'assurer la cohésion collective nécessaire pour atteindre l'optimum. Elles servent à recréer de l'unité lorsque le marché, en raison de l'incertitude, ne parvient pas à éliminer l'altérité. Cette

faculté des organisations économiques tient précisément au fait qu'elles sont organisées sur un modèle inverse à celui du marché. Elles sont en effet centralisées et hiérarchiques.

321- La justification des organisations

“Le besoin de codes mutuellement compréhensibles au sein des organisations impose une condition d'uniformité au comportement des participants. (...) Par suite, l'organisation elle-même sert à modeler le comportement de ses membres.” (Arrow, 1974c, p.70).

Les organisations sont conçues comme des "ensembles de règles", éthiques et langagières, qui permettent aux individus de surmonter leur incertitude en se réfugiant dans un espace strictement balisé. Cet ensemble de règles limite en effet les possibles, et réduit donc d'autant les choix des individus⁴⁷. L'incertitude liée aux doutes quant au comportement des autres s'en trouve donc réduite, l'information peut circuler plus efficacement, et l'avenir devient mieux prévisible.

Selon Arrow, les organisations produisent donc de la conformité. Elles assurent l'existence à la fois d'un code éthique commun, assurant la coopération de l'ensemble des agents, et d'un langage commun, assurant leur coordination. Elles tirent leur efficacité d'une mise en commun des intérêts des individus que le marché ne permet pas. Elles permettent, parce qu'elles suspendent la compétition en leur sein, le développement de relations de confiance. Les individus unissent ainsi leurs intérêts de manière à améliorer leur capacité de traitement de l'information, pour minimiser l'incertitude et solidariser les pertes éventuelles. Cependant, l'efficacité de l'organisation repose sur une division du travail, une répartition des tâches. Elle implique que chacun des individus s'efface derrière la définition d'objectifs communs "organisationnels". De ce fait, les structures communes s'autonomisent et transcendent cette réunion d'individualités pour constituer un ordre collectif formatant les individus. C'est pour cette raison que les organisations représentent des mécanismes sociaux efficaces en situation d'incertitude. Elles *imposent* aux individus des objectifs collectifs que l'univers marchand leur fait perdre de vue.

La force de l'organisation s'explique donc par le fait que cet "ensemble de règles" transcende les individus qui les constituent, et finit par avoir une relative autonomie décisionnelle. Les organisations ne sont donc pas seulement des ensembles d'individus, elles ont une dimension verticale qui rend impossible leur décomposition en unités

⁴⁷ En termes de théorie des jeux de coordination, notamment étudiés par Schelling [1960] et Lewis [1969], l'organisation apparaît donc comme un moyen de résoudre une indétermination. Les codes internes représentent autant de point focaux permettant aux individus de se repérer. La limitation des possibles réduit l'incertitude quant au comportement des autres.

individuelles. L'organisation doit d'une part coordonner en son sein les échanges d'information entre un grand nombre d'individus, et, d'autre part, prendre des décisions collectives s'appuyant sur ces informations. L'utilisation d'un langage commun est dans cette optique une étape nécessaire mais non suffisante : il reste en effet à comprendre comment les individus sont coordonnés, afin de permettre une diffusion efficace des informations, et comment ils parviennent à s'accorder sur une décision collective.

322- Autorité versus consensus

Le problème qui est posé à l'organisation est donc celui de l'articulation de la rationalité individuelle et de la rationalité collective. La question de savoir comment déterminer une décision collective à partir du choix d'individus rationnels est précisément l'objet de la théorie du choix social⁴⁸. Arrow (1974c, p.86) rappelle donc qu'il existe (en dehors du marché) deux types de réponse "satisfaisante" à ce problème : le consensus et l'autorité⁴⁹.

Le premier terme de cette alternative est la prise de décision collective, démocratique, sur le modèle du vote. C'est ce type de procédure que le théorème d'impossibilité touche de plein fouet. Il faut, pour que cette procédure fonctionne bien, qu'il y ait une stricte convergence de vues entre les individus, et une structure d'information identique. Autrement dit, le consensus n'existe que si les préférences d'un individu sont aussi celles de tous les autres. Ce qu'Arrow résume ainsi : "*Dans quels cas le consensus peut-il être un substitut adéquat à l'autorité ? Une organisation dont les membres ont des intérêts identiques et la même information est une organisation dans laquelle le consensus serait efficace*" (Arrow, 1974c, p.87).

Le point de départ de l'organisation est, dans la problématique d'Arrow, l'existence d'une communauté d'intérêt entre les acteurs : tous ont intérêt à mettre en commun leurs informations et à se prémunir contre le futur. Cependant cette communauté d'intérêt est en elle-même fragile : il importe que chacun puisse en permanence la percevoir, et qu'elle perde même lorsque la situation change. Or, des événements imprévus ou des

⁴⁸ Arrow indique, à l'occasion de la réédition de son article consacré au paradoxe du choix collectif (Arrow, 1950), qu'il a été amené à se consacrer à cette question du choix social à l'occasion d'une réflexion sur l'analyse de la firme dans l'ouvrage de Hicks "*Value and capital*". Il cherchait alors à comprendre le mécanisme de la question de la prise de décision collective dans une firme.

⁴⁹ Arrow associe de manière un peu obscure consensus et agrégation des préférences : "*Par consensus j'entends tous les moyens raisonnables et acceptables d'agréger les intérêts individuels*" (Arrow, 1974c, p.86). Il rappelle ensuite que pour que l'agrégation fonctionne, il faut en fait qu'elle soit triviale c'est à dire que tous les individus soient d'emblée d'accord. C'est peut-être pour cette raison qu'il ne mentionne pas la décision par "convention", qui constitue dans Arrow (1951b) le troisième mode de décision collectif envisagé.

conséquences complexes difficiles à percevoir peuvent être sources de conflit et de rupture de cette communauté d'intérêt. Le conflit d'intérêt ne disparaît pas avec la mise en place des organisations, il est présent en permanence, sous-jacent à toutes les relations entre individus. La tentation de l'opportunisme ne disparaît absolument pas avec le constat des bénéfices qu'amène la coopération, au contraire⁵⁰. Arrow indique ainsi d'emblée : *"j'écarte l'idée qu'une unité absolue, une identité complète de sentiment entre les contextes collectifs et individuels soient possible. (...) Ainsi, on ne peut donner à la rationalité collective une signification absolument cohérente. Il y a, à un certain moment, une relation de pur pouvoir."* (Arrow, 1974c, p.26,28). Pour être efficaces, il est donc nécessaire que les organisations imposent certaines décisions. Il faut dans une certaine mesure que les individus substituent leurs intérêts privés aux intérêts collectifs de l'organisation. Cette substitution nécessite la contrainte.

L'autre aspect qui rend le consensus inatteignable est l'existence d'une même information entre les participants. Ce point soulève un problème intéressant. Il met en effet l'accent sur la double contrainte des organisations : spécialisation et mise en commun. Comme on l'a noté, le principe de spécialisation conduit les individus à se consacrer à un seul aspect de la réalité. Par suite, chacun des individus (ou groupe d'individus) aura une vision différente de l'avenir ou de ce qu'il conviendrait de faire pour assurer à chacun un paiement maximum. Le consensus est donc en contradiction avec le fonctionnement même de l'organisation. Il est nécessaire qu'une structure centrale gère la disparité de point de vue qui naît de la spécialisation.

L'autorité représente donc un substitut efficace au consensus :

- D'une part parce qu'elle résout la tension entre intérêts individuels et collectifs⁵¹.
- D'autre part parce qu'elle permet une centralisation de l'information : plutôt que d'informer tous les autres, chaque individu ne doit informer que le centre de décision.

⁵⁰ Il est toujours plus payant de faire défection lorsque tous les autres coopèrent, ce dont atteste, par exemple le dilemme du prisonnier (même lorsqu'il est répété un nombre fini de fois). Sur ce point également on peut se reporter à Cordonnier [1997].

⁵¹ Cette justification de l'autorité est très clairement une conséquence logique du théorème d'impossibilité étudiée dans la section précédente. Si on ne postule pas *a priori* une stricte identité de toutes les préférences individuelles, la seule solution rationnelle à la prise de décision collective est la dictature.

323- Une description de l'autorité.

“Nous dirons que B exerce une autorité sur W si W accepte que B sélectionne x [un ensemble de tâches à réaliser]. En général W accepte l'autorité uniquement si x_0 , le x effectivement choisi par B, appartient à un ensemble donné (l'ensemble des tâches acceptées par W) ”. (Simon, 1951, p.264).

Le principe de la délégation d'autorité a été initialement formalisé par Simon [1951], et Arrow reprend dans les grandes lignes, et sans les développer, l'esprit de ce modèle. Selon la définition de Simon, la relation d'autorité correspond à un contrat dans lequel l'individu qui entre dans l'organisation accepte, en échange d'un salaire, de réaliser un éventail de tâches entre lesquelles il laisse toute latitude de choix à l'autorité. C'est donc l'organisation qui, en fonction des besoins, choisit de lui assigner une des tâches correspondant à l'éventail plus ou moins large sur lequel l'individu s'est engagé. L'individu troque donc de la sécurité contre une forte diminution de son libre arbitre et de sa liberté. Cette forme de relation, distincte de la relation marchande habituelle, constitue une modalité de gestion de l'incertitude des acteurs. Le participant à l'organisation s'engage pour une période de temps longue, dans la mesure où on lui assure un paiement quel que soit l'état du monde qui se réalise. De l'autre côté, l'organisation se réserve le droit de prescrire des tâches adéquates à l'individu, en fonction de ses prévisions et de l'état du monde effectivement réalisé.

Ainsi, la relation d'autorité constitue une modalité de gestion de l'incertitude. Chacun des participants s'abrite de la contingence en se mettant à disposition de l'autorité. C'est en définitive la direction de l'organisation qui, seule, est amenée à gérer l'incertitude.

323- La centralisation de l'information

Cette tâche de l'autorité requiert une organisation centralisée des flux d'information. *“Puisque la transmission de l'information est coûteuse, dans la mesure où elle utilise des ressources, il est plus avantageux et plus efficace de transmettre tous les éléments d'information une fois pour toutes à un organe central, plutôt que de faire connaître chacun d'eux à tous les individus. Pour les mêmes raisons d'efficacité, il peut être plus avantageux pour un décideur ou un organe central de prendre la décision collective et de la transmettre, plutôt que de transmettre toute l'information sur laquelle elle repose.”* (Arrow, 1974c, p.85, c'est moi qui souligne). Ainsi, l'autorité, parce qu'elle est investie par les participants d'une mission de gestion de la contingence, prend seule les décisions, et n'a pas à les motiver.

Ce mode de fonctionnement de l'organisation donne aux dirigeants une entière liberté et une position absolument dominante. En effet, le centre de décision est le seul à détenir l'information nécessaire, issue de la spécialisation de l'ensemble des membres. N'ayant pas à redistribuer l'information qui lui parvient, mais uniquement les conclusions qu'il en tire, le centre dirigeant dispose d'un formidable moyen pour assurer son autorité. Chacun des autres membres de l'organisation ne dispose que d'informations partielles. Or, chaque morceau d'information parcellaire ne vaut rien s'il n'est pas mis en rapport avec les autres. En étant le seul à disposer de l'ensemble des morceaux, le pôle de décision réalise la synthèse dont chaque individu a besoin. Il devient donc un élément de référence indispensable pour chaque groupe. Il devient l'unique lien entre l'individu et "tous les autres". Il installe chaque individu, ou chaque groupe dans un rapport exclusif vis-à-vis de lui. Ce type de rapport exclusif lié à une centralisation de l'information existe dans un autre mode théorique de représentation de la coordination : le modèle Arrow-Debreu.

Ainsi, au-delà de l'apparente horizontalité de l'organisation, fondée sur l'existence de codes éthiques et langagiers communs, la structure même de l'organisation que dessine Arrow est caractérisée par sa verticalité. La solution au problème de coordination que pose l'incertitude, passe par la mise en place d'une structure isolant à nouveau chaque individu de tous les autres. Dans le modèle Arrow-Debreu, les relations interpersonnelles sont médiatisées par le biais du système des prix, qui établit un rapport, exclusif de tout autre, entre l'individu et le centre du marché. L'incertitude, en fragilisant le rapport exclusif individu-système des prix, fait resurgir le problème de l'altérité. L'organisation résout ce problème en remodelant les individus autour d'un code et d'un langage commun. Ainsi le lien entre chacun des individus et "tous les autres" est de nouveau médiatisé par une relation exclusive entre l'individu et le centre de décision. L'individu est déchargé de la question de la coordination. Il peut donc être instrumentalement rationnel, dans la limite des règles propres à l'organisation à laquelle il participe.

33- La structure hobbenne des organisations

"Les arguments en faveur de l'autorité sont simples et familiers. Hobbes les a énoncés sous une forme classique quoique extrême, en faisant référence au cas particulier du gouvernement. En l'absence de toute autorité il y a une guerre de "tous contre tous" et, par suite, "la vie de l'homme est misérable, désagréable, bestiale et brève". L'élément de vérité de la pensée de Hobbes peut recevoir une formulation plus générale, sinon

beaucoup moins dramatique : l'autorité est nécessaire pour réaliser la coordination des activités des membres de l'organisation". (Arrow, 1974c, p.84).

En faisant de manière explicite référence à Hobbes, au moment d'intégrer la dimension nécessairement verticale des organisations et le rôle crucial de la relation d'autorité, Arrow délivre une des clefs de lecture de son analyse. Trois points permettent en effet de qualifier de "hobbien" le programme de recherche proposé par Arrow :

- 1- La justification de l'autorité se fait, comme chez Hobbes, en référence à l'opportunisme des agents.
- 2- Une des dimensions essentielles de la solution organisationnelle est, comme chez Hobbes, la création d'un langage adapté.
- 3- Le contrat qui unit l'individu à l'organisation est de même nature que le contrat social hobbien.

Le premier point est explicite dans le passage cité ci-dessus. Cependant le rapprochement prend encore plus de sens si l'on cherche à reconstituer le cheminement théorique qui aboutit à cette justification de l'autorité. Arrow part du constat de l'incomplétude du système marchand, et de l'incertitude qui en découle. Cette incertitude oblige l'individu standard à sortir de la relation exclusive qui le lie au système des prix, et à chercher de l'information auprès des autres individus. Ainsi les individus se rencontrent sans pour autant être dotés d'un mode de rationalité différent du mode instrumental. Cette découverte de l'altérité pose un problème de coordination lié au fait que des individus instrumentalement rationnels deviennent nécessairement égoïstes et opportunistes lorsqu'un tel comportement est profitable. C'est de cette situation de conflit et d'absence de confiance mutuelle que naît le besoin d'organisations structurées autour d'une relation d'autorité. L'autorité est une solution parce qu'elle isole chaque individu de ses semblables, le rend conforme à elle-même, et le place dans un rapport exclusif et inégal avec le centre de décision. Cette séquence théorique qui relie l'existence d'incertitude à une forme autoritaire de coordination est tout-à-fait semblable à celle qui, chez Hobbes, relie l'existence d'un Léviathan au caractère invivable de l'état de nature. Dans un cas comme dans l'autre l'autorité hiérarchique naît de l'incapacité des individus à s'entendre et de la nécessité, en conséquence, de les isoler les uns des autres en mettant chacun sous le regard de "tous les autres", entité personnifiée par une autorité centrale omnisciente. La liberté pour chaque individu d'user librement de sa rationalité instrumentale sans avoir à s'imposer de lui-même des contraintes liées au respect des autres s'obtient au prix du respect par tous d'une autorité centrale.

Le second point se situe dans le prolongement du premier. L'arme suprême de l'autorité réside dans sa création d'un langage propre qu'elle seule maîtrise parfaitement. Le Léviathan, chez Hobbes, tire sa force de sa capacité à nommer chaque chose de manière non équivoque. Dans l'état civilisé, la paix peut régner parce qu'il n'existe plus d'ambiguïté sur la dénomination des objets. Le langage du Léviathan est le langage universel et ne permet aucune confusion. Ce nominalisme strict, caractéristique de Hobbes, n'est pas absent de la conception du langage organisationnel qui transparait dans l'analyse d'Arrow. Dans l'organisation qu'il présente, chaque individu connaît un domaine strict, et utilise les mots propres à ce domaine. Le centre de décision centralise toutes les informations et en déduit un mode d'action qu'il communique à tous. Ainsi, seul le centre utilise et crée le langage commun, et ce langage commun est "construit" pour être non-ambigu et efficace. Les individus ne connaissent qu'une petite part des connaissances globales et ne maîtrisent qu'un aspect du langage de l'organisation. Pour cette raison, ils sont tributaires des décisions de l'organe central de décision. Comme chez Hobbes, l'organe de décision, en créant et en maîtrisant le langage, s'assure le contrôle des hommes.

Cette puissance de la hiérarchie ne va cependant pas sans contrepartie. Hobbes montre combien les individus ont intérêt à s'y soumettre plutôt que de sombrer dans l'état de nature caractérisé par le règne de la mort violente. De même, pour Arrow, les individus tirent bénéfice du contrat qu'ils passent avec l'organisation, dans la mesure où celle-ci les protège de la contingence. En se mettant sous l'autorité de la direction, les individus obtiennent la garantie d'une certaine stabilité de leur environnement. Leur perte de liberté est en quelque sorte compensée par l'intérêt qu'ils ont à être déchargés d'une part du poids que représente la coordination avec les autres, d'autre part du risque qu'il y a à agir sans connaître le futur. Comme chez Hobbes, cette capacité de l'organisation à protéger ses membres vient de sa capacité à prendre des décisions collectives rationnelles, fondées sur des informations complètes⁵².

Ce rapprochement entre les théories d'Arrow et Hobbes doit cependant être nuancé. Arrow, en effet, s'attache bien plus à montrer dans quelle mesure les organisations sont nécessaires pour pallier les déficiences du marché qu'à étudier de manière approfondie ces organisations. Il ouvre ainsi, en partant de la notion d'incertitude, un espace théorique

⁵² Ce dernier point éclaire le statu de l'analyse standard de la rationalité dans l'incertain. Comme l'indique Suppes [1981], la conception de Savage s'applique davantage à une entité disposant d'une quantité d'information et d'une puissance de déduction supérieure à celle d'un individu unique. En définissant l'attitude rationnelle dans l'incertain, Savage visait sans doute davantage à définir l'activité d'un centre de décision organisationnel que d'un individu isolé. Comme le note Arrow [1958, p.64], "*Le problème de la décision statistique est une partie du problème du choix social. C'est sous cette forme que De Finetti et Savage ont posé ce problème*".

que la théorie de l'équilibre général laissait dans l'ombre : mais il n'investit pas réellement cet espace. Arrow n'est pas un institutionnaliste. Il reste un théoricien attaché à l'analyse formelle du marché. Son travail critique vise, d'une certaine manière, à circonscrire le sens et la portée des théorèmes mathématiques formels attachés au modèle Arrow-Debreu. Ils ne constituent pas en eux-mêmes des solutions permettant d'aller au-delà de ces théorèmes. Il reviendra à d'autres, et en particulier à Williamson, de franchir le pas et de proposer un véritable programme de recherche néoinstitutionnaliste permettant d'analyser plus finement les aspects internes des organisations, et les rapports entre organisation et marché.

Il n'en reste pas moins que le lien fondamental "incertitude->altérité->conflit->autorité" qu'Arrow met à jour est fondamental pour la perspective néoinstitutionnaliste. Or, elle est profondément du même type que le lien "état de nature->guerre->Léviathan" qu'Hobbes utilise pour justifier le pouvoir absolu du gouvernement civil. Ainsi, il nous semble légitime d'affirmer qu'Arrow propose une vision hobbenne des organisations : l'individu ne peut être "libre" et "autonome" qu'en acceptant la contrainte du Léviathan qui le libère de la menace de tous les autres individus. C'est à ce prix que l'hypothèse de rationalité instrumentale peut être sauvée.

Conclusion

“ Il résulte de ces remarques que, du point de vue de l’efficience comme du point de vue de la justice distributive, il faut quelque chose de plus que le marché ” (Arrow, 1974c, p.24).

Arrow pense que le marché est sans doute *“ l’un des moyens, peut-être le meilleur, par lequel n’importe quel système de valeur social peut être réalisé ”* et que *“ l’analyse de son fonctionnement est une des entreprises intellectuelles les plus remarquables ”* (Arrow, 1967, p.65,66). Il ne remet donc pas en cause le résultat auquel il est parvenu avec Debreu, et qui constitue aujourd’hui encore la base de l’analyse économique moderne, au moins dans sa version néoclassique.

Mais ce constat ne l’aveugle pas. Il affirme, avant même sa démonstration du théorème d’existence, et de plus en plus clairement par la suite, que la connaissance du marché n’épuise pas l’objet de l’économie. Le marché n’est qu’un moyen d’atteindre des fins déterminées hors de lui-même, et c’est un moyen insuffisant.

Le théorème d’impossibilité met en évidence les raisons pour lesquelles le marché ne peut pas constituer un mécanisme permettant de déterminer automatiquement un choix collectif satisfaisant. Pour percevoir le fondement du choix collectif, il faut s’intéresser aux valeurs sociales qui guident les choix des individus. Or, les valeurs sociales ne peuvent s’exprimer au travers des préférences individualistes, égoïstes, instrumentales, qui caractérisent les choix marchands. Le choix collectif est donc nécessairement déterminé en dehors du marché. Voilà donc précisée par Arrow une première limite au système de marché : le marché n’est qu’un moyen d’atteindre le choix collectif. Il ne peut se passer des règles et des valeurs collectives que parce qu’il est un instrument collectif dont les fins sont fixées par des règles et des normes sociales.

Mais Arrow doute aussi du marché comme moyen exclusif. Dès lors que l’on introduit de l’incertitude, le système de marché se fissure. Les individus ne peuvent plus s’engager, le système ne leur donne plus suffisamment d’information ni de garanties. Ils cherchent donc à s’informer en dehors du système de prix. Ce faisant, le principe d’unité du marché, qui reposait sur l’exclusivité du lien entre chaque individu et le système des prix est remis en cause. Le marché n’assure plus correctement la coordination des individus qui, entrant en contact, constatent leur égoïsme mutuel. C’est pourquoi il est économiquement nécessaire qu’existent, en plus du rapport marchand, des liens plus forts entre les individus. C’est le rôle des codes éthiques, des valeurs, mais aussi et surtout des

organisations qui imposent ces normes et valeurs à leurs membres et, par ce biais, recréent de l'unité. L'existence des organisations est donc indispensable pour que le marché puisse être efficace.

A travers ces deux lignes critiques, Arrow effleure la question de la règle : comment les individus parviennent à s'accorder sur le cadre collectif qui détermine l'exercice solitaire de la rationalité instrumentale ? Mais cette question n'est soulevée que pour être immédiatement résolue : l'accord collectif est toujours déjà là, non problématique, non sujet à débat. La rationalité instrumentale est "encastrée" dans des normes sociales et des règles organisationnelles qui ont en commun de transcender l'individu, de s'imposer à lui. C'est en ce sens que les règles qu'envisage Arrow sont d'essence hobbienne : elles sont nécessaires, et d'une nécessité d'ordre supérieur. Cette nécessité est imposée par le primat de l'efficacité collective.

Pour que l'étude du marché puisse être débarrassée du problème du choix des règles, il faut donc supposer qu'il existe une unité de la société constituée autour de certaines valeurs morales, et que ces valeurs morales sont imposées aux individus par le biais des organisations. Selon ce schéma, les organisations sont donc en quelque sorte les outils que la société utilise pour atteindre son objectif. Les règles organisationnelles sont entièrement déterminées en fonction de cet objectif d'efficacité.

Ainsi, de l'autre côté du marché, règne la contrainte collective. Arrow fait apparaître la contrepartie indispensable du primat de la rationalité instrumentale dans l'univers marchand. Il montre qu'il n'existe pas véritablement de solution "marchande" à la question de la règle. Cette solution qui prétend reposer sur la liberté des individus n'existe en réalité que sur fond d'une contrainte de type hobbien, incarnée par les organisations. C'est parce qu'elles existent que la question du choix des règles peut être évacuée.

La "critique d'Arrow" nous amène donc aux frontières de la "modalité organisationnelle" d'analyse du problème de la règle.

Seconde partie

*La modalité
organisationnelle*

L'étude de la modalité marchande nous a permis de faire apparaître, derrière le marché, le rôle des organisations. Le projet néoclassique vise à démontrer que le marché est un mécanisme qui, parce qu'il isole chaque individu de "tous les autres", rend inutile toute règle. Cependant, la démonstration *logique* de l'efficacité du marché est établie à un niveau qui ne permet pas de juger de l'efficacité *pratique* de ce mécanisme. Sur ce point, les réflexions critiques d'Arrow montrent précisément que le marché ne se suffit pas à lui-même : il requiert pour fonctionner l'existence de règles non marchandes. Arrow pointe alors un autre mécanisme permettant de prendre en charge la question du choix et de la mise en place de ces règles : l'organisation. Nous étudierons ici la "modalité organisationnelle" de traitement du problème de la règle.

Dans la tradition néoclassique, la prise de conscience du rôle des règles organisationnelles est relativement récente. Elle n'a donné lieu à un véritable mouvement théorique qu'à partir du milieu des années soixante-dix, lorsqu'Oliver Williamson a redécouvert le concept de coût de transaction, proposé initialement par Coase [1937]. C'est donc à partir de l'analyse du "coût" que représente la gestion marchande des échanges, que l'organisation a pu apparaître comme un univers économique essentiel. Ainsi, comme chez Arrow, l'organisation apparaît à partir de la mise en évidence des failles du modèle logico-mathématique de marché.

Cependant, Arrow se contente d'indiquer quels sont les éléments qui rendent possible l'existence d'un marché autorégulé. Il reste résolument attaché à une conception formelle de la science économique. S'il dégage les conditions pratiques du formalisme, il ne cherche pas à comprendre comment ces conditions peuvent être *effectivement* remplies.

Cette question nécessite en effet de quitter la temporalité purement logique qui caractérise la théorie de l'équilibre général pour prendre réellement en compte une forme de temps plus pratique. Ce n'est que lorsque l'on prend en compte cette forme de temps que l'on peut comprendre le fonctionnement des "*institutions*". C'est précisément le pas que franchit Williamson [1975] en jetant les bases du mouvement "néoinstitutionnaliste". Ce pas est essentiel. Prenant acte de l'échec de la "modalité marchande" du traitement de la règle, il consiste à étudier d'autres manières de faire face à la nécessité de coordonner dans le temps des individus instrumentalement rationnels. Il apporte donc une nouvelle réponse à la question non seulement du rôle, mais aussi du choix, des règles économiques.

Or, comme le souligne déjà Arrow [1951b], dans la perspective instrumentale, qui est celle de la théorie économique néoclassique, il n'existe que trois mécanismes de choix collectif : la "dictature", la "convention", le "consensus par agrégation". Le troisième mode de décision, dont le modèle économique est le marché, est un mode de décision qui nie le rôle de l'institution. Si le choix collectif est directement déductible des choix individuels, il n'est nul besoin de construire des institutions dont le rôle est, précisément, d'indiquer aux individus une direction commune. Les solutions institutionnelles, qui naissent de l'échec de l'agrégation logique marchande, sont donc au nombre de deux : la dictature et la convention. On repérera précisément ces deux modèles derrière les deux "concepts" néoinstitutionnalistes que nous étudierons : *l'organisation hiérarchique* de Williamson, et la *convention lewissienne* utilisée en théorie des jeux.

Ces deux modèles se répondent. Ils ont une caractéristique commune : pour préserver l'existence d'une forme de rationalité individuelle instrumentale, il est nécessaire de "suspendre" le temps. La hiérarchie et la convention sont deux modalités qui imposent une certaine uniformité aux individus. Elles règlent donc de fait la question du temps puisqu'elles en éliminent le problème de l'altérité. L'uniformité est cependant obtenue de façon tout-à-fait différente dans les deux cas.

Williamson montre que l'efficacité des règles organisationnelles hiérarchiques repose sur la *coercition*. L'organisation est efficace parce qu'elle impose aux individus un ordre déterminé et conçu en dehors d'eux-mêmes : sur le modèle de l'artifice hobbien. Cette solution repose sur l'*opposition* entre le niveau de la rationalité individuelle et celui de la rationalité collective, prise en charge par l'organisation. C'est la raison pour laquelle cette représentation des règles collectives ne nous permet pas de comprendre comment les individus construisent ces règles. Les règles organisationnelles émergent à partir d'une logique qui transcende les individus. Elles viennent d'en haut : mais d'où ? Une conception holiste de l'ordre social est inconciliable avec la perspective individualiste

néoclassique. L'obscur origine des organisations représente en particulier un obstacle pour comprendre leur dynamique.

C'est pourquoi la notion de convention, proposée initialement par le philosophe David Lewis (qui l'emprunte à Hume), représente une alternative intéressante. Elle permet en effet d'articuler rationalité individuelle et collective sans faire apparaître d'opposition nette entre ces deux niveaux. La convention est une modalité de l'ordre collectif qui émerge directement des interactions individuelles, puis s'autonomise et s'impose comme un référent collectif. C'est pourquoi la théorie des jeux, dans sa version "néoinstitutionnaliste", utilise ce concept pour rendre compte de l'émergence des règles collectives. Mais pour que ce concept prenne véritablement sens, il faut le réintroduire dans un temps historique, ce que ne permet pas l'hypothèse de rationalité instrumentale. La théorie des jeux "standard" ne parvient donc pas véritablement à expliquer *l'émergence* des conventions : elle ne rend compte que de ce qui est déjà là. L'hypothèse classique de rationalité représente un obstacle pour rendre compte de la genèse des règles : la théorie des jeux évolutionnistes cherche à le lever. Mais se débarrasser de l'hypothèse de rationalité pose un problème : on ne dispose plus d'aucune forme de causalité permettant d'expliquer l'évolution des règles. Celles-ci semblent donc être douées d'une forme de dynamique "interne" qui demeure mystérieuse. Pour aller plus loin, il faut non pas supprimer mais modifier ou élargir l'hypothèse de rationalité individuelle. Ce sera l'objet de la partie suivante.

Chapitre 3

L'analyse de Williamson

Oliver Williamson a initié, dans les années soixante-dix, une véritable école d'analyse des organisations économiques articulée autour de l'étude de la minimisation des "coûts de transaction" inhérents au système économique. Cette école d'analyse n'a cessé dès lors de se développer, dans des directions diverses, jusqu'à constituer, avec d'autres écoles plus ou moins proches, le paradigme "néoinstitutionnaliste". L'objectif de ce paradigme est de parvenir à intégrer le concept d'institution dans le champ théorique de l'économie standard. Certes, cette intégration nécessite des adaptations, des modifications du cadre d'analyse néoclassique, mais l'essentiel est conservé : il s'agit d'intégrer l'idée d'institution à partir de l'étude du comportement d'un individu instrumentalement rationnel, et donc potentiellement égoïste.

Ce mouvement théorique se nourrit directement de la littérature théorique décrivant les limites du mode de coordination marchand, dont Arrow est un des contributeurs les plus notables. Le constat, selon lequel le marché ne se suffit pas à lui-même et requiert, pour fonctionner, la présence d'autres modes de coordination mieux adaptés à certains aspects des transactions économiques, rend en effet nécessaire l'étude de ces modes de coordination alternatifs. Williamson tente donc de répondre à la question théorique : "quelles règles non marchandes faut-il mettre en place pour combler les lacunes du marché ?".

Cette question est une question propre au modèle néoclassique. Comme on l'a constaté dans le chapitre précédent, le concept de "faille" du marché ne prend sens qu'au regard de la démonstration de la théorie de l'équilibre général. Williamson se situe donc d'emblée dans l'univers néoclassique. Mais sa réponse consiste à intégrer dans cet univers certaines idées des institutionnalistes américains, et en particulier leur conception de l'entreprise capitaliste. Il convient en effet de ne pas considérer cette organisation uniquement comme

une *fonction de production*, mais aussi et surtout comme une *institution* remplissant une fonction de coordination que le marché ne saurait assumer. Williamson tente donc d'opérer une synthèse entre les postulats de base de l'économie néoclassique et les constats empiriques des institutionnalistes.

Cette attitude théorique conduit Williamson à mettre l'accent sur l'origine humaine des coûts de fonctionnement du marché, qu'il nomme "*coûts de transaction*". Son analyse repose donc sur une étude minutieuse des données comportementales qui occasionnent des dysfonctionnements du mode de régulation marchand. Cette analyse constitue, en quelque sorte, une anthropologie de l'*homo-æconomicus*, lorsque que celui-ci n'agit plus dans les conditions d'information parfaite sur lesquelles repose le modèle standard.

Une fois ces données comportementales éclaircies, Williamson entreprend de répondre à la question du choix des règles. Il postule que ce choix doit être guidé par un objectif unique : la recherche de l'efficacité. Or, les agents économiques sont, selon "l'anthropologie" williamsonienne, cognitivement limités et opportunistes. De ce fait, la réponse adéquate ne fait pas de doute : il faut mettre en place des règles coercitives dissuadant l'opportunisme et centralisant l'information. Le modèle de ce type de règle est la règle hiérarchique. Williamson justifie ainsi le mode d'organisation classique des entreprises capitalistes.

L'analyse de Williamson constitue donc une réponse cohérente au problème de la règle tel qu'il apparaît, par exemple, chez Arrow. Mais, s'appuyant sur les "défauts" des individus, il produit une théorie articulée autour de l'opposition entre rationalité individuelle et rationalité collective. L'organisation s'impose à l'agent, elle n'en émane pas. Une telle théorie présente le double défaut, dans une perspective orthodoxe, d'être à la fois "holiste", puisque les organisations collectives semblent avoir une logique propre, et "statique", puisque Williamson s'intéresse exclusivement au niveau microanalytique de l'agent, qui précisément ne perçoit pas la logique qui préside à l'organisation et ne peut donc participer à son évolution.

1- La synthèse williamsonnienne

Williamson [1975, 1985] cherche avec constance à se situer au confluent de deux traditions économiques : l'institutionnalisme et la tradition néoclassique. Ces deux courants sont, pourtant, opposés sur le plan méthodologique. L'école institutionnaliste américaine du début du siècle s'attache à mettre en évidence les mécanismes socio-historiques par lesquels une collectivité se dote des institutions à même de réguler l'économie. Elle produit un discours synthétique, qui est peu formalisé et s'appuie sur un fort contenu empirique. A l'inverse, la tradition néoclassique analyse l'étude du comportement d'un individu formellement rationnel, situé dans un univers marchand de nature mathématique. Elle produit un discours analytique, fortement formalisé et ayant très peu de contenu empirique. Pour parvenir à rapprocher ces deux courants, le pari que fait Williamson est de se saisir du vide théorique que fait apparaître le constat des "failles" du marché, pour réintroduire à l'intérieur de la théorie néoclassique les principaux concepts institutionnalistes. Il ambitionne ainsi de réaliser une synthèse entre ces deux courants théoriques en donnant aux concepts institutionnalistes un fondement analytique leur permettant de trouver place dans la science économique moderne.

Cette ambition trouve sans doute son origine dans l'insatisfaction ressentie par l'auteur face au discours antitrust dominant dans les années soixante-dix. Pour contrer ce discours, Williamson mobilise la conception institutionnaliste de l'entreprise. Il opère toutefois un lien direct entre cette conception et la littérature consacrée aux "failles" du marché, mises en évidence par Arrow en particulier. Il trouve ainsi dans la théorie orthodoxe des "questions" qui lui permettent de faire apparaître son effort de synthèse théorique comme une "réponse" adéquate.

11- Deux révélateurs : les lois antitrust et l'article de Coase.

"Les inspireurs de la politique publique ont entretenu une attitude extrêmement méfiante vis-à-vis des formes non standard et inhabituelles d'organisation économique. (...) Cette tradition inhospitalière explique l'animosité de la politique publique vis-à-vis de l'intégration verticale et des conglomérats. Les spécialistes de l'organisation industrielle ont été encouragés à découvrir des "distorsions", souvent imaginaires, aux dépens d'une compréhension plus profonde de la firme moderne. (...) [A l'inverse] cet essai tente de fournir une vue cohérente de la firme moderne (...) [notamment] en substituant à la

présomption selon laquelle les innovations organisationnelles ont des objectifs anticompetitifs, la présomption selon laquelle les innovations organisationnelles servent à économiser des coûts de transaction .” (Williamson, 1981, p.1539-40).

L'influence puissante du modèle d'équilibre général durant les premières années d'après-guerre aux USA apporte des arguments aux tenants des lois antitrust. Durant cette période, le prestige du mode de coordination marchand est en effet suffisamment important pour que tout écart délibéré par rapport au modèle de la concurrence pure et parfaite apparaisse suspect et contre-productif. Le bien-être collectif semble aller de pair avec la lutte contre la collusion des entrepreneurs et pour le maintien de relations fortement concurrentielles entre les entreprises. Cet état d'esprit bloque la compréhension de ce qu'est la firme capitaliste. La "boîte noire" est encore à ouvrir.

Williamson n'est cependant pas le premier à chercher à ouvrir cette "boîte noire". Durant les années trente, Ronald Coase a fait paraître un article précurseur dans lequel il donne toute sa place à l'idée d'entreprise ou d'organisation : *"A l'extérieur de la firme, des mouvements de prix déterminent la production, qui est ainsi coordonnée à travers une série de transactions sur le marché. A l'intérieur de la firme, ces transactions marchandes sont éliminées, un entrepreneur-coordonateur se substitue à ces complexes structures marchandes, c'est lui qui dirige la production. Il est clair qu'il s'agit là de deux modes alternatifs de coordination de la production."*(Coase, 1937, p.388, c'est moi qui souligne). Coase identifie l'élément qui détermine le choix entre ces deux modes alternatifs : *"La principale raison pour laquelle il peut être profitable de créer une entreprise est qu'il existe un coût d'usage du système des prix."* (p.390). Ainsi le rôle de l'entrepreneur et des entreprises est de minimiser ces coûts liés à l'organisation des transactions marchandes. Une relation de long terme est moins coûteuse qu'une multitude de transactions ponctuelles. L'entreprise prend donc en charge les transactions jusqu'au point où la prise en charge d'une transaction supplémentaire s'avère plus coûteuse que le recours au marché. A compter de ce moment, on entre dans l'espace *"des rendements décroissants du management"* (p.392). L'existence de ce seuil explique que le marché et l'entreprise coexistent, et que la production ne soit pas le fait d'une seule grande organisation. Coase propose donc une conception tout-à-fait aboutie du rôle de l'entreprise, et de sa spécificité en regard du marché.

Pourtant ces intuitions resteront lettre morte durant une trentaine d'années. L'auteur lui-même le reconnaît amèrement : son article est *"beaucoup cité et très peu utilisé."* (Coase, 1972, p.63). Il souligne plus loin, en faisant référence à la tradition antitrust déjà évoquée : *"une conséquence importante des problèmes inhérents au monopole est que lorsqu'un économiste trouve quelque chose qu'il ne comprend pas, il produit une*

explication en termes de pratique monopolistique” (p.67). L’initiative théorique de Coase n’a donc pas immédiatement porté ses fruits. Les premiers écrits de Williamson vont précisément contribuer à populariser l’intuition de Coase, en donnant un contenu concret au concept de “coût de transaction” qui souffrait jusqu’alors d’une certaine imprécision.

Le manque de légitimité de l’organisation, en tant que mode de coordination alternatif au marché, sert de déclencheur à l’initiative théorique de Williamson¹. L’analyse de Coase lui fournit les armes nécessaires.

12- Un retour limité aux institutionnalistes.

Les années soixante-dix sont cependant une période de renouvellement général de la théorie néoclassique. La prise de conscience des failles du marché et de l’importance de l’information et de l’incertitude conduit les théoriciens néoclassiques à modifier les conceptions microéconomiques qui sous-tendent la théorie de l’équilibre général. Williamson n’est donc pas, de ce point de vue, un cas isolé. Mais, tandis que la plupart des théoriciens orthodoxes vont chercher à mettre l’accent sur la complexité des contrats, en augmentant leurs exigences en terme de rationalité individuelle pour parvenir à conserver le caractère central du mécanisme de coordination par les prix, Williamson va exhumer le concept institutionnaliste d’une organisation conçue comme un mécanisme de coordination alternatif au marché permettant précisément d’assouplir les exigences portant sur la rationalité individuelle.

Cette spécificité par rapport à des travaux plus fidèles à la théorie néoclassique apparaît clairement lorsque l’on compare l’analyse de Williamson à celle que proposent la théorie des droits de propriétés (Coase [1960], Alchian et Demsetz [1972, 1973]) et de la théorie de l’agence qui s’en inspire (Jensen et Meckling [1976], Pratt et Zechauer [1985]). L’objectif de ces deux courants de recherche est en effet de parvenir à intégrer l’idée d’institution exclusivement par le biais de contrats marchands complexes. Cette stratégie d’intégration interprète donc l’institution, et les organisations, comme des ensembles de contrats sans spécificité particulière par rapport au marché. Comme l’indiquent Alchian et Demsetz dans leur article fondateur : “ *Il est courant de caractériser l’entreprise à travers le pouvoir qu’elle aurait de régler des conflits en imposant une solution, en s’appuyant sur*

¹ Comme le rappelle Brousseau [1989, p.125], Williamson a travaillé à la division antitrust du département de la justice américaine, ce qui a sans doute été un élément important dans le choix des ses recherches futures.

l'existence d'une relation d'autorité ou encore en prenant des mesures disciplinaires supérieures à celles disponibles sur le marché classique. C'est une illusion. L'entreprise ne dispose pas de tous ces inputs. Elle n'a pas le pouvoir de décréter, pas d'autorité, pas de mesures disciplinaires différant à quelque degré que ce soit d'un contrat marchand ordinaire entre deux individus". (Alchian et Demsetz, 1972, p.777).

Ce courant d'analyse prend acte, comme Williamson, des failles de la régulation marchande et développe des analyses parfois tout-à-fait convergentes². Mais il ne reconnaît pas la spécificité des organisations par rapport au marché et met l'accent sur l'étude du contrat. Ainsi, les failles du marché sont supposées pouvoir être comblées par la mise au point de contrats prévoyant *ex ante* tous les dysfonctionnements pouvant survenir *ex-post*. Les agents sont donc supposés pouvoir être suffisamment rationnels pour surmonter d'eux-mêmes les failles informationnelles et techniques du système de marché.

A l'inverse, Williamson pense qu'il est nécessaire de concevoir des dispositifs institutionnels permettant de gérer la transaction durant tout son déroulement parce que les individus ne sont pas suffisamment rationnels pour prendre en compte toutes les possibilités et concevoir des contrats trop complexes. Selon lui, l'organisation collective doit donc venir combler ces lacunes de la rationalité individuelle. C'est en ce sens que son optique est institutionnaliste : sa conception de l'individu implique la nécessité de l'insérer dans des structures collectives. Williamson (1975, p.1-4 ; 1985, p.2-7) dresse ainsi l'inventaire de ses emprunts théoriques aux institutionnalistes et aux théoriciens de l'organisation. De Knight, il retient la nécessité de concevoir "*la nature humaine telle que nous la connaissons*" avec ses limites cognitives et son caractère opportuniste. De Commons, il retient que "*la transaction doit être considérée comme l'unité de base de l'analyse*" et que le rôle des institutions est d'ordonner un univers de conflit en mettant au point des "*règles adéquates de l'action collective*". Conformément à ces deux auteurs, l'économie des coûts de transaction vise donc à étudier les différentes formes de transactions que passent entre eux des individus conçus de manière "réaliste" et tente de proposer, pour chacune d'elles, une structure collective permettant de la gérer efficacement.

Williamson réalise un premier rapprochement entre ce champ institutionnaliste et la théorie des organisations (Barnard, Simon, Chandler), pour faire finalement apparaître

² Dans son second ouvrage, Williamson tente d'intégrer, de manière un peu confuse, certaines de ses analyses à la nouvelle économie institutionnelle, sous la catégorie, floue, de "branche de la mesure". Il place son travail dans l'autre branche : la branche de la gouvernance. Sous le terme "économie des coûts de transaction" et "courant néoinstitutionnaliste", nous ne désignons que cette "branche de la gouvernance" qui nous semble être, en fait, la seule véritablement homogène. (Williamson, 1985, p.27-29).

l'organisation comme l'institution économique fondamentale. Il se réfère tout particulièrement à Barnard, à qui il emprunte une description de l'organisation comme un mode d'action collective permettant de gérer des savoir individuels de nature "idiosyncrasiques" non transférables. Les structures collectives de coordination vont donc être organisées autour d'un impératif de gestion efficace de cette opacité du travail et du savoir humain.

Cette suite de références que Williamson égrène avec attention désigne clairement le chemin théorique qu'il souhaite suivre. Williamson oppose le marché et l'organisation. Il cherche à montrer que les organisations ne sont pas réductibles à des formes de contrats marchands complexes, mais constituent des univers régulés de manière différente et adaptés à certaines spécificités des individus. Tandis que l'univers marchand est régulé par le biais d'une multitude d'échanges individuels décentralisés, dans lequel les agents ne s'engagent dans aucune relation de long terme, les organisations reposent précisément sur des échanges de long terme, centralisés, permettant aux individus d'établir une relation simplifiant considérablement les transactions. Les organisations sont le théâtre d'une action collective qu'elles rendent possible par l'existence d'une série de règles collectives facilitant le bon déroulement des transactions. C'est cette "mise en commun" qui distingue l'organisation du marché et explique son efficacité. Certains types d'échanges qui se déroulent dans le long terme nécessitent la présence de règles non marchandes qui permettent aux individus de pérenniser leur relation et de réduire leur incertitude.

Williamson fait apparaître l'organisation comme une "institution" adaptée aux spécificités de la "nature humaine". En cela, il se situe dans la lignée des institutionnalistes, bien qu'il intègre ces éléments théoriques dans un cadre analytique néoclassique qui leur est étranger.

13- La filiation Arrow-Williamson.

Williamson [1987a] place son entreprise théorique dans la lignée des travaux d'Arrow concernant l'organisation, l'économie de l'information et les "failles" du marché. Le caractère critique de ces travaux, que nous avons cherché à mettre en lumière dans le chapitre précédent, justifie cette filiation. Ainsi, il ne nous semble pas interdit de voir en Williamson le continuateur d'une grande partie du programme de recherche établi par Arrow [1974c] dans son ouvrage sur les organisations.

Whitaker note à propos de cet ouvrage : "*le livre lui-même est très allusif et elliptique,*

de plus, il semble n'avoir pas reçu l'attention qu'il méritait et n'a pas donné lieu à de nouveaux développements – ni par Arrow lui-même, ni par d'autres auteurs” (Whitaker, 1987, p.565). Cette remarque n'est qu'en partie pertinente. Le livre d'Arrow est effectivement avant tout un livre programme, dans lequel l'auteur dresse une série de chantiers théoriques liés à l'organisation, sans les développer entièrement. Il est également exact qu'Arrow n'a jamais tenté de développer ses propres “intuitions”. Cela l'aurait en effet obligé à sortir du cadre formel de la théorie standard, ce qu'il s'est toujours refusé à faire. Ce projet théorique n'est cependant pas pour autant resté lettre morte. Plusieurs courants de recherche se réclament des analyses d'Arrow en matière d'information, d'incertitude et d'organisation.

C'est par exemple le cas de la théorie de l'agence, qui s'appuie explicitement sur les concepts de hasard moral et de sélection adverse. Cependant la théorie de l'agence s'est développée dans une direction théorique visant à accentuer les capacités rationnelles des individus, de manière à prévoir les contrats optimaux permettant de dissuader toute déviance *ex ante*. Or, les réflexions théoriques d'Arrow ne visent précisément pas à pallier les défaillances du marché en augmentant les exigences en matière de rationalité individuelle. Au contraire, Arrow souligne, plus de dix ans après son ouvrage que : “*la rationalité n'est pas l'attribut d'un individu seul, bien qu'on la présente généralement ainsi. Elle prend non seulement toute sa force, mais aussi toute sa signification, lorsqu'elle est rapportée au contexte social dans lequel elle est encadrée. L'hypothèse de rationalité est ainsi plus plausible dans certaines conditions idéales. Quand ces conditions ne sont plus réunies, elle devient restrictive et possiblement contradictoire. Elle implique certainement une capacité à traiter l'information et à calculer qui est bien au-delà du faisable.*” (Arrow, 1987, p.69). Il précise un peu plus loin, citant Chamberlin : “*Il faut que la concurrence soit non seulement pure mais aussi parfaite pour que les hypothèses de rationalité aient toute leur portée*”. Cette réflexion souligne bien à quel point, chez Arrow, les problèmes rencontrés par la régulation marchande sont collectifs et ne peuvent être surmontés par une augmentation des capacités rationnelles individuelles³. C'est bien à l'étude des *structures collectives* permettant de pallier les défaillances du marché que convie sa réflexion critique.

Pour cette raison, le véritable héritier de ce programme de recherche nous semble être Williamson. Comme les auteurs précédents, Williamson emprunte à Arrow ses catégories de *hasard moral* et de *sélection adverse* qu'il réunit sous un concept unique : l'opportunisme (*ex post* et *ex ante*). Il emprunte également à Arrow sa définition des

³ Arrow note d'ailleurs dans un bilan critique de la théorie de l'agence : “*on ne trouve pas de relations aussi complexes dans la réalité.*” (Arrow, 1985, p.48)

⁴ Arrow note d'ailleurs dans un bilan critique de la théorie de l'agence : “*on ne trouve pas de relations aussi complexes dans la réalité.*” (Arrow, 1985, p.48)

coûts de transaction, comme “*coûts de fonctionnement du système économique*” (Arrow, 1969, p. 48), qui découlent de l’opportunisme des agents. Mais Williamson, à la différence des théoriciens néoclassiques plus “standards”, met également l’accent sur les limites cognitives des individus, dans une optique très proche de celle d’Arrow. Ce faisant, Williamson saisit l’esprit des remarques d’Arrow : les failles du marché ne peuvent être comblées par un surplus de rationalité, elles requièrent des règles organisationnelles qui les pallient. L’analyse critique d’Arrow constitue donc une “ouverture” dans la rhétorique marchande néoclassique, dans laquelle Williamson s’engouffre pour réintroduire des éléments théoriques institutionnalistes.

Toutefois, cette filiation ne va pas sans quelques écarts. Williamson sort du cadre néoclassique, en dehors duquel Arrow ne s’aventure jamais véritablement, même s’il en perçoit les limites. Le constat des limites d’un cadre d’analyse est en effet un projet distinct de la construction d’un autre cadre d’analyse visant à l’élargir. Williamson se révèle donc être un interprète très radical des réflexions d’Arrow, dépassant sans doute les visées de ce dernier.

Cette radicalité se traduit par une nette différence concernant la place que les deux auteurs font jouer à la dimension éthique ou normée du comportement. Arrow insiste à diverses reprises sur l’efficacité des bornes sociales mises au comportement intéressé, et semble parfois suggérer que la morale peut, à elle seule, sauver la régulation marchande. Tel n’est pas du tout le cas chez Williamson qui accentue à la fois les limites cognitives et le manque de fiabilité des agents, et n’envisage pas que l’éthique puisse avoir une quelconque efficacité. Son objectif est de proposer une théorie du fonctionnement des organisations qui puisse précisément s’affranchir des références à la confiance, aux codes éthiques, au respect des normes. Le secours ne vient pas, chez Williamson, d’un “aspect oublié” de la nature humaine mais de la mise au point de structures construites de façon intentionnelle pour limiter les choix possibles de l’agent et lui fournir l’information nécessaire pour guider ce choix. De manière beaucoup plus nette qu’Arrow, Williamson assume la nécessité de construire des organisations répondant aux échecs du marché. Cette radicalisation par rapport au discours d’Arrow a deux conséquences importantes : elle hisse la catégorie “organisation” au même niveau que la catégorie “marché”, et elle implique de la part de l’économiste de concevoir des structures collectives supplantant les agents. Or, une telle attitude conduit à adopter une position constructiviste en rupture avec le projet pseudo-naturaliste de l’économie néoclassique.

Il n’en demeure pas moins que cette filiation entre Arrow et Williamson est essentielle. Elle constitue en effet le ciment que Williamson utilise pour incorporer des propos institutionnalistes dans un cadre théorique qui reste proche de l’univers néoclassique.

L'analyse d'Arrow donne à Williamson un moyen d'articuler l'idée hétérodoxe d'organisation avec la représentation orthodoxe de l'individu maximisateur.

La synthèse de Williamson s'appuie donc sur trois piliers : l'approche de Coase, qui lui donne avec le concept de "coût de transaction" une clef permettant d'opposer marché et organisation ; l'approche des institutionnalistes, qui lui apportent une représentation théorique des organisations ; l'approche d'Arrow, qui lui permet de faire le lien entre institutionnalisme et théorie néoclassique.

2- “Contractual man” : l'individu économique selon Williamson.

“Les principales différences entre l'approche développée dans cet ouvrage et la littérature qui le précède sont les suivantes :

- Je porte une bien plus grande attention que jusqu'ici aux ramifications de l'hypothèse de rationalité limitée.

- J'introduis expressément la notion d'opportunisme et je m'intéresse à la manière dont le comportement opportuniste peut être influencé par l'organisation économique.

- Je mets l'accent sur le fait que ce ne sont pas l'incertitude et/ou l'existence d'un petit nombre d'échangistes qui occasionnent les défaillances du marché, mais c'est l'association de ces deux facteurs avec, d'une part, l'existence d'une rationalité limitée, et, d'autre part, l'existence d'opportunisme, qui perturbe l'échange marchand.”
(Williamson (1975, p.7)

Dans son second ouvrage Williamson consacre un chapitre au *contractual man*, pendant de l'*homo-œconomicus* dans l'économie des coûts de transaction. Au-delà de ce néologisme, qui révèle toute l'ambition du projet de Williamson, ce chapitre souligne le caractère central de la représentation de l'être humain qui sous-tend la théorie des coûts de transaction. En décrivant l'individu williamsonien, on est donc conduit à décrire les raisons pour lesquelles il existe des coûts de transaction que seule une structure organisationnelle peut minimiser.

Deux caractéristiques permettent de distinguer le *contractual man* de l'*homo œconomicus* classique. Sa rationalité est limitée et il est opportuniste. Cependant, une étude attentive de ces deux caractéristiques permet de constater qu'elles sont directement liées au contexte de l'action marqué par la *spécificité des actifs*. *Contractual man* est identique, au fond, à l'individu standard, mais il agit dans un contexte différent, et ce contexte fait apparaître des aspects cachés de l'hypothèse de rationalité standard.

Cependant, cela ne signifie sûrement pas que dans la théorie de Williamson, le contexte “technologique” prime sur l'aspect “humain” des transactions. En effet, la spécificité des actifs est directement liée au mode de savoir et d'apprentissage des agents. Cette caractéristique cognitive inscrit les agents dans une temporalité, irréversible, qui n'est plus celle du modèle Arrow-Debreu. C'est ce changement de temporalité qui fait apparaître certains aspects problématiques du comportement rationnel justifiant l'existence des organisations.

21- Un individu à rationalité limitée ?

“La rationalité limitée (bounded) est l’hypothèse cognitive sur laquelle s’appuie l’économie des coûts de transaction. C’est une forme semi-forte de rationalité dans laquelle les acteurs économiques sont supposés être «intentionnellement rationnels mais seulement de façon limitée» (Simon, 1961, p.xxiv)” (O.E. Williamson, 1984, p.197).

Williamson affirme dès son premier ouvrage, et dans presque l’ensemble des articles qu’il consacre à défendre ou appliquer le projet théorique de l’économie des coûts de transaction, que l’hypothèse de rationalité limitée, empruntée à Simon, est absolument centrale dans son analyse. Cette référence à Simon est cependant problématique. En effet, l’idée de rationalité limitée peut recouvrir deux sens bien différents. Soit elle désigne une forme de rationalité différente de la rationalité standard, le terme adéquat est alors celui de rationalité procédurale. C’est la position de Simon. Soit c’est une expression mettant l’accent sur le fait que certaines limites cognitives des agents apparaissent lorsque l’environnement est complexe. Ils maximisent donc toujours, mais, en quelque sorte, “moins bien”. C’est la position plus “orthodoxe” défendue par Stigler ou Arrow. Or, Williamson, bien qu’il se réfère à Simon, semble en réalité plus proche du second terme de cette alternative.

211- Une rationalité maximisatrice mais limitée ?

Williamson [1975, p.69 ; 1984, p.197] insiste sur l’importance des *deux* éléments de la “définition” qu’il emprunte à Simon. Tout d’abord, l’individu est “*intentionnellement rationnel*”, ce qui signe, selon Williamson, l’appartenance de *contractual man* à la famille plus vaste de l’*homo-œconomicus*. Il s’agit donc bien d’un mode de rationalité classiquement instrumental, calculatoire, et même potentiellement maximisateur. Mais l’individu n’est rationnel que de manière “*limitée*”. Or, la nature de cette limite n’est pas clairement explicitée par Williamson. On trouve cependant, dans son premier ouvrage, quelques indications à ce sujet : “*Les limites de la rationalité ne sont intéressantes, bien sûr, que dans la mesure où elles sont atteintes, c’est-à-dire dans des conditions d’incertitude ou de complexité. La rationalité limitée est en relation avec les conditions de l’environnement à l’origine des problèmes économiques*” (Williamson, 1975, p.22). L’origine de ces limites se situe donc dans l’environnement de l’individu. L’individu n’est pas moins rationnel, ou rationnel de façon différente, mais il est confronté à une

situation nouvelle dans laquelle son environnement n'est pas simple et transparent. Ces limites apparaissent alors parce que l'individu ne peut pas "tout savoir". Autrement dit, il n'a pas une capacité infinie de recherche, de traitement et de stockage de l'information.

Cette conception de la rationalité limitée fait l'objet de nombreuses critiques. D'une manière générale, adversaires et partisans de l'hypothèse standard de rationalité se retrouvent derrière l'idée selon laquelle il ne convient pas ici de parler de rationalité limitée, mais de connaissance ou de capacité de communication limitée (Hart [1990], Langlois [1990], Schlicht [1990], Simon [1991]). Alchian et Woodward [1988, p.66] résumant ainsi la proposition de Williamson : "*par rationalité limitée, Williamson entend indiquer que les individus ont des informations limitées, et une capacité limitée à la traiter*".

En fait, l'utilisation même du concept de "limite" pose problème : dans la théorie standard, la rationalité porte sur des critères formels de cohérence interne des choix individuels. Or, il ne saurait y avoir de "limite" en matière de cohérence logique. On ne peut produire un jugement qui serait "transitif mais seulement de façon limité". S'il y a une limite, c'est donc nécessairement en dehors de ce qui fait qu'un individu est rationnel au sens standard : par exemple, dans la quantité d'information qu'il peut amasser. L'idée de rationalité limitée est donc ambiguë, parce qu'elle peut désigner soit une autre forme de rationalité, soit une limite exogène laissant intacte l'hypothèse standard de rationalité. Cette distinction fait précisément l'objet d'un débat entre Arrow et Simon, deux auteurs dont s'inspire Williamson.

212- Arrow-Stigler ou Simon ?

Pour faire apparaître les limites, évidentes, que chaque individu rencontre en matière de traitement de l'information deux stratégies théoriques sont donc possibles : la première consiste à modifier la conception de la rationalité en lui donnant un contenu pratique, la seconde consiste à représenter l'information sous la forme d'un "coût", ce qui permet de l'introduire dans le programme de maximisation de l'agent. Dans le premier cas, on change de concept de rationalité, dans le second on conserve une hypothèse standard et on met l'accent sur le contexte informationnel de l'agent. Cette seconde stratégie est celle d'Arrow. Elle s'oppose à la première, celle de Simon.

Arrow admet l'existence de limites cognitives chez l'individu et dénonce les théoriciens oubliant ces limites (Arrow [1987]), mais il n'adhère pas pour autant à l'idée de rationalité limitée proposée par Simon. Il souligne ainsi : "*Je ne sais pas quoi faire avec cette théorie*" (Arrow in Feiwel [1987b], p.231). Arrow vise la même réalité empirique que Simon à travers l'idée que l'information est coûteuse. Mais il reste attaché à l'idée

d'optimisation comme catégorie générale de la rationalité. Il considère en conséquence que la théorie simonienne, dont il considère qu'elle se "*désintéresse du coût de l'information*", est "*partielle*" et donc inadéquate. L'attitude d'Arrow est en fait conforme à la théorie du "search" que propose Stigler [1968]. Cette théorie vise en effet à intégrer le problème de l'information en lui donnant un coût, et en faisant entrer ce problème dans la détermination du programme de maximisation de l'individu à partir d'une analyse "coût/avantage" de la recherche d'information.

C'est précisément pour éviter que son concept de "*rationalité limitée*" puisse être ainsi récupéré et maintenu "*dans le domaine de la rationalité calculable*" (Simon, 1978, p.453), que Simon a progressivement changé de terminologie. Ainsi, à partir de la seconde moitié des années soixante-dix, il abandonne le terme "*rationalité limitée*" et lui substitue le concept de rationalité "*procédurale*". Il entend ainsi préciser qu'il se situe dans le premier des deux termes de l'alternative exposée ci-dessus. La rationalité procédurale n'est pas une forme de rationalité maximisatrice limitée par l'existence d'un environnement complexe. C'est une autre forme de rationalité⁵.

A la lumière de cette évolution, il semble tout-à-fait significatif que Williamson s'en tienne à la première terminologie de Simon sans jamais faire référence à l'idée de rationalité *procédurale*. Williamson invoque en effet constamment l'idée d'un individu "*rationnel mais seulement de façon limitée*". Or cette définition appartient à la *première* période de l'œuvre de Simon. Elle se situe donc à une date à laquelle Simon n'avait pas encore eu à démarquer sa théorie des développements plus orthodoxes consacrés à l'information. L'utilisation de cette définition aplanit donc les différences, pourtant profondes, entre ces deux courants. La volonté de faire disparaître ce clivage apparaît encore plus nettement lorsque Williamson [1987b] cherche à établir une filiation entre ses travaux et ceux d'Arrow. Il souligne ainsi : "*Arrow est tout-à-fait conscient que les agents humains ont une capacité cognitive strictement limitée*" (Williamson, 1987b, p. 592). Et il en déduit que sa propre conception de la rationalité individuelle n'est pas différente de celle d'Arrow. Entre Arrow-Stigler et Simon, Williamson se refuse donc à trancher clairement.

Simon s'en charge pour lui. Dans un article récent, il récuse en effet la conception "*williamsonnienne*" de la rationalité limitée : "*On a pu constater que ces théoriciens introduisent une sorte de comportement à rationalité limitée, mais le caractère exogène des limites de la rationalité leur permet de maintenir la théorie dans le domaine magique de la maximisation de l'utilité et du profit*" (Simon, 1991, p.26). De fait, il semble bien que la rationalité limitée dont parle Williamson ne soit pas la rationalité simonienne. Williamson

⁵ Nous reviendrons plus en détail sur cette théorie et sur la controverse Stigler/Simon, dans le chapitre 6.

prend acte des limites cognitives et informationnelles des individus, mais la maximisation reste la norme d'un comportement rationnel⁶.

La référence de Williamson à Simon est donc problématique. Sa conception de la rationalité individuelle reste en effet assez classique, même s'il met l'accent sur les limites cognitives des individus de façon plus appuyée que ne le fait le théorie micro-économique standard. Une telle conception tient sans doute plus d'Arrow que de Simon. L'hypothèse de rationalité qui sous-tend l'économie des coûts de transaction reste donc instrumentale, calculatoire et optimisatrice, même si elle est limitée par l'existence d'un environnement complexe qui fait apparaître les limites cognitives des agents.

22- Un individu opportuniste

“Le principe de l’opportunisme étend l’hypothèse traditionnelle selon laquelle les agents économiques sont guidés par des considérations d’intérêt personnel de manière à prendre en compte les comportements stratégiques. Avec cette conception, on considère donc que la recherche de son propre intérêt inclut la tromperie”. (Williamson, 1975, p.26).

La définition de Williamson souffre d'un manque de précision : s'agit-il de distinguer *contractual man* de l'*homo œconomicus*, ou de tirer de nouvelles conséquences d'un aspect déjà connu de l'*homo œconomicus* ? Comme en ce qui concerne la rationalité, la position de Williamson est ambiguë. Son approche constitue en effet une avancée dans la prise de conscience du caractère potentiellement dévastateur de l'action d'individus égoïstes, lorsque ceux-ci sont amenés à se coordonner directement, sans la médiation d'un système les isolant les uns des autres. Mais, elle ne constitue en rien une innovation dans l'anthropologie de l'agent économique néoclassique. L'opportunisme de *contractual man* n'est qu'une conséquence de la rationalité de l'*homo œconomicus*.

221- Dans la lignée d'Arrow

Arrow, à travers son analyse de l'échec des contrats contingents, avait initié une conception du comportement marquée par l'existence de tromperie, avant et après le

⁶ Cette conception “limitative” de l'idée de rationalité limitée nous semble être proche de celle que défend Selten [1990], ou encore Conlisk [1996].

contrat. Avec le concept d'opportunisme, Williamson propose de réunir sous un même concept les deux problèmes que sont la sélection adverse et le hasard moral.

Dans l'analyse standard, la catégorie importante est l'hypothèse de rationalité. Le calcul des offres et demandes permet de spécifier entièrement le comportement de l'acteur. L'individu standard ne se soucie que de son propre intérêt, et, de ce point de vue, peut-être considéré comme un égoïste. Toutefois, dans le modèle Arrow-Debreu, ce qualificatif a très peu de sens, dans la mesure où les individus s'ignorent tout-à-fait, n'ayant affaire aux autres que sous la forme agrégée d'une demande ou d'une offre excédentaire se traduisant par des mouvements de prix. Ce n'est qu'à compter du moment où les individus se rencontrent que leur rationalité débouche sur un comportement égoïste, posant le problème de la confiance mutuelle. C'est ce que note Arrow lorsqu'il étudie le principe théorique des marchés contingents, puis son impossibilité pratique.

Williamson [1984, p.198-199 ; 1985 p.64-67] affirme qu'il va plus loin : cette fois explique-t-il, l'individu va chercher à transgresser les règles communes, à tricher. Il accuse ainsi l'analyse néoclassique standard de supposer que les individus ne font preuve que d'une forme "semi-forte" d'opportunisme, dans la mesure où ils acceptent le fonctionnement du système des prix, les règles de la propriété privée, les dotations initiales, sans chercher à voler, tromper, trahir. C'est la raison pour laquelle, selon lui, cette analyse est aveugle en ce qui concerne les mécanismes institutionnels que sont les organisations, et dont l'objet est précisément de décourager l'opportunisme. Si les individus étaient effectivement loyaux envers les règles collectives, l'incertitude et la rationalité limitée ne poseraient pas problème. Comme le remarque Williamson [1985, p.66], l'existence d'une "clause générale" prévoyant les dispositions à prendre en fonction des états du monde réalisés, éliminerait tout problème de coordination si chacun indiquait fidèlement l'état du monde réalisé et respectait ses engagements. C'est parce que les agents, *a priori*, mentent quand ils y ont intérêt que l'existence d'organisations les contrôlant est nécessaire.

222- Une généralisation

Il n'y a toutefois là rien de profondément neuf au regard de la théorie standard. Williamson généralise les remarques d'Arrow plus qu'il n'innove. En effet, les remarques de ce dernier à propos de l'impossibilité des marchés contingents mettent précisément en évidence l'impossibilité d'une "clause générale", comme le fait Williamson.

La nouveauté de Williamson se situe donc à un autre niveau : il n'attribue pas l'opportunisme à des failles ponctuelles du système marchand mais en fait un problème

systématique que le système de marché ne peut pas résoudre. Ce faisant, Williamson généralise la critique d'Arrow. Sa réflexion ne touche pas le système marchand à la marge mais en plein cœur. Cette généralisation est liée à la démarche de Williamson qui remet en cause l'ensemble du cadre abstrait Arrow-Debreu, et cherche à proposer une conception alternative, empiriquement juste, de la manière dont se déroulent les échanges. L'univers des individus n'est donc plus supposé être d'emblée logiquement organisé : il est à organiser. C'est en cela que la théorie de Williamson est institutionnaliste.

Nos remarques visant à limiter le caractère novateur de l'individu "williamsonnien" ne signifient donc absolument pas que les précisions de Williamson sont inutiles. Bien au contraire : cet auteur tire au clair les conséquences d'hypothèses théoriques trop souvent utilisées de façon partielle. On attend ainsi des individus rationnels qu'ils maximisent leur propre bien-être sans pour autant tricher sur la qualité de la marchandise échangée (ce qu'a noté Akerlof [1970]), ou encore sans tricher sur la nature des états du monde qui se réalisent (comme le note Arrow [1974b]). A l'inverse, Williamson fait clairement apparaître ce problème. Il lève le voile que constitue le cadre logique Arrow-Debreu. Il montre que si les individus sont rationnels et égoïstes, comme le prévoit la théorie, il n'y a aucune raison de supposer qu'ils respectent les règles du jeu. Comme nous le verrons plus loin, c'est dans le même esprit que Williamson rejette les solutions fondées sur l'existence de normes morales, de relations de confiance, de prescriptions éthiques. En rappelant que les individus sont opportunistes, Williamson tire simplement les conséquences logiques de l'hypothèse de rationalité instrumentale lorsqu'elle s'applique à un univers dont les rapports humains ne sont pas exclus. Il met en évidence, de façon appuyée, un aspect de l'individu économique qui n'était que latent tant que son comportement était cantonné à un marché anonyme de type Arrow-Debreu. Il ne change donc pas d'individu, il change de contexte. Ce changement de contexte est induit par la mise en évidence du problème de la spécificité des actifs.

23- La spécificité des actifs

“Les trois dimensions sur lesquelles l'économie des coûts de transaction s'appuie pour caractériser les transactions sont : (1) leur fréquence, (2) le degré et la forme de l'incertitude à laquelle elles sont sujettes, (3) le niveau d'investissement en actifs spécifiques qu'elles requièrent. Toutes sont importantes, mais la plupart des déductions logiques de l'économie des coûts de transaction s'appuient principalement sur cette dernière. (...) L'implication la plus importante de l'existence d'investissements en actifs

spécifiques est la suivante : alors que l'identité des contractants est sans importance dans le cadre néoclassique, l'identification des parties prenantes d'une transaction nécessitant un investissement dans des actifs spécifiques est essentielle.” (Williamson, 1991, p.94).

La spécificité de l'analyse de Williamson tient au contexte de la relation d'échange qu'il étudie. Or, l'élément dominant de ce contexte est le caractère spécifique des actifs engagés dans la transaction. Cette spécificité inscrit en effet la transaction dans une perspective temporelle irréversible. L'existence d'incertitude et le degré de fréquence d'une même transaction sont des conséquences de cette irréversibilité.

Cependant, on aurait tort de penser que Williamson propose ce faisant une analyse marquée par le primat de la technologie sur les aspects humains de l'échange. Si les actifs sont spécifiques, c'est en effet en raison du mode de connaissance des agents. En définitive donc, et c'est là la véritable innovation anthropologique de Williamson, le mode de connaissance des agents inscrit la relation d'échange dans une perspective temporelle qui n'est plus formelle mais pratique ou historique.

231- La notion d'actif spécifique

“La spécificité des actifs fait référence au degré selon lequel un actif peut être redéployé pour un autre usage et par un autre utilisateur sans sacrifier sa valeur productive.” (Williamson, 1990, p.12).

Williamson reprend à Commons l'idée selon laquelle la transaction est l'unité de base de l'analyse, et cherche à préciser cette idée en distinguant les différentes caractéristiques des transactions. Il en distingue trois : la spécificité des actifs, l'incertitude, la fréquence. Ces trois caractéristiques sont cependant d'importance inégale. La première d'entre elles est de loin la plus importante, dans la mesure où elle détermine largement le sens des deux autres. En effet, un actif spécifique est un actif non réutilisable et non revendable. L'achat ou la production d'actifs spécifiques ne peut donc se concevoir que si la transaction qui nécessite cet investissement doit se répéter plusieurs fois, instituant ainsi une relation de long terme. Ce type d'investissement requiert donc une certaine confiance en l'avenir, c'est-à-dire une maîtrise satisfaisante de l'incertitude. Si, au contraire, la transaction est standard et ne requiert aucun investissement spécifique, elle peut être sans lendemain : peu importe alors le degré d'incertitude et de fréquence future de cette transaction.

La notion d'actif spécifique est donc essentielle puisqu'elle inscrit la relation d'échange dans une temporalité radicalement nouvelle par rapport à l'univers néoclassique. En effet, l'impossibilité de redéployer les actifs spécifiques fait entrer les contractants dans une forme de temps irréversible. Or l'apparition d'une temporalité irréversible entre en

contradiction avec le comportement instrumentalement rationnel des agents. Le caractère limité de leurs connaissances ainsi que leur opportunisme rendent en effet impossible tout engagement définitif. Ainsi, la spécificité des actifs fait basculer les agents dans un univers qui ne peut plus être marchand. Le mode de transaction marchand, fondé sur l'anonymat et sur la parfaite réversibilité du temps, échoue à réguler de manière satisfaisante des relations se déroulant nécessairement dans le long terme et selon un processus irréversible. Il faut autre chose que le marché pour sécuriser ce type de relation.

La nécessité de l'organisation naît donc du problème de la spécificité des actifs. Williamson le souligne clairement : "*le principal facteur auquel j'attribue la décision d'intégration est une condition de spécificité des actifs*" (1985, p.86). Cette spécificité est l'élément constitutif d'un univers dans lequel les caractéristiques de l'agent néoclassique, mises en lumière par Williamson, deviennent problématiques. Dans cet univers marqué par l'irréversibilité, la régulation marchande, fondée sur l'anonymat et la fugacité de la relation, est mise en échec. En effet, lorsque la transaction est parfaitement réversible, quelque soit la malhonnêteté ou la courte vue des contractants, la régulation marchande est satisfaisante : tout ce qui est fait peut en effet être défait sans coût. Mais lorsque l'engagement doit être conçu comme définitif, il est nécessaire que les parties se connaissent. Les contractants ne sont plus interchangeables, ils se lient pour le long terme. Il faut donc qu'ils se fassent confiance. Or la confiance ne naît pas spontanément entre des individus opportunistes et cognitivement limités. Ainsi, c'est le cumul des *trois* éléments (opportunisme, limites cognitives et spécificité des actifs) qui met en échec le marché. Les agents sont pris dans une séquence temporelle dans laquelle, pour agir, ils auraient besoin de connaître avec certitude le futur, ce que leurs limites cognitives et le manque de fiabilité de leurs partenaires ne permettent pas.

232- La théorie de la transformation fondamentale

L'apport de Williamson ne se limite pas à constater que *certaines* transactions requièrent des investissements spécifiques. Il démontre en effet que la *plupart* des transactions, même si elles sont initialement standard, deviennent progressivement spécifiques à mesure que les deux contractants développent une relation mutuelle marquée par l'existence de routines et de services spécifiques non standardisables fondés sur l'existence d'une compréhension tacite. C'est là le cœur de la théorie de "*la transformation fondamentale*" (Williamson, 1985, p.61-63). Cette théorie constitue une analyse originale de la manière dont une transaction "standard" tend à basculer dans l'irréversibilité. Williamson assoit cette démonstration sur une analyse du mode d'apprentissage et de communication des agents.

Il s'appuie pour se faire sur la théorie de la connaissance et du langage de Michael Polanyi [1962]. Polanyi s'intéresse à la connaissance "personnelle", qui distingue par exemple les techniques spécifiques de l'artisan Stradivarius des méthodes systématiques théorisées par l'ingénieur Taylor : certains "savoir-faire", certaines connaissances ne font pas l'objet d'une explicitation suffisante pour être standardisés⁷. Ils ne peuvent pas être séparés de l'agent et constituent donc des éléments cognitifs difficiles à communiquer. Les transactions mettant en jeu ce type de compétence se font donc dans un mode "idiosyncrasique", adapté aux deux contractants, qui se comprennent selon des règles et des mots qui leur sont propres. On est ici à l'opposé des transactions anonymes marchandes. Il n'existe pas de définition standard du "bien" échangé, et le prix ne peut plus être l'unique lien entre les contractants⁸. Les agents développent donc ce que Williamson (1979, p.233), s'appuyant sur Polanyi, qualifie de langage "spécialisé". Ils mettent en place des codes facilitant la compréhension mutuelle et permettant de désigner des besoins et services tout-à-fait spécifiques. Or la mise en place de ces codes est indissociable de la durée et de la fréquence de la relation d'échange. Les transactions mettant en jeu des connaissances idiosyncrasiques requièrent en effet l'existence d'une relation durable permettant la mise en place d'une bonne connaissance mutuelle. C'est uniquement dans le long terme que l'apparition d'une forme de compréhension tacite marquée par l'existence de codes communs, de routines, d'habitudes spécifiques permet d'améliorer la relation et de la rendre plus efficace : "*des codes spécialisés apparaissent dans un contexte de transactions récurrentes, et cela est source d'économie*". Autrement dit, plus une relation d'échange se déroule dans le long terme, plus elle permet la mise en place de "codes" facilitant cette relation.

Cette apparition d'une forme de langage commun est au fondement du *processus* de la *transformation fondamentale*. Ainsi, lorsqu'un contact est noué entre deux contractants, même dans un univers concurrentiel normal, la relation tend à devenir spécifique, et cesse donc d'être concurrentielle : "*Bien que l'existence d'un grand nombre de contractants est fréquente lors des premiers appels d'offre, les transactions de type idiosyncrasiques (qui impliquent des investissements spécifiques) sont de celles pour lesquelles la relation entre acheteur et offreur se transforme progressivement en monopole bilatéral*" (Williamson, 1979, p.241). Une relation d'échange s'inscrivant dans un horizon temporel suffisamment lointain est donc souvent plus efficace qu'une succession de transactions de court terme, marquées par un changement de contractants. A compter du moment où une

⁷ "*Même dans l'industrie moderne, le savoir "indéfinissable" est une part de la technologie*" (Polanyi [1962], cité par Williamson, 1979, p.242).

⁸ On retrouve ici certains aspects caractéristiques des "relations de service" formalisées par Gadrey [1998]. Ce qui confirme, comme le souligne l'auteur, que l'existence d'une dimension "servicielle" de la relation d'échange n'est pas une spécificité du secteur tertiaire.

transaction s'installe dans la durée, le service rendu par l'offreur devient plus précis et mieux adapté aux besoins du demandeur. Dès lors cette relation tend à être protégée de la concurrence puisque le service rendu devient "unique", "spécifique". Le mode de coordination des agents, le processus par lequel s'établit une compréhension mutuelle efficace extirpent inexorablement la relation d'échange de la sphère marchande. Le mode de connaissance et de communication des agents est donc au fondement d'un processus qui tend à l'internalisation des transactions dans un univers organisationnel hors-marché.

L'anthropologie williamsonnienne se clôt ainsi avec la mise en évidence du *contexte* dans lequel les aspects classiques du comportement individuel rationnel, opportunisme et caractère limité des compétences cognitives, deviennent problématiques. La spécificité des actifs rend ces caractéristiques problématiques parce qu'elle rend nécessaire l'existence d'une relation de long terme. Dès lors, pour que la relation soit efficace, il faut combler les lacunes cognitives des individus et dissuader leur opportunisme. Ce *contexte* n'est cependant pas un élément exogène par rapport au comportement des individus : au contraire, la spécificité naît du mode de connaissance des agents. C'est finalement à ce stade que se situe la réelle innovation "anthropologique" de Williamson.

24- La nécessité des règles

Williamson [1990, p.13] affirme que sa théorie "*est plus micro-analytique que celle de l'orthodoxie*" et s'appuie sur une conception critique de l'individu économique. L'économie des coûts de transaction serait donc en quelque sorte fondée sur une nouvelle anthropologie de l'individu économique, dont les diverses caractéristiques seraient mises à jour et dont les modes de coordination seraient étudiés en tant que tels, et non à partir d'une position globale d'équilibre comme dans le cas du marché Arrow-Debreu. Il est vrai que les coûts de transaction ont essentiellement une origine humaine. Pour autant, l'individu williamsonnien n'est pas différent de l'*homo œconomicus* néoclassique. Il est simplement situé dans un univers différent, dans lequel ses caractéristiques font problème.

Sa rationalité n'est pas drastiquement plus limitée que celle de l'individu marchand. L'individu williamsonnien calcule et optimise de façon intertemporelle. Mais le système auquel il participe ne lui délivre plus toute l'information dont il a besoin. C'est pourquoi la question de ses capacités cognitives se trouve posée. Le fait qu'elles soient limitées est une évidence que personne ne songerait à nier, à moins de considérer que l'*homo*

œconomicus est Dieu.

Son opportunisme n'est pas plus développé que dans le modèle de base, où il était de toute façon maximal. Mais, il peut s'exercer, ce qui n'était pas le cas dans l'univers Arrow-Debreu. Là encore, c'est donc le contexte qui révèle une caractéristique de l'agent.

L'élément dominant de ce contexte est la spécificité des actifs, parce qu'elle inscrit la relation dans une temporalité de nature pratique. Cette temporalité est en effet caractérisée par le caractère irréversible des investissements spécifiques. Or, la nécessité de cet investissement est directement liée au mode de connaissance des agents, fondé sur une "expérience" non standardisable. C'est là un point essentiel : Williamson met à jour le lien entre le mode de connaissance des agents, et les failles du système marchand.

En effet, l'opportunisme et les compétences cognitives limitées peuvent tout-à-fait être gérés par un système marchand donnant de l'information gratuitement et isolant les individus les uns des autres. Mais cette solution repose sur une absence de temps, ou, au moins, une forme de temps parfaitement réversible. Or, à compter du moment où des éléments de connaissance pratique sont en jeu, le temps de la relation n'est plus réversible. Williamson montre précisément que les transactions s'appuient sur des connaissances qui nécessitent des investissements spécifiques et qui inscrivent la relation dans une temporalité qui n'est plus formelle. C'est pour cette raison que les caractéristiques comportementales de l'individu font problème. Cette fois, en effet, il est nécessaire que les individus se "comprennent". Comme le note Williamson (1991, p.94) lorsqu'il mentionne la conséquence principale de la spécificité des actifs, la relation marchande ne peut plus être anonyme. Le marché est donc mis en échec : les prix ne suffisent plus aux individus pour se coordonner. Il est nécessaire qu'ils se parlent, et donc qu'ils disposent d'un langage commun. Or, l'efficacité du marché tient justement au fait qu'il rend tout langage inutile : lorsque le temps est réversible et que les individus sont tous formellement identiques, les prix peuvent remplacer le langage. Lorsque ces conditions ne sont pas réunies, le système des prix est insuffisant pour assurer la bonne coordination des agents, le langage devient alors nécessaire et le marché est mis en échec.

La nécessité des organisations naît donc en définitive de la nécessité d'un langage commun. Il est nécessaire de mettre en place des règles assurant une bonne compréhension mutuelle. Ces règles suppléent l'organisation marchande, mise en défaut par l'existence d'une forme de temps irréversible. Cette forme de temporalité est directement liée au mode de connaissance des agents. La connaissance pratique repose en effet pour partie sur l'acquisition d'une "expérience" qui n'est pas transmissible,

revendable ou redéployable. Ce mode de connaissance n'est donc pas valorisable par le biais du système de marché, il requiert, pour être efficace une relation durable fondée sur la confiance et l'intercompréhension. C'est pour cette raison qu'il entre en contradiction avec la rationalité instrumentale du *contractual man*, la rationalité instrumentale ne lui permettant pas de nouer une relation stable avec les autres.

En définitive, l'anthropologie williamsonnienne opère donc un lien essentiel entre le mode de connaissance des agents et la nécessité des règles organisationnelles palliant les défauts de leur type de rationalité. L'organisation résout, en quelque sorte, une contradiction interne à l'individu marchand⁹. Le besoin de règles émerge de l'incomplétude de la rationalité instrumentale.

⁹ Il est logique que l'étude des failles du marché mène finalement à la découverte d'une contradiction interne à l'individu marchand. On a vu en effet dans le premier chapitre que l'individu marchand n'était pas distinct du système lui-même. Williamson mène ici à son terme la critique initiée par Arrow.

3- La règle hiérarchique

“L'économie des coûts de transaction contribue à la mise en place d'un ordre satisfaisant et d'arrangements viables à travers l'étude de la gestion (Governance) des relations contractuelles. La mise en ordre est réalisée en affectant les transactions à des structures de gestion (governance) efficaces en matière de réduction des coûts.” (Williamson, 1993a, p.44).

Williamson, à travers l'étude du “*contractual man*” et celle du contexte des transactions, fait apparaître les raisons pour lesquelles le marché ne peut pas gérer toutes les transactions. Le comportement instrumentalement rationnel des individus devient problématique dans un univers dont la temporalité n'est plus réversible. Ce comportement doit donc être régulé : les individus ont collectivement besoin de l'existence de règles non marchandes pour pouvoir se coordonner.

Williamson, en se saisissant de ce problème, lui donne un sens plus précis que l'on pourrait formuler ainsi : “le monde est à ordonner en fonction d'un critère unique d'efficacité collective”. Dans son étude de la question du choix des règles collectives, Williamson prend donc la place d'un régulateur surplombant le jeu des interactions entre acteurs, et cherchant à les ordonner de manière efficace, comme un urbaniste ordonnerait l'organisation routière d'une grande ville. Ainsi les règles organisationnelles qu'étudie Williamson n'émanent pas des acteurs eux-mêmes. Elles se situent à un niveau de rationalité qui leur échappe. Elles sont mêmes précisément conçues pour contrer les “défauts” des acteurs.

Cet impératif qui guide l'entreprise williamsonnienne se décline selon les deux types de problème à résoudre. Il faut d'une part donner une information suffisante à des agents cognitivement limités, et, d'autre part dissuader leur opportunisme naturel, ce qui implique non seulement de mettre en place des dispositifs incitatifs, mais plus encore de surveiller et de punir les agents déviants. Dans cette optique, la structure de régulation la plus efficace est celle qui rassemble les agents tout en guidant autoritairement leur conduite : l'organisation hiérarchique.

Reste à savoir qui est le plus à même de guider cette organisation. La réponse de Williamson est nette : le propriétaire doit décider seul, car lui seul confond vraiment ses intérêts et ceux, globaux, de l'organisation qu'il possède. Williamson justifie donc, au final, la forme de l'organisation capitaliste qui apparaît être la conséquence d'une recherche rationnelle de la meilleure efficacité possible.

31- Un seul critère de choix : l'efficacité.

“Arrow préfère étudier les organisations dans un esprit rationnel. (...) Mais je voudrais mettre l'accent sur un autre aspect de l'esprit rationnel : l'efficacité. Toute étude de l'organisation économique qui ne fait pas systématiquement appel à l'efficacité concevra certainement de manière erronée tout ou partie des raisons expliquant la variété des formes organisationnelles. Quiconque s'apprête à étudier sérieusement la variété des organisations est donc bien avisé d'être animé par, voir de s'appuyer en permanence sur, un esprit rationnel.” (Williamson, 1987, p.590)

Dans son étude du choix des règles adaptées aux caractéristiques comportementales des agents, Williamson utilise un critère unique : l'efficacité. Cet impératif souligne nettement quelle doit être la tâche de l'économiste : rationaliser le monde économique.

Cette position se distingue de la position néoclassique en ce qu'elle est plus explicitement constructiviste. Au concept néoclassique d'une rationalité attachée à l'individu, Williamson fait en effet correspondre l'esprit rationnel de celui qui dessine les *institutions du capitalisme*, c'est-à-dire les organisations. Le critère de choix entre les diverses règles possibles doit être celui de l'efficacité, de manière à ce que le système économique en lui-même, et non pas seulement les individus qui le composent, soit “rationnel”. L'analyse de Williamson cherche donc à concevoir les règles permettant de mettre en place cette “rationalité collective” que le système du marché, comme on l'a vu avec Arrow, ne permet pas d'atteindre. De telles règles existent déjà : il s'agit des organisations capitalistes, mais il convient, en théorie, d'en expliciter le rôle, afin d'éviter qu'une politique défendant le caractère naturel du marché ne les détruise¹⁰.

Cette position représente donc un passage, discret, à un discours de niveau supérieur à celui de l'analyse néoclassique à propos du marché. On a étudié dans le premier chapitre, le mode de justification qu'utilisent Arrow et Hahn [1971] : ils se réfèrent à l'existence d'une forme naturelle de coordination, le marché, qu'il convient de comprendre et, le cas échéant, de relativiser. Leur analyse du marché se déroule ensuite comme une étude des implications du caractère maximisateur des agents : elle ne quitte à aucun moment le statut de l'analyse descriptive du marché.

Williamson déploie un raisonnement différent, dans la mesure où il étudie les moyens institutionnels permettant de réguler ce même comportement d'agents maximisateurs,

¹⁰ Le sous-titre du premier ouvrage de Williamson (*analyse et implications concernant les lois antitrust*) indique clairement qu'un de ses objectifs essentiels est d'éviter que les lois antitrust ne détruisent la réponse capitaliste aux problèmes que pose la coordination économique.

lorsque le marché est en défaut. Il devient ainsi le prescripteur des “bonnes règles”. Ce faisant, il tranche le lien, ténu, qui rattache encore l'économie à la nature dans la théorie néoclassique. Le phénomène du marché peut en effet être regardé comme étant d'essence naturelle¹¹. Ce n'est pas le cas des “structures de gouvernance” qui sont là pour *protéger* la collectivité contre les aspects néfastes de la *nature* humaine. La nouveauté du discours de Williamson tient précisément à cette opposition entre les structures de gouvernance et certaines caractéristiques de l'agent économique. Au cœur de l'analyse de Williamson, il existe donc une distinction entre la rationalité collective, que vise la mise en place des organisations adéquates, et la rationalité individuelle, qui ne permet pas d'atteindre l'efficacité collective¹². Cette opposition est le propre d'un discours constructiviste visant à dessiner la machine institutionnelle qui permettrait de surmonter les contradictions de la nature humaine.

La recherche de l'efficacité nécessite donc, de la part de l'architecte et du gestionnaire des organisations économiques, une rationalité supérieure, globalisante, qui leur permette de percevoir l'intérêt collectif et de le défendre contre la multitude des intérêts particuliers qui le mine. C'est, implicitement, le rôle que se donne Williamson. L'efficacité des organisations dont il étudie et détermine le fonctionnement transcende la rationalité des individus qui les composent.

32- *Informer, inciter, surveiller*

“Le critère pour organiser les transactions commerciales est strictement instrumental : il s'agit d'économiser des coûts. Cela prend essentiellement deux formes : économiser dans le domaine de la production, et économiser dans le domaine des coûts de transaction. Nous mettons ici l'accent sur ce dernier objectif.” (Williamson, 1981, p.1547)

Pour Williamson, le rôle des organisations n'est donc pas essentiellement de produire efficacement. Elles visent aussi, et d'abord, à répondre aux problèmes spécifiques que pose la forme de rationalité des agents économiques. Les variables de ce problème sont

¹¹ C'est explicitement le cas dans les théories classiques de Smith, Ricardo, Malthus. Le rapport du modèle walrassien à la nature est plus complexe. Arrow et Hahn [1974] placent cependant la théorie de l'équilibre général dans la lignée de la théorie naturaliste smithienne. La théorie moderne du marché est construite autour de l'idée selon laquelle il existe de façon naturelle chez les individus une “*propension à trafiquer, troquer, échanger*” selon la célèbre expression de Smith.

¹² D'une certaine manière, l'on pourrait dire que Williamson tire toutes les conséquences du théorème d'impossibilité formulé par Arrow.

toujours les mêmes : les limites cognitives des individus et leur opportunisme. L'objectif de minimisation des coûts de transaction se décline en deux impératifs : "*L'impératif organisationnel qui émerge est celui-ci : organiser les transactions afin d'économiser sur la rationalité limitée tout en les garantissant simultanément contre les risques de l'opportunisme.*" (Williamson, 1985, p.32, voir aussi : Williamson, 1979, p.246)

Le premier des deux objectifs que doit remplir une "*structure de gouvernance*", selon l'expression de Williamson, rappelle très clairement l'analyse qu'Arrow propose des problèmes d'information. Williamson s'en inspire directement lorsqu'il met en avant la nécessité d'une mise en commun de l'information et de l'existence d'un langage commun permettant de faciliter la communication. L'organisation simplifie la tâche des individus en leur délivrant les informations nécessaires, et seulement celles-là. Williamson met cependant l'accent sur un aspect nouveau : le problème du manque d'objectivation des tâches des salariés, en raison de leur imparfaite maîtrise du langage (Williamson, 1975, p.24).

Le problème principal qui se pose au "législateur de l'efficace" est le caractère obscur, idiosyncrasique, incommunicable du travail. Pour contourner cette difficulté, qui rend complexe la prescription et la surveillance des tâches, il est nécessaire de mettre en place un code langagier permettant de retrouver une certaine objectivité dans la définition du rôle de chacun. Le langage organisationnel ainsi mis en place permet de simplifier les contacts entre agent, et donc "d'économiser de la rationalité limitée", mais aussi de les surveiller plus efficacement, et donc de réduire l'opportunisme.

Ce second objectif est nettement plus développé chez Williamson : "*Un problème chronique auquel est confrontée l'organisation économique est : comment contenir l'opportunisme ?*" (Williamson, 1975, p.98). Combattre l'opportunisme des agents implique de prendre deux séries de mesures : la première consiste à inciter les agents à adhérer aux objectifs de l'entreprise, c'est un champ connexe à celui de la théorie des incitations. Le second consiste à mettre en place des dispositifs de surveillance suffisamment efficaces pour dissuader l'opportunisme et le rendre coûteux. La spécificité de l'analyse de Williamson tient à sa volonté de ne pas réduire la gestion des conflits entre les contractants à une discussion préalable permettant *ex ante* de prévoir toutes les difficultés possibles. Le contrat est nécessairement incomplet et doit donc, selon cet auteur, faire l'objet d'une surveillance non seulement *ex ante* mais aussi *ex post* : l'organisation doit être capable de faire face aux défections imprévues et de les résoudre. C'est la raison pour laquelle l'économie des coûts de transaction met en avant l'importance des structures organisationnelles contre la représentation bilatérale des

modèles principal/agent. Le rôle de ces structures consiste non seulement à prévoir des dispositifs permettant de faciliter "l'arrangement privé" mais aussi, le cas échéant, à mesurer les fautes et responsabilités de chacun. C'est sur ce dernier point en particulier que l'organisation est plus efficace que le marché.

La structure de gouvernance a donc trois fonctions : informer, inciter et surveiller. Ces trois fonctions nécessitent deux éléments communs. Il est nécessaire que les agents se comprennent, c'est-à-dire parviennent à inscrire leur relation dans le long terme et dans un univers facilitant l'intercompréhension. C'est l'avantage de l'organisation interne sur le marché. Mais il est également nécessaire que cette compréhension serve les intérêts de l'entreprise, et non pas leurs intérêts propres. Il faut donc que la motivation et les informations données aux contractants par la structure de gouvernance servent ses objectifs. La courroie de transmission nécessaire entre ces objectifs généraux et les objectifs propres des agents est l'existence d'une relation hiérarchique. C'est la clef de voûte du modèle de Williamson.

33- La "démonstration" de la supériorité du mode hiérarchique

"Les modes les moins hiérarchiques, en terme à la fois de contractualisation et de prise de décision ont les plus mauvaises propriétés d'efficacité. Le groupe des pairs et la relation d'autorité reposent considérablement sur une hiérarchie de la prise de décision – ce qui contribue pour beaucoup à expliquer la performance de chacun. Il peut y avoir des types de hiérarchie plus ou moins souhaitables ; mais la hiérarchie en elle-même est inévitable à moins d'accepter des sacrifices d'efficacité." (Williamson, 1985, p.231).

Williamson [1975 et 1985] développe successivement deux justifications de la hiérarchie fondées sur une comparaison avec d'autres modes organisationnels, en terme d'efficacité comparée. La première prend la forme d'une fable historique, la seconde est de nature plus empirique.

331- La supériorité "historique" de la hiérarchie.

La première justification (Williamson, 1975, chap.3) s'appuie sur un "conte"

historique¹³. Partant du principe “qu’au départ il n’y avait que le marché”, Williamson étudie comment émerge une organisation communautaire, qu’il nomme “groupe des pairs”, puis comment celle-ci cède progressivement la place à l’organisation hiérarchique. Le schéma général est le suivant : pour économiser de la rationalité limitée, il est préférable de se regrouper afin de rentabiliser les investissements spécifiques. C’est la raison pour laquelle le groupe des pairs l’emporte sur le marché.

Mais, ce groupe des pairs souffre de deux maux. Tout d’abord, la collégialité des décisions est contre-productive car elle est mal adaptée au caractère cognitivement limité des individus. Il vaut mieux, en effet, que les plus “doués” prennent les décisions pour l’ensemble de la collectivité : “On servira l’efficacité en réservant le rôle de centralisation de l’information et de prise de décision à ceux ou celui des individus qui ont un pouvoir supérieur de traitement de l’information et d’exceptionnelles capacités oratoires dans le domaine de la décision. Il en résultera une sorte d’élite, que le processus de sélection placera dans une relation asymétrique vis-à-vis de tous les autres.” (Williamson, 1975, p.52).

Le second problème, beaucoup plus essentiel, est lié au mode de rémunération des membres de l’organisation : ils se partagent équitablement les revenus, indépendamment de l’effort de chacun. Une telle méthode résiste mal à l’opportunisme des agents, et le groupe des pairs est miné par le problème du “passager clandestin”. Williamson, reprenant sur ce point les réflexions d’Alchian et Demsetz [1972], en conclut qu’il est nécessaire qu’un membre supervise la contribution de chacun. C’est la naissance de la hiérarchie : “La hiérarchie simple peut faire tout ce que fait le groupe des pairs, et même davantage” (Williamson, 1975, p.54). Elle lui succède donc, logiquement, comme mode de gouvernance.

332- La supériorité empirique de la hiérarchie

Dans son second ouvrage, Williamson [1985, chap 9] utilise un mode de justification plus comparatif. Son étude vise à démontrer, contre les radicaux américains, qu’on peut expliquer le choix du mode hiérarchique sans faire appel à des relations de pouvoir et de lutte sociale. Il entreprend donc de comparer, objectivement, l’efficacité relative de chacun des modes d’organisation envisageables, à partir d’un test et d’une pondération relativement simple.

¹³ Cette expression n’est pas de nature péjorative. C’est presque explicitement sur le modèle du conte philosophique que Williamson construit cette première justification. Les deux hypothèses selon laquelle, d’une part, le marché est l’état naturel de départ, et, d’autre part, les firmes autogérées représentent le modèle qui succède au marché, sont évidemment deux hypothèses dénuées de tout fondement historique. La forme historique de l’argumentation de Williamson est purement pédagogique, elle n’a pas de prétention à être vraie.

Le test porte sur six modes d'organisation. Deux d'entre eux sont de type "entrepreneurial" : la *sous-traitance* et le mode *fédératif*. Deux sont du type "communautaire" : l'un repose sur un partage collégial des profits (le "*groupe des pairs*"), l'autre fonctionne sans redistribution collégiale des profits ("*chacun pour soi*"). Enfin, les deux derniers modes testés sont de type capitaliste : il s'agit de la *contractualisation interne* et de la *relation d'autorité*. La démarche de Williamson consiste à sérier trois critères de performance : (1) *Gestion du flux des produits*, (2) *Affectation de ces flux*, (3) *Incitation au travail*. Ces critères sont, au total, subdivisés en onze sous-critères ((1) *dépenses de transport / stock tampons / pente d'interface* ; (2) *Postes de travail / leadership / Maintenance* ; (3) *Intensité du travail / Utilisation des équipements / Adaptabilité locale / Adaptabilité Globale / Innovation* . Le score des différents modes de production est déterminé selon un mode binaire. Ainsi, pour chacun de ces critères, le mode étudié est satisfaisant et il marque un point, ou il ne l'est pas et il ne marque rien. Il faut ensuite agréger les performances pour obtenir le score sur onze¹⁴.

Le résultat de ce test est sans ambiguïté :

| Mode d'organisation | Score (sur 11) | Rang |
|----------------------------|----------------|------|
| Communautaire | 4 | 6 |
| sous-traitance | 5 | 5 |
| fédératif | 5 | 4 |
| Contractualisation interne | 6 | 3 |
| Groupe des pairs | 8 | 2 |
| Relation d'autorité | 9 | 1 |

Williamson en tire la conclusion suivante : "*le mode communautaire est le moins efficace. (...) Les modes les moins hiérarchiques ont les plus mauvaises propriétés en termes d'efficacité (...).*" (Williamson, 1985, p.231).

Le succès de la relation d'autorité s'explique par trois de ses caractéristiques : il est le plus hiérarchique, à la fois en matière de contractualisation (le propriétaire régit tous les contrats), et en matière de décision (il n'y a pas de sous-traitance interne) ; et il est fondé sur une contractualisation de long terme. La *hiérarchie* de contractualisation dissuade l'opportunisme, à la différence des modes communautaires. La *hiérarchie* décisionnelle permet une plus grande efficacité et cohérence de la décision. La contractualisation de *long terme* permet de mieux gérer la régulation interne de l'entreprise.

Le fait que la relation d'autorité se classe première de ce test démontre que le mode

¹⁴ Cette méthodologie fait l'objet d'une présentation critique par Gabriel et Jacquier [1994].

hiérarchique est choisi pour son *efficacité*. L'organisation hiérarchique des organisations ne doit donc rien, selon Williamson, à l'histoire sociale ou à l'émergence d'un rapport de force particulier. Il récuse ainsi les critiques formulées par les économistes "radicaux" américains (Goldberg, Marglin, Braverman) contre le mode d'organisation capitaliste hiérarchique. Le mode hiérarchique n'est pas la résultante d'une lutte des classes. Il n'est pas la résultante d'un rapport de force. Il ne fait même pas véritablement l'objet d'un choix social : il *s'impose* nécessairement parce qu'il est le plus efficace.

34- Les raisons de l'efficacité de la hiérarchie

L'argument de Williamson en faveur de la hiérarchie est donc simple : c'est le mode de gouvernance le plus efficace. Il faut cependant préciser davantage les raisons de cette efficacité. Elle s'explique par la nature des problèmes que la structure de gouvernance doit résoudre.

Arrow a déjà souligné l'efficacité de la hiérarchie en matière de gestion de l'information. Elle permet une centralisation, un tri et une prise de décision plus efficace que les modes organisationnels fondés sur une décision collégiale. Williamson met également l'accent sur la nécessité de gérer l'opportunisme. La hiérarchie représente un mode de surveillance efficace des individus et évite qu'ils poursuivent des objectifs propres contraires aux objectifs généraux de l'organisation. Toutefois, la hiérarchie n'est efficace que dans le cadre d'une gestion interne des transactions. L'existence d'un "marché interne du travail", selon l'expression que Williamson emprunte à Doeringer et Piore [1971], constitue donc un préalable à l'existence d'une relation hiérarchique.

341- La nécessité d'une structure de gouvernance

Williamson [1975, chap 4] souligne en effet que, prise hors de toute référence aux structures internes de l'organisation, la relation d'autorité ne résout rien. Il faut que le donneur d'ordres puisse comprendre et vérifier la nature du travail réalisé par le salarié, et, à l'inverse, que le salarié soit capable de comprendre ce que l'on attend de lui. Cela nécessite l'existence d'une certaine proximité entre les deux contractants : il est nécessaire que ceux-ci s'insèrent dans un ensemble plus vaste dont ils se reconnaissent comme deux éléments, et qui donne à leur relation l'épaisseur nécessaire. L'existence d'une *structure de gouvernance* permet aux contractants de se référer à un univers commun. Ils ont à leur disposition des habitudes, des modes de compréhension, des facilités d'observation qui

pérennisent leur relation. Pour utiliser une expression classique, la structure de gouvernance crée le "Common Knowledge" qui leur faisait défaut dans le cadre marchand. La structure organisationnelle permet de surmonter le barrage "communicationnel" que constitue le caractère idiosyncrasique des tâches de chacun, elle objective la relation des agents à l'organisation à laquelle ils participent. L'organisation interne apparaît en quelque sorte comme une forme de société parallèle, disposant de son langage, ses codes, ses habitudes. C'est cependant une forme spécifique de société, construite autour de l'impératif d'efficacité. Les codes communs, le langage, les habitudes ont en effet comme but unique de rendre les relations de travail plus efficaces.

C'est seulement à ce stade que la relation d'autorité devient nécessaire : il faut en effet éviter que les agents "oublient" de servir l'efficacité collective et ne se consacrent plus qu'à la réalisation d'objectifs propres. Une fois que cet univers organisationnel est bien en place, que le rôle et l'activité de chacun sont "objectivés", l'incommunicabilité ou l'incompréhension entre les agents, qui était à l'origine d'une opacité source d'opportunisme, disparaît. Les moyens de surveillance et d'incitation classiques peuvent alors opérer. D'une part, l'agent perçoit l'avantage qu'il peut avoir à jouer le jeu de l'organisation et à être efficace en étant associé aux résultats (par exemple par un salaire indexé à la productivité), d'autre part il est sous le regard de tous les autres, et en particulier de ses supérieurs qui le surveillent et rendent la défection coûteuse.

342- Langage et coercition

Le lien entre existence d'un espace commun et efficacité de la hiérarchie est particulièrement net dans l'analyse du langage que propose Williamson : "*Le marché manque d'un langage commun suffisamment riche. Le problème du langage est particulièrement important lorsque les jugements qui doivent être faits sont subjectifs. L'avantage de la hiérarchie dans ce domaine est particulièrement net si l'individu qui est le plus familier avec les caractéristiques de l'exécutant, c'est-à-dire son superviseur immédiat, est aussi celui qui évalue son travail.*" (Williamson, 1975, p.78). Le lien entre hiérarchie et langage commun avait déjà été mis en évidence par Arrow, pour des raisons d'efficacité de la transmission des informations et de conformation de l'individu à l'entreprise. Williamson va plus loin en montrant que le langage commun, associé à la hiérarchie, permet de *contrôler* les travailleurs de manière efficace. Il permet non seulement de guider l'agent dans une direction conforme aux intérêts collectifs de l'organisation, mais également de mieux le surveiller. L'existence d'un langage commun permet notamment à des informations concernant l'efficacité personnelle de l'agent de parvenir au centre de direction. Le langage, associé à la hiérarchie, est donc très

clairement mis en avant comme un instrument d'orientation et de contrôle, c'est-à-dire de pouvoir.

Cette réflexion sur le langage dans l'organisation éclaire la raison principale pour laquelle la hiérarchie est efficace selon Williamson : sa fonction coercitive. Cette dimension peut être repérée à deux moments dans son raisonnement. La fonction coercitive de l'organisation est tout d'abord l'argument que Williamson utilise pour expliquer pourquoi la hiérarchie succède "historiquement" au groupe des pairs. Elle dissuade davantage l'opportunisme, dans la mesure où elle permet de mettre chacun sous le regard d'un supérieur direct. Par ailleurs ce rôle de surveillance et d'évaluation de la hiérarchie se traduit également à travers les fonctions "*quasi-judiciaires*" que Williamson (1975, p.30) attribue la gouvernance. L'économie standard estime généralement qu'il est du ressort des tribunaux de régler les conflits survenant entre principal et agent durant l'exécution des contrats. Or, selon Williamson, ce recours au droit est coûteux et moins efficace que le règlement par un "arrangement privé" interne à l'entreprise. Or, ce type de règlement sera plus efficace s'il existe des relations hiérarchiques permettant de trancher le différend. Le *règlement autoritaire*¹⁵ des conflits est en effet un mode de règlement très efficace et très économique. Or, des conflits surviennent fréquemment entre agents. L'existence d'une forme de "justice" interne à l'entreprise représente donc une source importante d'économies. L'organisation hiérarchique est ainsi une réponse adaptée au phénomène de l'opportunisme parce qu'elle sait non seulement surveiller, mais aussi punir.

L'intérêt de la réflexion de Williamson est le lien qu'il opère ainsi entre la fonction socialisatrice de l'organisation et la fonction coercitive de la hiérarchie. Arrow avait, avant lui, mis l'accent sur la fonction informationnelle et décisionnelle des structures hiérarchiques. Il avait également mis l'accent sur le rôle coordinateur que pouvaient jouer les codes et règles internes à l'organisation. Mais Williamson va plus loin en insistant sur la manière dont l'organisation permet de reconstruire un langage et un univers permettant une meilleure évaluation de l'ensemble des tâches idiosyncrasiques, intimement liées au travailleur. L'obscurité qui résulte de cette non-standardisation des tâches est au fondement de la possibilité de l'opportunisme. L'organisation hiérarchique tire précisément son efficacité de sa capacité à éliminer cette source d'opportunisme, en permettant une évaluation plus objective du travail de chacun. L'existence d'un univers commun, d'un "*Common Knowledge*" construit de manière *intentionnelle*, sert donc directement la fonction coercitive de l'organisation. En mettant ainsi en évidence cette

¹⁵ Williamson (1975, p.101) utilise le mot "*Fiat*" pour désigner une prise de décision autoritaire. Le règlement par *fiat* stoppe le débat et le conflit par un acte d'autorité. Il met fin à la discussion.

fonction "régaliennne" de la structure de gouvernance, Williamson tranche avec le rôle neutre que la théorie fait généralement jouer aux "règles collectives". La gouvernance n'a plus un rôle strictement *coordinateur*. Elle ne sert pas à coordonner les intérêts d'individus divers, mais à imposer un comportement conforme à l'intérêt collectif¹⁶. Ici Williamson met à jour l'ultime conséquence de l'hypothèse de rationalité instrumentale : l'opportunisme qui en découle nécessite la mise en place de dispositifs de surveillance. Il lie ainsi rationalité individuelle et nécessité collective de la coercition.

La disjonction qui apparaît ici clairement entre les objectifs de chaque individu et ceux de l'organisation pose un problème. Comment parvenir à trouver un type d'acteur suffisamment désintéressé pour qu'on puisse lui confier le rôle consistant à prendre les décisions stratégiques ? Williamson répond également à cette question en désignant une catégorie d'acteurs "désintéressés", ou, plus exactement, exclusivement intéressés à la réussite de l'organisation collective. C'est à eux que doit revenir, *in fine*, le pouvoir.

35- *Le pouvoir aux propriétaires.*

Le choix des agents occupant les postes élevés de la hiérarchie trouve son origine dans les deux caractéristiques du *contractual man* : son opportunisme et ses limites cognitives. D'une part, les tâches de gestion opérationnelle doivent être confiées aux individus "les moins limités". D'autre part, le pouvoir de direction stratégique doit être assuré par les seuls agents dont les intérêts se confondent avec ceux de l'entreprise : les actionnaires-propriétaires.

351- *Une hiérarchie naturelle*

"Ceux dont la rationalité est moins limitée que celle des autres sont des candidats naturels pour assurer les tâches techniques, administratives, et les positions politiques : ce qui revient à dire qu'une hiérarchie peut émerger sur cette base". (Williamson, 1975, p.24)

L'existence de limites cognitives est au fondement d'une méritocratie interne à l'organisation. Selon Williamson, la hiérarchie, grâce à la possibilité de promotion qu'elle implique, constitue un système efficace pour révéler les agents les plus aptes à occuper des postes de direction. L'analyse de la hiérarchie débouche donc sur une justification

¹⁶ C'est un rôle qui, comme on l'a vu, demeure tacite dans l'analyse d'Arrow.

naturaliste et évolutionniste de l'existence des asymétries de pouvoir. Ce sont les agents les moins limités qui dirigent, ou, plus exactement, qui font respecter les directives de l'organisation.

Comme on l'a déjà noté lors de la présentation de la justification de la hiérarchie, cette "sélection naturelle" se joue principalement au niveau de la maîtrise du langage. La tâche des supérieurs est en effet de parvenir à motiver les individus dans une direction conforme à l'intérêt organisationnel. Ils doivent donc être capables de centraliser l'information, de la traiter, et d'expliquer les décisions qu'ils prennent à partir de ces informations. Le fait de posséder "*d'exceptionnelles capacités oratoires*" (Williamson, 1975, p.52) est donc un pré-requis à l'accession aux postes de commandement. Cette conception de la hiérarchie est en parfaite cohérence avec l'analyse des raisons qui expliquent l'efficacité de la hiérarchie. Le rôle de la structure hiérarchique n'est pas seulement de gérer efficacement l'information, mais aussi de surveiller et punir. Il faut donc que les "managers" sachent "trancher" autoritairement les différends qui surviennent. Il faut pour cela être capable de "stopper" la discussion en faisant apparaître l'intérêt commun. Il faut convaincre les employés de faire prévaloir les intérêts de l'organisation sur ce qui pourrait être à court terme leur intérêt propre. Pour convaincre, expliquer, ordonner, il est nécessaire de maîtriser le langage organisationnel¹⁷.

352 - Le pouvoir décisionnel aux actionnaires

"Doter les actionnaires d'une capacité à contrôler les activités de la firme et à remplacer la direction en cas de crise devrait faciliter le recours à un financement par augmentation des capitaux propres. Pour cette raison le conseil d'administration devrait être considéré principalement comme l'instrument de gouvernance des actionnaires.." (Williamson, 1985, p.324)

Le problème principal n'est cependant pas tant de savoir qui transmet et fait appliquer les décisions, que de déterminer qui fixe les directives. Sur ce point, le problème principal à résoudre est celui de l'opportunisme. Il faut en effet éviter que les décideurs cherchent à fixer des directives qui soient plus conformes à leur propre intérêt qu'à l'intérêt collectif. Pour savoir qui doit avoir le pouvoir de décision, il faut donc identifier des agents dont la relation à l'organisation constitue une profonde désincitation à lui nuire. C'est le cas des propriétaires, qui investissent leur capital dans l'organisation. Les revenus qu'ils espèrent en tirer dépendent en effet de sa bonne "santé".

¹⁷ Ce point est important dans la mesure où l'impératif de la maîtrise du langage est un aspect typiquement "hobbien" de l'analyse de l'ordre social. Celui qui maîtrise les mots a le pouvoir, dans une optique où il n'existe pas de discussion réellement démocratique.

Williamson affirme donc que le droit de vote au conseil d'administration devrait être réservé aux seuls actionnaires. L'argument qu'il développe est double. D'une part, la firme a besoin du financement des actionnaires. D'autre part, les actionnaires, qui prennent tous les risques, doivent pouvoir décider seuls du devenir de leur investissement. Le pouvoir doit donc être donné aux propriétaires des capitaux. Ce sont en effet les seuls dont les intérêts propres ne sont pas distincts des intérêts collectifs. Ils sont, structurellement, amenés à être les interprètes de la volonté collective.

353- Une justification des structures capitalistes ?

La théorie de Williamson se termine donc sur une justification de la structure capitaliste des règles organisationnelles. Il est nécessaire qu'elles soient hiérarchiques et qu'elles soient établies et dominées par les propriétaires. Cette justification du mode capitaliste de production est qualifiée par Dow [1987, p.32] de *logique "à la Pangloss"* : "tout est pour le mieux dans le meilleur des mondes possibles !" ¹⁸. Selon cet auteur, Williamson serait en effet trop fonctionnaliste, et ne dévoilerait pas suffisamment les rapports sociaux sous-jacents au système capitaliste. Dow traduit ces critiques en reprochant à Williamson de ne pas envisager de défections de la part des dirigeants, de la même manière qu'il l'envisage pour les salariés. La réponse de Williamson [1987b, p.623] est offusquée : "*c'est une façon de brouiller les pistes*".

De fait, la logique de Williamson se veut tout à la fois objective, et étrangère à l'idée de rapport de force et de lutte pour le pouvoir. Il récuse d'ailleurs absolument l'emploi de ce dernier terme "*qui est si pauvrement défini qu'il peut servir à expliquer à peu près n'importe quoi*" (Williamson, 1985, p. 237-238). Sa justification du mode organisationnel capitaliste ne s'appuie, selon lui, que sur la recherche de l'efficacité. Or, pour être efficace :

- Il faut d'abord que ceux qui président aux destinées de l'organisation n'aient aucun objectif propre différent de ceux de l'organisation.
- Il faut ensuite que l'ensemble des agents se déchargent de tous leurs droits et objectifs propres au profit de la collectivité organisationnelle, afin de ne plus être tentés de jouer leurs propres cartes au lieu de se consacrer à la maximisation du bien être collectif.
- Il faut enfin que les individus chargés de veiller au bon ordre soient suffisamment doués pour percevoir des bribes de la rationalité globale qui régit le système.

C'est à ces trois conditions que l'organisation peut être efficace. Les actionnaires ne sont donc chargés de veiller au bien-être collectif que parce qu'ils n'ont pas les mêmes

¹⁸ Williamson [1987a] s'en défend fortement sans pour autant présenter d'arguments contre cette logique.

objectifs que les employés. Ces derniers souhaitent maximiser leur propre bien-être, avant de penser à l'organisation. Les actionnaires cherchent eux à maximiser les profits de l'organisation¹⁹.

A travers cette justification, Williamson fait clairement apparaître une altérité, tout-à-fait étrangère à l'analyse néoclassique. Il montre que la gouvernance de l'entreprise fait l'objet d'un mode de rationalité et de désir différent de celui de la plupart des agents. C'est finalement cette altérité qui explique, à ses yeux, l'efficacité du capitalisme. Ce mode d'organisation de la production se justifie entièrement par le fait qu'il est le plus efficace. Il est inutile de chercher à fournir d'autres explications socio-historiques pour rendre compte de son émergence et de son développement.

La réponse de Williamson au problème de coordination que pose la nature humaine telle qu'il la décrit tient donc en deux mots : organisation hiérarchique. Les règles organisationnelles doivent être construites pour lutter efficacement contre les attributs "gênants" de la nature humaine. L'organisation hiérarchique est de ce point de vue la meilleure, dans la mesure où elle permet à la fois d'économiser sur le stockage et la transmission de l'information et de dissuader l'opportunisme en le surveillant de façon efficace. Ce faisant, Williamson fait apparaître un lien essentiel entre d'une part le mode de *rationalité* des agents économiques, et, d'autre part, la nécessité d'une organisation *coercitive* de leurs transactions. L'opposition qu'il dresse entre marché et hiérarchie est de ce point de vue tout à fait édifiante : en dehors du système "libre" de marché, l'individu économique doit être soumis à une organisation coercitive²⁰. Autrement dit, à compter du moment où le problème de la règle se pose, l'unique réponse à lui apporter doit être guidée par l'exigence d'efficacité, ce qui implique la coercition.

¹⁹ Un autre problème, que Williamson n'évoque pas et qui sort du cadre de notre analyse, est de savoir dans quelle mesure les actionnaires, enclins à rechercher un profit à court terme, sont toujours aptes à prendre des décisions stratégiques sur le long terme.

²⁰ On retrouve donc chez Williamson, en filigrane, le dilemme existant entre liberté et rationalité collective pointé par Arrow dans le théorème d'impossibilité.

4- Force analytique et faiblesse pratique du projet Williamsonien

Le projet que poursuit Williamson, à travers le développement d'une "*Nouvelle Économie Institutionnelle*" fondée sur l'analyse des moyens d'économiser les coûts de transaction, est de réintroduire le concept d'institution dans l'analyse économique, à partir d'une réflexion strictement micro-analytique. Ce programme de recherche vise donc à justifier l'existence des institutions économiques, mais aussi à en expliquer le fonctionnement et la genèse en se fondant exclusivement sur l'analyse du comportement d'un agent conçu comme instrumentalement rationnel. En quelque sorte, toute institution économique devrait peu ou prou pouvoir être réduite à un calcul économique. Cette attitude réductionniste permet de faire table rase des autres dimensions explicatives, historiques et sociales, pour produire une théorie strictement analytique des institutions. Par cette attitude théorique, Williamson cherche à éviter de tomber dans les "travers" d'une tradition institutionnaliste qui, à ses yeux, a trop souffert du pointillisme de ses réflexions et n'est pas parvenue à proposer une analyse logique convaincante des organisations.

Cette posture théorique radicale fait l'objet de critiques plus ou moins virulentes. On souligne que Williamson oublie le rôle de la "confiance" (Orléan [1994], Karpik [1998]) ; celui des règles éthiques (Alchian et Woodward [1988]) ; celui des normes sociales qui permettent d'expliquer *l'engagement* de l'individu (Granovetter [1985], Ménard [1995, 1994a et 1994b], North [1986, 1991], Simon [1991]). On lui reproche finalement de situer son analyse à un niveau d'abstraction qui ne lui permet plus de représenter le fonctionnement concret des organisations, ni de rendre compte de leur émergence ou de leur dynamique. Même l'élément qui paraît, *a priori*, être le plus directement "empirique" de l'analyse, le concept de "coût de transaction", est défini de manière trop abstraite pour être véritablement mesurable (Brousseau [1989, 1993]). La théorie "institutionnaliste" de Williamson s'avère donc, paradoxalement, avoir une grande force analytique mais peu de contenu pratique.

41- Le refus de l'éthique : l'exemple de la confiance

"La calculabilité (calculativeness) est la condition générale que j'associe à l'approche économique et à l'extension progressive de l'économie dans les sciences sociales qui lui

sont reliées. (...) Si les relations fondées sur le calcul sont mieux décrites en termes calculatoires, alors les termes verbeux, qui peuvent avoir des significations doubles, comme «la confiance», doivent être bannis autant que possible. “ (Williamson, 1993b, p.456 et 469)

Williamson se méfie des mots et de l'analyse "littéraire" des rapports humains. Il consacre ainsi un texte relativement récent à récuser l'usage du terme "confiance" dans l'analyse économique. Cette charge contre l'usage de ce terme est, selon nous, emblématique de la volonté de Williamson de s'interdire tout appui sur des éléments socio-institutionnels étrangers à l'univers calculatoire des coûts de transaction. Nous nous accordons sur ce point avec la remarque de Karpik : "*Williamson entend écarter un concept qui lui semble menacer non seulement l'unité du modèle néoclassique mais aussi, dès lors que la confiance est assimilée, fût-ce secondairement, à un mécanisme de coordination des activités économiques irréductibles au marché comme à l'organisation, la cohérence de la théorie des coûts de transaction qui reste, malgré la multiplication des variantes et des hybrides, fondamentalement dualiste*" (Karpik, 1998, p.1044). C'est bien en effet la cohérence de l'économie des coûts de transaction que Williamson veut sauver, et, sur ce point, on ne peut que lui donner raison.

411- L'absence de fiabilité de "contractual man".

Williamson est en effet suffisamment explicite sur le choix de ses hypothèses comportementales pour ne pas avoir à justifier son rejet de concepts semi-éthiques comme celui de confiance. Comment, en effet, des "*contractual men*" pourraient-ils se faire confiance ? Le *contractual man* ne fait pas partie d'une société humaine. Il ne se définit nullement par le respect de normes ou de repères socio-historiques communs. Son rapport à l'autre est exclusivement instrumental : les autres ne sont là que comme des instruments potentiels de ses propres objectifs. Sa capacité à être rationnel, au moins dans une certaine mesure, épuise quasiment la description du *contractual man*, le rapprochant par là même de la machine plutôt que de l'homme.

Il y a cependant un aspect qui le différencie profondément de la machine : *contractual man* est profondément égoïste, au point de chercher par tous les moyens possibles à tourner les règles collectives à son avantage, à tromper ses partenaires, à rompre ses engagements lorsque cela lui est utile. Sachant que, dans son univers, tous ses partenaires partagent les mêmes caractéristiques comportementales, pourrait-il être suffisamment stupide pour leur faire confiance ? L'unique possibilité pour qu'un agent ait confiance en l'action d'un autre agent réside dans le fait que cette action soit contrainte, ou parfaitement

conforme aux intérêts de l'autre. Une telle forme de confiance ne doit donc rien à la fiabilité accordée à l'autre agent, mais repose sur un pur calcul permettant de prévoir son comportement. Or, comme le note Williamson (1993b, p. 463) : *“une confiance calculée est une contradiction dans les termes”*.

S'appuyant sur cette idée, Williamson cherche alors à dévoiler, sous les prétendus rapports de confiance qu'entretiendraient les agents économiques, la présence d'un calcul rationnel : un agent ne place sa confiance dans le comportement d'un autre agent que dans la mesure où cet autre agent a rationnellement intérêt à agir selon ses attentes (éventuellement parce qu'il y est contraint). Les seules attentes réciproques qui existent ici sont donc celles qui portent sur la rationalité commune. Pour Williamson, il n'existe donc pas dans le monde économique de relation de confiance qui ne soit pas réductible à un calcul : seuls les liens familiaux, strictement privés, sont en dehors du calcul et s'appuient sur une confiance réciproque.

412- Une confiance institutionnalisée

Cette affirmation ne signifie pas, cependant, que la rationalité soit suffisante pour fonder des anticipations croisées stables. Les individus ne parviennent pas forcément à percevoir leur intérêt commun. Ils n'identifient pas forcément les bénéfices de long terme qui peuvent être attachés à un comportement coopératif. Ils ne peuvent pas toujours exploiter les opportunités que la situation d'interaction leur offre²¹. Les institutions ont précisément pour objet de restaurer la confiance en donnant aux individus des repères communs fiables. L'organisation que décrit Williamson donne ainsi aux individus des moyens de se comprendre, elle facilite l'émergence de la coopération en limitant leurs marges de manœuvre. L'origine de cette “confiance” réside donc non pas dans la confiance mutuelle que se prêtent les deux agents, mais dans la confiance que chacun d'eux place dans l'efficacité coercitive de la structure institutionnelle gouvernant leur relation. Pour cette raison, Williamson qualifie cette forme de confiance de “confiance institutionnelle” et précise : *“La confiance institutionnelle réside dans le contexte social et organisationnel dans lequel les contrats sont insérés. Dans la mesure où les caractéristiques pertinentes de ces institutions sont exogènes, la confiance institutionnelle peut paraître d'essence non calculatoire. Mais, en réalité, les transactions sont toujours organisées en référence au contexte institutionnel où elles se déroulent. En conséquence le calcul réapparaît.”* (Williamson, 1993b, p.486).

Ainsi, si les individus peuvent, parfois, se passer de calcul et se faire apparemment

²¹C'est l'idée contenue dans les jeux de coordination, d'une part, et le jeu du dilemme du prisonnier d'autre part, que nous présenterons dans le chapitre suivant.

confiance, c'est en réalité parce qu'ils se conforment à des règles institutionnelles. Or ces règles elles-mêmes ont été conçues, calculées, pour faire naître cette "pseudo-confiance". Le calcul réapparaît donc à un niveau supérieur à celui des agents, celui d'une sorte d'ingénieur social qui conçoit, de façon calculatoire, les institutions adéquates. Il est inutile de parler de "confiance" pour désigner une telle relation, qui est simplement une relation correctement gérée par une structure institutionnelle bien calculée. Les acteurs peuvent, éventuellement, avoir l'illusion qu'ils se font confiance. Pas l'ingénieur social.

La force de la position de Williamson s'apprécie surtout au regard des utilisations de la notion de confiance dans l'analyse économique standard. Williamson s'oppose ici aux économistes qui, dans la lignée d'Arrow, cherchent à développer l'idée que des "*institutions invisibles*" comme des normes, des codes éthiques ou des relations de confiance sont nécessaires au déroulement des transactions économiques²². Williamson veut ainsi éviter de sortir de l'univers calculatoire et donc de l'analyse économique. Sa position a le mérite de la cohérence. Les agents économiques de la théorie néoclassique comme ceux qu'étudie Williamson ne sont en effet pas conçus pour faire autre chose que du calcul. Dès lors, selon Williamson, faire entrer dans la théorie des notions telles que l'éthique et la socialisation représente un grand risque : celui de constituer une théorie composite, manquant de cohésion, et de retomber ainsi dans les travers de l'ancien institutionnalisme. En effet, si les individus peuvent se faire confiance, s'ils respectent des règles éthiques, s'ils suivent des normes sociales, c'est qu'ils ne sont plus tout-à-fait tels que l'économie des coûts de transaction les perçoit. En quelque sorte Williamson estime que l'on ne peut convoquer de notions extérieures à l'idée de calcul sans modifier entièrement les fondements de la théorie. En effet, tout comme l'*Homo œconomicus*, le *contractual man* n'a ni sens social, ni sens éthique. Cette posture théorique est conséquente : Williamson reste fidèle jusqu'au bout à ses hypothèses comportementales, ce qui permet d'en mettre à jour toutes les implications logiques²³.

²² Williamson [1987b], à l'occasion de sa "reconnaissance de dette intellectuelle" à l'égard d'Arrow, précise nettement que la notion de confiance lui semble être un héritage bien peu utilisable. Il s'oppose aussi explicitement à Kreps à qui il reproche de ne pas s'en tenir à l'idée de réputation qui suffit à expliquer l'attitude coopérative dans certains contextes organisés en réseaux.

²³ Cette attitude le distingue par exemple d'Alchian et Woodward qui, dans leur revue du second ouvrage de Williamson, lui reprochent d'oublier les "limites sociales" imposées au comportement opportuniste : "*Nous pensons qu'il est très important de reconnaître la force, dans les relations d'affaires, que peuvent avoir l'éthique, la réputation, les normes définissant un comportement "adapté, correct, raisonnable."*" (Alchian et Woodward, 1988, p.77).

42- L'organisation est-elle la seule institution ?

“Il y a des ambiguïtés concernant la relation entre institution et organisation. Ces confusions remontent peut-être à Marshall.(...) Plus récemment Oliver Williamson a publié, dans un laps de temps très court, deux ouvrages dont le contenu couvre à peu près les mêmes territoires. Le premier porte sur les “organisations”, le second sur les “institutions”. Moe [1991], qui est très fortement influencé par la littérature sur les coûts de transaction, va plus loin en affirmant que les institutions **sont** des organisations.” (Ménard, 1995, p.174).

Claude Ménard, dans cet article récent, utilise la distinction de North entre *environnement institutionnel* et *arrangement institutionnel* pour distinguer d'une part le niveau des “institutions”, d'autre part les arrangements institutionnels que sont le marché ou l'organisation. L'institution apparaît comme la catégorie la plus générale, au niveau de laquelle se déterminent les “règles du jeu”, fixées de manière impersonnelle sur le long terme, et incarnées dans le droit, la coutume, les habitudes. Le marché et les organisations s'insèrent dans ces règles du jeu et constituent donc des arrangements *internes* et *spécifiques* à cet environnement institutionnel. L'organisation apparaît ainsi comme un ensemble de règles construit de manière consciente et volontaire et fondé sur la coopération et l'autorité. Cette nette distinction permet, par contraste, de souligner ce qui chez Williamson s'apparente à de “l'impérialisme organisationnel” : l'organisation tend en effet à devenir le modèle général de l'institution.

421- Un modèle abstrait ...

La réflexion de Ménard souligne le flou qu'entretient Williamson en ce qui concerne la distinction entre la notion d'institution et celle d'organisation. Cette absence de distinction conceptuelle dissimule en fait une volonté de faire se recouvrir les deux notions. Ce faisant, Williamson tend à substituer la notion d'organisation à celle, plus générale, d'institution. Alchian et Woodward le lui reprochent en notant : “*les institutions du capitalisme ne sont pas uniquement des firmes, des marchés et des relations contractuelles. (...) Les mutuelles, les clubs sociaux, les “joint venture”, et la famille, pour suggérer quelques autres institutions capitalistes, sont également essentielles.*” (Alchian et Woodward, 1988, p.76). Dans la même lignée critique, North [1986,1991] souligne que l'analyse de Williamson ne prend pas suffisamment en compte la superstructure institutionnelle qui explique la forme particulière que prennent les

infrastructures organisationnelles. Il souligne ainsi : “*les travaux [de l'économie des coûts de transaction] négligent le sujet principal de cet essai : expliquer les performances variées des économies à travers les âges et dans le monde entier*”. En quelque sorte, Williamson propose un modèle unique et anhistorique qui ne prend pas en compte le rôle des institutions sociales transcendant l'organisation capitaliste.

Cet aspect de la théorie de Williamson est directement lié à son souci analytique. Williamson se refuse à introduire dans sa théorie des éléments hétérogènes qui ne relèveraient pas de l'univers calculatoire des coûts de transaction. Cette volonté de proposer une théorie homogène se prolonge donc en une sorte de généralisation abusive. En quelque sorte, le projet de Williamson tend à considérer la “structure de gouvernance” des transactions comme l'unique forme institutionnelle économique. Williamson cherche par là à proposer un mode d'explication de la régulation des interactions en dehors du marché qui soit suffisamment abstrait pour constituer un modèle général. Il forge une théorie indépendante de toute forme de particularisme local. C'est la raison pour laquelle, dans son analyse, les “structures de gouvernance” sont quasi indépendantes de toute superstructure sociale, culturelle, ou historique²⁴. Williamson souhaite ainsi émanciper la théorie des institutions économiques de la réflexion sociologique, politique et philosophique. Appliquer une méthodologie analytique fondée sur le calcul nécessite en effet d'isoler des caractéristiques objectives, universelles, et invariantes.

422- ... dans la suite logique de la théorie de l'équilibre général

En cela Williamson se situe dans la suite logique de la théorie de l'équilibre général. Ainsi, le fait que sa justification “historique” de la forme de gouvernance hiérarchique s'appuie sur l'hypothèse rhétorique selon laquelle “*au début il y avait le marché*” nous délivre un indice essentiel quant à l'esprit de son entreprise théorique. Sa théorie ne peut se comprendre qu'en référence à la théorie néoclassique du marché. Elle est dans le prolongement d'un modèle d'équilibre général qui a séparé l'agent économique de la société. Williamson prend comme point de départ de son analyse cette formalisation qui lui donne l'élément fondateur de sa propre problématique : celle d'un agent conçu pour agir sur un marché pur et parfait, se retrouvant en butte à des problèmes d'information et

²⁴ Baudry [1999], dans un article récent, défend une analyse sensiblement différente de la théorie de Williamson. Il est nécessaire de préciser que Williamson ne nie pas qu'il existe des institutions qui sont extérieures à l'organisation et qui lui imposent des règles du jeu, comme par exemple les institutions juridiques. Il tente même de préciser l'articulation organisation/institution dans un article relativement récent (Williamson [1993a]). Mais la présence de ces règles institutionnelles n'intervient jamais vraiment dans la théorie des coûts de transaction lorsqu'il s'agit d'analyser les raisons du choix de la structure de gouvernance, comme par exemple la relation d'autorité. Elles ont un rôle extrêmement marginal et sont en permanence tenues à distance du cœur de l'argumentation.

de coordination que le marché ne peut résoudre. Le principe même des “failles” du marché ne prend sens que par rapport à la théorie analytique du marché que propose l'économie néoclassique

Or, nous avons, dans notre premier chapitre, indiqué le caractère purement formel de l'agent néoclassique qui sous-tend le modèle d'équilibre général. La théorie néoclassique a promu la figure d'un agent rationnel qui n'a pas d'autre univers que le marché. C'est donc un individu hors société²⁵. Williamson reprend à son compte la réduction opérée par Walras puis par Arrow-Debreu : il supprime toute référence à la sphère non économique. Il fait “table rase” de la sphère institutionnelle et sociale. L'individu rationnel passe ainsi du marché à l'organisation sans jamais rencontrer la société²⁶. Pour retrouver la dimension sociale et culturelle de l'individu rationnel, il faudrait, en quelque sorte, revenir en aval de la tradition walrassienne d'analyse du marché. Ce n'est pas l'objectif théorique que se fixe Williamson puisqu'il cherche au contraire à formuler une théorie située au même niveau de généralité que la théorie néoclassique. Ainsi, l'ambition de Williamson est sans doute de parvenir à mettre sur pied une théorie de l'organisation qui soit le pendant de la théorie du marché proposée par Arrow et Debreu [1954]. Ce projet théorique ne lui permet pas de s'appuyer sur l'existence d'institutions distinctes de la catégorie “organisation”. Sa volonté de disposer d'une solide base analytique nuit donc forcément au contenu pratique de sa théorie.

43- Un problème pratique

“Les tentatives de la nouvelle économie institutionnelle cherchant à expliquer le comportement organisationnel en s'appuyant uniquement sur les concepts d'agence, d'asymétrie d'information, de coût de transaction, d'opportunisme, et sur d'autres concepts tirés de l'économie néoclassique, négligent des mécanismes organisationnels fondamentaux comme l'autorité, l'identification et la coordination, et sont donc sérieusement incomplètes”. (Simon, 1991, p.42).

La théorie williamsonienne s'appuie sur une solide base analytique. Cependant, sa définition abstraite du “contractual man”, si elle lui permet de définir clairement le rôle de la structure de gouvernance, l'empêche en revanche de comprendre comment les organisations peuvent émerger. En effet, pour disposer d'une théorie de la genèse, ou de

²⁵ Frydman [1992] note à juste titre que cet agent est en fait non pas “sous-socialisé” mais “sur-socialisé” puisque son mode de comportement se déduit du système global.

²⁶ C'est précisément le reproche que Granovetter [1985] adresse à Williamson dans son célèbre article défendant l'idée de l'encastrement des formes économiques dans la société.

la dynamique des organisations, il serait nécessaire d'expliquer d'une part pourquoi les individus s'engagent et coopèrent à l'organisation, et d'autre part d'opérationnaliser la notion de "coût de transaction" en lui donnant un contenu pratique. Or ceci est par construction impossible : dans la théorie de Williamson, la "solution" vient d'une structure qui transcende *nécessairement* les individus. La genèse des organisations est donc par définition une question insoluble dans un cadre théorique individualiste.

431- *Le problème de l'engagement*

Herbert Simon, pourtant consacré par Williamson comme l'un des inspirateurs de l'économie des coûts de transaction, a nettement pris ses distances vis-à-vis de ce courant. Il regrette ainsi, dans un article fortement critique, l'optique "coercitive" et "calculatoire" de cette analyse. Simon explique qu'une telle optique ne peut pas rendre compte de tous les aspects du fonctionnement des organisations et, en particulier, ne peut expliquer que : "*dans la plupart des organisations, les employés contribuent beaucoup plus à atteindre les objectifs collectifs que ce qu'une supervision visant à la bonne application des (vagues) termes du contrat d'embauche ne pourrait obtenir d'eux.*" (Simon, 1991, p.32).

Dans la même optique, Claude Ménard (1994a) souligne qu'il existe un "dilemme organisationnel" entre rationalité instrumentale et coopération, et s'interroge : "*peut-on expliquer ce qui amène les agents à coopérer dans les organisations ?*". Il souligne que la réponse de Williamson à cette question constitue une avancée essentielle dans la mesure où elle met en avant l'importance de la structure collective nécessaire à l'émergence de la coopération. Mais il indique également que "*l'accent mis sur la hiérarchie dans l'analyse de Williamson (...) rejette dans l'ombre cette autre dimension fondamentale, celle de l'organisation comme tissu de relations sociales. (...) Il faut chercher la solution dans l'organisation fondée sur le consentement mutuel autant que sur la contrainte.*" (Ménard, 1994a, p.199).

On retrouve cette idée, à travers un cheminement différent, dans un autre article de Claude Ménard soulignant l'importance de l'autorité. L'auteur insiste en effet sur la nécessité de distinguer conceptuellement *l'autorité* et la *hiérarchie*. L'autorité est une faculté attachée à la personne. La hiérarchie est une modalité de régulation fondée sur le statut : "*un individu a de l'autorité et il est dans une position hiérarchique*" (Ménard, 1994b, p.233). C'est la combinaison de ces deux éléments qui permet aux organisations de fonctionner. La hiérarchie seule n'aboutirait donc à rien. Or l'autorité, comme la coopération, repose en partie sur le consentement des agents.

Ces deux auteurs mettent donc l'accent sur un aspect essentiel du fonctionnement des

organisations : l'engagement volontaire des individus, qui se traduit par un respect de l'autorité. Ces remarques critiques éclairent un aspect de l'organisation qui n'apparaît que très peu chez Williamson : l'adhésion de l'employé à la structure organisationnelle. Il apparaît ainsi que l'employé travaille davantage que ne pourrait en exiger de lui la structure parce qu'il adhère aux objectifs généraux de l'organisation. Il respecte les mots d'ordre de la hiérarchie, même lorsque son activité est invérifiable, parce qu'il accepte l'autorité de ses supérieurs hiérarchiques. En quelque sorte, dans leurs analyses Simon [1991] et Ménard [1994a, 1994b] retournent l'optique williamsonnienne en éclairant le rôle moteur de l'*acteur* et non plus de la *structure*, dans la bonne marche du contrat. La mécanique institutionnelle que dessine Williamson ne fonctionnerait pas si les agents n'adoptaient pas un comportement plus coopératif que ne permettent de l'envisager les hypothèses comportementales qu'il adopte. Pour que l'organisation fonctionne, il faut que les employés s'engagent volontairement et cet engagement n'est pas envisageable dans une optique strictement coercitive et conflictuelle²⁷.

Williamson manque ainsi un premier élément essentiel permettant de comprendre la genèse et la dynamique des organisations.

432- *L'impossible mesure des coûts de transaction.*

“D'une part, il est très difficile de mesurer et même d'approximer les coûts de transaction ; d'autre part, un recensement et une typologie de l'ensemble des arrangements institutionnels possibles manquent. Il en résulte une grande difficulté à établir une formalisation logique des liens existant entre coûts de transaction et nature des contrats ou organisation des institutions.” (Brousseau, 1989, p.138).

On pourrait arguer du fait que Williamson, à travers son concept de coût de transaction, propose une justification pratique du choix de la structure de gouvernance. Les agents mesureraient, en quelque sorte, les avantages de chacune des structures institutionnelles possibles. Le concept de coût de transaction est donc tout à la fois un moyen d'unifier la théorie et de proposer une représentation pratique du mécanisme de choix de la structure adéquate. Le problème qui se pose cependant est de savoir comment mesurer de tels coûts. Or, précisément, comme le montrent les analyses de Brousseau (1989,1993), ils ne sont pas mesurables. Williamson le reconnaît d'ailleurs lui-même à demi-mot : *“ils sont souvent difficiles à quantifier. (...) La recherche empirique sur les coûts de transaction n'essaie presque jamais de les mesurer directement.”* (Williamson,

²⁷ Ménard [1995] souligne que le caractère *volontaire* de l'engagement de l'agent dans l'organisation est un des traits qui distinguent l'organisation de l'institution.

1985, p.21-22). Les coûts de transaction sont difficiles à identifier en raison de leur double origine. Le concept de coûts de transaction regroupe en effet deux types de coûts différents : les coûts effectifs (liés à l'organisation du travail), et les coûts d'opportunités (liés aux pertes que constitue le choix d'une structure qui s'avère sous-optimale par rapport à une autre). Cette dualité est source de confusion, et rend le concept de coût de transaction difficile à saisir²⁸. Dans ces conditions on peut se demander quel est le critère permettant de choisir la bonne structure.

Williamson (1985, p.22) affirme que l'impossibilité de mesurer directement les coûts de transaction ne constitue pas un problème grave. Il suffit en effet selon lui de pouvoir comparer entre eux les systèmes possibles : le plus efficace est aussi celui qui minimise les coûts de transaction. Cette réponse indique à l'évidence que le choix de la forme de gouvernance, y compris le choix de l'alternative entre marché et hiérarchie, ne peut pas faire l'objet d'une décision calculatoire des agents eux-mêmes. Seul un individu omniscient peut connaître l'ensemble des structures possibles et à la capacité de mesurer leur efficacité relative. Toutefois cette conception du choix de la structure adéquate se justifie dans le cadre d'une réflexion instrumentale sur le choix du meilleur mécanisme collectif. Williamson adopte donc le point de vue d'un ingénieur social construisant les mécanismes collectifs qui assurent une coordination optimale. Mais, même de ce point de vue, comme le note Éric Brousseau, il faudrait pouvoir disposer d'une liste complète des modes d'organisation possibles. La théorie des coûts de transaction ne nous fournit pas une telle liste, elle ne s'appuie que sur ce qui existe déjà. L'argument de l'optique "comparative" ne fonctionne donc qu'à "structures organisationnelles données". Il ne répond pas à la question de l'origine de ces structures. L'optique comparative ne donne donc pas à la théorie le caractère opératoire qui lui permettrait de comprendre comment les organisations émergent et évoluent.

433- L'obscure origine des organisations

Ces remarques critiques visent à promouvoir une conception plus réaliste du phénomène organisationnel. De notre point de vue, elles permettent surtout de mettre en évidence une faiblesse théorique fondamentale de l'analyse de Williamson : son absence de dimension pratique. Williamson propose en effet une justification convaincante du rôle des organisations, mais son mode de justification abstrait ne lui permet pas de décrire leur émergence et leur évolution dynamique. La question fondamentale qui se pose dans l'optique micro-analytique de Williamson, puisque la "*transaction est l'unité de base de*

²⁸ Goldberg (1985, p.395) indique ainsi : "*Le concept de coût de transaction a prouvé son inutilité et son caractère trompeur, et il serait préférable de l'abandonner*".

l'analyse”, est en effet de comprendre comment des agents égoïstes et limités parviennent à percevoir l'avantage qu'ils auraient à agir ensemble et à constituer des règles communes. Comment expliquer qu'ils acceptent de se mettre sous la coupe d'un seul d'entre eux qui supervise leur travail. D'où vient qu'ils s'engagent à accepter l'autorité de l'un des leurs ?

Il existe bien sûr de multiples réponses à cette question : les acteurs “perçoivent” les bénéfices d'une coopération mutuelle, ils sont culturellement et moralement enclins à se faire confiance et à agir ensemble, ils ont été éduqués pour respecter l'autorité... Le problème est que ces explications de nature pratique ne peuvent trouver place dans l'univers williamsonien. La justification de l'organisation réside d'une part dans l'égoïsme fondamental des agents, ce qui exclut tout rapport de confiance, et d'autre part dans leur manque de rationalité, ce qui exclut toute capacité à comprendre l'avantage de l'action collective. Elle nie, de plus, le rôle des institutions ou des “normes” permettant de conformer l'individu aux organisations.

Williamson oppose deux “niveaux” de rationalité : il existe une altérité fondamentale entre le mode de raisonnement des agents de la structure et l'esprit rationnel sur lequel cette structure elle-même est fondée. La rationalité qui fonde l'organisation est distincte de celle des agents : elle la transcende. La structure de gouvernance fixe le langage commun. Elle met en place des relations de confiance entre les agents. Elle permet l'émergence d'une rationalité collective qui s'oppose au comportement contre productif et conflictuel des agents. Comment dès lors, les agents pourraient-ils sortir d'eux-mêmes une telle structure ? La mise en opposition des deux concepts de rationalité individuelle et collective interdit la formulation de tout processus “spontané” d'émergence des organisations.

L'analyse de Williamson ne permet donc pas de comprendre comment les organisations émergent et évoluent. Les agents que décrit Williamson sont en effet incapables de s'accorder et de construire eux-mêmes les structures de gouvernance. La rationalité qui préside à la construction de celles-ci leur est en effet étrangère. C'est une forme de rationalité supérieure, collective, qui se définit précisément contre les défauts des agents. L'analyse des coûts de transaction peut sans doute justifier ce qui existe. Elle ne peut pas du tout, en conservant ses hypothèses comportementales, expliquer comment ce qui existe a pu émerger et comment cela pourrait évoluer. La disjonction entre le niveau de rationalité individuelle, qui guide l'agent, et le niveau collectif, qui guide les organisations, constitue un fossé infranchissable, une altérité fondamentale qui empêche la mise au point d'une théorie pratique de l'émergence et de la dynamique des règles organisationnelles.

Cette altérité, le rôle essentiel que joue le langage dans la théorie williamsonienne, ainsi que l'accent mis sur la fonction coercitive des organisations font de l'analyse de Williamson une théorie typiquement hobbienne.

5- Une théorie typiquement hobbiennne.

“La référence à la relation d'autorité comme moyen de combattre l'opportunisme constitue une redécouverte de l'analyse hobbiennne, circonscrite ici à la sphère économique. (...) Cette analyse [de Williamson] repose sur une articulation d'hypothèses de sur-socialisation et de sous-socialisation des agents identique à celle que l'on trouve dans *Léviathan*. L'efficacité du pouvoir hiérarchique à l'intérieur de la firme est accentuée, comme l'Etat du souverain est sursocialisé. Le marché ressemble à l'état de nature de Hobbes”. (Granovetter, 1985, p.494)

La théorie des règles organisationnelles que propose Williamson nous semble en effet être profondément hobbiennne. Elle l'est dans l'esprit : cette théorie de l'institution est d'essence calculatoire, analytique, anhistorique, et se veut définitive. Elle l'est dans la forme : l'argumentation déployée par Williamson met l'accent sur l'opportunisme des agents et sur leur rationalité limitée pour justifier l'existence d'une structure coercitive s'appuyant sur un langage et un mode de rationalité qui transcende les acteurs.

On sait que la théorie hobbiennne s'oppose à la tradition classique et scolastique d'analyse des règles collectives. Hobbes reproche à cette tradition d'être confuse et de ne pas à se dégager suffisamment de la contingence et de la complexité des rapports humains. Il faut ordonner le monde pour sortir de cet état chaotique. Or, on ne connaît, selon le précepte nominaliste de Hobbes, que ce que l'on construit : Hobbes propose donc de construire de toute pièce un état civilisé. Cet état sera ordonné parce qu'il reposera sur un langage construit selon une logique et un objectif distinct de la confusion propre aux interactions des individus. L'artifice qu'il convient de construire fait donc table rase de toutes les institutions existantes (y compris le langage) et propose une théorie calculatoire fondée sur un langage adapté aux objectifs de cet artifice. Chaque objet reçoit une qualification nette et univoque, et les objets sont ordonnés entre eux selon des rapports logiques indiscutables.

Le projet de Williamson est un calque du projet constructiviste hobbiennne, appliqué à l'économie. Williamson, comme Hobbes, rompt avec une tradition soucieuse de la diversité du réel, la tradition institutionnaliste et propose une théorie de l'ordre fondée exclusivement sur l'analyse logique des caractéristiques des individus.

Hobbes, on le sait, considère avec dépit la nature humaine comme encline au conflit. L'état de nature est un état de guerre de tous contre tous, dans lequel chacun est sous la menace de la mort violente. Le *contractual man* n'a rien à envier à l'individu hobbiennne. Il

est tout aussi opportuniste, tout aussi limité, et donc tout autant incapable de s'accorder avec les autres. L'état marchand, considéré comme un état sans règle, un état de nature, est marqué par l'absence de transactions de long terme, chacun ayant peur de l'avenir. Le marché n'est pas viable.

Pour pallier ces défauts de la nature humaine, Hobbes propose deux grands principes : chaque individu se défait de ses droits au profit du souverain qui, en échange leur assure paix et prospérité. En acceptant la loi du souverain les individus acceptent également son langage et sa logique qui ordonnent le monde sur le modèle d'une science calculatoire. C'est également le cas dans l'analyse williamsonienne : les individus doivent accepter d'obéir à l'organisation à laquelle ils se soumettent, celle-ci leur assure en échange un salaire adapté à leur effort. En entrant dans l'organisation, ils se plient à une logique et à un langage construits de façon à rendre optimale l'organisation interne. Williamson, comme Hobbes, placent la coercition comme première vertu du nouvel ordre ainsi créé. C'est parce que l'organisation contraint chaque individu que tous coopèrent. La confiance est restaurée, artificiellement, grâce à la peur de tous les individus face au Léviathan ou, chez Williamson, aux structures de gouvernance.

La logique du contrat que passent les individus avec le Léviathan repose sur un échange. Le Léviathan abandonne la poursuite de sa propre richesse et reçoit en échange le pouvoir. Dans l'état civil, les individus renoncent à tous leur droit de nature en faveur du souverain, qui en échange assure aux individus la paix civile, condition nécessaire de leur enrichissement. Le souverain troque la paix contre le pouvoir. Les individus troquent la liberté contre la richesse. Cet échange est collectivement bénéfique. l'enrichissement de ses sujets augmente la puissance du prince, sa capacité à faire la guerre, ce qui à son tour favorise leur enrichissement

Le même type d'échange existe dans l'organisation que décrit Williamson. Dans son analyse, le Léviathan prend la forme du groupe d'individus gouvernant l'entreprise, les actionnaires. Les autres salariés se plient à leurs décisions, et peuvent en échange profiter des moyens de production mis à leur disposition pour valoriser et tirer profit de leur force de travail. Les actionnaires, eux, ne veillent pas à leur propre rémunération mais à l'enrichissement de l'organisation. Leur pouvoir est en effet déterminé par la puissance de l'organisation qu'ils dirigent, et cette puissance dépend de la volonté qu'ont les membres de l'organisation de s'enrichir eux-mêmes en s'impliquant au maximum. La puissance de l'organisation augmente sa capacité à gagner des parts de marché dans la guerre qu'elle livre aux autres entreprises.

Cette mise en évidence de la structure profondément hobbiennne de la problématique de Williamson nous permet de mettre à jour le problème central auquel se trouve confrontée

son analyse. L'analyse hobbenne se donne comme une justification de l'absolutisme et du pouvoir centralisé et coercitif. De la même manière, la théorie de Williamson justifie la structure capitaliste des organisations. Cependant Hobbes ne parvient pas à expliquer comment l'état civil peut émerger : le contrat social qu'il prône ne peut en effet apparaître que lorsque l'état civil est déjà constitué et que les individus peuvent comprendre sa logique. Cette logique leur est inaccessible lorsqu'ils sont dans l'état de nature. Hobbes démontre la nécessité de l'état civil à partir des "caractéristiques" des individus. Mais il ne parvient pas à expliquer comment cet état pourrait émerger à partir d'un comportement volontaire de ces mêmes individus. Comment les individus pourraient-ils rédiger un contrat les faisant passer de l'état de nature à l'état civil, alors même qu'ils ne peuvent devenir rationnels qu'une fois qu'ils sont situés dans l'Etat civil ? Le basculement d'un état à l'autre est fondamentalement obscur. En cela la théorie de Hobbes manque d'un volet "pratique" : sa démonstration est, en pratique, incorrecte. On a précisément cherché à pointer un problème identique dans l'analyse de Williamson : comment expliquer que les individus se plient à l'organisation ?

Cette caractéristique hobbenne de la théorie de Williamson pose un problème théorique dans son articulation à l'économie politique moderne, c'est-à-dire la théorie néoclassique. Hobbes propose en effet une théorie fondée explicitement sur la nécessité d'imposer aux individus une force transcendante, coercitive, centralisée. Sa théorie est de ce point de vue à l'antithèse du projet de l'économie politique, qui prétend justement faire émerger l'ordre à parti du choix libre d'agents rationnels. Williamson en marchant dans les pas de Hobbes, ne prolonge donc pas du tout le projet de l'économie politique : il le clôt. Il met à jour une contradiction fondamentale entre l'égoïsme des individus et la nécessité de coopérer, ou, tout au moins, de se coordonner. Cette mise à jour est une pierre jetée dans le jardin de l'analyse économique néoclassique. Pour que cette théorie de l'institution économique soit acceptable, il faudrait en effet que l'institution émerge du choix des agents eux-mêmes. Dans le cas inverse, on est en face d'une contradiction profonde avec le principe méthodologique de l'individualisme.

Conclusion

Williamson propose une théorie qui prend au sérieux les critiques d'Arrow et répond très précisément à la question du choix des structures institutionnelles adéquates. C'est donc une théorie qui non seulement met en évidence le rôle des règles collectives, mais qui, de plus, décrit des règles d'une nature particulière : les règles organisationnelles. Les règles organisationnelles naissent de la nécessité non seulement de la *coordination* des individus, mais plus encore de leur *coercition*. Elles constituent donc une réponse optimale au problème que constitue la rencontre d'individus instrumentalement rationnels. Pour que cette forme de rationalité n'ait pas de conséquences collectivement désastreuses, il est nécessaire que les individus soient placés dans un cadre limitant leur marge de manœuvre. L'organisation crée de l'ordre en leur imposant un comportement précis, profitable à la collectivité.

Compte tenu du primat de l'efficacité que Williamson impose comme unique critère du choix des règles collectives, il n'existe pas d'autre alternative. La règle organisationnelle est donc de type particulier, puisqu'elle est d'une certaine manière nécessaire. C'est une forme de loi. Ce n'est cependant plus une loi "naturelle" qui s'imposerait aux individus, mais une loi qu'il faut construire et faire advenir en construisant les institutions adéquates. Williamson assume donc une position constructiviste. Il démontre que la collectivité, si elle souhaite être efficace, n'a pas le choix. L'organisation hiérarchique est la structure institutionnelle nécessaire. Cette nécessité élimine à nouveau la question du choix des règles. La réponse de Williamson à la question du choix des règles collectives est une réponse conçue comme définitive. Une fois mise en place, l'organisation s'auto-régule. Elle produit les règles permettant d'assurer sa pérennité et par voie de conséquence l'efficacité collective.

Cette théorie de la règle souffre cependant d'un défaut majeur aux yeux de l'économiste : la règle ne peut ici émerger qu'à partir de structures collectives transcendant l'individu. L'agent ne produit donc pas les règles, il s'y plie. Or, cette conception de l'articulation de l'individuel et du collectif est en contradiction avec le projet fondamental de l'économie politique qui est de faire émerger l'ordre à partir des interactions des agents. La modalité williamsonienne d'intégration du concept de règle dans l'économie, bien qu'elle respecte les hypothèses fondamentales du paradigme néoclassique, le fait néanmoins éclater en introduisant une disjonction entre la rationalité collective, qui préside au choix des règles de coordination, et la rationalité des individus qui, seule, mènerait au chaos. Elle ressort le démon hobbiien de la "boîte noire" où l'économie politique avait cherché à le contenir.

Il existe cependant une autre modalité de règles collectives qui ne repose pas sur une opposition de l'individuel et du collectif, et qui semble être compatible avec les hypothèses comportementales "standard". Cette modalité est précisément celle que Hume oppose à l'idée hobbiennne d'un contrat social rationnel : c'est le concept de convention. Il a été reformulé dans les termes modernes de la logique formelle par le philosophe David Lewis. Il a été ensuite "utilisé" par la théorie des jeux dans une optique néoinstitutionnaliste. L'idée est ainsi de chercher à résoudre le problème de Hobbes, qui resurgit dans la théorie de Williamson, en faisant de la convention une modalité de l'ordre collectif émergeant *directement* des interactions individuelles sous la forme d'un équilibre de Nash.

Chapitre 4

*L'analyse stratégique de
l'émergence des règles*

L'analyse économique standard s'est construite autour de l'étude du comportement d'un individu rationnel. L'intégration du concept de règle institutionnelle dans son corpus théorique requiert donc qu'il existe une théorie expliquant l'émergence et la dynamique des règles institutionnelles à partir de l'action rationnelle d'individus libres. On a pu constater que l'analyse de Williamson ne respectait pas cette partie du "cahier des charges" : Williamson étudie le rôle des règles hiérarchiques, mais n'explique pas pourquoi ces règles, qui sont nécessaires au niveau *collectif*, pourraient émerger à partir de l'action d'*individus* économiquement rationnels.

L'objet de ce chapitre est d'étudier dans quelle mesure la théorie des jeux permet de pallier cette lacune. L'étude du rôle et de l'émergence des institutions est en effet au centre d'un ensemble de travaux qui utilisent la théorie des jeux afin d'étudier les mécanismes reliant l'existence de règles collectives à l'action stratégique d'individus rationnels. Nous concevons donc ces travaux comme étant complémentaires du programme de recherche de l'économie des coûts de transaction. Toutefois, il s'agit plus d'un ensemble de recherches convergentes que d'un mouvement théorique organisé.

Le "problème de la règle" peut être formulé en termes d'interactions stratégiques. Nous étudierons quelques-uns des cas "paradigmatiques" permettant de traduire, en termes stratégiques, la question du rôle et du choix des règles collectives. Sur cette base, on peut distinguer deux types d'attitude. La règle institutionnelle peut être analysée comme un moyen collectif mis en place afin de sélectionner une issue plus intéressante, du point de vue de la rationalité économique individuelle, que celle(s) qu'un comportement strictement solitaire nous permettrait d'atteindre. C'est la position fonctionnaliste que défend Ullmann Margalit. Mais on peut également aller plus loin, en

considérant que l'institution émerge réellement de l'interaction stratégique sous la forme d'un équilibre stratégique. C'est en particulier cette seconde position qui nous intéressera ici.

La plupart de ces travaux font référence au concept de *convention*, redéfini par le philosophe David Lewis. Le concept de convention permet en effet de rendre compte des règles collectives sans faire appel ni à la contrainte, ni au contrat social, deux moments théoriques "hobbiens" difficiles à intégrer dans un paradigme individualiste. Lewis montre que la nécessité de la règle apparaît spontanément à partir de l'action rationnelle d'individus libres et égoïstes. Mais son analyse se situe à un niveau logique qui exclut toute temporalité et, par là même, toute véritable théorie de l'émergence.

L'utilisation du concept de convention comme fondement des normes et institutions collectives nécessite donc des adaptations. Il faut, d'une manière ou d'une autre, intégrer une dimension temporelle permettant de saisir le processus de l'émergence d'une convention particulière. C'est ce que tentent de faire, de façon différente, Schotter, Choi, Axelrod ou Young. Nous étudierons leurs théories : toutes, à des degrés divers, cherchent à expliquer l'émergence de la règle en supposant *a priori* l'existence d'une certaine conformité chez les agents rationnels. Or, une véritable théorie de l'émergence devrait nous permettre précisément de comprendre comment cette "unanimité" est obtenue.

1- L'analyse de l'institution en théorie des jeux.

“En tant qu'économiste, je pense que nous devons simultanément expliquer à la fois la structure institutionnelle et l'ensemble des valeurs constituant un équilibre dans le cadre de cette structure institutionnelle. Se contenter d'étudier un ensemble de prix d'équilibre qui émergent dans un cadre institutionnel donné ne saurait être suffisant. (...) L'intention originale de Von Neumann et Morgenstern était précisément de créer une théorie mathématique des institutions. La question à laquelle ils se confrontèrent dans leur ouvrage : “Theory of Games and Economic Behavior”, est de savoir quel type d'institution était susceptible d'émerger comme ensemble de relations stables à partir d'une situation d'interaction sociale décrite sous la forme d'un jeu.” (Schotter, 1983, p.675,676).

Le plaidoyer de Schotter en faveur d'une analyse des institutions à partir de la théorie des jeux opère clairement un lien entre les deux spécificités de la théorie des jeux au regard de la théorie néoclassique du marché : l'analyse des interactions directes entre individus et un concept de solution conçu comme l'adoption d'une règle de comportement optimale. La théorie des jeux cherche en effet à comprendre comment les individus agissent dans le cadre d'interactions spécifiques, représentées comme des situations de jeu. De ce fait, la “solution” de ces interactions peut être analysée comme constitutive d'une “régularité de comportement” adoptée par des individus rationnels, ce qui est une autre manière de définir l'institution¹. Dans cette première optique l'institution est un équilibre de Nash qui émerge à partir d'interactions directes entre agents rationnels.

Cependant, le concept d'institution n'a d'intérêt que lorsqu'il désigne autre chose que le comportement strictement solitaire d'un individu rationnel. Un comportement est “institutionnalisé”, dans le cadre de la théorie des jeux, lorsqu'il désigne le choix par un ensemble d'individus d'un comportement collectif leur permettant d'atteindre une issue plus intéressante que ce que chacun d'eux aurait obtenu en luttant seul contre tous les autres. On peut donc faire apparaître le problème de la règle institutionnelle dans une classe particulière de jeux : les jeux à somme non nulle, qui nécessitent l'adoption d'une règle de comportement commune permettant d'exploiter au maximum les potentialités de la situation d'interaction. C'est à Schelling que revient le mérite d'avoir mis en évidence l'intérêt de ces jeux. Après avoir présenté son analyse, nous tenterons d'isoler quelques uns des cas paradigmatiques symbolisant les problèmes de la coopération et de la

¹ Schotter propose en fait de substituer le terme d'institution aux termes exacts utilisés par Von Neumann et Morgenstern [1947, p.43] qui sont : “standard of behaviour” et “orders of society”.

coordination, avant de proposer une expression générale du problème que résout l'adoption d'une règle de comportement institutionnel dans un cadre stratégique.

Ullmann Margalit [1978] propose précisément une typologie des différents types de normes associées aux problèmes de coordination (au sens large). Cette typologie permet de faire apparaître le problème de la coercition des individus, qui se pose en particulier lorsqu'un problème de coopération vient s'ajouter à celui de la coordination. Cependant, Ullmann Margalit ne cherche pas à expliquer *comment* les règles émergent, elle propose une réflexion fonctionnaliste visant à démontrer qu'il "faut" qu'une norme existe pour régler un problème qui, précisément, dépasse le simple cadre de l'interaction stratégique. La dimension coercitive des normes assurant la "coordination" des individus semble de fait rendre impossible une analyse de leur émergence "spontanée". Sur le plan logique et atemporel, la norme semble donc devoir être "imposée" aux acteurs.

Cependant, l'introduction d'une dimension temporelle dans le processus d'interaction permet de dénouer les conflits. L'attitude coopérative est plus rentable à long terme, elle devient une des stratégies viables. Or, si le problème de la coopération se dénoue, seul celui de la coordination continue de se poser. C'est sur cette base que la théorie des jeux pourra se ressaisir de la question de l'émergence des règles, afin d'endogénéiser ce qu'Ullmann Margalit situe résolument à l'extérieur de la théorie. Dans ce processus d'endogénéisation, le phénomène central est la formation d'attentes réciproques, et de croyances communes.

11- Le cadre d'analyse

Nous étudierons la réponse que les théoriciens des jeux apportent à la question de la genèse de l'institution dans un cadre *non coopératif*. Dans ce cadre, le problème de l'institution est analysé à travers l'étude d'un certain type de jeu que Schelling [1960] qualifie de "jeux de négociation". Dans ces jeux, en effet, les joueurs ont des intérêts communs, et sont donc amenés à se poser la question de la coordination de leurs actions.

111- Une approche non coopérative

On distingue traditionnellement les jeux non-coopératifs et les jeux "coopératifs". Les jeux coopératifs sont des jeux dans lesquels les acteurs peuvent passer des "*accords*

contraignants"². Dans ces jeux, les individus sont donc rationnels au sens standard, mais ils ont de plus la possibilité de passer des accords que chacun *doit* par la suite respecter. Cette distinction précise ne doit cependant pas masquer la tendance hégémonique de l'analyse dite "non-coopérative". En effet, dans l'esprit même de Nash, qui en a pourtant proposé le principe (Nash, 1950 et 1953), l'approche coopérative n'est qu'une étape préalable à l'étude des processus non coopératifs menant à une issue coopérative³. L'axiomatique permettant de définir une solution collectivement rationnelle devrait ensuite pouvoir être "endogénéisée" pour apparaître comme le fruit d'une négociation interne au jeu. Telle est en tout cas, dans la lignée du "programme de Nash", la voie qu'ont tracée Harsanyi [1966], Aumann [1987] et Harsanyi-Selten [1988]⁴. La théorie des jeux coopératifs est donc fortement concurrencée par une "approche non-coopérative de la coopération"⁵.

En ce qui concerne la question de l'émergence des règles collectives, il semble bien que l'approche coopérative ne présente pas un intérêt supérieur à l'approche non coopérative. La spécificité et l'apport de l'optique purement coopérative tiennent à l'utilisation d'une démarche axiomatique qui permet de poser le problème de l'articulation de la justice et de l'efficacité collective. Cependant, une telle démarche ne nous permet pas de comprendre comment en pratique la solution du jeu peut émerger. La réflexion du modélisateur consiste donc à définir un certain nombre d'axiomes permettant de caractériser une solution acceptable par les deux joueurs. Cette solution devrait logiquement constituer l'issue prévisible du marchandage. Le modélisateur "mime" en quelque sorte la négociation entre les joueurs, en cherchant à se mettre à leur place et à définir un plan de partage que tous devraient accepter. De ce fait, la théorie des jeux coopératifs est essentiellement normative et a très peu de contenu "pratique" : le processus

² On doit cette définition à Aumann [1987] qui l'attribue à Harsanyi [1966]. Harsanyi et Selten [1988,p.1] indiquent que cette distinction est plus nette que celle de Nash qui incluait également la possibilité de communiquer comme critère distinctif.

³ Le programme de Nash est par exemple présenté par Binmore et Dasgupta [1987, chap 1]. Pour une présentation plus générale des rapports entre jeux coopératifs et non coopératifs, on peut également se reporter au survey proposé par Binmore (Binmore et Dasgupta, 1986, chap 1).

⁴ Contre cette approche dominante, Moulin [1995] défend le principe d'une spécificité des jeux coopératifs qui reposeraient sur un modèle de joueur "coopératif" différent du joueur standard des jeux non coopératifs. Sa démarche semble être porteuse d'un réel renouvellement de la figure du "joueur" en théorie des jeux.

⁵ On distingue en fait plusieurs concepts de solution en théorie des jeux coopératifs qui sont plus ou moins réductibles à une approche non-coopérative. On peut ainsi distinguer [Moulin (1981), Mas-Colell (1989), Shubik (1991) Myerson (1995)] : le cœur, la valeur de Shapley, la solution de Nash, les ensembles stables ("stable sets"), l'ensemble de négociation, le noyau, le nucléolus. Mas-Colell (1989) propose une typologie de ces différentes solutions selon qu'elles mobilisent le principe stratégique de domination (la solution qui émerge est celle qui assure le paiement maximum aux joueurs) ou qu'elles reposent sur une axiomatique définissant une issue raisonnable. Ce sont les solutions axiomatiques, du type "valeur de Shapley" ou "solution de Nash", qui sont spécifiquement coopératives.

de négociation est le plus souvent ignoré⁶ et c'est la réflexion du théoricien qui remplace l'analyse des rapports de négociation des acteurs. La règle collective qui permet d'atteindre une issue stable est donc entièrement déduite d'une série d'axiomes posés *a priori*, elle n'émerge pas de la confrontation entre individus rationnels. La théorie des jeux coopératifs ne prétend pas expliquer *comment* les règles émergent, elle cherche uniquement à prévoir *quelles* règles de partage *devraient* émerger. Elle ne traite donc nullement de la question pratique de la constitution des règles collectives, comme processus émergeant à partir de la confrontation d'individus rationnels. Dans notre optique, il ne nous a donc pas semblé nécessaire de nous départir de l'approche "standard" non-coopérative.

Le choix de l'approche non-coopérative entraîne, d'une certaine manière le choix d'un cadre excluant *a priori* une coordination par le langage. Dans un cadre non-coopératif, les paroles n'ont aucun poids. Il n'est donc pas utile de considérer que les individus peuvent se parler⁷. D'une part cette hypothèse représenterait une perte de généralité, puisque dans bien des cas la communication directe est impossible. Ensuite, elle reviendrait à éliminer le problème premier que pose la coordination, c'est-à-dire la constitution d'un langage commun non équivoque. Le choix du cadre non-coopératif entraîne donc naturellement l'hypothèse d'une absence de communication pas le langage. Cette attitude théorique permet d'isoler précisément le problème de la coordination et du choix des règles, sans supposer qu'il est déjà en partie réglé, ce qui serait le cas si les individus disposaient d'un langage commun assurant parfaitement l'intercompréhension. On posera donc le problème de l'émergence des règles institutionnelles, non seulement dans un cadre non coopératif, mais également dans un cadre où les individus ne peuvent pas se coordonner en se parlant. C'est précisément l'absence de langage qui permet de poser le problème du choix d'un comportement permettant de s'accorder avec les autres. Dans ce cadre on mettra l'accent un type particulier de jeux dont les caractéristiques permettent de poser le problème de la règle : les jeux de coordination dont Schelling [1960] a initié l'étude.

112- L'analyse de Schelling

"Les jeux à somme non nulle sont des jeux dans lesquels, bien que le conflit représente l'élément dramatique, la dépendance mutuelle est intégrée à la structure logique et requiert donc une forme de collaboration ou d'accommodement. (...) La théorie des jeux traditionnelle a appliqué aux jeux caractérisés par l'existence d'une forme de

⁶ C'est l'argument que développent Harsanyi et Selten [1988, p.9] contre la branche "coopérative" de la théorie des jeux. On peut également se reporter à Schotter [1983, p.677-683].

⁷ On reprend ici l'argument qu'utilise Schelling [1960, p.67], ou encore Harsanyi et Selten [1988, p. 31,359] pour justifier cette hypothèse.

dépendance mutuelle les méthodes et concepts qui s'étaient avérés efficaces dans le cadre des jeux de pur conflit. (...) Cette étude tente d'élargir le cadre de la théorie des jeux en faisant apparaître les jeux à somme nulle comme un cas limite." (Schelling, 1960, p.83)

Schelling [1960] propose dans un ouvrage de redéfinir la ligne de partage qui s'est rapidement instaurée entre les jeux à somme nulle et les jeux à somme non nulle. Il s'oppose ainsi à la prédominance de l'étude des jeux à somme nulle, dans lesquels les intérêts des joueurs sont strictement opposés. L'intérêt des jeux à somme nulle est qu'ils se prêtent à un traitement mathématique. La stricte divergence des intérêts des joueurs permet en effet de mettre au point des techniques de résolution visant à éliminer le problème de l'*interdépendance* des joueurs. C'est le cas par exemple de la stratégie minimax : "*Avec une solution minimax, un jeu à somme nulle est réduit à un problème tout à fait unilatéral (...). C'est un moyen d'expurger le jeu de tous ses aspects non mathématiques et d'enlever aux joueurs toute possibilité de communication*" (Schelling, 1960, p.105). Dans les jeux à somme nulle, l'interdépendance des acteurs est un problème que le modélisateur cherche à éliminer en "paramétrisant" la situation stratégique à laquelle est confronté l'acteur. Dans une situation de pur conflit, l'individu rationnel doit s'abstenir de compter sur son partenaire. Il agit donc indépendamment de l'autre. Il considère que l'attitude de l'autre lui est donnée, compte tenu de la situation simple de "pur conflit". En ce sens, l'individu dans le cadre des jeux à somme nulle adopte une attitude "paramétrique". L'équilibre de Nash symbolise précisément une situation dans laquelle la stabilité du comportement des deux joueurs est directement liée au fait que chacun considère l'attitude des autres comme une donnée.

Schelling considère que les jeux à somme nulle devraient être considérés comme un cas extrême. Il lui oppose un autre extrême : le cas où les intérêts des joueurs sont strictement convergents. Il considère que les jeux représentant ce type de situation peuvent être qualifiés de "*jeux de coordination*". Cette bipolarisation permet de faire apparaître un continuum de jeux dans lesquels les individus ont des intérêts communs, mais sont également en conflit en ce qui concerne le choix entre les issues conjointes envisageables. Ces jeux mêlant conflit et intérêts communs sont les plus courants selon Schelling. Il les qualifie de "*jeux de négociation*". A travers cette redéfinition des différents types de jeux Schelling tente de faire apparaître un problème oublié de la théorie des jeux "standard", parce qu'il n'est pas d'essence mathématique. En effet, en dehors des jeux de pur conflit, tous les autres jeux représentent des situations dans lesquels les joueurs ont des intérêts communs. Or, lorsqu'il existe une similitude d'intérêt entre les acteurs, il devient absurde de ne pas tenter de se *coordonner* avec l'autre joueur pour profiter de cette opportunité.

Le problème de la coordination, qui consiste à atteindre une issue conjointe, apparaît dans tous les jeux à somme non nulle. Mais les “jeux de coordination” permettent de l'exprimer dans toute sa pureté, puisqu'ils excluent tout conflit⁸. Les jeux de coordination se caractérisent par le fait que bien qu'il existe des équilibres de Nash, aucun d'entre eux ne se distingue des autres. Les joueurs, ne disposant pas de stratégie dominante ni de stratégie *maximin* leur permettant de sélectionner un des équilibres, n'ont aucune façon “naturelle” de jouer pour les atteindre : “*Ce type de jeux admet de nombreux équilibres, et la théorie ne nous aide pas à savoir si l'un d'entre eux est la solution, et, le cas échéant, lequel c'est.*” (Kreps, 1990b, p.97)⁹. Les joueurs sont donc dans une situation d'incertitude complète, à ceci près que chacun sait qu'il a intérêt à faire comme l'autre (quel que soit le choix de l'autre). Dans un tel contexte, l'application de stratégies issues de situations de pur conflit est inefficace puisque ces stratégies consistent à nier l'interdépendance des acteurs, alors même qu'il convient ici de l'exploiter. En effet, lorsqu'il existe une convergence d'intérêt, même partielle, l'interdépendance devient source de gains mutuels des acteurs. Pour “gagner”, il faut alors parvenir à atteindre une issue commune mutuellement bénéfique. Il faut donc comprendre l'autre joueur, se mettre à sa place, voire l'inciter à jouer comme nous le souhaiterions. Il existe cette fois une volonté de se *coordonner* avec les autres.

L'interdépendance des joueurs pose un problème spécifique dans un cadre stratégique, puisque pour agir, il est nécessaire, en quelque sorte, de savoir déjà comment l'autre a agi. Or, le comportement de notre adversaire/partenaire dépend étroitement de la manière dont celui-ci anticipe notre propre comportement. Ainsi, ce que l'on découvre dans le regard de l'autre, c'est le reflet de notre propre hésitation. L'interdépendance, dans un cadre stratégique, nous projette donc dans un jeu de miroirs, sur le modèle de la mise en abîme. Chacun se reflète à l'infini dans le regard de l'autre.

Schelling [1960] estime pourtant que l'on peut s'extraire de cette spirale sans fin grâce à l'existence de points de repères communs aux joueurs. Ainsi, des amis perdus dans une ville identifient en général un point de rencontre dont ils pensent qu'il fait l'objet d'une

⁸ L'intérêt des jeux de *pure coordination* est d'isoler un problème, celui de la coordination, qui se retrouve dans tous les jeux à somme non nulle : “*les jeux de coordination au sens strict, bien qu'intéressants en eux-mêmes, sont surtout utiles pour l'éclairage qu'ils apportent quant à la nature des jeux avec intérêts mixtes.*” (Schelling, 1960, p.99).

⁹ Luce et Raïffa [1957, p. 106-107] expriment le problème ainsi : aucun des équilibres de Nash n'est solution du jeu, puisqu'ils sont tous équivalents (aucun n'assure à tous les joueurs un paiement supérieur à un autre) mais pas interchangeables [(c1,r1) et (c2,r2) sont des équilibres, mais pas (c1,r2) et (c2,r1)]. Kreps [1990a, p. 411] rappelle la distinction essentielle entre l'idée d'équilibre et celle de solution. Sa propre terminologie, que nous reprenons ici, vise à montrer qu'une “façon naturelle de jouer” inclut forcément un équilibre de Nash, mais que l'inverse n'est pas vrai. Cette terminologie permet de distinguer l'existence de l'équilibre, et la manière d'en faire une solution, c'est-à-dire de l'atteindre, ce qui pose le problème de son accessibilité et de sa stabilité.

anticipation convergente (par exemple : la gare centrale). Ce sont ces issues, possédant un magnétisme intrinsèque, que Schelling qualifie de “*points focaux*” (Schelling, 1960, p.57). Le point important de son analyse est de montrer que la solution repose ici au moins autant sur les facultés d'imagination et d'observation des acteurs que sur leurs capacités logiques : “*la logique est utile, mais pas avant que l'imagination nous ait permis de sélectionner un point particulier s'appuyant sur les particularités concrètes de la situation*” (ibid, p.58). On ne s'extirpe d'un problème de coordination qu'en faisant preuve de sens *pratique*. La logique rationnelle ne suffit pas pour agir efficacement.

Avec ce concept de point focal, Schelling met donc en évidence l'importance du *contexte* de jeu, et opère un lien essentiel entre ce contexte et la structure des croyances des joueurs. Les anticipations des joueurs s'appuient sur leur perception de la situation, qui n'a aucune raison d'être identique pour tous. La capacité des joueurs à se coordonner dépend alors de l'existence de “repères communs” qui leur permettent d'avoir des “croyances” et donc des “anticipations” communes. L'analyse de Schelling tend donc à faire dépendre l'existence d'une solution dans les jeux de coordination de l'existence d'une structure collective permettant d'harmoniser les anticipations. Cette intuition théorique est à l'origine du mouvement par lequel la théorie des jeux s'est réappropriée la question du rôle des institutions : “*des jeux de coordination se situent probablement derrière la question de la stabilité des institutions sociales*” (Schelling, 1960, p.91).

Schelling propose, de plus, une conception dynamique des points focaux. Il décrit ainsi le mécanisme des “coups stratégiques” dont l'esprit est bien résumé par l'idée suivante : “*un coup stratégique est un coup qui influence le comportement de l'autre en un sens qui nous est favorable, en modifiant les anticipations de l'autre à l'égard de notre propre conduite. On contraint ainsi son partenaire en se contraignant soi-même*” (Schelling, 1960, p.160). Avec ce concept, Schelling entend rompre avec une représentation “passive” de l'individu rationnel. Il n'y a aucune raison évidente pour considérer que les agents ne cherchent pas, à travers leur “coup”, à modifier la situation de jeu à leur avantage, en modifiant les anticipations des autres joueurs plutôt que de considérer celles-ci comme des données immuables. A travers leurs coups, les joueurs parviennent ainsi à communiquer entre eux. C'est à travers cette communication “par les actes” que les points focaux peuvent évoluer. Schelling lie le fait de reconnaître l'importance du contexte de jeu à la nécessité de prendre en compte l'attitude active des joueurs visant à modifier ce contexte¹⁰.

Schelling a mis l'accent sur les situations dans lesquelles la question de la coordination, et donc de l'institution, se trouvait posée. Cependant, son analyse met

¹⁰ Les analyses de Kreps (1982, 1984, 1990) en terme de réputation, de menace, de signalement s'inscrivent dans l'optique de Schelling. De même Ponssard [1994, 1995] développe une théorie de la “forward induction” dont l'objectif est de mettre en avant l'ancrage empirique des jeux.

surtout l'accent sur le "contexte" du jeu et l'attitude des joueurs "en dehors de leur rationalité". Schelling situe donc l'institution en dehors de ce qui constitue le jeu proprement dit. Or, le projet de la théorie des jeux, tel qu'il est exprimé par Schotter [1983], vise précisément à *endogénéiser* l'institution. Schelling [1960] formule, en quelque sorte, un programme de recherche. Il identifie un comportement qui est "gagnant" dans une large classe de jeu et que la théorie n'est pas encore parvenu à intégrer.

12- Les situations paradigmatiques qui nécessitent une règle

Schelling montre que, dans le cadre de la théorie des jeux non coopératifs, la question de la règle peut être abordée à travers l'étude de certaines situations dans lesquelles l'adoption d'une attitude rationnelle et solitaire conduit à une issue pareto-dominée. Dans certains cas, il serait individuellement rationnel de s'entendre avec les autres, pour tirer parti de toutes les potentialités de la situation. Il est possible de préciser les définitions générales de Schelling en identifiant quelques situations typiques posant chacune un problème spécifique.

Parmi ces figures paradigmatiques, deux s'imposent comme des cas polaires représentant d'une part un pur problème de coordination et d'autre part un pur problème de coopération. Ce sont respectivement le "*jeu du rendez-vous*" et le "*dilemme du prisonnier*". Cependant au-delà de ces deux cas polaires, les situations les plus courantes mêlent les problèmes de coordination et de coopération. On le verra en particulier à travers l'étude du jeu de "*la guerre des sexes*". On cherchera donc à proposer une qualification commune de l'ensemble de ces situations qui, toutes, posent la question de l'accord collectif sur l'adoption d'une règle de comportement commune.

121- Le problème de la coordination : l'exemple paradigmatique du "jeu du rendez-vous"

L'exemple typique du premier problème est le "*jeu du Rendez Vous*", sur lequel s'appuie en particulier Lewis [1969] et qui est, à l'origine, proposé par Schelling [1960, p.78] : un mari et une femme perdent contact l'un avec l'autre dans un grand magasin. Comment vont-ils se retrouver ? Lewis [1969] formalise ainsi (sous forme normale) les différentes issues de ce jeu :

| | | ELLE | |
|-----|------------|---------|------------|
| | | légumes | spiritueux |
| LUI | légumes | 1 | 1 |
| | spiritueux | 0 | 1 |

fig.1 : un jeu de pure coordination

Le jeu est ici représenté de manière à ce que l'incertitude sur le lieu de rendez-vous soit réduite à deux issues possibles : le rayon "légumes" ou le rayon "spiritueux". Or, même dans ce cas, chacun de nos deux époux risquent d'arpenter longuement les couloirs de leur supermarché favori à la recherche l'un de l'autre. En effet, rien, d'un point de vue strictement rationnel, ne permet *a priori* d'établir une stratégie pure gagnante dans ce jeu : il n'existe pas de stratégie dominante. Pourtant, les désirs des joueurs sont ici parfaitement convergents : peu importe où ils se retrouvent, dès lors qu'ils se retrouvent. Chacune des issues conjointes forme un équilibre de Nash. Les deux équilibres de Nash sont ici parfaitement équivalents, que ce soit du point de vue de la rationalité individuelle ou du point de vue de la rationalité collective. Lewis [1969] parle pour cette raison de *jeu de pure coordination*.

L'intérêt de ce jeu est en effet d'isoler un problème particulier : celui de la coordination. Il n'y a pas ici de "conflit" entre les deux joueurs, ni de lien entre l'égoïsme de chacun et la difficulté à atteindre une issue commune. Ce n'est pas de l'irréductible conflit des objectifs de chacun que naît le problème, mais simplement d'un manque de repère collectif. A première vue le problème n'est donc pas économique, il n'y a pas de problème de "partage" des ressources. Pourtant, il faut bien admettre que si les individus agissent uniquement de manière économiquement rationnel, il est fort probable qu'ils n'atteignent *aucun* des deux équilibres de Nash. Une attitude rationnelle ne suffit donc pas à atteindre une des issues non pareto-dominées. La rationalité économique n'assure pas à elle seule l'efficacité collective.

122- Le cas de la coopération : l'exemple du dilemme du prisonnier.

Le second cas "paradigmatique" est celui du célèbre "dilemme du prisonnier", proposé à l'origine par le mathématicien A.W.Tucker.

Manquant de preuve pour établir le forfait de deux brigands (Roger et Marcel), un juge

machiavélique propose à chacun d'eux le contrat suivant : s'il avoue sa participation au forfait commun, il sera relaxé et son comparse sera lourdement condamné (10 ans). Au cas où les deux comparses avoueraient, ils seraient également condamnés, mais ils bénéficieraient de la clémence (5 ans chacun). Au cas où ils se tairaient tous les deux, ils seraient condamnés pour port d'arme, soit une condamnation très légère (2 ans chacun). En forme normale, le jeu peut-être représenté ainsi :

| | | Roger | |
|--------|-------|----------|----------|
| | | Nie | Avoue |
| Marcel | Nie | -2 -2 | -10 0 |
| | Avoue | 0 -10 | -5 -5 |

fig.2 : le dilemme du prisonnier

La conclusion de ce dilemme est bien connue : la seule issue qui soit un équilibre de Nash est celle dans laquelle les deux complices avouent. Or, celle-ci est pareto-dominée par l'issue correspondant à la double négation. Les deux brigands auraient tous les deux intérêt à nier ensemble, c'est-à-dire à coopérer, mais la structure même du jeu ne le leur permet pas.

Plaçons nous du point de vue de Roger. Pour déterminer son comportement, Roger doit réfléchir à l'attitude que devrait adopter Marcel (qui est par définition un individu rationnel). Or, Marcel n'a pas à hésiter. S'il pense que Roger est prêt à nier, Marcel doit avouer pour être libéré. S'il pense au contraire que Roger va avouer, Marcel doit là aussi avouer, pour éviter, cette fois, d'être condamné seul et donc lourdement. Lorsque Roger fait ce raisonnement, il réalise donc que, quoique présume son complice Marcel quant à son propre comportement, ce dernier avouera. Dans ces conditions, Roger n'a pas non plus à hésiter : il doit lui-même avouer. En termes techniques : la stratégie de l'aveu est non seulement un équilibre de Nash, mais correspond également au choix de sa *stratégie strictement dominante* par chacun des deux joueurs. Dans une telle situation, les joueurs rationnels ne peuvent avoir aucune hésitation : l'issue "non coopérative" de l'aveu est la seule possible.

Le dilemme tient bien sûr au fait que cette issue n'est souhaitée par aucun des deux joueurs. S'ils pouvaient choisir de concert l'une des deux issues conjointes (la négation commune ou l'aveu commun), leur choix rationnel se porterait sur l'issue coopérative (la négation). Mais ils ne parviennent pas atteindre cette issue commune parce que chacun d'entre eux préfère encore avouer lorsqu'il est sûr que l'autre va nier. L'issue solitaire de

la défection domine nettement la stratégie conjointe de la coopération. L'équilibre conjoint de l'aveu commun est donc la conséquence malheureuse et inévitable d'un comportement rationnel et solitaire. La rationalité individuelle, en quelque sorte, se retourne contre elle-même¹¹.

Ce second cas paradigmatique met en évidence le problème de la *coopération*. En effet, les acteurs n'ont pas, *a priori*, de problème de coordination à résoudre, puisqu'il existe un seul équilibre de Nash, soutenu de plus par les stratégies dominantes de chacun des deux joueurs. En fait, ils préféreraient même ne pas se coordonner. Ils ont en effet des objectifs strictement opposés. L'issue préférée par chacun des deux joueurs est celle où l'autre purge seul la peine. Cependant, cette issue ne peut être atteinte par aucun des deux joueurs, s'ils sont tous deux rationnels. Les joueurs ne coopèrent pas parce qu'il est plus rationnel de tenter de faire défection, bien qu'ils sachent que cette attitude les mène inévitablement à une issue pareto-dominée. C'est seulement une fois que l'échec d'une attitude purement conflictuelle apparaît que peut se poser le problème de coordination. Puisqu'il ne peut spolier l'autre, chacun préférerait, faute de mieux, atteindre l'issue collective la plus favorable. Mais c'est impossible. La suspicion d'une volonté de coopérer (de nier) chez l'autre partenaire fait immédiatement renaître la possibilité de faire défection (avouer tandis que l'autre nie)¹².

Ainsi, il n'est pas possible d'atteindre l'issue conjointe la plus bénéfique parce qu'elle est dominée par la défection unilatérale. Pour atteindre l'issue qui est mutuellement la plus intéressante, il faut d'une certaine manière que chacun des deux joueurs accepte de sacrifier une part de ce que la situation lui permettrait d'obtenir et s'engage à coopérer de manière *irrévocable*. On entend donc ici par "coopération" le fait de se sacrifier pour atteindre une issue collectivement bénéfique (ce qui n'est envisageable que parce que l'issue naturelle du jeu est pareto-dominée). Pour que l'issue mutuellement souhaitée puisse être atteinte, il faut donc que les deux joueurs renoncent chacun de leur côté à faire cavalier seul en tirant parti de l'attitude coopérative de l'autre. Mais le renoncement au gain facile que représente la défection, dans un cadre où les autres coopèrent, n'est pas un comportement rationnel. Le dilemme du prisonnier isole donc d'une certaine manière le problème de la coopération. Si les joueurs pouvaient s'engager à être coopératifs, ou si on pouvait rendre impossible la défection, la coordination serait triviale puisqu'une des deux issues collectives domine l'autre.

¹¹Pour une étude complète de la signification du dilemme du prisonnier, on peut se référer à Cordonnier [1997].

¹² Le problème n'est donc pas non plus un problème de confiance en l'autre. Au contraire, si l'un des deux joueurs a suffisamment confiance en la volonté de coopérer de son partenaire, cela ne fera que renforcer sa volonté de faire défection.

123- Un troisième cas paradigmatique "intermédiaire" : la guerre des sexes.

Les deux premières situations "paradigmatiques" sont différentes l'une de l'autre. La première pose un problème de pure "coordination", la seconde pose le problème de pure "coopération". Dans la première en effet, les individus ont des intérêts strictement identiques, et, une fois atteinte, l'issue mutuellement satisfaisante est stable (c'est un équilibre de Nash). Dans la seconde au contraire, l'issue mutuellement satisfaisante est instable, puisqu'elle est toujours dominée par la tentation de duper son partenaire. Pour surmonter cette divergence d'intérêt, il faut que l'on s'engage à agir contre son propre intérêt, en renonçant à l'attitude de défection.

La situation est moins "pure" dans le troisième jeu paradigmatique connu sous le nom de "guerre des sexes" (pe : Luce et Raïffa [1957, p.90]). Un couple cherche à agrémenter sa soirée. Monsieur est plutôt attiré par un ballet, Madame par un match de football, mais tous les deux souhaitent avant tout sortir ensemble. Dans ce cas, les deux joueurs ont un intérêt à se coordonner, mais leurs préférences divergent quant au choix de l'équilibre de coordination. C'est donc un jeu de coordination, puisque chacune des issues possibles est un équilibre de Nash, mais avec un élément interne de conflit concernant le choix de l'un des deux équilibres. On peut le représenter sous la forme normale suivante :

| | | LUI | |
|------|----------|----------|----------|
| | | Football | Ballet |
| ELLE | Football | 2 1 | -1 -1 |
| | Ballet | -1 -1 | 1 2 |

Fig. 3 : Guerre des Sexes

Ce jeu pose un troisième problème qui est distinct à la fois de celui de la pure coordination et celui de la pure coopération. Car ici le problème n'est pas *seulement* de se coordonner, puisque il y a conflit quant à l'issue à atteindre. Il n'est pas non plus *seulement* de coopérer, puisqu'il y a deux équilibres de coordination également stable possibles. On notera toutefois qu'un comportement coopératif (qui correspondrait ici à l'attitude : "faisons ce qui te plaît mon(ma) cher(e)") résout de fait le problème de coordination. De même l'existence d'un repère mutuellement identifié (du type : "il n'y a pas de match ce soir") résout non seulement le problème de coordination mais encore celui de la coopération. La "guerre des sexes" est l'archétype de ce que Schelling qualifie de "jeux de négociation".

Ce jeu brouille donc la distinction entre *coordination* et *coopération* que les deux premiers cas paradigmatiques avaient permis d'opérer. Il n'est pas le seul. Il existe en fait un très grand nombre de *jeux à somme non nulles*, du "jeu du croisement" au "jeu du désarmement", mêlant des problèmes de coopération et de coordination et qui posent chacun des questions spécifiques¹³. Ainsi au-delà de la différence, importante, qui sépare les deux cas "paradigmatiques" que sont le dilemme du prisonnier et le jeu de pure coordination, il importe de relever les deux éléments communs à ces deux jeux "polaires", et à tous ceux qui se situent dans une position "intermédiaire" : il existe une issue "commune" nous rapportant plus que le paiement minimum garanti. Autrement dit, on ne peut pas y jouer en se fondant exclusivement sur la rationalité économique, alors même qu'il y a bien un intérêt économique en jeu. On fait individuellement et collectivement mieux en cherchant à s'accorder qu'en cherchant à jouer solitairement. C'est donc précisément le problème du choix d'une règle que la structure de ces jeux permet de poser. Dans ces situations, suivre une règle institutionnelle, qu'elle sélectionne une issue meilleure que l'équilibre de Nash ou qu'elle nous permette de sélectionner un des équilibres atteignables, est plus efficace que se fier exclusivement à sa rationalité. C'est à partir de ce constat qu'Ullmann Margalit s'est saisie de la théorie des jeux pour proposer une approche fonctionnaliste des règles et normes communes.

13- La typologie d'Edna Ullmann-Margalit

On peut opérer un premier lien explicite entre l'analyse des jeux de coordination au sens large (incluant le conflit) et le rôle des normes à travers l'analyse d'Edna Ullmann-Margalit. Cette philosophe s'est en effet appuyée sur les typologies que la théorie des jeux permet de mettre au point pour mettre en évidence le rôle des normes collectives. Son étude constitue donc une étape importante dans la mise en place d'une théorie néoinstitutionnelle des règles collectives.

131- Un projet structuraliste et fonctionnaliste

Edna Ullmann Margalit propose en effet une théorie générale des normes visant à mettre en évidence les structures dans lesquelles les normes émergent : "*Notre*

¹³ Cordonnier [1997, p.27-57] donne de nombreux exemples de jeu qui posent le problème de la coordination et de la coopération. On peut également se reporter à Walliser [1987] ou encore à Kreps [1990, p.95-107] qui passent en revue les cas typiques de situation où il existe "*trop d'équilibres, et aucun moyen de choisir*".

reconstruction rationnelle vise à décrire les caractéristiques formelles des situations d'interaction sociale dans lesquelles les normes peuvent émerger, ou encore dans lesquelles il y a de grandes chances qu'une norme émerge" (Ullmann-Margalit, 1977, p.1). Cette reconstruction rationnelle ne vise pas à expliquer comment, en pratique, les normes émergent. L'auteur renonce à cette explication "empirique" qu'elle juge trop complexe. Elle souligne que son étude se situe sur le plan de la logique formelle et que c'est pour cette raison qu'elle mobilise la théorie des jeux. Elle souligne cependant : "le cadre de la théorie des jeux est trop étroit, et donc inadéquat, pour expliquer comment les normes sont générées" (op.cit., p.6). Ullmann-Margalit ne s'intéresse donc qu'au plan de la logique formelle. Son analyse de l'émergence est donc *structurelle* et non pas *génétique* : "De la même manière qu'il y a eu un glissement dans la théorie du langage d'une explication génétique à une explication structuraliste, il doit y avoir un glissement identique dans l'étude des normes" (op.cit., p.8).

Dans cette optique, elle propose une typologie visant à spécifier trois situations "caractéristiques" dans lesquelles des individus rationnels sont confrontés à un problème de *coordination*. Ces trois situations correspondent aux trois cas paradigmatiques que nous venons de présenter. Ullmann Margalit considère en effet que ces trois situations ont en commun une structure qui "appelle" la présence d'une norme collective. C'est cette nécessité de la règle collective, de type normatif, qui caractérise dans son optique le problème général de la *coordination*. A ce degré de généralité, il n'est plus nécessaire de distinguer les problèmes de la *coopération* et de la *coordination*. La norme est alors conçue comme ayant une fonction précise et générale de résolution d'un problème collectif : en cela la théorie d'Ullmann-Margalit est *fonctionnaliste*. En revanche, une norme spécifique est associée à ces trois types de situation. La structure du problème détermine donc une norme de nature particulière : en cela l'analyse d'Ullmann Margalit est également *structuraliste*.

132- Trois types de normes

Les trois "types" d'interaction appelant la présence d'une norme sont : le jeu de pure coordination (type "jeu du rendez-vous"), le dilemme du prisonnier, et le jeu du marchandage (type "guerre des sexes"). Ces trois situations ont en commun la nécessité qu'une norme vienne régler les doutes des agents concernant le choix de la stratégie adéquate. L'auteur propose donc de définir la norme en trois points (Ullmann-Margalit, 1977, p.12) :

- 1- La norme constitue une *obligation* d'agir lorsqu'il existe une demande

générale pour la conformité. Elle impose une forte pression pour pénaliser la déviance.

2- La norme est nécessaire pour le maintien de la vie sociale, ou au moins de certaines de ses caractéristiques essentielles.

3- La conduite requise par la norme entre généralement en conflit avec ce qu'un individu souhaiterait faire dans son propre intérêt.

La norme dont il est question ici est donc dotée d'une dimension *coercitive* essentielle et résout les situations en allant parfois contre l'intérêt de certains des joueurs. C'est un aspect essentiel, puisque la dimension coercitive de la règle est souvent problématique dans un cadre individualiste qui entend faire surgir la règle du libre choix d'agents rationnels. C'est en raison de cette dimension coercitive qu'Ullmann-Margalit ne pense pas que la théorie des jeux puisse expliquer la façon dont les normes émergent.

Cependant cette définition générale ne s'applique pas forcément à toutes les situations. En fait, seul le dilemme du prisonnier requiert une norme ayant les trois caractéristiques énumérées ci-dessus. La présence d'un équilibre non coopératif et la menace permanente de la défection nécessite la mise en place d'une norme pour assurer une issue pareto-optimale (condition 2) ; stabilisée par la sanction de la défection (condition 1) , et allant contre l'intérêt immédiat des individus qui est de faire défection lorsque tout le monde coopère (condition 3). L'auteur qualifie cette norme de "PD norme". Elle la rapproche de la vision hobbenne de l'ordre social.

La norme résolvant les situations de négociation n'a pas la seconde caractéristique. Le problème qu'elle résout n'est pas un problème de pareto-optimalité, mais un conflit strictement privé entre deux individus ayant des intérêts en partie opposés. Elle doit cependant comporter une dimension coercitive de manière à stabiliser un des deux équilibres possibles. En effet, bien que chacune des deux issues "communes" soit un équilibre de Nash, l'auteur indique paradoxalement qu'elles sont toutes deux "*stratégiquement instables*" (Ullmann-Margalit, 1977, p.163). Cette expression est un peu malheureuse dans la mesure où l'équilibre de Nash représente par essence la stabilité stratégique. Par cette expression, l'auteur entend cependant souligner que les joueurs qui estiment être spoliés lorsque le choix collectif se porte sur l'issue qui leur est la plus défavorable tenteront toujours de faire advenir l'autre équilibre de Nash. Cet argument est cependant en lui-même problématique. En effet, si les joueurs n'ont pas les moyens de s'influencer mutuellement, une fois qu'un équilibre est atteint, il est stable. A chaque nouveau tour, les joueurs n'ont pas d'autre choix que de rejouer l'équilibre atteint précédemment, même s'il les défavorise. L'argument d'Ullmann Margalit, qui est en pratique tout-à-fait évident, ne trouve pas véritablement place dans le cadre stratégique

dans lequel elle situe son étude. En quelque sorte, c'est un argument de nature pratique qui ne trouve pas place dans un exposé qui s'est donné comme objet l'étude des structures logiques du jeu. Néanmoins, l'auteur affirme que, pour que cette norme fonctionne et soutienne un des deux équilibres, il faut qu'elle soit coercitive. Il est nécessaire qu'elle dévalorise et rende non viable l'autre équilibre de Nash. Pour cette raison, Ullmann-Margalit qualifie la norme qui permet de stabiliser un des deux équilibres en pénalisant la déviance par rapport à ce point de "statu quo" de "*norme de partialité*". Elle la rapproche de la vision que propose Locke du maintien, dans l'état civil, des inégalités présentes dans l'état de nature.

Seuls les problèmes de pure coordination peuvent être résolus sans que la norme collective soit coercitive. La norme de coordination ne respecte en effet que la seconde caractéristique. Elle est nécessaire à la vie sociale, mais naît à partir d'une parfaite coïncidence entre les intérêts des individus (ce qui rend sans objet les conditions 1 et 3). C'est donc un type particulier de norme qui peut émerger spontanément et qui ne devient que *progressivement* déontique. Selon Ullmann-Margalit, la solution au jeu de coordination *évolue en norme* au sens où l'attente réciproque des individus, concernant le respect de l'issue commune préalablement atteinte, cesse progressivement d'être le fruit d'une réflexion rationnelle pour s'imposer comme une obligation sociale.

En fait, cette forme particulière de norme est une "*convention*" telle que Lewis [1969] l'a définie. Nous étudierons ses propriétés dans la section suivante. Mais on peut d'ores et déjà souligner deux caractéristiques essentielles de ce type de règle collective. Elle est la seule à ne reposer que sur le consentement volontaire des agents et elle comporte un aspect que doivent avoir toutes les normes : le principe de l'auto-renforcement. Les individus auront tendance à dupliquer en permanence les solutions qu'ils ont observées par le passé lorsqu'ils ont été confrontés à un problème de coordination. Dans le cas limite de la pure coordination cette tendance suffit à assurer l'émergence d'une norme commune, tandis que dans les autres cas, la présence d'un conflit rend plus délicat le mécanisme de la duplication.

133- Coopération et coercition.

La typologie d'Ullmann Margalit peut être organisée selon une échelle allant de la coercition maximale à l'absence de coercition. Or, on peut faire un parallèle entre le caractère plus ou moins nécessaire de la coercition, et le caractère plus ou moins présent du problème spécifique de la coopération, en plus de celui de la coordination.

L'étude d'Ullmann Margalit fait ainsi apparaître qu'en l'absence de temporalité réelle, sur le plan strictement logique, tout problème de coordination, même le plus simple,

requiert la présence d'une norme. Elle montre également que lorsque le problème de la coopération vient s'ajouter à celui de la coordination, la norme doit être *coercitive* parce qu'elle impose un renoncement de l'acteur à certains de ses intérêts propres.

La typologie d'Ullmann-Margalit permet ainsi de tirer au clair les obstacles auxquels se heurte la question de l'émergence des normes et plus généralement des règles collectives dans le cadre de la théorie des jeux. Sur le plan logique, la résolution des jeux qui posent le problème de la coordination nécessite de faire référence à un élément extérieur à la rationalité des acteurs. La norme est donc ici conçue comme ce qui pallie les manques du principe de l'équilibre de Nash. Elle vient combler l'incomplétude de la rationalité individuelle. Cette théorie situe donc la question de l'émergence des repères collectifs en amont de l'étude stratégique du jeu.

Mais son étude comporte une distinction encore plus fondamentale : lorsque le problème à résoudre comporte un problème de *coopération* en plus de celui de la *coordination*, il faut que la norme soit *coercitive*. A l'inverse, dans le cas où la norme règle un problème de pure coordination, elle n'implique pas de coercition et peut émerger "spontanément". Cette distinction est essentielle car la théorie des jeux ne peut espérer rendre compte que des normes qui émergent "spontanément" à partir des interactions des agents. En quelque sorte, Ullmann Margalit limite la capacité d'explication de la théorie à la question de l'émergence des règles de coordination. Or, ce problème, qui exclut tout conflit entre acteurs, n'est pas le plus représentatif du point de vue économique. Dans le domaine économique, en effet, le conflit est toujours présent. Ullmann Margalit semble donc grandement limiter l'espoir de formuler une théorie générale de la genèse des institutions s'appuyant exclusivement sur la théorie des jeux.

L'analyse d'Ullmann Margalit fait apparaître d'une part la nécessité *fonctionnelle* des règles, pour répondre à un problème général de coordination, et, d'autre part le caractère plus ou moins coercitif de ces règles, qui dépend de la *structure* de l'interaction étudiée. Plus la situation pose un problème de coopération (en plus du problème de coordination), plus la coercition paraît nécessaire. Cependant, l'analyse d'Ullmann Margalit se situe délibérément dans un cadre logique qui exclut tout processus temporel. Or, lorsque le jeu prend une dimension temporelle, la structure du jeu lui-même est modifiée. Cette modification ne change pas le besoin fonctionnel de règle, mais peut remettre en cause le lien entre coercition et coopération : dans le long terme, en effet, la coopération peut s'avérer être une attitude parfaitement rationnelle. Autrement dit, avec l'introduction du temps, l'espoir de disposer d'une théorie de l'émergence spontanée des règles collectives peut renaître.

14- La question de la genèse des institutions

L'attitude fonctionnaliste d'Ullmann Margalit n'est pas la seule possible. On peut en effet, comme le souligne Schotter [1983], tenter de montrer que l'institution émerge du jeu lui-même. Mais pour cela il faut parvenir à montrer que l'attitude coopérative peut dans certain cas être compatible avec une rationalité stratégique standard. Or, ceci n'apparaît qu'à long terme, lorsque le jeu est répété. A partir de là, le problème principal à régler est celui de la coordination des agents. Cette question porte sur l'harmonisation des croyances des agents comme le montre Binmore [1987].

141- Répétition versus coercition.

Lorsque le dilemme du prisonnier est répété un nombre infini de fois, il n'existe plus de stratégie dominante, et le nombre d'équilibres possibles se multiplie¹⁴. La défection systématique figure parmi les stratégies d'équilibre, mais ce n'est pas la seule : l'attitude coopérative en est une autre¹⁵. Le *Folk Theorem* indique ainsi que toute stratégie coopérative peut apparaître comme équilibre d'un jeu répété de manière infinie¹⁶. En effet, à long terme il devient rationnel de tenter de profiter des "bienfaits" de la coordination.

La possibilité de coopérer est directement liée à l'impossibilité de "flouer" systématiquement son adversaire, si celui-ci est rationnel. En effet, une attitude coopérative peut être prise au dépourvu une fois par une attitude agressive. Mais, dans le long terme, la défection systématique n'est plus rentable, puisqu'elle se voit systématiquement opposer une attitude similaire de la part de l'adversaire. Il est donc plus rentable de tenter de coopérer à long terme que de pratiquer la défection systématique qui condamne l'individu à ne toucher, dans le long terme, que les faibles gains correspondant à l'équilibre sous optimal de la "défection commune".

Nous reviendrons dans la troisième section de ce chapitre sur l'étude des jeux répétés et évolutionnistes. Mais, d'ores et déjà, il est important de souligner que ces jeux dénouent la tension inhérente au problème de la coopération que repère Ullmann Margalit. Si la coercition est indispensable sur le plan logique, il est tout à fait imaginable, lorsque

¹⁴ Le principe de l'induction rétrograde fait apparaître qu'à compter du moment où les joueurs peuvent identifier le terme du jeu répété, la défection redevient la seule stratégie possible. En effet, au dernier tour, il est toujours rationnel de faire défection sachant que l'autre coopère. Or, chacun des joueurs prévoyant cela cherchera à faire défection dès l'avant-dernier tour. Le raisonnement rétrograde enclenche donc un processus dans lequel chacun fait défection tout de suite, comme dans le cadre statique. Pour éviter ce problème, on se situe là dans un cadre où l'on ne connaît pas le terme du jeu.

¹⁵ Une stratégie coopérative est une stratégie qui implique de ne jamais faire défection le premier.

¹⁶ Voir Fudenberg et Maskin [1986], et Aumann [1989].

l'on introduit une dimension temporelle, que le conflit entre les individus se dénoue de lui-même, du fait des bénéfices importants que permet d'obtenir l'attitude coopérative. C'est l'intuition que suivront Schotter [1981] et plus encore Axelrod [1984].

Ainsi, à long terme, le problème de la coopération devient moins prégnant et celui de la coordination devient plus central. On peut donc espérer parvenir à une explication de la coordination des agents par l'émergence d'une norme collective ne reposant sur aucun dispositif transcendant les agents. C'est justement ce que semble promettre l'analyse de Lewis [1969].

142- Les règles collectives comme support des croyances.

Lorsque le jeu se déroule sur une période de temps suffisamment longue, la question qui se pose est celle de savoir comment évoluent les croyances des agents. Les anticipations qu'ils forment sur le comportement des autres joueurs déterminent en effet leur propre attitude. La règle institutionnelle émerge alors comme un mécanisme permettant d'harmoniser les croyances. C'est donc en étudiant les mécanismes de formation et d'évolution des croyances des agents que la théorie des jeux peut parvenir à saisir le mécanisme d'émergence des règles et normes collectives.

Ce point est remarquablement mis en évidence par les analyses critiques de Binmore [1987, p.142-147 et 176-179]. Cet auteur propose en effet une critique très précise de la "*doctrine Harsanyi*"¹⁷ qui constitue le pivot de la représentation "standard" du jugement de probabilité des joueurs.

Cette doctrine définit une forme très stricte de rationalité cognitive, qui assure qu'à chaque instant les anticipations des joueurs sont convergentes. Les individus rationnels assignent nécessairement les mêmes probabilités aux mêmes événements. Cette position théorique sert de fondement aux nombreux "raffinements" de l'équilibre de Nash qui consistent à décrire un processus séquentiel de convergence vers l'équilibre : il est possible, par le processus de l'induction rétrograde, d'éliminer tous les coups reposant sur de fausses croyances, ou des croyances non rationnelles, et ainsi de sélectionner un seul équilibre de Nash parmi ceux qui existent¹⁸. Au centre de cette représentation

¹⁷ Aumann [1976] nomme ainsi la thèse d'Harsanyi [1967-68] selon laquelle il est logique de considérer que les individus ont une même distribution de probabilité au départ du jeu. De ce fait tout écart concernant les probabilités attribuées aux chances de réalisation des événements doit être attribué à une différence d'information. Aumann [1976] souligne que dans un cadre interactif l'action des joueurs révèle les informations dont ils disposent, il ne saurait donc y avoir durablement d'écart informationnel, les joueurs assignent donc toujours les mêmes probabilités aux mêmes événements. Ceci fait de plus l'objet d'un "Common Knowledge" chez tous les joueurs. (voir Gérard-Varet [1995]).

¹⁸ Voir sur ce point la théorie "générale" d'Harsanyi et Selten [1988]. Walliser [1995, p. 316] souligne que la mise au point d'une telle rationalité cognitive, associée à la traditionnelle rationalité instrumentale,

“rationnelle” des croyances se trouve une conception “bayésienne” qui suppose que les individus :

- 1- Ont la même structure d'information au départ.
- 2- Découvrent les mêmes informations à propos de la structure du jeu.
- 3- Réagissent de la même manière aux nouvelles informations selon une procédure de révision bayésienne des croyances.

La conséquence de la doctrine Harsanyi est qu'il devient impossible de supposer que les individus ont, à aucun moment, des anticipations divergentes. Il devient, selon l'expression d'Aumann [1976], impossible de “s'accorder sur des désaccords”.

Dès lors, comme le notent Binmore et Samuelson [1994, p. 49], il est nécessaire d'abandonner la technique Harsanyi de sélection des équilibres de Nash, pour concevoir que “cette question de la sélection des équilibres est aussi la réponse à la question de l'origine des normes sociales”. Binmore a ainsi entrepris de desserrer cet étau en dévoilant le caractère “naïf” de la conception bayésienne qui sous-tend la doctrine Harsanyi : “Un bayésien naïf suppose qu'il est suffisant d'assigner des croyances *a priori* à un décideur pour oublier ensuite d'où ces croyances proviennent” (Binmore, 1987, p.142). Ainsi, en postulant *a priori* l'existence d'un accord sur les croyances initiales, la doctrine Harsanyi ferme toute possibilité de compréhension des mécanismes de formation de cet accord. La critique de Binmore vise, à l'inverse, à remettre en question cette conception en arguant du fait que la question de l'origine des croyances est essentielle.

La technique d'Harsanyi consiste à se donner un modèle d'individu vierge de toute influence sociale : “les individus sont supposés entrer dans le monde avec un esprit fondé sur le principe de la table rase” (Binmore, 1987, p.144). Cet individu vierge est simplement doté d'une règle de révision bayésienne des croyances qui lui permet de traiter l'information qu'il reçoit. Mais, selon Binmore, une telle conception est inadéquate dans la mesure où elle revient à supposer que tous les individus disposent d'une même technique de traitement de l'information, avant même de disposer de la moindre information. Cela interdit d'envisager la manière dont l'individu “apprend” à gérer l'information au fur et à mesure qu'il la reçoit. Une règle de révision bayésienne n'existe qu'à compter du moment où l'individu rencontre l'information, elle n'est pas une faculté antérieure à tout processus de formation des croyances. Les individus n'ont donc pas *a priori* les mêmes croyances, ils jugent toujours du présent en s'appuyant sur une expérience passée qui peut différer d'un individu à l'autre¹⁹. C'est cette diversité des

permet d'internaliser le “commissaire priseur Nashien” dont le rôle tacite, dans le cadre de l'équilibre de Nash sous forme stratégique, est de montrer aux joueurs quel équilibre choisir.

¹⁹ On retrouve ici le type de critique que Kahneman et Tversky adressent à la théorie du choix rationnel.

croyances et des histoires possibles qui explique que, contrairement à ce que prédit la théorie, certaines issues coopératives puissent être atteintes.

Il existe toujours une suspicion selon laquelle l'autre joueur ne perçoit pas la situation de jeu de la même manière que nous. Cette suspicion ouvre un espace dans lequel il est possible de rechercher les bases d'un alignement des croyances sur des "normes" ou des règles collectives" en vigueur. On peut par exemple supposer, dans un jeu du type du dilemme du prisonnier, que le joueur rationnel qui nous fait face n'est pas absolument sûr que nous sommes sûrs qu'il ne va pas chercher à coopérer. Il cherchera donc peut-être à nous faire comprendre qu'il fonde son comportement sur une norme de réciprocité du type : "je coopérerai tant que tu coopéras"²⁰. C'est là l'origine possible d'une "norme de réciprocité".

Les remarques critiques de Binmore permettent de préciser un point important à propos de l'analyse de la règle proposée par la théorie des jeux : le rôle des règles collectives est essentiellement "épistémique". Les règles collectives servent à harmoniser les croyances des joueurs afin d'orienter leurs anticipations dans un sens identique, et, le cas échéant, dans un sens favorable à l'issue coopérative. C'est précisément en cherchant à percer les mécanismes par lesquels des individus rationnels parviennent à harmoniser leurs anticipations que le philosophe David Lewis a proposé une définition du concept de convention, qui est à même de servir de base à une théorie "rationnelle" de l'institution.

²⁰On peut, à titre d'exemple, citer le jeu du mille-pattes de Rosenthal [1981], version séquentielle du dilemme du prisonnier. Une application de la théorie de l'équilibre parfait par induction arrière donne comme solution unique la défection au premier tour, alors qu'une légère incertitude sur la "rationalité", ou le profil cognitif, de l'autre joueur rend possible un paiement final cent fois plus important pour les deux joueurs.

2– *L'analyse de Lewis*

“Le langage n'est qu'une des nombreuses activités reposant sur des conventions que nous ne pourrions créer par un accord, et que nous ne pouvons décrire.” (Lewis [1969, p.3])

David Lewis a proposé une définition novatrice de la convention, reposant sur le principe de la connaissance commune, ou du Common Knowledge (désormais : CK). Son analyse de la coordination par convention est progressivement devenue un passage “obligé” pour ce qui est de l'analyse économique des institutions : elle semble en effet permettre d'intégrer l'idée d'institution, de norme, de règle, à travers le principe de la convention dont la caractéristique est de ne reposer que sur la libre décision d'individus égoïstes et rationnels. La plupart des manuels de théorie des jeux, la plupart des articles traitant de l'analyse économique des institutions, règles et conventions indiquent, en référence, l'ouvrage de Lewis [1969]. Pourtant Lewis n'est pas un théoricien des jeux, ce n'est pas non plus un économiste, ni même un auteur de philosophie politique. C'est un logicien. L'utilisation de la théorie des jeux dans son étude est fortuite. Lewis la considère comme un outil pratique, un artifice de présentation efficace, mais affirme qu'il aurait aussi bien pu s'en passer²¹. La théorie des jeux, en particulier dans sa branche néoinstitutionnelle, a pourtant “récupéré” l'analyse de Lewis et s'appuie sur son analyse pour intégrer l'idée d'institution. Ainsi une théorie logique se retrouve utilisée pour analyser les problèmes de coordination et de coopération, problèmes traditionnellement traités dans le champ de la philosophie politique.

Pour mesurer la distance qui existe entre le propos initial de Lewis et ce qu'en a retenu l'analyse économique, nous présenterons le débat abstrait dans lequel sa théorie de la convention s'insère. Nous présenterons ensuite sa définition de la convention, qui s'avère être un élément essentiel du cadre d'analyse de la théorie des jeux néoinstitutionnaliste. Une troisième partie permettra précisément de mesurer les difficultés liées à l'utilisation du concept de convention comme fondement des institutions économiques.

²¹*“La théorie des jeux n'est, au final, qu'un outil. J'aurais tout à fait pu l'exposer sans m'en servir” (Lewis, 1969, p.3).*

21- Un problème logique quant à la constitution du langage

Pour comprendre la théorie de Lewis et mesurer la pertinence qu'elle peut avoir pour le sujet qui nous préoccupe, il est nécessaire de s'arrêter un moment sur le débat dans lequel elle s'insère.

Lewis construit son analyse, et son concept de convention, en réponse à la théorie de Quine, dont il est un disciple critique. C'est la raison pour laquelle Quine signe la préface de son ouvrage. Or, la théorie de Quine se situe dans le champ de l'empirisme logique, et participe d'une controverse qui s'y noue à partir des travaux de Wittgenstein [1921]. Cette controverse porte en particulier sur le statut des propositions analytiques.

Contre Quine, Lewis cherche à redonner un consistence à l'idée de *vérité analytique*. La clef de la possibilité d'existence d'une vérité de type analytique réside dans la défense d'un fondement *conventionnel* du langage. C'est par une redéfinition "astucieuse" du concept de convention que Lewis pense pouvoir s'émanciper de l'empirisme strict et austère de Quine.

211- Les deux dogmes de l'empirisme logique

L'empirisme logique s'est construit, notamment avec les théoriciens du cercle de Vienne²², à partir d'une interprétation des thèses que Wittgenstein [1921] a exposée dans son premier ouvrage : le *Tractatus Logico Philosophicus*. Dans le *Tractatus*, Wittgenstein cherche à établir quelles formes doivent nécessairement prendre le monde et le langage, afin que le langage puisse décrire le monde. Sa théorie repose sur une distinction entre les notions de *sens* et de *vérité*.

Selon Wittgenstein, pour qu'une proposition puisse décrire le monde, il faut :

- 1- Qu'il existe une relation biunivoque entre les mots du langage et les objets du monde (les "choses" que le langage désigne).
- 2- Qu'il existe une communauté de structure *logique* entre les propositions linguistiques (des assemblages de mots) et les états de choses (des assemblages d'objets dans le monde réel)²³.

²² On désigne par l'expression "théoriciens du cercle de Vienne" les participants à un groupe de recherche, regroupé initialement autour de Verner Schlick. Ce groupe a été rendu célèbre et identifié comme une véritable école de pensée à partir de la parution en 1929 du "manifeste du cercle de Vienne" qui expose une conception scientifique du monde et jette les bases de l'empirisme logique. Les trois rédacteurs de ce manifeste sont : Carnap, Neurath, Hahn. Ce sont également les membres les plus représentatifs de ce mouvement philosophique auquel ont également participé, par exemple, Gödel et Menger.

²³ Une proposition est un assemblage logique de mots. Un état de chose est un assemblage logique d'objets.

Sur cette base, Wittgenstein distingue :

1- Les propositions "*sensées*" qui sont logiquement construites et sont à l'image d'un état de chose. Elles peuvent être *vraies* ou *fausses*, selon qu'elles désignent un état de chose qui existe réellement ou, au contraire, un état de chose qui n'existe pas. L'ensemble des propositions sensées contient plusieurs mondes possibles²⁴, dont un seul est "actuel" : celui dans lequel nous vivons. L'ensemble des propositions vraies constitue les sciences de la nature (Wittgenstein, 1921, aphorisme 4.11).

2- Les propositions "*dépourvues de sens*" qui bien que logiquement construites ne sont pas à l'image d'un état de chose. *Toujours vraies* ou *toujours fausses*, elles ne nous apprennent rien sur le monde dans lequel nous vivons. Les lois de la logique ont ce statut (Wittgenstein, 1921, aphorisme 6.11).

3- Les propositions qui sont des "*non sens*" parce qu'elles ne sont pas logiquement construites et qu'en conséquence elles ne peuvent être l'image d'un état de chose.

La distinction entre les propositions sensées et les propositions dépourvues de sens renouvelle la distinction kantienne entre propositions *synthétiques* et propositions *analytiques*.

Sur cette base, Wittgenstein établit un mécanisme de décomposition des propositions qui permet de "vérifier" si elles sont vraies ou fausses. Une proposition sensée est une articulation logique de plusieurs "propositions élémentaires" qui désignent chacune un "état de chose" directement observable. Pour établir la vérité ou la fausseté d'une proposition sensée, il suffit donc de la décomposer et d'observer le caractère véritable ou fallacieux des "propositions élémentaires" qui la composent. L'empirisme logique reprendra ce principe "*réductionniste*" et élaborera sur cette base une théorie du savoir scientifique.

Cependant, l'empirisme logique se heurte à la question du statut de la logique. Les lois logiques ne reposent pas, en effet, sur des observations empiriques. Elles ne peuvent être vraies. Wittgenstein donne pour cette raison un statut particulier à la logique : elle est transcendantale. Cette transcendance des règles logiques qui organisent le langage interdit que l'on puisse dire quoique ce soit à propos du langage. On ne peut pas le dominer, on ne peut pas en faire une *théorie*. Sur ce point, la position des théoriciens du cercle de Vienne, regroupant les tenants de l'empirisme logique, est radicalement différente de celle de Wittgenstein. Ceux-ci affirment au contraire qu'il est possible de construire une théorie du langage et de la structure des énoncés linguistiques. Pour ce faire, il faut faire une

²⁴ Lewis et Kripke reprendront et développeront cet aspect de la philosophie de Wittgenstein à travers la "sémantique des mondes possibles".

place à l'idée d'une vérité qui ne serait pas d'ordre empirique : une vérité *analytique*. C'est ce que fait en particulier Carnap qui affirme : *“dans tout domaine scientifique, il est seulement possible de parler dans les énoncés de ce domaine ou au sujet des énoncés de ces énoncés ; et ainsi, les seuls énoncés qu'on peut formuler sont soit des énoncés d'objet soit des énoncés syntaxiques.”* (Carnap, cité par Meyer [1994], p.229).

Cependant, pour pouvoir dire d'un énoncé qu'il est analytiquement vrai, il faut éclaircir l'origine de cette vérité qui ne saurait être empirique, comme le montre Wittgenstein, ni transcendantale, car dans ce cas elle ne pourrait faire l'objet d'une théorie scientifique²⁵. C'est précisément à ce stade qu'apparaît la notion de convention. Pour les tenants de l'empirisme logique, les propositions analytiques sont dites vraies et nécessaires par *“convention.”* Le langage scientifique repose ainsi sur des conventions nécessaires qui font *“autorité”*.

212- Le radicalisme de Quine

C'est précisément contre cela que s'élève Quine, qui, en deux textes importants, brise la distinction entre vérité *analytique* et *synthétique*.

Il poursuit ce faisant une entreprise de *“nettoyage”* conceptuel qui dont l'objet est, selon lui, de *“rendre explicite ce qui a été laissé tacite, et de rendre précis ce qui a été laissé vague, (...) d'exposer et de résoudre les paradoxes, de raboter les aspérités, de faire disparaître les vestiges des périodes transitoires de croissance, de nettoyer les bidonvilles ontologiques.”* (Quine, 1960, p.377). Cette clarification passe selon Quine par la suppression des notions confuses que sont *“l'analyticité”*, la *“convention”*, la *“nécessité”*.

Dans un premier texte, Quine [1936] critique la notion de convention. Son argument est simple et dévastateur : la vérité ne peut reposer sur des conventions, sauf à supposer que ces conventions sont nécessaires, et, dès lors, ce ne sont plus des conventions. Comme le note Bouveresse, Quine *“tente d'enfermer le conventionnalisme dans le dilemme suivant : ou bien les conventions linguistiques auxquelles on se réfère n'ont aucun caractère délibéré et explicite, et il est difficile de leur attribuer une valeur explicative quelconque, ou bien l'on suggère que la signification pourrait avoir été réellement déterminée et communiquée par l'intermédiaire de stipulations explicites”* (Bouveresse, 1987, p.97). Dans ce dernier cas, indique Quine, il faut présupposer l'existence d'un langage sur lequel fonder la définition de ces conventions premières : on

²⁵ Wittgenstein pour sa part se refusait à dire quoique ce soit des propositions logiques dont la vérité était supposée être transcendantale : *“Sur ce dont on ne peut parler, il faut garder le silence”* (Wittgenstein, 1921, aphorisme 7).

entre ainsi dans une régression infinie²⁶. Quine récuse donc l'usage du concept de convention, et préfère parler de "régularité" et de "vérité *a priori*", qui ne sont pas, et ne peuvent pas être, objet de connaissance.

Dans le second texte, Quine expose une conception qui radicalise les thèses de l'empirisme logique. Il débarrasse l'empirisme de ces deux "dogmes" que sont d'une part le principe de la décomposition analytique (le "réductionnisme"), et d'autre part celui de la distinction entre les vérités de type analytique et celles de type synthétique. Selon Quine, en effet, les vérités logiques n'ont pas de statut particulier, elle sont aussi peu "nécessaires" que nos observations empiriques. Elles peuvent être révisées et atteintes par une expérience qui les invalide²⁷. Elles n'ont donc ni le caractère transcendant que leur donnait Wittgenstein, ni le caractère "nécessaire par convention" que leur donnait Carnap. Elles perdent tout-à-fait leur statut particulier. Elles deviennent aussi peu solides et tout autant révisables que nos connaissances empiriques.

De ce fait selon Quine, on ne peut pas "décomposer" les propositions scientifiques pour tester les énoncés élémentaires un à un. La science ne se présente pas comme un plancher dont il suffirait de changer quelques lattes défectueuses. Elle est d'un seul tenant. Quine souligne ainsi : "*l'unité de signification empirique est la totalité de la science. (...) Il est absurde de parler des composants linguistiques et factuels de la vérité d'un énoncé individuel. Prise collectivement, la science est dans une double dépendance à l'égard du langage et de l'expérience ; mais on ne peut pas suivre cette dualité à la trace dans les énoncés de la science, pris un à un.*" (Quine, 1951, p.108). Quine rejette donc d'un seul mouvement la distinction entre vérités analytiques / vérités synthétiques, et la méthode réductionniste. Tout notre savoir est révisable, il n'a pas une structure duale constituée d'une partie solide et inaltérable (la logique) et d'une partie plus friable et aisément révisable (les connaissances empiriques). Quine défend donc une conception "holiste" de la science, et adopte une attitude d'un relativisme radical.

Quine explicite cette position "radicale" grâce à une métaphore qu'il emprunte à Neurath²⁸. Ce dernier compare le savant à un marin. Lorsqu'il est embarqué en pleine mer, le marin ne peut pas isoler son bateau de l'eau dans lequel il baigne. Il ne peut que le reconstruire planche par planche, de l'intérieur. Quine étend cette métaphore au philosophe : il ne peut pas fonder la science sur un discours qui lui serait extérieur. Il est

²⁶ "Si la logique doit procéder médiatement des conventions, la logique est nécessaire pour inférer la logique à partir des conventions." (Quine [1966, p.97]).

²⁷ Ainsi comme le note Gochet : "pour Quine, la philosophie est une science empirique qui ne diffère des autres que par son extrême généralité. S'il n'y a pas d'expérimentation en philosophie, ce n'est pas parce que cette discipline nous parle d'un monde transcendant, mais parce qu'elle est reliée à l'expérience par des voies indirectes, comme la logique qui peut, elle aussi, subir le choc en retour de l'expérience" (Gochet, 1978, p.16)

²⁸ Cité par Gochet (1994, p.340).

tributaire du “bateau” dans lequel il est. Le seul contact avec la vérité s’effectue cependant au contact physique avec la nature, la manière dont on a mathématiquement construit ce contact physique est une donnée que le philosophe ne peut que constater, sans la fonder. Le langage mathématique que nous utilisons pour retranscrire cette vérité doit être construit de façon “canonique” pour être le plus conforme possible à l’organisation logique de la nature. Il peut donc s’avérer, à l’usage, être inadapté et doit être révisé lorsque les faits prouvent son caractère inadéquat. Certes, il est nécessaire qu’il ne soit pas *constamment* remis en cause par l’expérience, mais ceci se traduit simplement par la nécessité de conserver notre outil principal de connaissance. Cela ne lui donne pas le statut de vérité. Simplement, on ne peut pas changer de bateau en pleine mer²⁹.

Notre connaissance est ainsi tout d’une pièce, tissée à *partir* des observations empiriques effectuées par la physique selon des règles syntaxiques que le philosophe doit définir pour qu’elles “miment” le mieux possible la réalité³⁰. Dans ce système quinién, il ne peut donc y avoir de vérité *analytique*, puisqu’il ne saurait y avoir de *convention* fondant l’analyse. Le refus de la notion de convention s’inscrit dans le refus d’une *nécessité* qui ne serait pas d’ordre empirique ou naturel.

213- La critique de Lewis

“Certains philosophes ont fortement mis l’accent sur la distinction entre les vérités de nature analytique qui ne sont vraies que par convention linguistique, et les vérités synthétiques qui disent quelque chose de substantiel quant à la réalité. C’est un cas caractéristique de la manière dont on fait appel à la notion de convention dans une situation où l’on ne peut penser la notion d’autorité. Les philosophes en question estiment en effet qu’il existe une vérité logique analytique, et il y aurait donc un risque de circularité si l’on prenait la notion de convention dans son sens premier [qui implique celui d’autorité]. (...) Nous avons devant nous une étude dans laquelle Lewis tente de rendre la notion de convention libre de tout fait ou fiction d’autorité” (Quine, Préface à Lewis [1969, p.XI-XII], c’est moi qui souligne)³¹.

²⁹ Quine utilise également une autre métaphore pour préciser sa conception du savoir : *“L’ensemble de la science est comparable à un champ de forces, dont les frontières seraient l’expérience. Si un conflit avec l’expérience intervient à la périphérie, des réajustements s’opèrent à l’intérieur du champ. Il faut alors redistribuer les valeurs de vérités à certains de nos énoncés. La réévaluation de certains énoncés entraîne la réévaluation de certains autres, à cause de leurs liaisons logiques— quant aux lois logiques elles-mêmes, elles ne sont que des énoncés situés plus loin de la périphérie du système...”* (Quine, 1951, p.232).

³⁰ *“La recherche d’une notation canonique qui ait la structure la plus simple et la plus claire ne doit pas être distinguée d’une recherche des catégories dernières ou d’un effort de reproduction des traits les plus généraux de la réalité.”* (Quine, 1960, p.232)

³¹ Nous avons choisi de traduire le terme anglais “*convening*” utilisé par Lewis, par le terme français “*autorité*”. Il n’existe pas d’équivalent français direct du terme anglais “*convening*”, et le terme “*autorité*”

La préface de l'ouvrage de Lewis, signée par Quine, cible nettement les enjeux de cet ouvrage et souligne le caractère "interne" de la critique que propose Lewis du système philosophique de Quine. Lewis relève le défi de Quine en cherchant à réhabiliter la notion de convention. Son "astuce" théorique consiste à montrer qu'il est "nécessaire" qu'il y ait des conventions, même si la convention choisie n'est pas, en elle-même, "nécessaire".

Cette réhabilitation est le premier pas d'une entreprise philosophique visant à restaurer la possibilité d'une vérité analytique. Lewis considère que notre langage particulier n'est qu'un exemplaire parmi plusieurs langages possibles répondant tous à certains critères logiques. La vérité analytique se situe dans ces règles logiques communes à tous les langages possibles. Elle n'est donc pas attachée à un langage particulier, même canonique ou parfait, qui transcenderait tous les autres langages imparfaits, mais s'actualise dans un langage particulier qui a été choisi par convention, et qui aurait tout aussi bien pu être autre. Ainsi : "*l'analyticité est la vérité dans tous les mondes possibles*" (Lewis, 1969, p.208). A partir de là, Lewis [1986] a développé une sémantique des mondes possibles fondée sur la logique modale. Selon cet auteur, qui retrouve ce faisant un aspect important du *Tractatus* de Wittgenstein [1921], notre monde n'est qu'un exemplaire parmi plusieurs mondes possibles, tous réels mais dont un seul, le nôtre, est actuel. De ce fait, comme le note Engel (1994, p.357) : "*l'univers de Lewis est aussi éloigné que possible de l'univers de Quine et semble aussi peuplé que celui de ce dernier était désert*". La redéfinition de la convention dans un sens qui lui permet d'échapper au paradoxe mis en évidence par Quine constitue le premier pas de cet itinéraire théorique.

La stratégie de Lewis est parfaitement résumée par Quine. Le problème fondamental attaché à la notion de convention est celui de sa justification : pourquoi une convention fait-elle autorité ? C'est une question dévastatrice pour quiconque souhaite faire de la convention le début de la connaissance et du langage. Lewis construit sa théorie de manière à esquiver cette question : la convention est *nécessaire*, mais parfaitement *arbitraire* et *sans fondement*. Elle aurait pu être autre, l'important étant qu'elle soit, non pas pour elle-même, mais en tant que *réfèrent* commun.

Lewis procède de manière à démontrer la nécessité logique de l'existence d'une convention, tout en montrant que le choix de la convention est arbitraire et ne procède de rien de plus que de la rationalité des individus et de leur intérêt commun à se coordonner. Il utilise, pour sa démonstration, la théorie des jeux de pure coordination de Schelling. L'emploi de cet outil, et plus encore le minimalisme de cette théorie qui n'a besoin de prendre en considération que des individus rationnels et intéressés, explique l'intérêt de l'analyse économique pour cette thèse de logique modale. En cherchant à relever le défi de

nous semble permettre de souligner l'idée selon laquelle Lewis rend à la convention son caractère "arbitraire".

son professeur, W.O. Quine, Lewis a fortuitement donné à la théorie économique l'espoir d'une théorie des institutions qui ne reposerait à la base sur "*aucun fait ou fiction d'autorité*", et qui échapperait donc absolument à la contrepartie coercitive, théorisée par Hobbes, repérée chez Williamson, et que toute théorie de l'organisation économique semblait devoir comporter. Comme le note Dupuy [1989, p.380], la théorie de Lewis semble nous promettre une parfaite "*transparence du collectif*", et parle directement au mythe fondateur de la main invisible et au "*paradigme de la rationalité*" qui en a découlé.

On mesure cependant la distance entre l'objectif de Lewis, qui s'inscrit dans l'analyse *logique* de la constitution du langage et dans le champ du *positivisme logique* dans sa version quinienne, et ce que la théorie économique cherche à lui faire dire des institutions, et qui s'inscrit dans le champ *pratique* de la *philosophie politique*. Cette distance est source d'une interprétation erronée de la théorie de Lewis qui en dit, en fait, beaucoup moins que ce que l'on est tenté de lui faire dire. Elle n'a, en particulier, aucun contenu pratique expliquant *comment* les conventions émergent. Lewis se consacre exclusivement à démontrer *pourquoi* il est nécessaire qu'elles existent. C'est dans cette optique qu'il distingue la convention du contrat, de la promesse, ou encore de la règle coercitive.

22- La définition de la convention chez Lewis

Pour parvenir à exprimer le sens que revêt son concept de convention, Lewis choisit de situer son étude dans le cadre de l'étude des "jeux de coordination", qu'il emprunte à Schelling [1960].

L'archétype d'une situation de pure coordination est le "*jeu du rendez-vous*" que nous avons présenté dans la première partie (fig.1). Un tel jeu comporte deux équilibres de Nash parfaitement symétriques et entre lesquels les joueurs sont parfaitement indifférents. Ce jeu est sans solution classique, dans la mesure où il n'existe ni stratégie dominante ni stratégie maximin permettant de discriminer les deux équilibres de Nash possibles. Ce cadre est parfaitement adapté aux deux idées que Lewis souhaite défendre :

1- Il faut que les individus disposent d'une règle commune pour sélectionner un des deux équilibres.

2- Le choix particulier de cette règle est parfaitement arbitraire. La règle choisie ne repose sur aucune forme de nécessité particulière : c'est une convention.

Ces deux idées apparaissent clairement à travers l'analyse que Lewis propose des situations de "pure coordination".

221- *“Putting ourselves in the other fellow's shoes”*³²

La stratégie gagnante dans un jeu de coordination consiste à choisir la même stratégie que l'autre joueur. Cette situation pose donc le problème de l'anticipation des actions de l'autre : il faut *“se mettre à la place de l'autre”*. Cette anticipation croisée entraîne les deux joueurs dans un jeu de miroir sans fin, chacun renvoyant à l'autre sa propre image. Toute anticipation du comportement de l'autre nous renvoie à un degré d'anticipation supérieur, intégrant notre propre comportement. L'interdépendance des acteurs provoque ainsi un phénomène d'auto-réflexion infini, dont la logique seule ne permet pas de s'évader : je dois anticiper l'autre en train d'anticiper mon *raisonnement*. Or, en toute logique, il sait que je fais ce *raisonnement*. Je dois donc passer au stade supérieur, c'est à dire *“anticiper la manière dont l'autre anticipe la façon dont j'anticipe la façon dont il anticipe mon raisonnement . Et ainsi de suite de manière a priori infinie”* (Lewis , 1969, p.27-28).

Lewis souligne cependant que *“nous faisons rarement des anticipations d'un degré supérieur à deux”* (Lewis, 1969, p.32). Il constate pourtant, comme l'avait fait Schelling [1960] avant lui, que cette indétermination logique est aisément surmontée dans la vie de tous les jours. La raison en est qu'il existe des points de convergence des anticipations qui permettent aux individus d'échapper à la spirale des regards croisés. Cette fois, ce que l'on découvre dans le regard de l'autre, ce n'est pas l'incertitude qu'il nous prête quant à la manière dont nous percevons ses intentions, mais au contraire la certitude que nous avons qu'il va aller vers un point extérieur particulier, pour la simple raison qu'il nous anticipe en train de penser qu'il va lui-même s'y rendre.

222- *Un exemple de coordination réussie*

Supposons un cas simple. Je souhaite rencontrer un ami, et ce désir de rencontre est réciproque, mais nous n'avons pas fixé de rendez-vous précis. Cet ami aime se rendre le lundi à dix-huit heures à la brasserie qui fait face à la gare pour se remettre de sa longue journée de travail et se désaltérer... On peut sans peine imaginer que, comme il sait que je sais qu'il aime se rendre à dix-huit heures dans cette brasserie, il s'y rendra effectivement (puisqu'il souhaite lui aussi me rencontrer). La probabilité de le retrouver à dix-huit heures dans cette brasserie devient donc très élevée. La présence d'un repère extérieur, dont nous avons tous deux connaissance et dont nous savons tous deux que l'autre a également connaissance, nous a permis de surmonter l'indétermination de la logique pure. L'acte de se rendre à la brasserie est, pour mon ami, suffisamment indépendant de ma

³² Lewis [1969,p.27).

³³ Lewis [1969,p.27).

propre action pour pouvoir servir de point de repère extérieur et nous permettre de sortir de l'infernale spirale de l'auto-réflexion.

Cependant, l'indépendance de cet acte, indispensable dans un premier temps pour servir de point de repère, pourrait s'avérer être un obstacle s'il conduisait mon ami à ne pas se rendre, pour un motif quelconque, dans sa brasserie préférée. Pour que l'action réussisse, il faut que l'acte de s'y rendre devienne, dans un second temps, dépendant de notre volonté commune de se coordonner. Ce qui sera le cas, si mon ami est soucieux de me rencontrer et qu'il est capable d'anticiper mon raisonnement, parce qu'il sait que je connais son habitude. Notre intérêt commun à nous rencontrer *et* notre rationalité nous font percevoir l'estaminet comme un lieu de rencontre probable et renforcent la probabilité que mon ami, et donc moi-même, nous y rendions effectivement.

Deux choses marquent cette coordination réussie : l'existence d'un point de repère commun, mais aussi, dans un second temps, notre rationalité et notre intérêt commun. On a vu que Schelling, avec la notion de point focal, mettait l'accent sur le premier aspect du mécanisme de coordination. Lewis s'appuie lui presque exclusivement sur le second. Ce second moment est celui de l'institutionnalisation de la croyance commune en convention mutuelle.

223- La solution de Lewis

Lewis [1969] se saisit donc des intuitions développées par Schelling mais va mettre l'accent sur l'importance de la rationalité, davantage que sur le contexte et l'interaction directe entre les joueurs. Son objectif est de généraliser l'idée qu'il existe des solutions "naturelles", même entre des gens qui ne se connaissent pas, et que ces solutions ne font pas particulièrement appel à des qualités d'observation, de mémoire ou de valeurs partagées, mais uniquement à la force du précédent et de la répétition, qui transforment une solution obtenue par hasard en "façon évidente de jouer". Le pivot de cette généralisation est l'idée de "*préférence conditionnelle pour la conformité : chacun veut se conformer aux autres, à condition que les autres le veuillent aussi*" (Lewis, p. 39).

Lorsque une même situation se répète, il devient plus aisé d'anticiper sur la conduite que les individus adoptent en règle générale face à cette situation. Chacun ayant intérêt à valider les anticipations que tous les autres font sur son propre comportement, l'attitude traditionnelle à adopter devient publiquement connue. Lewis propose, pour asseoir cette idée d'une connaissance qui serait partagée par tous, le principe de la "connaissance commune" :

"X fait l'objet d'une connaissance commune [Common Knowledge] dans la population P si et seulement si la situation A est réalisée de telle manière que :

(1) tout le monde dans P a raison de penser que l'on se trouve bien dans le cas de figure A .

(2) A indique à tout le monde dans P que tout le monde dans P a raison de penser que A est réalisé.

(3) A indique à tout le monde que X ” (Lewis, 1969, p.56).

Cette notion de croyance collective est absolument centrale en théorie des jeux. Elle permet de désigner le principe d'une connaissance parfaite, par tous les joueurs, des caractéristiques d'un jeu³⁴. C'est à ce premier titre que l'analyse de Lewis est constamment citée par les théoriciens et les manuels. Comme le note Dupuy [1989], dans sa critique de la notion de CK, cette conception de la connaissance revient à supposer une parfaite transparence de la croyance collective. Tout le monde sait que tout le monde sait que tout le monde sait ... de manière infinie. Cette hypothèse d'une rationalité parfaite, à un degré infini, est au centre de la définition de la convention.

En effet, dans un jeu de coordination, l'apparition d'une régularité de comportement constitue un point d'ancrage des anticipations des joueurs. Chacun des individus sait que les autres attendent de lui un comportement identique à celui qu'il a régulièrement observé par le passé. Ainsi se constitue une *convention*, constitutive d'un repère collectif général, objet de connaissance commune, et permettant d'assurer la réussite de la coordination.

La caractéristique principale de la convention est son arbitraire³⁵. Cet arbitraire est essentiel aux yeux de Lewis, qui souligne que si le choix entre les diverses issues coopératives du jeu de coordination n'était pas arbitraire, il n'y aurait pas de problème de coordination et donc pas de convention. A côté de la convention R qui fonctionne, existe toujours une autre issue collective possible et équivalente, R' . Le choix entre R et R' est arbitraire. Il l'est tout au moins dans un premier temps, car une fois le processus du CK mis en place, il devient rationnel de suivre R plutôt que R' pour réussir la coordination. L'existence d'une régularité de comportement enclenche un processus auto-renforçant³⁶. A partir du moment où il est *probable* que chacun suive R , il devient *rationnel* pour chacun de suivre R , et R se réalise alors *effectivement*. Comme chacun a accès à ce type de raisonnement, une probabilité se transforme rapidement en certitude. Ainsi une régularité R devient rapidement une convention stable par la seule force de la rationalité et de l'intérêt commun.

³⁴ La démonstration formelle de cette hypothèse de CK est réalisé par Aumann [1976]. Elle est présentée et critiquée par exemple par Binmore [1987] ou Hargreaves Heap et Varoufakis [1995].

³⁵ “Toute convention est arbitraire parce qu'il existe une autre régularité de comportement qui aurait pu faire office de convention à la place de celle qui a été finalement choisie.” (Lewis, 1969, p.70)

³⁶ La convention est auto-renforçante puisque chacun préfère une situation à laquelle tout le monde coopère, soi-même compris, à une situation dans laquelle tout le monde coopère, sauf soi-même. (Ullmann-Margalit, 1978, p. 118).

Lewis (1983, p.165) propose la définition générale suivante de la convention :

“Une régularité R dans l'action ou les croyances est une convention dans une population P si et seulement si dans P les six conditions suivantes sont vérifiées :

(1) Tous se conforment à R.

(2) Tous s'attendent à ce que tous les autres se conforment à R.

(3) Cette croyance que les autres se conforment à R donne à tous une bonne et décisive raison de se conformer soi-même à R.

(4) Il y a une préférence générale pour la conformité générale à R plutôt que pour une conformité un peu moindre que générale – en particulier une conformité de tous sauf un.³⁷ (...)

(5) R n'est pas la seule régularité possible remplissant les deux dernières conditions. Il existe au moins une autre régularité R' telle qu'une croyance en le fait que les autres suivent R' donnerait à tous une bonne et décisive raison, pratique ou épistémique, de se conformer à R'. (...)

(6) Les conditions (1) à (5) font l'objet d'un Common Knowledge. (...) Cette connaissance commune peut être seulement potentielle. Elle assure la stabilité.”

Ainsi, la priorité à droite, la reconnaissance des billets ou des pièces comme support de valeur, les logos représentant des personnes handicapés, sont autant de conventions courantes, comprises et respectées par tous dans un espace donné. Elles sont parfaitement arbitraires. On sait par exemple, dès lors que l'on a eu l'occasion d'effectuer un voyage outre-manche, que la priorité à gauche constitue une convention tout aussi efficace que la priorité à droite.

Ainsi le cadre d'analyse stratégique que Lewis utilise lui permet de démontrer la *nécessité* des conventions (sans convention le jeu est indéterminé) ; leur *arbitraire* (une autre issue collective aurait été équivalente), et leur *absence de caractère coercitif* (l'efficacité de la convention dépend exclusivement de la rationalité des joueurs).

23- Les caractéristiques de la convention lewissienne

La théorie de Lewis permet de montrer comment une coordination peut réussir entre des individus égoïstes sans que cette coordination repose en aucune manière sur l'existence d'une obligation limitant leur liberté. Le collectif n'est pas contractuel, il n'est

³⁷ Cette condition distingue un jeu de coordination dans lequel chaque issue commune est un équilibre, des jeux, du type “dilemme du prisonnier, dans lequel l'issue commune est menacée par la déviance.

pas non plus transcendant, il ne repose pas sur une intersubjectivité harmonieuse ou un accord social entre acteurs : il n'est que la somme de désirs égoïstes rationnellement assouvis, comme dans la plus pure tradition smithienne.

231- La convention : avant les mots

L'analyse individualiste de l'émergence des institutions pose un problème logique : on parvient rarement à éviter de présupposer ce que l'on cherche à démontrer. Lewis reprend à son compte le paradoxe mis en évidence par Quine [1936] : pour qu'il y ait convention, il faut qu'il y ait accord, et, pour qu'il y ait accord il faut présupposer l'existence du langage³⁸.

Cependant, Lewis évite cet obstacle parce que son concept de convention lui permet de montrer que l'accord tacite des individus sur des croyances, et donc des actes communs, ne nécessite pas que ces croyances et ces actes soient définis avant d'exister. L'explicitation de la convention n'est nécessaire à aucun stade de son émergence³⁹. Les individus s'accordent tacitement, sans avoir à s'entendre sur une définition de leurs actes avant de les effectuer : la compréhension s'effectue dans le cours de l'action. La convention est là avant les mots, avant l'accord contractuel. Ce type de connaissance, largement intuitive et tacite, n'est pas nécessairement générale. Il suffit qu'elle soit de type local, circonscrite à une situation et non généralisable. Les individus qui se coordonnent par convention ne sauraient parfois expliquer les raisons générales de leur action. Ils ne disposent pas systématiquement d'énoncés généraux, indépendants de la situation dans laquelle ils se trouvent, et qu'ils appliqueraient par déduction. Ils ne disposent que d'un savoir-faire, qui leur dicte l'acte qui convient relativement dans la situation précise à laquelle ils sont confrontés⁴⁰.

La convention est donc une modalité particulière de règle collective, ou de norme, parce qu'elle est d'emblée tacite, et que son existence peut à chaque instant être expliquée par son efficacité en matière de coordination : elle ne s'élève jamais à un niveau supérieur à celui de l'accord tacite, dans le cours de l'action. Donc, elle n'a pas besoin d'être mise

³⁸ Harsanyi et Selten [1988] opposent le même type de réticence à l'égard des concepts de solution dans les jeux coopératifs, et dans l'analyse des jeux faisant référence au "pre-play games". Là encore ces auteurs estiment que s'appuyer sur l'existence d'une solution coopérative, sur l'existence d'un accord passé avant le jeu, revient à supposer le jeu déjà solutionné avant même de l'étudier.

³⁹ Lewis souligne ainsi : "*Like or not, we have plenty of knowledge we cannot put into words.*" (Lewis, 1969, p.64)

⁴⁰ Cette connaissance est "sensu diviso" par opposition à une connaissance "sensu composito" (Lewis, p.65-68). La différence peut être illustrée par l'exemple suivant : "Je sais que tous les conducteurs français roulent à droite, et je sais que M. Dupont est un conducteur et qu'il est français. Donc M. Dupont va rouler à droite." (sensu composito) et "Je sais que M. Dupont, lorsqu'il prend son véhicule, roule à droite" (Sensu diviso). Favereau [1989] s'appuie sur ce type de connaissance pour mettre en évidence le rôle des DCC qui répondent à un "savoir comment" sans disposer du "savoir pourquoi".

en mots pour exister.

232- La convention n'est pas contractuelle

Cet aspect de la convention la distingue radicalement du contrat, dans lequel la parole et l'engagement précèdent l'acte. La force du concept de convention est précisément de ne tenir que par l'intérêt commun qui lie les acteurs. Ainsi Lewis souligne : *“une convention peut très bien avoir été initiée par un accord mais une convention qui débute par un accord ne peut devenir une convention, selon ma définition, qu'à compter du moment où l'influence de l'accord a eu le temps de s'effacer. C'est pourquoi une convention initiée par le biais d'un accord n'est pas différente d'une convention initiée d'une autre manière : elle ne porte aucune trace de ses origines”* (Lewis, 1969, p.84) . La mécanique conventionnelle gomme les conditions de sa naissance, à la différence du contrat qui agit comme un référent permanent permettant de spécifier les droits et devoirs de chacun. Une action collective devient conventionnelle à compter du moment où les participants se coordonnent non plus par respect d'un accord, mais en raison de leur intérêt commun à perpétuer l'habitude. Une régularité de comportement devient une convention lorsqu'elle s'autonomise par rapport aux conditions de son apparition. De ce fait, une action collective non conventionnelle peut, avec le temps, le devenir, parce que les raisons initiales de cet acte sont devenues obscures et apparaissent extérieures à la volonté individuelle. La convention réalise donc l'accord des volontés de manière tacite, sans qu'aucun des participants ne puisse justifier son action par un accord ou un engagement de sa part.

233- La coordination sans obligation

La convention échappe à la parole, et donc à l'accord avant l'acte, mais elle échappe aussi à la sanction, et donc à la question du respect de l'accord. *“Une promesse solennelle (forceful promising) est un moyen de se débarrasser des problèmes de coordination et non pas de les résoudre”* (Lewis, 1969, p.35). En effet, la structure même des problèmes de coordination suppose qu'il existe une alternative à la convention. C'est l'essence même du concept de convention d'être arbitraire, et d'avoir tout aussi bien pu être autre. C'est la raison pour laquelle les comportements effectués par respect d'une règle, ou par peur d'une sanction, ne peuvent être considérés comme conventionnels. L'existence d'une sanction pénalisant le non-respect d'un engagement élimine en effet l'alternative. S'engager, par le biais d'un accord ou d'un contrat, c'est s'interdire d'agir autrement. C'est la seconde raison pour laquelle la convention se distingue du contrat social

hobbien⁴¹. A l'origine du contrat se trouve l'engagement de tous les individus qui acceptent de se soumettre à l'autorité du souverain. Cette dimension coercitive est absente de la convention : agir une fois par convention n'implique aucun engagement de long terme. On ne "s'engage" pas en respectant une convention. La convention est un mode de coordination qui repose à *chaque instant* sur la volonté libre des individus.

En cela, elle se distingue également de la règle contraignante. Certaines règles sont des conventions, mais la plupart ont une dimension coercitive. La déviance par rapport à cette règle est sanctionnée, de manière à dissuader la possibilité de déviance en la rendant plus coûteuse que la coopération. L'acceptation de la règle entraîne une restriction de la liberté individuelle et le refus de toute autre alternative. Or la convention ne tient pas grâce à une hiérarchie soumettant l'individu, elle tient par la seule force de sa rationalité et de l'intérêt personnel.

Enfin la convention ne s'appuie pas non plus sur une obligation déontique, intériorisée par les individus, et les incitant à la respecter pour elle-même. Pour Lewis, l'acceptation d'une règle de conduite normative repose sur une inversion du rapport entre règle et action qui prévaut dans la convention. L'acteur, agissant par respect de la norme, accepte la règle avant d'agir et aligne son action sur cette règle. A l'inverse la régularité de comportement, l'action préexistent à la convention qui n'est en définitive que la résultante de pratiques déjà existantes. Ces pratiques, ces actions, ne sont pas apparues par conformité à une règle quelconque mais par la force de la volonté et de la rationalité individuelle. De même elles ne se maintiennent pas parce qu'elles sont intrinsèquement bonnes, collectivement souhaitables ou individuellement souhaitées : une convention ne vaut pas pour elle-même mais en ce qu'elle constitue un moyen de se coordonner. La convention n'est donc pas normative⁴².

La convention est un voile dont la trame est constituée d'individus libres de faire défection quand ils le désirent. Cette mosaïque interindividuelle n'a pas d'existence propre. Elle n'est pas une réalité distincte de la réunion, à chaque instant, de toutes les volontés individuelles.

Lewis fonde son analyse de la coordination par convention sur deux grands principes. Le premier est l'idée que l'acte de coordination précède l'apparition de la convention, et ne nécessite à aucun moment son explicitation. Cet aspect le distingue de la tradition contractualiste. Le second est qu'une convention peut se maintenir par la seule force de l'intérêt et de la rationalité. La convention n'implique donc aucune subordination ni

⁴¹ "Un contrat social ne peut être une convention parce qu'il lui manque une alternative." (Lewis, 1969, p.95)

⁴² Même si elle peut le devenir, comme on l'a vu avec l'analyse d'Ullmann Margalit [1978].

coercition. La théorie de Lewis peut donc apparaître comme une solution au problème récurrent que rencontre l'analyse économique individualiste à savoir : comment expliquer l'existence de règles collectives, sans pour autant subordonner l'autonomie de l'individu à un ordre collectif transcendant ? La réponse de Lewis séduit par sa simplicité : parce que tout le monde y a intérêt, et que, pour que de telles règles émergent, il n'est nul besoin de se mettre explicitement d'accord.

24- De Hume à Lewis : la disparition du temps historique.

Lewis est un disciple de Quine. Mais c'est à Hume qu'il emprunte le concept de convention. Il souligne ainsi, dans l'introduction de son ouvrage, qu'il propose une théorie *"dans la lignée de celle de Hume, dans sa discussion sur l'origine de la justice et de la propriété."* (Lewis, 1969, p.3).

Cependant, Lewis utilise le concept Humien dans un cadre analytique fondamentalement différent du cadre pratique et historique dans lequel Hume observait le rôle et l'efficacité des conventions. La théorie de Lewis intègre le concept de convention dans un cadre logique qui exclut le temps. Voilà sans doute, au fond, la raison pour laquelle sa théorie "parle" directement aux théoriciens des jeux. Elle semble parvenir à intégrer un concept issu de la philosophie politique, dans un cadre purement logique, proche de celui de l'économie théorique.

241- L'héritage de Hume.

"Cette convention n'est pas de la nature d'une promesse, car les promesses elles-mêmes, comme nous le verrons plus loin, naissent des conventions humaines. Ce n'est qu'un sens général de l'intérêt commun, sens que les membres de la société s'expriment les uns aux autres, et qui les conduit à régler leur comportement selon certaines règles. J'observe qu'il sera de mon intérêt de laisser autrui en possession de ses biens pourvu qu'il agisse de même avec moi. Il a conscience d'avoir un intérêt semblable à régler sa conduite. Quand ce sens commun de l'intérêt est mutuellement formulé et qu'il est connu des deux, il produit une résolution et un comportement qui lui correspondent. Et cela on peut l'appeler avec assez d'à propos une convention entre nous, bien que fasse défaut la médiation d'une promesse, puisque les actions de chacun d'entre nous font référence à celles de l'autre et sont accomplies en supposant que de l'autre côté, on doit aussi accomplir quelque chose. Deux hommes qui tirent aux avirons d'une barque le font selon

un accord ou une convention, bien qu'ils n'aient jamais échangé de promesses." (Hume, 1740, vol III, 2, 2, p.90)

L'introduction de l'ouvrage de Lewis débute par cette citation de Hume, dans laquelle ce dernier définit son concept de convention, l'oppose à la promesse et le fonde sur l'intérêt commun. Ce que Lewis emprunte à Hume est cette idée essentielle selon laquelle la convention découle exclusivement de l'intérêt commun.

Hume [1740, p.94-96] oppose clairement son concept de convention à la théorie du contrat social hobbesien⁴³. Il souligne ainsi que l'état de nature est une "*pure fiction qui n'est pas différente de celle de l'âge d'or inventée par les poètes*". Cette fiction méthodologique est d'ailleurs explicitement conçue comme telle par Hobbes. Elle est destinée à asseoir sa démonstration rationnelle de la supériorité de l'état. Or, dans la théorie humienne, la solution politique à "*l'égoïsme et la générosité limitée*" (Op cit., p.95) ne résulte pas d'une construction abstraite à partir du principe de la *table rase*, mais émerge progressivement au fil des interactions humaines, chacun percevant peu à peu que l'intérêt commun réside dans le respect de certaines conventions. Il existe donc bien un artifice humain⁴⁴, mais cet artifice n'est pas construit et ne fait pas l'objet d'un contrat : il émerge à partir de la puissance de l'intérêt commun bien compris.

Cette opposition entre Hume et Hobbes naît d'une différence d'appréciation de l'efficacité de la rationalité calculatoire. Pour Hobbes, l'état civil est construit et géré selon les *calculs* du souverain, qui cherche à obtenir un enrichissement de son royaume et donc de ses souverains. La justification de l'état civil se fait également par un *calcul* : Hobbes démontre que aux individus qu'ils ont "intérêt" à respecter le contrat social que leur propose le souverain. La construction de l'artifice a précisément pour but d'établir un univers logiquement organisé, qui puisse être décrit de façon mathématique et permette l'application d'un mode de rationalité calculatoire. Selon Hobbes, l'efficacité du contrat social dépend donc de la possibilité de "calculer" son intérêt. Le scepticisme que Hume entretient vis-à-vis de la rationalité calculatoire le conduit au contraire à privilégier la force de la coutume et de la tradition contre l'idée d'un état civil construit de façon abstraite et instauré par un calcul rationnel. Hume ne pense pas qu'il soit possible de s'extirper ainsi du domaine pratique pour instaurer un artifice mathématique. Il ne croit pas à la possibilité d'appliquer aux affaires humaines un langage mathématique, construit hors de l'action

⁴³ Sur l'opposition Hume/Hobbes, on peut consulter, outre l'introduction de Saltel [1993] à l'édition de poche du troisième tome du *Traité de la Nature Humaine*, l'étude de Simone Goyard Fabre [1994, p.62-64].

⁴⁴ "*le remède dès lors ne provient pas de la nature mais de l'artifice.*" (Hume, 1740, p.89)

elle-même⁴⁵. Chez Hume, l'ordre ne naît pas d'un raisonnement abstrait opérant par un arrêt momentané, une mise entre parenthèse, du temps et l'action. Au contraire, Hume met en évidence la force de l'*Histoire* qui tient ensemble les individus parce qu'ils ont *intérêt* à respecter les habitudes sociales communes qui les lient. L'Histoire commune donne aux acteurs la possibilité de se comprendre tacitement, spontanément. L'ordre naît de la coutume. Il n'est pas le fruit d'un calcul mathématique situé hors du temps.

Lewis en reformulant la théorie humienne dans les termes modernes de la théorie des jeux revitalise donc une tradition alternative au contrat social. Mais il intègre le concept de convention dans un cadre abstrait, logique, calculatoire, instrumental. C'est sans doute la raison pour laquelle la théorie de Lewis constitue un moment important de la réflexion néoinstitutionnaliste. Lewis redéfinit le concept humien dans les termes modernes de la théorie des jeux. Il donne au concept de Hume une dimension analytique qu'il n'avait pas. Ce faisant, la théorie de Lewis semble proposer une conception de l'ordre social qui s'affranchit du caractère coercitif du contrat social hobbién, tout en conservant une assise rationaliste conforme à la démarche méthodologique de l'économie théorique. Lewis, d'une certaine manière extrait le concept de convention de la temporalité historique, irréductible à la logique dans laquelle Hume l'avait conçu. Cependant, Lewis ne propose pas, en réalité, de théorie de l'ordre social. Son projet philosophique se situe, comme on l'a vu, à un tout autre niveau. C'est précisément parce que Lewis cherche à résoudre un problème purement abstrait qu'il peut situer le concept de convention dans un cadre "anhistorique" exclusivement défini par la rationalité instrumentale des agents. Lewis fait jouer à ce concept un rôle absolument différent de celui que lui faisait jouer Hume.

Ainsi, de Hume à Lewis, un glissement s'opère vers une conception calculatoire, et non plus coutumière, de la convention. Mais, avec ce glissement, le concept de convention change de statut : il ne désigne plus une modalité *pratique* de coordination, mais un concept abstrait permettant d'exprimer l'idée de nécessité et de restaurer par là même le principe de "l'analyticité". C'est sans doute là l'origine d'une certaine forme de quiproquo entre ce que cherche à dire Lewis, et ce que la théorie des jeux cherchera à lui faire dire. En effet, il y a un point pour lequel la théorie de Lewis, en raison même de son objet, ne peut être d'aucune utilité : c'est la question de l'émergence. Or, c'est cette question qui préoccupe, précisément, les théoriciens des jeux.

⁴⁵ Hume précise sa conception du langage dans un court passage (issu d'un ouvrage dans lequel il reprend de manière concise les thèmes développés dans le livre III du *Treatise in Human Nature*). Il souligne : "Ainsi le langage, les mots et la langue sont fixés par convention et accord mutuel entre les hommes." (Hume, 1751, p.230-231).

242- Une théorie logique de la convention.

La théorie de Lewis ne peut rien dire de la question pratique de l'émergence. Lewis construit même explicitement sa théorie de manière à éviter cette question de la genèse.

L'objectif de Lewis est en effet de répondre au défi de Quine concernant l'impossibilité de parler d'un fondement "conventionnel" du langage. Lewis situe donc son analyse sur le plan strictement logique : la structure logique des jeux de coordination nous montre qu'*a priori* il y a plusieurs issues possibles équivalentes. Il faut donc nécessairement qu'il existe une convention pour que seule une de ces issues s'impose comme la "solution". Cette réflexion logique étant opérée, il reste à comprendre comment des individus se mettent d'accord pour le choix d'une convention de langage : c'est sur ce moment de la genèse que porte le paradoxe mis en évidence par Quine.

Lewis règle cette question en niant sa pertinence et en développant une théorie de la diffusion et du maintien de la convention qui est censée rendre sans objet la question de la genèse. L'analyse de Lewis est ainsi toute entière construite contre l'idée que la question des "origines" est importante : la convention fonctionne par oubli des origines. Il n'existe pas, contrairement à ce que l'on constate dans le cadre du contrat social hobbesien, un moment mythique où l'on passerait d'une absence de convention à une convention générale commune. L'émergence de la convention est progressive et il est illusoire de vouloir fixer un moment précis où elle apparaîtrait *ex nihilo*, par opposition à un avant où elle n'existait pas encore. C'est donc précisément en mettant l'accent sur le processus auto-renforçant de la convention, que l'analyse de Lewis parvient à la fois à démontrer la nécessité logique de la présence d'une convention et à déjouer le piège tendu par Quine aux conventionnalistes. En quelque sorte la mise en évidence du *processus* de diffusion et de maintien de la convention permet à Lewis de supprimer toute théorie de la *genèse*.

Cette force du concept de convention avait déjà été mis en évidence par Hume, mais ce dernier s'appuyait sur une conception *sociale* de l'action humaine, se déroulant dans un processus *historique* lent absolument irréductible à un raisonnement rationnel. La convention fonctionne précisément chez Hume comme un palliatif à la rationalité. Or, ce n'est plus tout-à-fait le cas dans la théorie rationaliste lewissienne. Certes, on peut dire que la convention vient "au secours" de la rationalité dans la mesure où elle constitue un point d'ancrage des anticipations qui permet de résoudre le problème de coordination. Cependant qu'est ce qui, dans la convention de Lewis, n'est pas *in fine* réductible à de la rationalité ? Le processus fondamental durant lequel la convention se diffuse est entièrement réductible au mécanisme rationnel de la "préférence pour la conformité". Son maintien est du au phénomène du *Common Knowledge* par lequel les croyances de

chacun des individus sont comme projetées sur la place publique et rendues transparentes grâce à la rationalité infinie de tous les acteurs. C'est bien parce qu'ils sont rationnels, et non parce qu'ils respectent les us et coutumes de leur société, que les individus parviennent à se sortir du problème de coordination qui se pose à eux. La convention lewissienne repose donc exclusivement sur la rationalité des agents. Elle ne s'appuie pas sur leur socialisation.

Un moment requiert cependant une faculté distincte de la rationalité : l'instant précis où l'individu, confronté à un problème de coordination, repère une convention naissante et y adhère. Ce moment peut comporter deux aspects. L'individu se retourne sur ses actions antérieures pour étudier comment il a pu, par le passé, résoudre un problème de coordination similaire. Ce mécanisme de duplication des solutions passées peut être qualifié de "force du précédent". Mais il n'y a pas toujours un précédent heureux, sauf à considérer que la convention est toujours déjà là et qu'il n'y a donc jamais de problème de coordination. Il existe nécessairement un moment où l'individu n'a pas encore appris à respecter la convention, parce qu'il est pour la première fois confronté à ce type de situation, ou bien encore parce que la situation à laquelle il est confrontée est nouvelle pour tous les protagonistes et ne fait donc pas encore l'objet d'une convention de comportement. Il existe donc un mécanisme d'apprentissage et de découverte de la convention que Lewis passe sous silence. Il ne peut pas en être autrement : Lewis ne traite pas d'une question pratique mais cherche seulement à faire apparaître le rôle logique que jouent les conventions.

On touche là au problème fondamental concernant cette théorie de la convention : l'absence d'Histoire. Pour répondre à Quine, Lewis doit nécessairement situer son analyse sur le plan logique, et faire "table rase" de l'environnement de l'individu rationnel. Il procède ainsi de façon hobbienne, en dégageant l'individu de tout ancrage social afin, de faire apparaître le besoin de règle commune (ici, la convention). Or, la convention de Hume s'appuie justement sur l'Histoire. C'est parce qu'ils se situent dans un temps historique que les individus que dépeint Hume sont capables d'activer des repères communs et, le cas échéant, de les adapter progressivement aux situations nouvelles. En ne retenant que l'aspect rationnel des individus, Lewis se coupe inévitablement de ce qui peut expliquer la genèse et le fonctionnement des conventions. Il ne pouvait sans doute pas faire autrement : le temps logique de la déduction exclut le temps historique de l'action. Son objectif justifie son choix théorique. Ce choix constitue cependant un obstacle de taille lorsque l'on souhaite, comme c'est le cas d'une partie des théoriciens des jeux "néoinstitutionnalistes", s'appuyer sur Lewis pour comprendre comment fonctionnent et émergent les institutions économiques.

243- Une convention sans dimension "pratique".

Lewis cherche à restaurer l'analyticité. Pour ce faire, il cherche à démontrer que les conventions sont nécessaires. Mais, pour réfuter la critique de Quine, il doit également montrer que la convention ne repose "sur aucun fait ou fiction d'autorité". Autrement dit, la convention est parfaitement arbitraire. Du point de vue logique, le choix d'une convention particulière est absolument sans importance. La nécessité ne tient pas au choix particulier d'une convention, mais au fait qu'une convention doit être choisie, quelle qu'elle soit.

Or, il se trouve que la question pratique de la *genèse* tient précisément aux choix d'une convention de nature *particulière*. La question qui se pose dans le cadre d'une théorie de l'émergence n'est jamais : "pourquoi une convention a-t-elle été choisie ?" mais : "pourquoi *cette* convention particulière a-t-elle été choisie ?". Or, cette seconde question n'a strictement aucun sens sur le plan logique. La seule idée que Lewis puisse défendre contre Quine, c'est qu'il existe une nécessité liée à la présence d'une convention. Il ne peut justement pas faire remonter la nécessité jusqu'au choix précis d'une convention particulière. Pour le dire de manière concise : Lewis montre que la convention est purement arbitraire. Or, la seule question qui vaille sur le plan de la théorie des institutions est précisément d'éclairer les raisons et les mécanismes précis qui ont donné lieu au choix d'une institution particulière. La problématique de la *genèse* et celle de la nécessité logique s'excluent inévitablement.

C'est précisément parce qu'il considère que la convention est arbitraire que Lewis situe son analyse dans un univers logique qui exclut l'histoire. L'exclusion du temps historique est intrinsèquement liée au fait que, précisément, Lewis ne se pose pas la question de la *genèse*. Sa théorie n'est donc pas "miraculeuse" : elle n'apporte aucune réponse rationaliste à la question de l'émergence des règles. Pour éclaircir cette question, il faut donc retrouver l'Histoire, ce qui revient à élargir la perspective rationaliste qui caractérise la théorie des jeux, ou retrouver la coercition, et accepter la description hobbenne de la mise en place de l'ordre social.

La théorie de Lewis ne peut à la fois se situer sur le plan *logique* de la polémique avec Quine et sur le plan *pratique* de la question du fonctionnement des institutions. Ces deux niveaux de réflexion ne s'excluent pas, mais ils sont irréductibles l'un à l'autre pour une raison simple : ils ne s'appuient pas sur une même notion du temps. L'argumentation de Lewis emprunte à Hume un concept, celui de convention, qui ne prend toute sa dimension que lorsqu'il est situé dans un temps historique. En le situant dans le cadre logique dont le temps est absent, Lewis lui fournit des bases rationnelles, mais lui donne

un caractère statique que ne peut avoir la convention. Utiliser Lewis pour donner un fondement rationnel aux institutions nécessite donc de parvenir à réintroduire une certaine forme de temps historique, ou de fournir d'autres explications au point aveugle de cette théorie lewissienne : la genèse de la convention. C'est là l'enjeu des développements menés en théorie des jeux à partir de l'analyse logique lewissienne.

3- La convention comme fondement des institutions : une impasse ?

L'analyse de Lewis a semblé rendre possible une analyse "rationnelle" du rôle et du fonctionnement des règles collectives, menée exclusivement à partir de l'étude du choix rationnel d'individus libres et égoïstes⁴⁶. Cette tentative a trouvé son domaine d'application privilégié dans la théorie économique des institutions. Pourtant, la théorie de Lewis n'est pas directement "applicable". Elle pose en effet un problème crucial parce qu'elle substitue en partie la "rationalité" à "Histoire". Lewis utilise en effet un concept humien s'appuyant sur l'Histoire, mais l'intègre dans le temps logique de la déduction et de la rationalité calculatoire qui repose sur une négation de l'Histoire. C'est donc un soubassement bien précaire pour fonder une théorie satisfaisante des institutions. Pour parvenir à rendre compte de l'émergence des règles institutionnelles à partir du concept de Lewis, il faut réintroduire une certaine forme de "coercition", de "socialisation" ou de "temporalité".

C'est à Schotter [1981] que revient le mérite d'avoir tenté, pour la première fois, d'expliquer l'émergence des institutions à partir d'un cadre stratégique. Schotter mobilise pour ce faire le concept de Lewis, mais lui adjoint une dimension coercitive qui le dénature en partie. De plus l'analyse de Schotter, si elle introduit une dimension temporelle repose sur une hypothèse d'unanimité des croyances qui exclut de fait toute "altérité". Ainsi, son modèle permet de concevoir un processus évolutif, mais dit peu de choses des véritables causes expliquant le choix d'une règle institutionnelle par un ensemble d'individus rationnels.

Choi [1994] a récemment tenté d'utiliser de manière différente l'idée de convention. Le pivot de l'étude de Choi est la mise en avant de la recherche par chaque individu de "l'approbation sociale". Cette conception "socialisée" de l'individu constitue un premier moyen de retrouver l'assise humienne de la convention. Mais il manque à cette étude une conception dynamique.

C'est le rôle que peut jouer une conception évolutionniste. La stratégie évolutionniste consiste à relâcher les hypothèses en matière de rationalité pour faire jouer un rôle plus important à l'habitude et à l'imitation. Les analyses pionnières d'Axelrod [1984] ont ainsi permis de représenter des processus évolutionnaires qui permettent de comprendre de

⁴⁶ Philip Pettit [1993, p.89] parle ainsi "*d'une explication de l'existence des normes par la main invisible : l'existence des normes est le résultat non intentionnel de l'agrégation des décisions individuelles dans des stratégies de coordination de l'action.*"

manière plus satisfaisante l'évolution des contextes stratégiques lorsque l'interaction se répète. Cette voie a été utilisée pour rendre une dimension dynamique au concept de convention par Sugden [1986] ou Young [1993]. Toutefois un simple *relâchement* des hypothèses de rationalité ne suffit pas à comprendre la dynamique des règles conventionnelles : il faudrait parvenir à *élargir* la notion de rationalité faute de quoi la théorie fait jouer au hasard un rôle explicatif qui paraît trop important.

31- L'essai de synthèse néoinstitutionnelle de Schotter

Le concept de Lewis sert de fondement au premier essai visant à formuler une théorie générale des institutions économiques fondée sur la théorie des jeux. Cette théorie est l'œuvre de Schotter [1981]. Elle propose une analyse mathématique de l'émergence des institutions à partir d'une conception "évolutionnaire-bayésienne" de l'équilibre de Nash.

311- Le projet de Schotter

La théorie de Schotter [1981] s'inscrit dans le cadre de la "nouvelle économie institutionnelle" que Williamson appelle de ses vœux. Schotter construit explicitement son étude contre l'analyse néoclassique centrée sur le marché, et rejette en particulier la modalité d'intégration de l'incertitude par le système des biens contingents. Il dénonce une forme de "rigidité" de la théorie néoclassique qui cherche à "*rationaliser un bouclage excessif de la théorie*". Schotter conçoit au contraire sa théorie comme un "*relâchement de cet aspect de la théorie*" qui " *vise à étendre son champ d'analyse et à la rendre plus adaptée à l'univers qu'elle décrit*" (Schotter, 1981, p.152). Schotter s'inscrit ainsi dans la lignée d'Arrow et Williamson. Il utilise, comme ces auteurs, le mode "informationnel" d'intégration de l'institution : "*Les institutions économiques et sociales sont d'abord des mécanismes informationnels qui complètent l'information contenue dans les prix marchands quand ces prix ne parviennent pas à coordonner entièrement les activités économiques*" (Schotter, 1981, p.143). Conformément à ce schéma, Schotter se donne donc comme objectif de définir des règles du jeu optimales : "*Nous considérons le gouvernement dans une économie mixte ou une économie de marché comme un agent dont le rôle est de prescrire un ensemble de taxes, lois et règles de comportement qui limitent de manière optimale le nombre de stratégies dont l'agent social peut disposer. L'essence de l'économie du bien-être est l'étude des règles du jeu optimales pour les situations économiques et sociales*" (*op cit.*, p.6).

Cette position "normative" s'accompagne cependant d'un objectif "positif" visant à restituer la manière dont les institutions émergent à partir du choix libre d'individus rationnels. Schotter construit sa théorie en deux temps. Elle débute par une étude des "structures" d'interaction nécessitant une norme, dont il tire une définition de l'institution inspirée de la convention lewissienne. Son étude se poursuit ensuite par la présentation d'un modèle "bayésien-évolutionnaire", fondé sur une version adaptée de la "tracing procédure" d'Harsanyi, qui permet selon lui de prévoir quelle institution émergera en fonction de la situation de départ.

312- Une "convention hobbienne" ?

Schotter s'inspire de l'analyse d'Ullmann-Margalit [1977], mais en modifie profondément l'esprit "structuraliste". Il cherche, en effet, d'une part à proposer une définition "commune" de l'institution qui puisse fonctionner dans tous les cas de figure, d'autre part à réduire la typologie d'Ullmann-Margalit à deux cas polaires : le "dilemme du prisonnier" et la situation de "pure coordination". A ces deux cas paradigmatiques, Ullmann-Margalit ajoutait un troisième cas, du type "guerre des sexes", dont elle pensait qu'il nécessitait l'existence d'une norme particulière : la norme de partialité préservant un *statu quo* favorable à l'un des deux joueurs.

Schotter pense que cette distinction n'est pas véritablement utile dans la mesure où le jeu de la négociation n'est qu'un sous-ensemble des jeux de coordination : "*la distinction entre des institutions préservant les inégalités et des institutions créées pour résoudre des problèmes de coordination est artificielle. (...) Une institution préservant les inégalités est au départ une institution créée pour régler un problème de coordination.*" (Schotter, 1981, p.27). Ainsi, pour Schotter, la "norme de partialité" ne préserve pas un *statu quo*, elle "tranche" un différend en faveur de l'un des joueurs. Schotter unifie donc la catégorie des jeux de coordination en la rabattant, en fait, sur le cas extrême de la pure coordination. Une telle généralisation est abusive dans la mesure où l'on ne peut en toute rigueur mettre sur un plan identique une situation dans laquelle les joueurs sont parfaitement indifférents quant à l'issue qui doit être choisie, et une situation marquée au contraire par un conflit quant au choix d'une des issues conjointes possibles. La généralisation de Schotter fait artificiellement disparaître la question des conflits entre individus, en ce qui concerne au moins la question de l'émergence des institutions. En effet, la situation mêlant conflit et intérêt commun tire sa spécificité de l'existence d'une divergence de vue concernant l'issue commune. Dans le cas de la *pure coordination*, les agents sont indifférents entre les deux issues conjointes possibles. Dans le cas du *dilemme du prisonnier*, les agents préfèrent tous les deux la même issue conjointe (même s'ils préféreraient en fait une des

deux issues non conjointes). En éliminant le cas particulier de la divergence d'intérêt, Schotter élimine donc, artificiellement, un premier problème lié au choix des équilibres de coordination, et donc à la genèse des institutions.

Cette modification de la typologie d'Ullmann-Margalit n'est pas cependant fondamentale au regard de la modification de la théorie lewissienne. Schotter fonde en effet sa définition de l'institution sur une conception révisée de la convention lewissienne. En effet, conformément cette fois aux réflexions d'Ullmann-Margalit, il prend en compte le problème de la défection dans le dilemme du prisonnier, ou du caractère "stratégiquement instable" à long terme (pour reprendre l'expression d'Ullmann-Margalit) des équilibres de coordination défavorisant une partie des joueurs. Il souligne ainsi qu'une convention à la Lewis est sous la menace permanente de la défection dans des cas où il n'y a pas stricte coïncidence des intérêts⁴⁷. Il adjoint donc à l'idée de convention le principe d'un dispositif de coercition stabilisateur : *"Une autorité externe doit être instituée pour préserver certains droits [comme les droits de propriété]. C'est le rôle de l'état. En conséquence, nous devons modifier la théorie de Lewis comme suit : une institution est une régularité de comportement social acceptée par tous les membres de la société, qui détermine leur comportement dans des situations spécifiques récurrentes, et qui peut s'auto-entretenir, ou être au contraire maintenue par une autorité externe"* (Schotter, 1981, p.11). Cette définition constitue le cadre institutionnel commun à l'ensemble des interactions dans lesquelles la présence d'une institution est nécessaire.

La convention de Lewis est donc associée à un dispositif stabilisateur et coercitif "externe". Schotter souligne que cet "acteur externe" est la plupart du temps l'Etat. Concrètement cette modification du concept de Lewis se traduit par l'ajout d'une clause particulière à la définition lewissienne de la convention : *"Si quelqu'un ne respecte pas R [l'institution], on sait que certains ou la totalité des autres joueurs ne la respecteront plus non plus et que les paiements associés à ces stratégies déviantes seront moins intéressants que le paiement assuré par R."* (Schotter, 1981, p.11). Or, l'ajout de cette clause ne constitue pas un simple amendement de la théorie de Lewis mais une modification radicale. Elle souligne tout simplement que la convention n'est plus auto-renforçante, ne tient plus par la force de l'intérêt commun mais par la peur de la sanction. Or une convention coercitive n'est plus une convention au sens de Lewis. Elle n'est plus libre de *"tout fait ou fiction d'autorité"*.

La seconde modification essentielle apportée par Schotter à la théorie de Lewis se situe non plus dans sa définition de l'institution, mais dans la présentation du problème à

⁴⁷ Les réticences que nous avons exposées au début de ce chapitre (lors de la présentation de la théorie d'Ullmann Margalit) quant au principe d'un "équilibre de Nash stratégiquement instable" s'appliquent donc aussi à la théorie de Schotter.

résoudre. Sa stratégie théorique consiste en effet à postuler l'existence d'une situation initiale "hors institution" qu'il qualifie "d'Etat de Nature" : "*Si la science économique souhaite véritablement étudier l'évolution et la fonction des institutions sociales, alors il lui sera fort utile de s'appuyer de façon claire et nette sur l'hypothèse d'un état de nature initial.*" (Schotter, 1981, p.51). Or, on l'a vu, l'hypothèse rhétorique de l'état de nature est une caractéristique de la conception hobbiennne de l'institution contre laquelle Hume s'est élevé. Lewis rappelle ainsi : "*Il faut se souvenir qu'une convention comporte nécessairement une alternative, alors qu'un contrat social ne requiert aucune autre alternative qu'un état de nature. L'état de nature n'est pas un état dans lequel nous pouvons atteindre un équilibre de coordination*" (Lewis, 1969, p.95)⁴⁸. Schotter, en se donnant comme point de départ un état de nature neutre de tout lien social fait donc un pas de plus vers une conception "hobbiennne" de la convention.

Ainsi, la "synthèse" que réalise Schotter en regroupant l'ensemble des "types de norme" envisagés par Ullmann-Margalit sous une définition unique fondée sur la convention opère comme un révélateur de la nature hobbiennne de la convention de Lewis lorsqu'elle est utilisée comme fondement des institutions. Lewis a intégré l'idée humienne dans un cadre rationaliste qui ne lui correspondait pas, mais l'a située à un niveau de généralité qui excluait la question du conflit. Schotter en cherchant à utiliser ce concept est contraint de lui donner finalement une assise hobbiennne caractérisée par la présence de la "coercition" et la référence à un "état de nature". Cependant, l'expression "convention hobbiennne" a tout d'un oxymore : l'analyse de Hobbes se caractérise en effet par l'existence d'un contrat, et, précisément, contre l'idée d'une convention émergeant progressivement. Schotter, en conservant le concept de convention tout en lui donnant une assise hobbiennne, vide de sens le principe de la convention. Car son concept d'institution désigne tout à la fois une règle instituée par *convention* et une règle instituée *autoritairement*. Cette dualité est une source de confusion que l'analyse mathématique de Schotter ne supprimera pas.

313– Un modèle évolutionniste-bayésien

Le second temps de l'étude de Schotter consiste à développer une théorie *mathématique* de l'émergence des conventions. Pour y parvenir, Schotter utilise une méthode évolutionniste. Il suppose que les jeux étudiés sont indéfiniment répétés. Les jeux de base, qui se jouent en un coup, deviennent ainsi les étapes répétitives d'un "super-jeu". C'est l'issue de ce super-jeu qui est étudiée et interprétée comme étant une institution.

⁴⁸ Voir aussi Ullmann-Margalit [1978, p.62-67).

L'analyse de Schotter s'inspire de la "*tracing procédure*" par laquelle Harsanyi cherche à raffiner la notion d'équilibre de Nash. Cette procédure est une des procédures de "raffinement de l'équilibre de Nash" permettant de sélectionner un équilibre unique. La stratégie d'Harsanyi consiste à supposer que chacun des joueurs émet un jugement subjectif sur les probabilités qu'ont chacun des autres joueurs de choisir chacune de leurs stratégies pures. Harsanyi démontre alors qu'il existe un "sentier" unique associé aux probabilités initiales et qu'il est donc possible de définir un équilibre unique⁴⁹. Schotter utilise cette technique mais l'adapte dans un sens "évolutionniste" : "*La procédure d'Harsanyi débute par une solution bayésienne naïve puis modifie le comportement prescrit jusqu'à ce qu'un ensemble de stratégies d'équilibre soit choisi. Dans la procédure de traçage, il n'y a pas de différence entre un comportement optimal et un comportement d'équilibre. La procédure que nous employons n'a pas besoin d'avoir cette propriété, puisqu'elle ne repose pas sur une utilisation unique de la procédure menant à la solution. On répète la procédure encore et encore jusqu'à ce que l'on obtienne une solution*" (Schotter, 1981, p.67). Les agents qu'il étudie ne mettent pas en place une stratégie à rebours à partir de l'équilibre, mais tâtonnent en permanence et cherchent à réajuster leurs croyances au fur et à mesure que le jeu se déroule. Ils ne déterminent pas de solution *a priori*.

Le concept de solution que mobilise Schotter est donc fondé sur le principe d'un état "absorbant" les chocs : "*Un état qui «absorbe les chocs» est un état dans lequel chacun des agents attribue la probabilité de «1» au choix, par les autres individus, d'une stratégie particulière. Cet état est tel que lorsque chaque agent observe ensuite le comportement effectif des autres individus, il se produit exactement ce qu'il avait prévu.*" (Schotter, 1981, p.53)⁵⁰. Un tel état constitue une issue "*stochastiquement stable*" c'est-à-dire qu'elle est soutenue par des anticipations convergentes qui prévoient son maintien : les actions sont en équilibre, et les anticipations prévoyant cette issue sont reconduites à chaque nouvelle période.

Le processus markovien⁵¹ de convergence vers une issue stable s'appuie sur un principe "bayésien" de révision des croyances : les individus enregistrent l'information que leur révèlent les coups immédiatement précédents de chacun et modifient ainsi leurs

⁴⁹ Pour une présentation de la procédure de traçage, voir par exemple : Shubik (1982, p.305). Il note ainsi : "*La procédure de traçage peut être assimilée à un processus intellectuel par lequel les joueurs rationnels s'arrangent pour coordonner leurs plans et leurs prévisions concernant les actions des autres joueurs afin de les faire converger vers un équilibre particulier qui est la solution du jeu.*" (Shubik, 1982, p.307-308).

⁵⁰ Schotter note la proximité de son raisonnement avec le principe des "anticipations rationnelles" lorsque celles-ci se "calent" sur un état stable "absorbant" les chocs. Les anticipations des agents sont auto-validatrices.

⁵¹ Le processus est dit "markovien" parce que la révision des croyances ne porte que sur le coup immédiatement précédent, et non pas sur l'ensemble des coups précédents. (Bicchieri, 1996, p.29).

anticipations sur les coups à venir, ce qui leur permet de choisir une stratégie optimale⁵². A un moment donné, l'ensemble des croyances des agents se trouve validé par les choix effectivement réalisés par chacun. On est alors à l'équilibre.

Schotter prend comme point de départ un "état de nature" marqué par "l'unanimité" des agents : avant que l'histoire du jeu ne commence, les individus ont tous les mêmes croyances. Ils utilisent par la suite tous le même procédé bayésien de révision des croyances. Enfin, ils ont tous accès aux mêmes informations. De ce fait, ils ont tous, toujours, les mêmes croyances. On reconnaît donc ici le principe de la "doctrine Harsanyi" critiquée par Binmore. Ainsi, "puisque nous démarrons l'analyse à partir d'une norme uniforme et que chaque joueur transforme sa propre distribution de probabilité de manière identique aux autres, la norme générée, $p=(p_1, p_2)$ a la propriété suivante : $p^t_1 = p^t_2$ pour tout $t=1, 2, \dots$ (ou "t" est l'indice correspondant à chaque joueur). En d'autres termes, chaque vecteur des probabilités pour chaque joueur est identique tout au long de l'analyse et, à chaque instant, la norme existante est un ensemble de vecteurs de probabilité tous identiques." (Schotter, 1981, p.72). Cette unanimité permanente des joueurs trouve un point d'ancrage stable lorsque les anticipations se calent sur une des issues représentant un des équilibres en stratégie pure dans la mesure où personne n'a individuellement intérêt à en dévier dans le long terme. Ainsi, dans le cadre du dilemme du prisonnier répété, une fois atteinte l'issue coopérative, chaque joueur réalise que toute déviation sera punie durant les périodes suivantes, et calcule donc qu'il a individuellement intérêt à continuer de coopérer. Il sait en effet que chacun faisant un raisonnement identique au sien, et chacun ayant intérêt à perpétuer la coopération plutôt que d'être indéfiniment puni, l'issue coopérative se maintiendra. L'issue coopérative est donc stochastiquement stable : c'est une "régularité de comportement", c'est-à-dire une institution, au sens de Schotter. La défection mutuelle peut constituer une autre institution, pour les mêmes raisons.

Comme le précise Schotter, cette description de l'émergence des institutions laisse une part importante au *hasard*. Le jeu peut se dérouler de différentes manières et donner différentes institutions. L'issue atteinte dépend étroitement du choix de la distribution de probabilité *a priori*, ainsi que du taux d'actualisation que les individus appliquent à leurs gains⁵³. Ainsi Schotter conclut : "l'ensemble des institutions existantes à un moment

⁵² Le mode de révision des croyances est basique : le choix effectivement réalisé au tour précédent par chacun augmente la probabilité que les autres lui prêtent de refaire ce choix au tour suivant. (Schotter, 1981, p.72).

⁵³ Un paramètre " α " permet de pondérer les prévisions de gain en fonction de la préférence des individus pour le présent. Il est clair, par exemple, qu'une préférence faible pour le présent est nécessaire à l'émergence d'une issue coopérative dans le dilemme du prisonnier répété.

donné est vraiment un accident historique et ce qui existe aujourd'hui aurait pu évoluer de manière différente. Tout ce que nous savons à propos de l'histoire institutionnelle de la société est ce qu'il est arrivé, non pas ce qui aurait pu arriver." (Schotter, 1981, p.79). Sa théorie vise donc à éclairer le processus d'harmonisation des croyances, mais l'issue de ce processus reste fondamentalement indéterminée.

314- Une explication partielle

Le grand mérite de Schotter est d'avoir cherché à construire une théorie de l'émergence des institutions. De fait, sa théorie constitue une avancée importante sur la voie de l'explication du phénomène de l'émergence des institutions. Schotter parvient à montrer que l'institution émerge à partir d'une répétition d'interactions identiques. Il met de plus en évidence le rôle des normes collectives dans le processus d'émergence de l'institution. Le concept de "norme" désigne en effet chez Schotter (1981, p.65) l'ensemble des probabilités que chacun des joueurs attribue respectivement à chacun des autres individus en ce qui concerne le choix de chacune des stratégies possibles⁵⁴. Or, Schotter montre que l'émergence d'une institution particulière est profondément liée à cet ensemble de probabilités. Dès lors : "*la norme est le fondement probabilistique informationnel à partir duquel les institutions sont créées.*" (Schotter, 1981, p.53). Schotter parvient donc à montrer qu'un individu rationnel peut parfaitement intégrer l'existence de normes collectives, les mobiliser pour anticiper le comportement des autres, et que c'est par ce biais que l'on peut parvenir à expliquer comment les institutions émergent. Il opère un lien important entre la structure des croyances des agents et l'émergence des institutions. Il parvient par là même à remplir une partie du programme qu'il s'est fixé : l'institution émerge bien à partir des interactions d'individus rationnels. Elle apparaît comme un équilibre, de type équilibre de Nash, en jeu répété.

Mais son analyse est incomplète. En effet, dans la théorie de Schotter, le choix de l'institution dépend de trois éléments : la norme de départ, les chocs stochastiques qu'elle subit, la structure de la situation d'interaction. Or, sa théorie *postule* l'existence d'une norme de départ. Elle attribue les chocs stochastiques au "hasard". Elle dissimule la diversité des structures possibles sous une définition unique de l'institution. Elle n'explique donc pas l'*émergence*, mais plutôt la diffusion ou l'évolution d'une norme en institution. C'est par ce biais que Schotter parvient à intégrer l'idée d'institution sans

⁵⁴ Dans le cadre d'un jeu à deux joueurs, la norme est un ensemble de deux vecteurs de dimension quatre. Chacun est en effet constitué des quatre probabilités que chacun des deux joueurs attribue à l'occurrence du choix d'une des quatre stratégies possibles par son partenaire.

modifier le comportement standard des agents : l'institution est déjà présente sous la forme d'une norme au départ du jeu. L'activité des agents ne fait que l'actualiser. En ce sens c'est une forme d'explication très particulière de l'émergence, puisque l'on doit supposer que l'institution était en quelque sorte déjà là au départ.

Schotter exploite cependant au maximum le cadre d'analyse "bayésien" qu'il se donne. En effet, ce cadre ne permet pas d'envisager la diversité des acteurs. La "doctrine Harsanyi" qui caractérise ce cadre d'analyse suppose en effet une parfaite unanimité des croyances des agents. Ce que Schotter ajoute tient à l'évolution de cette croyance collective. Il ne remet à aucun moment l'homogénéité des anticipations des agents. La temporalité du processus est mécanique, puisque la seule nouveauté susceptible d'intervenir est liée à "l'extérieur" du modèle. L'action des individus eux-mêmes n'influe pas sur le contexte de l'interaction. Or, dans le cadre d'une temporalité mécanique, rien n'émerge véritablement. Au mieux on parvient à montrer comment un élément se transforme. C'est bien le cas ici : Schotter n'explique pas comment l'institution émerge mais comment une norme initiale peut devenir une institution.

Il existe cependant chez Schotter un acteur capable de faire apparaître de la nouveauté : l'Etat. C'est finalement à lui que cette théorie s'adresse : *"la question qui importe n'est pas seulement de savoir quel équilibre conventionnel s'établira, mais combien cet équilibre mettra de temps à s'établir. (...) Si le passage à une situation optimale est prévu pour être très long, cela justifie une intervention du gouvernement pour créer explicitement cette institution"* (Schotter, 1981, p.80). Cette courte allusion prend en effet tout son sens lorsque l'on cherche à mettre bout à bout l'ensemble des questions que pose l'analyse de Schotter. D'une part, comme on l'a souligné, Schotter fait jouer un rôle important aux dispositifs coercitifs qui dissuadent la déviance, et que l'on doit supposer être déjà présents au départ de l'analyse. D'autre part, il se donne comme point de départ un "état de nature" caractérisé par l'unanimité des croyances des agents⁵⁵. Enfin, il souligne le rôle potentiel que peut jouer l'Etat dans la mise en place des institutions. Ces trois éléments dessinent une conception hobbienne des règles collectives.

L'analyse de Schotter établit en définitive à quelle condition un processus stochastique, fondé sur des anticipations bayésiennes et un comportement rationnel, peut produire un comportement institutionnel : il faut que les individus soient *unanimes*. Or, le mécanisme qui produit cette unanimité ne peut pas être mis à jour puisque Schotter utilise un cadre d'analyse qui présuppose l'identité des profils cognitifs des acteurs. Le seul acteur qui puisse produire de la conformité est l'Etat, placé dans une situation

⁵⁵ On peut noter que l'expression "état de nature" est ici maladroit, puisque dans l'état de nature, les individus suivent déjà des normes, et qu'il existe déjà des dispositifs coercitifs...C'est une "nature" déjà très "civilisée".

transcendant les interactions des agents.

Pour utiliser le concept de convention comme fondement d'une théorie de l'émergence des institutions économiques, il est nécessaire de mobiliser une forme de temporalité pratique, ou de désigner un réformateur social capable de trancher les différends et de désigner une solution à l'ensemble des acteurs. Schotter hésite, en quelque sorte, entre ces deux méthodes. Il introduit une forme de temps mécanique, qui exclut l'apparition d'une nouveauté autre qu'extérieure à la situation d'interaction (c'est le rôle de la norme). Il suppose également qu'un acteur, l'Etat, est capable de dissuader la déviance et éventuellement de hâter le processus de convergence. Sur cette base, il produit une théorie permettant de montrer comment une norme peut évoluer en institution. C'est indéniablement une avancée importante. Cette théorie pose néanmoins deux problèmes. D'une part elle fait jouer un rôle à l'Etat sans véritablement lui donner une existence concrète. L'Etat apparaît en quelque sorte ici comme un *deus ex machina*. D'autre part, elle ne résout pas une question fondamentale dans une optique individualiste : d'où vient l'unanimité des acteurs ?

32- La théorie "humienne" de Young Back Choi.

Young Back Choi [1994], dans un livre récent, a tenté de reprendre le défi auquel Schotter s'était confronté : "*J'explique non seulement "pourquoi" les comportements sont guidés par des normes et règles, mais aussi "comment" ces guides sont générés à l'origine.*" (Choi, 1994, p.8). Son analyse s'inscrit dans une perspective plus générale, critique et ambitieuse, visant à reconstruire un nouveau paradigme économique à partir d'une nouvelle conception du comportement individuel.

On peut distinguer trois étapes dans la construction théorique de Choi. L'auteur part d'une réflexion sur l'incertitude. Récusant la conception standard – c'est-à-dire l'axiomatique de Savage et les raffinements bayésiens – Choi développe l'idée selon laquelle l'individu utilise pour se repérer dans l'incertain une vision du monde progressivement construite au fil de ses expériences. Cette "vision du monde" constitue un "*paradigme*" interprétatif qui lui est propre. Son analyse se nourrit de la critique par Binmore de l'attitude bayésienne naïve, ainsi que des expériences menées par Kahneman et Tversky. Choi tente de synthétiser une conception alternative du comportement dans l'incertain en mobilisant le concept de paradigme, hérité de Kuhn. Il considère que le comportement dans l'incertain peut être comparé au mécanisme de la recherche

scientifique. Ainsi, *“Les individus décident et agissent dans une situation donnée en fonction d'un paradigme identifié et approprié à cette situation”* (Choi, 1994, p.37). Les agents pallient donc leur rationalité limitée, leur incapacité à prévoir correctement l'avenir, en se construisant un outil permettant de discriminer rapidement les informations. Cette attitude paradigmatique entraîne bien sûr une certaine rigidité ou régularité de leur comportement. Un paradigme doit, pour être abandonné, accumuler un certain nombre d'informations le réfutant.

La seconde étape consiste à étudier l'homme en société à partir d'une réinterprétation “sociologique” des concepts moraux de *“sympathie”* et de *“spectateur impartial”* qui sont au cœur de la Théorie des Sentiments Moraux d'Adam Smith. Choi, dans une rapide synthèse de l'œuvre smithienne, estime pouvoir utiliser ces concepts de philosophie morale pour décrire un comportement social caractérisé par la recherche de l'assentiment des autres. *“Face à l'incertitude, les individus recherchent l'approbation et fuient la désapprobation”* (Choi, 1994, p.63). Le paradigme qui guide les individus dans l'incertain est donc immédiatement “social” puisque chacun cherche à agir en conformité avec les autres. Le rapprochement de l'attitude *cognitive* des individus face à l'incertitude et de leur attitude *sociale* conduit presque immédiatement au concept de convention lewissienne. *“Ces conventions ressemblent aux paradigmes individuels puisqu'elles aident les individus à réagir face à l'incertitude. Mais elles en diffèrent puisqu'elles sont de simples conséquences inintentionnelles de l'action d'individus cherchant à forger leur propre paradigme, plutôt qu'un choix social délibéré”* (Choi, 1994, p.68).

L'analyse de Choi repose donc sur un procédé simple et efficace : faire dépendre l'attitude cognitive des individus de leur enracinement social. Ce rapprochement est réalisé par la mise en évidence de la volonté de “conformisme” des individus qui souhaitent s'intégrer socialement. A compter du moment où le conformisme est un des motifs d'action, la notion de convention devient un facteur d'explication très pertinent du comportement individuel et des régularités de comportement que l'on peut qualifier de “règle” ou “d'institution”. L'un des intérêts de cette étude est de parvenir à retrouver une conception conventionnelle de l'ordre collectif, à la manière de ce que cherchait à faire Schotter, sans pour autant mobiliser un appareil théorique hobbien. Choi distingue d'ailleurs très clairement cette conception conventionnaliste de l'analyse hobbienne : *“ Les solutions conventionnelles aux situations de type “dilemme du prisonnier” nous évitent de sombrer dans l'état de nature hobbien, une guerre civile, ou une solution “rationnelle” à ce dilemme. La convention est ce qui rend le dilemme non paradoxal dans un monde où il se répète”* (Choi, 1994, p.81). Il n'y a précisément pas d'état de nature, ni de piège de la rationalité, dans une société fondée sur la recherche de l'agrément des autres.

Conformément à l'esprit de Hume, Choi restaure donc l'idée de convention contre le rationalisme et contre l'idée d'un état ordonné surgissant *ex nihilo*. L'état de nature n'existe pas, précisément parce que les individus respectent les conventions faute de pouvoir gérer seuls leur incertitude.

L'intérêt de l'ouvrage de Choi est de mettre en lumière le lien intime qui unit, dans l'idée de convention, l'attitude cognitive de l'individu et son attitude sociale. La disjonction traditionnelle que la science économique opère entre la réflexion rationnelle des individus et leur ancrage social représente un obstacle à l'intégration de la notion de convention. C'est parce qu'il se situe d'emblée dans une société, que l'individu "humien" parvient à déjouer les pièges des interactions avec les autres individus. Une attitude rationnelle mais "asociale" le condamnerait aux pièges que repère la théorie des jeux⁵⁶. Choi retrouve donc, dans sa théorie, un premier fondement efficace de la notion de convention.

Pendant cette théorie souffre encore de deux imprécisions. La relecture de Smith proposée par Choi est un peu trop rapide pour être convaincante, et la conception "conformiste" de l'individu représente sans doute une conception sociologiquement naïve du comportement des acteurs. De ce fait, la théorie de Choi semble présupposer ce qu'elle cherche à démontrer : il est en effet tout-à-fait clair que, si l'individu cherche avant tout à se conformer aux attentes des autres, l'idée même de conflit disparaît⁵⁷.

Cette impression est due à l'absence, dans l'analyse de Choi, de l'autre assise essentielle de la convention : l'Histoire. Ce n'est en effet qu'en restituant l'idée selon laquelle le comportement individuel se situe dans un temps historique qu'il est possible de faire apparaître l'idée d'une conformité sociale qui ne se confond pas avec un conformisme au sens strict. Dans le temps long des processus historiques, les individus peuvent tenter d'influer sur les conventions, ou de faire évoluer le regard des autres, sans pour autant entrer directement en conflit. La perspective historique permet en quelque sorte de réconcilier la volonté rationnelle de faire triompher ses propres intérêts, et la nécessité de conserver un ancrage social caractérisé par une attitude conventionnelle. Comprendre les mécanismes par lesquels une convention apparaît, se diffuse, évolue ou disparaît implique donc de quitter l'étude de la *structure* d'interaction, ce que permet l'étude des jeux statiques, pour analyser sa *dynamique*. C'est précisément ce que cherche

⁵⁶ Choi tire en cela la leçon de l'analyse de Field [1984]. Cet auteur critique en effet l'approche des normes que propose la théorie des jeux en soulignant que la règle dénoue les pièges de la rationalité, et ne peut donc être le produit de cette même rationalité.

⁵⁷ Comme le souligne Ullmann-Margalit [1978, p.95-96], l'attitude conformiste, lorsqu'elle est généralisée (et que chacun sait que c'est le cas), conduit à considérer tout problème de décision comme un problème de pure coordination.

à faire la théorie des jeux évolutionniste.

33- *L'analyse d'Axelrod.*

“Dans quelles conditions la coopération peut-elle émerger dans un monde d’individus égoïstes ne comportant pas d’autorité centrale ? (...) La réponse la plus fameuse est celle que proposa Hobbes il y a plus de trois cent ans. Il était pessimiste. Selon lui, la coopération ne pouvait pas se développer sans autorité centrale. (...) La théorie de la coopération que propose cet ouvrage est fondée sur l’étude d’individus ne poursuivant que leur propres intérêts sans l’aide d’une autorité centrale.” (Axelrod, 1984, p.3-6).

C’est l’étude de Robert Axelrod qui a popularisé en science sociale l’étude des jeux évolutionnistes introduits en biologie par Maynard-Smith. Axelrod s’est appuyé sur ce type de jeux pour chercher à comprendre comment une “norme” de coopération pouvait émerger sans contrainte entre des individus égoïstes : il relève ainsi le défi de Hobbes. Axelrod cherche à montrer, en utilisant la théorie des jeux évolutionnistes, que les stratégies coopératives devraient être choisies par des individus égoïstes.

331- La théorie des jeux évolutionnistes

La théorie d’Axelrod étudie le dilemme du prisonnier dans le cadre de la théorie des jeux évolutionnistes. L’étude des jeux répétés a déjà permis d’observer que la coopération devenait envisageable, dans le cadre du dilemme du prisonnier, lorsque l’horizon des joueurs était suffisamment lointain et que le terme du jeu n’était pas connu avec certitude. Elle n’a cependant pas permis d’établir la supériorité de la coopération, puisque la défection systématique est également un équilibre de Nash dans le cadre de jeux répétés. La théorie des jeux évolutionnistes permet d’aller plus loin en changeant de méthode d’analyse et en faisant apparaître que certains types de stratégie coopérative, parce qu’ils sont les plus efficaces à long terme, devraient progressivement s’imposer.

La caractéristique de la théorie des jeux évolutionnistes est de remplacer le raisonnement intellectuel par lequel les individus prévoyaient, selon un processus *cognitif* introspectif, l’ensemble des coups menant à l’équilibre par un processus d’*actions effectives*. Les joueurs tâtonnent, selon un processus d’apprentissage de type

“essai/erreur”, jusqu’à ce qu’ils découvrent le comportement optimal⁵⁸. La technique utilisée pour sélectionner une “solution” du jeu consiste alors à considérer un ensemble de stratégies possibles parmi une population d’individus confrontés à un jeu de manière récurrente, et à observer quelles sont, au bout d’une certaine période de temps, les stratégies gagnantes. Ces jeux étaient initialement utilisés en biologie. Selon le principe de l’évolutionnisme darwinien les individus appliquant une stratégie gagnante se multiplient, tandis que les autres meurent et disparaissent. Ce principe est réinterprété en économie selon le principe du mimétisme : les individus utilisent les stratégies dont ils constatent, quand ils y sont confrontés, qu’elles sont plus efficaces que la leur. Les stratégies perdantes disparaissent donc progressivement, tandis que les stratégies gagnantes, assurant un paiement maximal en terme d’utilité, se multiplient. On atteint une forme d’équilibre lorsqu’une des stratégies possibles envahit la population et constitue dès lors une façon “stable” de jouer.

Le concept central de cette théorie est celui de stratégie évolutionnairement stable (ESS) proposé initialement par Maynard-Smith et Price (1973). Une stratégie p est évolutionnairement stable si, quelle que soit q (une autre stratégie) :

$$-1- pAp \geq qAp$$

$$-2- \text{Si } pAp = qAp, \text{ alors } pAq > qAp$$

Ainsi une stratégie “ p ” est ESS si aucune autre stratégie ne peut faire mieux contre “ p ” que “ p ” elle-même ; et que, lorsqu’une autre stratégie “ q ” fait aussi bien que “ p ” contre “ p ”, alors “ p ” fait mieux que “ q ” contre “ q ”. Ce concept correspond assez bien à celui de l’équilibre de Nash : face à p , le mieux à faire est de jouer p : p est donc stable.

Axelrod s’appuie sur ce concept pour définir le concept un peu moins exigeant de “stratégie collectivement stable”. Cette seconde définition ne nécessite que la première des deux conditions. Ainsi, une stratégie qui fait aussi bien que les autres contre elle-même est “collectivement stable”, mais pas forcément “ESS”.

332- L'analyse d'Axelrod

L’analyse d’Axelrod s’appuie sur une formalisation abstraite du dilemme du prisonnier. Coopérer rapporte R (Récompense) lorsque le partenaire coopère et S (le Salaire de dupe). Faire défection procure un gain de T (Tentation de l’égoïste) lorsque le partenaire coopère et P (Punition) lorsque celui-ci choisit également la défection. En forme stratégique, le dilemme est donc représenté sous la forme suivante :

⁵⁸ Il y a une “substitution progressive entre une spéculativité cognitive et une temporalité effective” (Walliser et Rullière, 1995, p.603).

| | | |
|-------------|-------------|-----------|
| | Coopération | Défection |
| Coopération | R R | S T |
| Défection | T S | P P |

Pour qu'il y ait "dilemme", il faut, selon Axelrod⁵⁹, que les deux conditions suivantes soit respectées :

$$(1) T > R > P > S$$

et

$$(2) R > (T+S) / 2$$

Ces deux conditions prises ensemble assurent que la seule stratégie rationnelle (en un coup) soit la défection systématique, et que cette stratégie donne lieu à l'issue (P,P) qui est pareto-dominée par l'autre issue conjointe (R,R). La rationalité individuelle s'oppose donc à la rationalité collective, et au final à elle-même.

Axelrod prend par ailleurs soin de rejeter toute simplification du dilemme, comme par exemple la possibilité de s'engager (il se situe donc dans un cadre non-coopératif), la possibilité de communiquer, le rôle des réputations (et de toute information spécifique sur les joueurs précédant l'interaction directe), la modification des paiements en fonction de l'altruisme des joueurs, ou le "goût" pour la coopération. Axelrod fait l'hypothèse que les joueurs adoptent d'une certaine manière l'attitude de saint Thomas : ils ne jugent de la fiabilité des autres joueurs qu'à travers leurs interaction directes.

L'unique hypothèse que requiert Axelrod est l'existence d'une "mémoire" des interactions passées. Progressivement, lorsque le jeu se répète, il se forme donc une histoire du jeu. Les individus se souviennent de l'attitude qu'ont adoptée chacun de leurs partenaires lors de leurs interactions précédentes. Cette "mémoire" introduit dans le jeu une réelle dimension temporelle : au moment de jouer, chacun sait que son attitude influencera l'autre joueur lors de leur future interaction. On ne joue donc pas une fois

⁵⁹ En réalité la première condition est suffisante pour définir le dilemme. La seconde condition appartient en propre à Axelrod et ne constitue pas une condition classique du dilemme du prisonnier (par exemple, Fudenberg et Tirole [1991] proposent une formalisation du dilemme qui ne respecterait pas cette seconde condition). Cette condition permet à Axelrod de faire en sorte qu'une "exploitation mutuelle" visant à alterner de façon concertée la défection et la coopération ne donne pas, dans le long terme, un paiement aussi satisfaisant que l'issue coopérative. C'est donc une forme de condition supplémentaire qui lui est utile pour défendre sa thèse, mais qui est de nature *ad hoc*.

pour toutes. L'interaction s'insère dans un horizon temporel infini⁶⁰.

Dans ces conditions, le dilemme ne peut connaître de solution satisfaisante, ou pareto-optimale, que dans le cadre d'un jeu répété⁶¹. La répétition multiplie les équilibres possibles, puisque la perspective temporelle des joueurs s'élargit et permet d'envisager les bénéfices de la coopération. Par exemple, la défection systématique est toujours un équilibre de Nash possible, mais la coopération conditionnelle en est un autre. Il peut donc être rationnel de tenter de coopérer, à condition que l'autre coopère lui aussi, et qu'il coopère de manière également conditionnelle. En effet, si notre partenaire fait défection, la coopération n'est plus rentable, on doit alors choisir de le "punir" en faisant nous aussi défection tant qu'il ne change pas d'attitude. La perte que notre partenaire subira consécutivement à cette "punition" l'incitera peut-être à coopérer de nouveau. Si, au contraire, notre partenaire coopère quoique nous fassions, alors il est parfaitement rationnel de l'exploiter et de faire systématiquement défection. Axelrod tente d'établir que la stratégie la plus "rudimentaire" qui consiste à ne coopérer que si l'autre coopère et à le punir *une fois* à chaque fois qu'il fait défection, est aussi la plus efficace. Il qualifie cette stratégie de "Donnant-Donnant" (désormais "DD")⁶². C'est la stratégie la plus efficace puisqu'elle permet d'engranger des récompenses de la coopération (R), tout en évitant le "salaire de dupe" (S).

La démonstration comprend deux temps. Axelrod montre tout d'abord que "DD" est la meilleure stratégie dans certaines "expériences de laboratoire" qui prennent la forme de tournois informatiques. Il décrit ensuite les raisons de ce succès et souligne que DD est une "ESS".

L'expérience menée par Axelrod consiste à faire s'affronter par le biais d'un tournoi informatique des joueurs ayant chacun des stratégies distinctes. Axelrod a ainsi organisé par deux fois un tournoi invitant des spécialistes et des néophytes à proposer une stratégie pour jouer à un dilemme du prisonnier destiné à être répété 200 fois. Un des participants

⁶⁰ Cependant, le présent est plus important que le futur, et Axelrod indique donc que les perspectives de gains futurs doivent être ramenées à leur valeur présente en étant multipliées par un "paramètre de réduction", selon l'expression d'Axelrod, ou, plus simplement : un taux d'actualisation.

⁶¹ Si l'on se situe dans le cadre analytique "classique" il faudrait préciser que le terme du jeu ne doit pas être connu par les joueurs (en raison du problème de l'induction rétrograde). Mais, Axelrod n'étudie précisément pas des joueurs rationnels au sens standard, qui pratiqueraient un raisonnement introspectif avant d'agir. Ici, les individus agissent puis adaptent leur stratégie en fonction des résultats qu'ils obtiennent. Ils ont une attitude beaucoup plus rudimentaire, qui leur permet de coopérer même en horizon fini.

⁶² "Donnant-Donnant" est la traduction habituelle de "tit for tat". On peut cependant noter que cette traduction suggère que l'on attend que l'autre face le premier pas ce qui, précisément, n'est pas le cas. "Oeil pour oeil" traduit peut-être mieux qu'on n'agresse jamais l'adversaire le premier, mais qu'on réplique par l'agression à l'agression. C'est d'ailleurs la définition de la coopération en théorie des jeux non-coopératifs. Nous nous en tiendrons cependant à cette traduction usitée.

(Anatol Rapoport) a proposé la stratégie "DD" : "commencer par coopérer puis faire ce que l'autre fait". La première étape comportait quinze concurrents, "DD" obtint les meilleurs résultats.

La seconde étape se déroula entre soixante concurrents, ayant tous pris connaissance des données de la première étape. Ces données indiquaient notamment qu'une stratégie moins susceptible que "DD" l'aurait emporté. Cependant, malgré ces indications, "DD" gagna à nouveau. Les stratégies nouvelles, fondées sur l'étude du tournoi précédent, n'obtinrent pas de bons résultats car elles se "contredisaient" mutuellement : les stratégies plus agressives que "DD" ont ainsi neutralisé les stratégies coopératives moins réactives que "DD"⁶³.

Axelrod ajoute à cette expérience empirique une simulation virtuelle permettant d'envisager comment la population des stratégies en présence lors du second tournoi pourrait évoluer à terme si le jeu se prolongeait en une succession de tournois. Cette simulation nous fait véritablement entrer dans la méthodologie évolutionniste : à chaque nouveau tour, les stratégies gagnantes sont multipliées, tandis que l'on diminue le nombre des stratégies perdantes. Sur cette base, on observe l'évolution de la population au fil des "générations" successives. Une fois encore, "DD" envahit la population au fil des générations successives. Ainsi, non seulement la coopération peut émerger, mais elle paraît être stable. Elle devient à terme une norme de comportement : la coopération émerge donc spontanément. Il ne s'agit cependant pas là d'une "démonstration" rigoureuse mais d'une observation empirique. Ce résultat n'est donc pas généralisable, il dépend étroitement des conditions particulières dans lesquelles est menée l'expérience. Or, le contexte étudié par Axelrod a une particularité importante : la défection systématique ne faisait pas partie des stratégies présentes au départ.

Axelrod commente cette expérience et donne quelques raisons du succès inattendu de cette stratégie coopérative dans le dilemme du prisonnier. Il souligne les caractéristiques de "DD". Elle est bienveillante, ce qui lui permet d'engranger les bénéfices de la coopération. Elle est indulgente, ce qui lui évite de rester bloquée dans une attitude inutilement répressive à la moindre erreur du partenaire. Mais elle est néanmoins susceptible, et donc elle ne peut être durablement exploitée. Enfin, elle est transparente : les adversaires la reconnaissent facilement et comprennent qu'ils n'ont contre elle rien de mieux à faire que de coopérer. Se dessine ainsi un aspect important de son étude : la coopération peut se développer parce que "DD" n'est pas exploitable et que cela se comprend vite lorsqu'on rencontre cette stratégie pourtant bienveillante. Sa transparence

⁶³ L'application des stratégies qui auraient bien fonctionné dans le contexte précédent modifie précisément le contexte de jeu et rend ces stratégies inopérantes. Il y a là un paradoxe de l'attitude "inductive" dans un environnement interactif sur lequel nous reviendrons.

pallie l'absence totale de communication (comme Schelling le soulignait à propos des coups stratégiques). Elle constitue donc un point de repère simple.

Dans la seconde phase de son analyse, Axelrod s'extirpe de ces résultats particuliers et tente de formuler quelques remarques générales. Après avoir montré, expérimentalement, que la coopération pouvait parfois *exister*, il cherche à déterminer les conditions théoriques de sa *stabilité*. Là encore, Axelrod procède en deux étapes. Il montre tout d'abord que DD est "collectivement stable" dès lors que la préférence pour le présent, mesurée par un paramètre "p" d'actualisation, n'est pas trop forte (*i.e.* : que "p" n'est pas trop petit). C'est cependant aussi le cas non seulement de toute stratégie bienveillante, mais encore de la défection systématique⁶⁴. Néanmoins, Axelrod souligne que la défection systématique, si elle résiste à un individu seul, peut être envahie par des groupes d'individus pratiquant la réciprocité : ils tirent entre eux les bénéfices de la coopération, sans pour autant être exploitables par des "défectionnistes". "DD" est donc un candidat sérieux à l'envahissement d'une population pratiquant la défection systématique, à condition que le groupe d'envahisseurs soit assez grand⁶⁵. A l'inverse, "DD" ne peut pas être envahi par des "défectionnistes" quelque soit la taille de leur groupe envahisseur. Ces résultats sont d'ailleurs intuitivement évidents. Entrevoir les bénéfices de la coopération suppose d'une part que l'on dispose d'un horizon temporel suffisamment long et d'autre part que l'on soit suffisamment nombreux. A l'inverse, la défection systématique est rationnelle à court terme, et ne tire aucun bénéfice de la rencontre d'un autre "free rider". La conclusion d'Axelrod est donc que la coopération, pour peu qu'elle repose sur une attitude vigilante de réciprocité, et qu'elle s'appuie au départ sur un groupe, même de taille restreinte par rapport à la population, peut se répandre dans une population égoïste, sans qu'il existe d'autorité centrale.

333- Critique du "modèle" d'Axelrod.

Le but du modèle étudié par Axelrod est de mettre en évidence *les conditions* dans lesquelles la coopération peut émerger d'elle-même dans une population égoïste sans que ce soit par la contrainte d'une autorité centrale. C'est une théorie de l'émergence d'un ordre collectif fondé sur l'application d'une règle collective de coopération émergeant spontanément.

Cependant, cette théorie souffre d'une certaine imprécision. La théorie de l'émergence ne peut se fonder sur des expériences aussi restreintes. Des tests menés de façon

⁶⁴ Tout équilibre de Nash symétrique est évolutionnairement stable.

⁶⁵ Les tests empiriques donnent une taille critique de 5% de la population.

identique selon le mécanisme de la “réplication dynamique” ont montré en effet que la coopération ne pouvait pas se maintenir à très long terme dans un environnement comportant la défection systématique⁶⁶. On sait de plus que “DD” ne parviendrait pas à s'imposer dans un environnement qui serait au départ massivement défectionniste. Pour pallier ce problème, Axelrod s'appuie sur un exemple empirique *réel* censé montrer l'efficacité de la coopération dans la pire des situations : la guerre. Nous chercherons à montrer que cette démonstration empirique n'est pas satisfaisante. Par ailleurs, il semble que la démonstration théorique de la “stabilité” de “DD” comporte aussi d'importantes failles. De ce fait, Axelrod est amené à faire jouer à l'Etat un rôle essentiel compensant les faiblesses de “DD”.

Le seul exemple concret convoqué par Axelrod à l'appui de sa thèse est celui du principe “vivre et laisser vivre” étudié par un sociologue américain (Ashworth, 1980). Durant la première guerre mondiale, les soldats français et allemands, sans contact entre eux, choisirent en effet durant un certain temps de ne plus se tirer directement dessus de tranchée à tranchée. Cette entente tacite a émergé, selon Axelrod, dans une situation typique du dilemme : les soldats ont “collectivement” intérêt à la paix, mais celle-ci n'est pas individuellement praticable. Un soldat choisissant d'être pacifiste en temps de paix risque de se faire tuer rapidement, il lui faut donc “tuer plutôt que d'être tué”. Tous font ce raisonnement, et cela les mène nécessairement à un équilibre de défection mutuelle sous-optimal dans lequel chacun se tire dessus. Pourtant, à bien y regarder, cette interprétation “stratégique” de la situation de guerre est critiquable. Elle fait en effet abstraction du fait que l'antagonisme entre soldats est uniquement le fait d'une hiérarchie militaire qui, précisément, est en crise à cette époque de la guerre. Dans ces conditions, il est tout-à-fait indifférent aux soldats d'un camp donné que les soldats du camp ennemi vivent ou meurent, dès lors qu'eux-mêmes ont la vie sauve. Ainsi, l'issue coopérative garantissant “R” n'est pas plus intéressante que la défection unilatérale garantissant “T”. De la même manière, la “punition” qui consiste à se tirer dessus mutuellement n'est pas pire que le “salaire de dupe” qui consiste ici à se faire tirer dessus sans réagir : dans les deux cas, c'est la mort. Dans les termes d'Axelrod, l'inégalité $T > R$ n'est donc pas vérifiée, et le salaire de dupe (S) n'est pas pire que la punition mutuelle (P). Contrairement à la définition générale du dilemme : $T = R$, $R > P$, et $S = P$. Dans une telle configuration, la défection n'est plus une stratégie dominante : elle rapporte, selon les cas,

⁶⁶ L'expérimentation menée par Nachbar [1992] a montré que les stratégies bienveillantes s'imposaient dans l'environnement durant les premières générations avant d'être rattrapées par la défection systématique. Ainsi, la défection systématique est fatalement l'issue du processus évolutif (ce qui est cohérent avec le principe selon lequel le D.P. répété un nombre fini de fois n'a qu'un unique équilibre : la défection systématique). Mais, l'enseignement important est que durant le processus la coopération émerge pendant un long moment (durant les 300 premières générations dans l'expérience de Nachbar [1992]).

T ou P et équivaut donc à la coopération qui rapporte R ou S. La coopération est donc plus probable, parce que le dilemme est beaucoup moins marqué⁶⁷. L'interprétation d'Axelrod passe en fait sous silence le trait principal de cette situation "absurde" de guerre : les acteurs refusent de jouer parce qu'ils ont la capacité d'influer sur "les règles" du jeu. Leur préférence pour la coopération est plus forte que dans le cadre d'un dilemme entre individus égoïstes et ils ont une possibilité d'action, en dehors du jeu, qui n'est pas envisageable dans le cadre théorique d'Axelrod. Cette "démonstration empirique" démontre donc bien peu de choses : il faut se satisfaire de la "démonstration théorique" de la stabilité de la règle de réciprocité.

Or, précisément, cette démonstration contient elle-même une faille importante. Lors de la présentation de son modèle, Axelrod souligne : "*il n'est pas nécessaire de supposer que les agents sont rationnels. (...) Les actions des joueurs ne sont pas nécessairement des choix conscients.*" (Axelrod, 1984, p.18). En cela, Axelrod est conforme à l'esprit de la théorie évolutionniste : la succession des actes remplace la réflexion cognitive et épistémique de la théorie traditionnelle (Walliser [1996]). Les joueurs agissent "à l'aveugle" et réutilisent les stratégies dont ils ont pu constater l'efficacité : "*ce qui marche bien a plus de chance d'être réutilisé alors que ce qui ne fonctionne pas risque fort d'être abandonné*" (Axelrod, 1986, p.1097). Les joueurs adoptent donc une attitude inductive, fondée sur l'expérimentation, et cette attitude simple ne nécessite pas la mise en place de procédés cognitifs complexes. Le modèle doit pouvoir fonctionner aussi bien pour des "bactéries" que pour des "humains".

Cependant, contrairement à ce que prétend Axelrod, cette hypothèse n'augmente pas la généralité de sa thèse. Le fait que les agents soient rationnels constitue en effet le cœur du dilemme⁶⁸. Ce qui fonctionne en supposant qu'ils ne sont pas rationnels ne fonctionne plus lorsqu'ils le sont ! Il est donc nécessaire de spécifier quel type de rationalité les joueurs mettent en œuvre pour donner un sens à la théorie. Or, Axelrod est extrêmement imprécis sur cette question. La seule certitude que l'on puisse avoir tient au caractère "paramétrique" de la rationalité des joueurs : ceux-ci réagissent aux choix des autres comme à une donnée extérieure à leur propre choix. Il se trouve pourtant que, dans un environnement interactif, cette attitude "paramétrique" entre en contradiction avec la

⁶⁷De plus Axelrod note que l'émergence de la coopération est ici liée à la capacité des divisions entre elles à se communiquer le mot de passe de la "non agression". Cette communication entre joueurs ne respecte pas la condition de "non connaissance initiale du contexte de jeu" qu'Axelrod énonce au début de son ouvrage.

⁶⁸C'est ce que souligne Aumann [1994, p.58] : "*la rationalité ne peut se suffire à elle-même : elle ne peut prendre sens que dans un domaine plus vaste comprenant de l'irrationalité*".

méthode "inductive" que suivent les joueurs, comme le montre Cordonnier [1993]⁶⁹. Les joueurs fondent en effet leur choix sur une connaissance qui n'est vraie que dans un environnement donné, en l'occurrence l'environnement du coup précédent. Or, le fait que tous les joueurs aient une attitude inductive, qui les incite à réviser leur stratégie à chaque tour, rend cette connaissance inutile. Ce qui était vrai au tour précédent ne le sera pas forcément au tour suivant, puisque tous les joueurs auront modifié leurs stratégies. L'attitude paramétrique des joueurs, qui consiste à supposer que les stratégies des autres sont données, tourne donc ici au "fiasco" : dans le pire des cas, chacun change de stratégie en même temps et les stratégies choisies ne peuvent donc jamais s'accorder. Les joueurs ont "toujours tort", parce qu'ils ne perçoivent pas les conditions d'efficacité de leur choix (sauf à l'équilibre stationnaire).

Ce problème trouve une application très précise dans le modèle d'Axelrod. Axelrod prouve que "DD" est collectivement stable, mais il ne prouve pas qu'elle est "évolutionnairement stable". On peut aller plus loin en montrant que "DD" *n'est pas* "évolutionnairement stable". Il existe en effet une stratégie aussi efficace que "DD" contre "DD", et aussi efficace que "DD" contre elle-même : c'est même le cas de toutes les stratégies coopératives. Pour cette raison DD peut être "envahie" par une stratégie moins réactive qu'elle-même et en particulier par une attitude systématiquement coopérative du type : "ne fait jamais défection". Lorsque le jeu évolue, la population peut donc devenir progressivement de moins en moins réactive. C'est même une évolution prévisible, dans un environnement massivement coopératif, l'attitude suspicieuse que représente la "loi du talion" ("DD") apparaît déplacée. La vigilance s'érousse. Il existera donc probablement un certain stade à partir duquel "ne fait jamais défection" aura envahi la population et pris la place de "DD" comme norme de comportement⁷¹. Or, "ne fait jamais défection" n'est ni "évolutionnairement stable" ni même "collectivement stable". Lorsque tout le monde coopère systématiquement, il devient possible et rentable de faire défection. Une stratégie purement coopérative peut être envahie à partir d'un individu faisant toujours défection. La coopération n'est donc pas nécessairement stable dans le long terme. La réciprocité peut dégénérer en gentillesse naïve et s'exposer à un "contre" de la défection ! Ainsi,

⁶⁹"Pour résumer, la conjonction d'un raisonnement inductif et d'une rationalité paramétrique plongés dans un cadre interactif produit de l'irrationalité ; les agents reproduisent toujours les mêmes erreurs" (Cordonnier, 1994, p.115). L'auteur prend l'exemple de l'innovation technologique : elle n'est efficace que tant qu'elle n'est pas imitée par tous. La duplication générale d'une innovation brise les conditions d'efficacité de cette innovation (c'est d'ailleurs son principe).

⁷⁰"Pour résumer, la conjonction d'un raisonnement inductif et d'une rationalité paramétrique plongés dans un cadre interactif produit de l'irrationalité ; les agents reproduisent toujours les mêmes erreurs" (Cordonnier, 1994, p.115). L'auteur prend l'exemple de l'innovation technologique : elle n'est efficace que tant qu'elle n'est pas imitée par tous. La duplication générale d'une innovation brise les conditions d'efficacité de cette innovation (c'est d'ailleurs son principe).

⁷¹ C'est même l'attitude prônée par la religion chrétienne. Le principe du pardon, ou l'enseignement du Christ qui "tend l'autre joue", représente typiquement le passage de "DD" à "Coopère toujours".

l'évolution probable de l'attitude des agents sera la suivante :

- 1- "DD" devient la norme dominante.
- 2- "Ne fait jamais défection" remplace "DD".
- 3- "Fait toujours défection" envahit "ne fait jamais défection".

Le concept d'équilibre "collectivement stable" est donc insuffisant. En réalité, sur le plan purement abstrait, "DD" n'est pas stable. La défection systématique a toutes les chances de s'imposer au final. De plus, "ne fait jamais défection" est, quant à elle, "évolutionnairement stable". Axelrod ne nous donne donc pas la "fin de l'histoire". On peut ajouter un paragraphe "dramatique" :

- 4- "Fait toujours défection" s'impose et est stable dans le long terme.

Mais Axelrod, bien qu'il n'envisage pas ce "problème", suggère tout de même une "solution" : la "société" doit mettre en place un dispositif permettant de maintenir la réciprocité en condamnant les individus qui ne punissent pas la déviance. Cette idée est déclinée par Axelrod sur deux modes différents : celui de "conseils aux réformateurs", qu'il développe dans la quatrième partie de son ouvrage de 1984, et celui d'une "méta-norme de répression" que l'on trouve dans un article postérieur (Axelrod, 1986). Dans les deux cas, l'objectif poursuivi par Axelrod est le même : pallier les défauts d'individus mimétiques qui, ne connaissant pas les raisons profondes du succès des stratégies "DD", oublient la nécessité de la réciprocité. Dans le premier cas le réformateur est donc invité à inculquer à ses "citoyens" le principe de la réciprocité à l'école, dans la vie courante, ou par le biais de la morale. La seconde solution met en évidence l'efficacité d'une "méta-norme" qui provoque l'obligation déontique de "*punir ceux qui ne punissent pas la défection*" (Axelrod, 1986, p.1100). Axelrod énumère divers modes d'application de cette norme, à travers le droit et le mécanisme de la réputation.

On peut donc, en définitive, tirer deux conclusions de l'approche évolutionniste d'Axelrod :

- 1- La coopération ne peut émerger que s'il existe un groupe minimal d'individus la pratiquant déjà entre eux. Il faut donc qu'elle soit toujours déjà là, et ce que l'on peut étudier est sa diffusion plutôt que son émergence.⁷²
- 2- Pour être stable, la norme de coopération doit être soutenue par un dispositif

⁷² Rappelons que la coopération, même du type "DD", ne peut pas apparaître dans un environnement hostile où tout le monde pratique la défection (ce qui n'est pas le cas des deux expériences d'Axelrod).

transcendant : une organisation sociale incarnée par l'état ou une méta-norme, dont les origines restent à déterminer.

En définitive, Axelrod ne dément donc pas la théorie hobbenne : il la conforte. La coopération ne peut émerger dans un environnement égoïste que si on l'introduit par le biais d'un groupe donnant l'exemple, et si on la maintient artificiellement par le biais de la sanction de la défection⁷³. Que la sanction soit le fait de l'ensemble des individus formés par l'Etat ou guidés par une méta-norme, et non pas d'un seul d'entre eux, élu pour la prendre en charge (le Léviathan), ne change pas grand chose à l'affaire. La coopération n'émerge pas spontanément, elle est construite et soutenue par une autorité transcendant les interactions entre agents. Comme il ne spécifie pas le mode de rationalité des agents, Axelrod est contraint de supposer qu'ils sont en quelque sorte "programmés" par l'Etat.

Axelrod ne règle donc pas le dilemme proposé par Hobbes. Il propose cependant une véritable innovation en décrivant de manière convaincante un processus de *diffusion* de la norme s'appuyant sur un comportement simple et réaliste des individus. Il ouvre ainsi une perspective évolutionniste radicalement différente de celle de Schotter dont l'utilisation des "raffinements bayésiens" supposait que les agents procèdent à des calculs extrêmement complexes. La théorie gagne donc en réalisme pratique. Toutefois, il manque à cette analyse une explicitation des raisons d'agir de l'individu, ce qui la conduit à faire dépendre la réussite de la coordination d'un mécanisme transcendant l'individu. Or, l'intérêt du concept de convention est précisément de rendre ce mécanisme inutile.

34- L'évolutionnisme conventionnaliste.

A la suite des travaux d'Axelrod, la théorie des jeux évolutionnistes (désormais TJE) a été utilisée pour donner un éclairage sur la manière dont des conventions, du type de celle de Lewis, pouvaient émerger et se stabiliser : "*Les conventions ne peuvent pas se comprendre si l'on utilise comme point de départ la théorie des jeux "classique" – des joueurs parfaitement rationnels dans des interaction uniques. Il peut être plus utile de mettre un peu moins l'accent sur la rationalité et d'analyser les conventions comme étant le produit d'un processus évolutionniste*" (Sugden, 1989, p.91).

⁷³ Sur le plan empirique, il apparaît que l'émergence des stratégies coopératives est effectivement très dépendante du contexte institutionnel : Güth, Marchand, Rullière [1998].

341- L'introduction d'un temps historique

Cette remarque de Sugden met très clairement en balance la rationalité des joueurs et le poids de l'Histoire que cherche à "formaliser" la théorie des jeux évolutionnistes. Cet objectif s'inscrit bien dans l'optique de Hume, dans la mesure où il permet de retrouver une assise non exclusivement rationaliste de la convention, ce qui évite de mobiliser une thématique hobbenienne d'état de nature et d'état central coercitif. Le projet de cette "théorie des jeux évolutionnistes" est d'éclairer, plus que ne le faisait Axelrod, les processus cognitifs qui, historiquement, donnent naissance à des conventions. Ce faisant, cette théorie se présente comme la première étape d'une théorie "générale" des institutions et normes économiques, dans la mesure où elle parvient à fonder toute règle collective sur un processus de type conventionnel.

Une convention est en effet définie comme : "*toute ESS dans un jeu qui a deux ou plusieurs ESS*" (Sugden, 1989, p.91). Cette attitude théorique s'applique donc aux trois cas paradigmatiques que nous avons présenté ("jeu de coordination", "dilemme du prisonnier", "jeu du marchandage"). C'est la raison pour laquelle la TJE a vocation à rendre compte de toutes les normes et institutions, et non seulement de celles qui règlent des problèmes de stricte coordination. La répétition des interactions permet, en quelque sorte, de gommer les différences de structures mises en évidence par Ullmann-Margalit.

Cette idée était déjà au cœur de l'analyse d'Axelrod. Toutefois celui-ci ne disposait pas du concept de convention et faisait donc jouer un rôle essentiel à des mécanismes transcendant les individus. L'idée de convention permet au contraire de décrire la connexion qui existe entre l'attitude cognitive des individus et leur "encastrement" social. Elle est à l'origine d'une capacité de la société à générer entre ses membres des attentes réciproques qui permettent de réussir la coordination, sans qu'une puissance transcendante n'entre en jeu. Ce concept dénoue donc, en partie, les contradictions repérées chez Axelrod. L'idée de l'importance de la "*préférence pour la conformité*" mise au point par Lewis et généralisée par Choi est ainsi au cœur de la TJE. La convention émerge et se stabilise en raison de sa capacité à représenter un point d'ancrage "*populaire*", selon l'expression de Sugden [1989, p.94]. Elle se répand et se développe grâce à sa simplicité et au mécanisme de l'*analogie*. Ainsi, pour Sugden, la convention "*premier arrivé, premier servi*" peut s'appliquer à la géopolitique, aux files d'attente, ou encore à la gestion du personnel à travers la règle "*dernier arrivé, premier dehors*" qui prime lors des licenciements. Cette capacité des conventions à fixer l'imaginaire des joueurs lui permet, à terme, de devenir une norme. Lorsque progressivement les joueurs s'attendent à ce que chacun respecte la convention, il devient *nécessaire* de respecter la

convention : “le mécanisme qui transforme les conventions en normes est le désir humain d’être approuvé par les autres” (Sugden, 1989, p.95). Le concept de “pression sociale” sert donc de point de passage entre la simple régularité de comportement et l’obligation déontique qui caractérise la norme. Dans le cadre de la TJE, l’attitude cognitive des individus est d’emblée conçue comme dictée par la nécessité d’une certaine conformité sociale.

La rationalité individuelle n’est pas pour autant absente de ce processus. Mais les individus sont conçus comme cognitivement “limités”, et cette rationalité n’est donc qu’un des critères de choix à côté de l’importance du “précédent” ou de “l’habitude”. La TJE mobilise ainsi une conception du comportement héritée de Schelling [1960] : à côté de la rationalité stratégique purement calculatoire, les individus doivent avoir la faculté de se repérer et de se comprendre en raison de leur expérience commune. Ils sont rationnels, mais ils sont aussi pris dans un temps historique : celui-là même qui rythme l’apparition des conventions. C’est précisément parce qu’elle semble à même de restituer cette temporalité historique qui dénoue la contradiction entre rationalité et convention, que la “TJE” apparaît comme une voie essentielle de compréhension des institutions. Elle éclaire historiquement les processus cognitifs par lesquels les agents en viennent à générer des conventions.

342- Le principe de la stabilité stochastique

Les modèles utilisés décrivent un processus markovien dans lequel les individus utilisent les informations contenues dans les interactions précédentes pour adapter leur réponse. Ils ont une attitude inductive déjà repérée chez Axelrod. Young [1997, p.108] souligne : “trois éléments définissent un jeu évolutionnaire dynamique : une interaction locale entre plusieurs individus, une perception limitée des réponses rationnelles à apporter à l’environnement stratégique, et des perturbations dues au hasard”. L’analyse consiste alors à étudier quelle stratégie, dans une interaction particulière, a des propriétés de stabilité, du type ESS⁷⁴. Il a ainsi été possible de démontrer que certaines issues ont des propriétés de stabilité indépendantes de leurs conditions d’apparition. La règle “moitié, moitié” est par exemple remarquablement stable dans de nombreuses situations, et elle a donc des chances d’émerger à chaque fois que ce partage est possible. Les conventions sont donc décrites comme des stratégies présentant une résistance particulière aux chocs extérieurs. Par “chocs extérieurs”, la TJE entend la volonté d’un groupe d’individus de

⁷⁴ Young [1993,1997] propose un concept plus exigeant de “stratégie stochastiquement stable” c’est-à-dire résistant non pas à un seul choc ponctuel mais à des chocs extérieurs persistants. C’est la résistance à cette agression *persistante* qui distingue le concept de stochastiquement stable de celui d’ESS.

changer de convention, à la manière des mutants envahisseurs qu'étudie Axelrod.

Young prend ainsi l'exemple du fermage. Trois partages sont possibles : $1/4$; $1/2$; ou $3/4$ de la récolte au propriétaire et le reste (respectivement $3/4, 1/2, 1/4$) à l'exploitant. Il se trouve que le passage des deux partages ($1/4$; $3/4$) au partage (50/50) est obtenu plus facilement que le passage de la convention (50/50) à l'un ou l'autre des partages "inégaux". L'auteur montre en effet que, quelque soit l'individu favorisé, le passage de la convention ($1/4$; $3/4$) à la convention (50/50) peut se produire à partir du moment où le "choc extérieur" atteint le seuil critique de " $1/3$ ". C'est-à-dire à partir du moment où, pour une raison ou pour une autre, le *tiers* des individus changent de stratégie et commencent à pratiquer le partage ($1/4, 3/4$). En revanche, le passage inverse de (50/50) à ($1/4, 3/4$) nécessite que la *moitié* de la population change brusquement de convention. De ce fait (50/50) est une convention plus stable et donc plus probable. L'analyse de la TJE repose donc exclusivement sur la notion de "résistance" aux "chocs extérieurs". Ces deux notions mobilisent une conception de l'Histoire qui mérite d'être questionnée.

343- Les hasards font-ils seuls l'Histoire ?

Comme le note André Orléan [1997, p.178] : "*dans la TJE, les jugements 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.*". L'idée qui sous-tend l'explication évolutionniste est en effet que les valeurs morales ne sont que des rationalisations *a posteriori* de régularités de comportement qui sont d'abord choisies parce qu'elles sont "stables". Les individus recherchent un paiement maximum, ce qui nécessite de mobiliser une convention commune pour parvenir à se coordonner dans le long terme. Lorsque cette convention commune est atteinte, ils estiment qu'elle est "juste" : "*quelque soit la norme qui émerge du jeu, elle sera considérée comme juste, après coup, par les participants*" (Binmore et Samuelson, 1994, p.47). Le même argument est utilisé pour réfuter l'idée qu'une convention émerge parce qu'elle est "pareto-efficace" : "*une convention peut être évolutionnairement stable, même lorsqu'elle n'est pas pareto-efficace*" (Sugden, 1989, p.89). Des conventions inefficaces peuvent émerger parce qu'elles ne sont pas collectivement "choisies" mais qu'elles sont le processus d'une polarisation des attentes qui évolue jusqu'à un point stable et arbitraire. Autrement dit, comme le soulignent Sugden [1986, 1989, 1995] ou Young [1993, 1996], l'idée centrale de cette tentative d'explication des normes et institutions est d'abandonner l'idée qu'une norme émerge parce qu'elle est juste, ou qu'une institution émerge parce qu'elle est efficace, mais de défendre l'idée inverse : c'est parce qu'une convention est stable qu'elle est, ensuite

considérée comme juste et efficace.

Cette forme d'explication de la genèse des institutions et des normes à partir d'une conception évolutionniste de la convention a le mérite de mettre en avant la nature "historique" des règles. Elle constitue en cela une forme d'explication de la dynamique des règles collectives conforme à l'esprit humien de la convention. Cette ouverture historique permet d'intégrer dans le modèle des faits empiriques et contextuels. C'est en particulier le cas des analyses qui cherchent à expliquer le changement d'un mode de régulation à un autre, comme le font par exemple Orléan et Boyer [1991, 1994, 1997]. L'importance pratique du contexte est donc mise en avant, contre une idée abstraite et axiomatisée d'optimalité ou de justice, à l'inverse de ce qui est habituellement le cas dans la théorie des jeux. La théorie évolutionniste constitue donc une avancée importante sur le plan de l'explication "pratique" des conventions. En outre, elle le fait sans remettre en cause les fondements du modèle économique orthodoxe, puisqu'elle étudie des interactions libres entre des agents rationnels, même de façon limitée, et égoïstes.

Il n'en reste pas moins que cette forme d'explication en cours d'élaboration pose encore de nombreuses questions. La plus importante nous semble être celle de l'articulation de la rationalité et de l'Histoire. Pour la TJE, bien que les individus agissent dans un temps historique, leur rationalité reste profondément "paramétrique". De ce fait le choix de la convention repose de manière centrale sur le hasard : l'Histoire se fait en effet en dehors des individus qui à chaque moment constatent et scrutent le passé, en le considérant comme un élément tout-à-fait extérieur à leur propre comportement.

L'exemple de Young dans son commentaire du choix de la conduite à droite permet d'illustrer ce problème. Young explique ainsi la genèse de ce choix en France : *"A la suite de la révolution française, il fût décrété que les carrosses devaient se tenir sur la droite. La tradition avait pourtant été de se tenir à gauche, tandis que les piétons circulaient sur la droite, face au trafic. L'abandon de la tradition était symbolique du nouvel ordre : aller à gauche était devenu politiquement incorrect parce que cela était le fait des classes privilégiées, circuler à droite était dans les habitudes de l'homme ordinaire et donc plus "démocratique"*" (Young, 1997, p.106). La conduite à droite s'est ensuite répandue en Europe notamment parce que Napoléon l'imposa à ses armées dans les pays qu'il conquiert. Cet exemple est doublement intéressant en ce qu'il montre à la fois l'intérêt et les limites de l'approche évolutionnaire. Son intérêt est de permettre une ouverture de l'économie sur le monde qui l'entoure. L'évolutionnisme libère l'analyse économique des règles du primat de l'efficacité. Mais (et c'est là sa limite) la TJE ne comble pas ce vide : l'Histoire se résume alors à une longue suite de "hasards". C'est en tout cas ce que conclut Young concernant le choix de la conduite à droite.

L'exemple pris par Young nous montre pourtant que, même dans les cas où le choix de la convention est tout-à-fait arbitraire (*a priori* rien ne permet de défendre une conduite à droite plutôt qu'à gauche), le choix de la convention repose sur des raisons qui le légitiment. Le hasard n'a pas une place primordiale dans le choix des institutions, sauf à décréter la fin de toute philosophie politique et morale. Comme le rappelait Weber [1995, p.64] : *“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”*. En ne faisant dépendre le choix des conventions que des “hasards” extérieurs aux individus, la théorie des jeux évolutionnistes donne donc une interprétation tronquée de la convention, et, par voie de conséquence, de la norme et de l'institution. Elle permet d'attirer l'attention sur l'importance des performances propres des conventions en termes de “stabilité”, mais elle ne peut pour autant prétendre résumer l'histoire des conventions et normes à ce seul concept de stabilité. L'idée de “stabilité” dépend en effet de la définition des “chocs extérieurs”, or on ne peut durablement considérer comme “extérieur” le changement d'attitude des individus. La théorie doit pouvoir rendre compte des mécanismes qui poussent les individus à refuser, ou à respecter, la convention.

La vision de l'Histoire qui sous-tend la TJE est problématique. Elle procède tout d'abord par une négation des réflexions menées en dehors de l'économie sur les mécanismes historiques. Il serait peut-être plus utile de chercher à les intégrer plutôt que de les placer dans la catégorie vide de sens de “choc extérieur”. La seconde est de nature plus interne et tient à la conception de l'individu qui sous-tend le modèle. Il est difficile d'admettre qu'un individu rationnel, même de façon limitée, oublie systématiquement qu'il participe lui-même à travers ses choix à la construction de l'Histoire. L'attitude paramétrique dans un temps historique est tout simplement absurde : qui croit que l'Histoire se fait indépendamment des choix des individus ? Au mieux la TJE s'adresse donc à des réformateurs sociaux soucieux de conserver une certaine harmonie sociale en faisant “oublier” aux individus que les règles collectives auxquelles ils participent n'ont rien de “nécessaires” et peuvent être discutées en vertu de principes “raisonnables”, incluant la rationalité et la justice. Elle éclaire donc de manière pertinente les ressorts de l'inertie sociale, mais c'est un résultat qui laisse dubitatif pour une théorie qui veut expliquer l'émergence des règles...

La TJE constitue une perspective intéressante d'explication de l'émergence des conventions, et des normes et institutions qui en découlent. Elle permet de rendre compte de l'importance de l'Histoire et du contexte dans le choix des conventions, relâchant par là même le poids excessif qui pèse sur l'hypothèse de rationalité dans la perspective

standard. De ce point de vue, la TJE constitue une synthèse, réussie, de l'analyse néoinstitutionnelle menée en théorie des jeux à partir des analyses de Schelling et de Lewis. Elle redonne un fondement et un sens Humien au concept d'ordre social conventionnel. Mais elle est profondément incomplète, puisqu'elle se prive de la force explicative du concept de rationalité et ne lui substitue aucun autre mode d'analyse des processus historiques qui donnent lieu à la convention. Elle ne modifie pas l'hypothèse de rationalité, elle la limite, faisant apparaître un vide dans sa conception du processus historique que la notion de hasard, et de stabilité stochastique, ne comble pas.

Conclusion

Une analyse néoinstitutionnelle de la règle s'est développée en théorie des jeux avec comme objectif de proposer une explication de la *fonction* et de l'*émergence* des règles collectives à partir de la confrontation d'agents rationnels et égoïstes.

Le premier objectif a été atteint. La théorie des jeux permet de concevoir la fonction des règles collectives à partir de la nécessité de sélectionner un des équilibres de Nash possibles lorsqu'il en existe plusieurs, ou d'atteindre une issue pareto-optimale lorsque l'équilibre ne l'est pas. Ainsi la règle apparaît comme un moyen de "dénouer" les pièges tendus par l'interaction sociale à un individu faisant preuve d'une attitude instrumentalement rationnelle. C'est en harmonisant les "croyances" des joueurs, en amont de leur rationalité, que l'institution leur permet de se coordonner et, le cas échéant, de coopérer.

Pour rendre compte de ce rôle des "règles", la théorie des jeux s'est saisie du concept de convention. La convention est un procédé d'harmonisation des anticipations qui ne repose ni sur le contrat ni sur la contrainte. C'est le philosophe David Lewis [1969] qui a redéfini ce concept humien dans les termes modernes de la théorie des jeux. L'idée de convention semble permettre de réconcilier "règle" et "rationalité" dans la mesure où, chez Lewis, la convention paraît émerger spontanément des interactions entre individus rationnels. Ce concept semble donc permettre de boucler le projet néoinstitutionnaliste : les agents génèrent eux-mêmes les institutions dont ils ont besoin. L'idée de "règle collective" n'entre donc plus en contradiction avec le projet "individualiste" de l'économie classique dès lors que les normes, institutions et règles, peuvent toutes être analysées comme étant initialement des conventions.

Pourtant, fonder l'ensemble des règles collectives sur la convention est problématique. L'idée de convention repose en effet sur une temporalité historique étrangère à la temporalité logique sur laquelle se fondent les choix rationnels. Pour que la convention ne se réduise pas à l'idée "creuse" d'un équilibre de Nash de pure coordination, il est nécessaire de lui donner un fondement historique et donc d'associer rationalité et "Histoire". Les premières analyses néoinstitutionnalistes post-lewissiennes d'Ullmann-Margalit et Schotter montrent clairement que la convention, si elle se situe dans un cadre exclusivement rationaliste, évolue nécessairement vers une forme hobbienne qui suppose une transcendance. Pour fonder l'institution sur la convention il paraît donc nécessaire de retrouver le lien qui unit le raisonnement cognitif des acteurs et leur ancrage social. Et cela ne peut se faire que dans une perspective historique.

Cette exigence est liée au second objectif de la théorie des jeux : rendre compte de l'émergence des règles. Pour ce faire, il est nécessaire de quitter le cadre statique traditionnel : dans ce cadre, en effet l'institution est ce qui permet au joueur de sélectionner un équilibre ou une issue pareto-optimale. Elle ne peut donc pas être à la fois le "moyen" et le "résultat" du choix... L'analyse rationaliste s'est donc attachée à étudier des jeux répétés et évolutionnistes. Ce cadre d'analyse a permis d'éclairer le mécanisme par lequel il devient rationnel, dans le long terme, de suivre les conventions, dès lors qu'il est nécessaire d'aligner ses croyances sur celles des autres. Il permet également de donner un sens à l'idée de stabilité de la convention. La convention devient une institution stable par le biais d'un processus stochastique durant lequel les croyances se polarisent. La convention qui émerge finalement dépend donc pour une part importante des caractéristiques de la situation de départ et du déroulement historique du processus.

Toutefois la théorie des jeux ne parvient pas à mobiliser une conception satisfaisante de l'Histoire, car elle ne peut pas analyser l'origine du changement autrement que par le concept, très pauvre, de "hasard". Une théorie de l'institution qui ne fait dépendre l'émergence et le changement que du hasard est partielle : il faudrait expliquer en quoi les individus, rationnels, contribuent à ce "hasard".

Pour cela, il serait nécessaire de mobiliser un concept plus riche de rationalité. En supposant que la rationalité est limitée, la théorie des jeux évolutionnistes constitue un appel d'air permettant d'intégrer "l'Histoire". Cependant, comme cette théorie ne spécifie pas le comportement de l'individu en dehors de la stricte rationalité instrumentale, l'Histoire reste fondamentalement opaque. La rationalité et l'Histoire coexistent, mais ne sont pas articulées de façon satisfaisante.

Troisième partie

*Vers une théorie pratique de
la règle ?*

Dans le cadre économique standard, le problème de la règle émerge à partir du constat de l'incomplétude du système marchand. La critique d'Arrow consiste ainsi à établir deux conditions institutionnelles indispensables à un fonctionnement efficace d'un système de marché "autorégulateur". Il faut tout d'abord qu'il existe un consensus collectif quant à l'état social désiré par la "société" (l'ensemble des individus). Ce consensus permet de déterminer la dotation initiale à laquelle le marché devrait faire correspondre automatiquement l'état collectif souhaité. Toutefois, le marché ne fonctionne efficacement que lorsque l'avenir est parfaitement prévisible. En effet, l'incertitude fait apparaître une *altérité* qui entrave le déroulement du processus marchand. Il est donc nécessaire que des organisations (qui peuvent éventuellement prendre la forme d'un ensemble de "codes éthiques") soient mises en place afin d'assurer une bonne "gestion" de l'information et réduire ainsi l'incertitude et l'altérité entravant le fonctionnement du système marchand. Les règles organisationnelles viennent "combler" les lacunes du système marchand en imposant une certaine "conformité" des comportements individuels relativement à l'optimum social recherché.

L'analyse néoinstitutionnaliste s'inscrit dans cette perspective théorique. Williamson démontre que, du point de vue de la rationalité collective, l'organisation hiérarchique constitue le meilleur remède aux défaillances avérées du marché. Les analyses utilisant la théorie des jeux permettent d'apporter une réponse symétrique en soulignant que, même du point de vue des individus, la règle est nécessaire. Suivre une règle commune constitue une réponse "individuellement rationnelle" aux problèmes de coordination qui apparaissent dans le cadre des interactions directes entre agents.

Ces deux formes de "réponse" au problème identifié par Arrow ne permettent cependant pas d'éclaircir les mécanismes de la genèse et plus généralement de la

dynamique des règles. La théorie "hobbienne" de Williamson repose sur une opposition entre rationalité individuelle et collective qui interdit que l'ordre collectif puisse être déduit des comportements individuels. La théorie des jeux tente de réduire cette fracture entre individuel et collectif en soulignant que la règle "émerge" comme une solution rationnelle aux problèmes de coordination que les individus rencontrent. Mais cette tentative se heurte à un problème simple : comment une règle qui pallie les "manques" de la rationalité individuelle peut-elle se déduire du comportement d'individus rationnels ? De fait la seule forme de réponse à la question de l'émergence des règles est la réponse évolutionniste, qui remet en cause le principe de rationalité individuelle. Mais elle ne qualifie pas véritablement les individus autrement que de manière négative, en "allégeant" l'exigence de rationalité. De ce fait les *raisons* pour lesquelles des règles particulières émergent demeurent obscures.

La perspective "standard" du traitement du problème de la règle souffre donc d'une certaine forme d'incomplétude. Elle permet de comprendre la nécessité, le rôle, mais pas la genèse ou la dynamique des règles collectives. Cet "échec" vient de la manière dont la question de la règle se trouve posée dans l'univers standard. La règle est analysée comme une réponse collective aux failles du système de marché. Or, l'individu standard reçoit son mode de comportement du système marchand. On ne peut donc attendre qu'il résolve de lui même les problèmes que rencontre le système dont il est une partie. Les failles du marché sont aussi des failles de la rationalité instrumentale. Les règles ne peuvent être à la fois une "solution" aux failles de la rationalité instrumentale et un produit de ce mode de rationalité. Pour rendre un fondement individuel à la règle, il faut donc nécessairement parvenir à représenter d'une autre façon le comportement des individus. Pour cela il faut "sortir" l'individu du système de marché. Ce n'est que dans la mesure où l'individu ne reçoit plus son mode de rationalité d'un système de marché lui préexistant qu'il peut être suffisamment autonome pour être à l'origine des règles qui permettent à ce système de fonctionner.

Ce constat est au point de départ d'une autre forme d'analyse du problème que constitue le choix des repères collectifs permettant aux individus de se coordonner. Cette analyse non standard trouve son origine chez Keynes, qui construit sa théorie contre la représentation "klassique" du marché sur laquelle s'appuient les théoriciens standard¹.

Keynes ne pense pas que l'incertitude, et la diversité des anticipations qui en découle,

¹Klassique est mis ici comme contraction de "classique au sens de Keynes" selon la formule proposée par Favereau [1985]. L'analyse keynésienne est antérieure à la théorie de l'équilibre général. Elle contient néanmoins une critique des fondements de la représentation "standard".

sont la conséquence de “failles” du système marchand. Il les considère comme des caractéristiques propres à l’action économique. En effet, sa théorie ne s’appuie pas sur une représentation abstraite du marché, mais sur une observation pratique des mécanismes qui régissent effectivement l’économie de marché. Pour cette raison, l’analyse keynésienne rencontre immédiatement les problèmes que posent l’imprévisibilité du futur et l’irréversibilité du temps. Elle s’appuie sur une conception du comportement individuel qui, d’emblée, prend en compte le fait que les décisions doivent être prises sur la base d’une connaissance imparfaite du futur, et dans un cadre marqué par l’interdépendance et l’altérité des acteurs. La rationalité selon Keynes a donc une dimension pratique que la rationalité standard n’a pas.

Cette représentation pratique permet d’envisager immédiatement le rôle des règles : le comportement des acteurs, même raisonnables, n’est pas régi par une loi qui assure leur homogénéité. Il faut des règles pour harmoniser les différents points de vue qui s’expriment et s’entrechoquent. Mais, il ne faut pas n’importe quelle règle : Keynes pense qu’il est nécessaire de mettre en place des règles cohérentes avec l’objectif collectif du plein emploi. Or, les acteurs sont eux-mêmes trop englués dans la contingence pour parvenir à s’entendre sur des règles satisfaisante. Le salut ne peut venir que d’un acteur de type supérieur capable d’agir non seulement *dans* les règles mais aussi *sur* les règles : l’Etat. C’est à lui en définitive que revient la tâche de réguler l’économie, sur le conseil du savant. En effet, pour “réguler” le système marchand de manière satisfaisante, il faut s’en extirper et le contempler de l’extérieur. C’est ce que fait l’économiste : il identifie alors les raisons pour lesquelles les règles produites par les acteurs ne sont pas “bonnes”. Cette “hauteur de vue” lui permet de conseiller l’Etat, l’Etat, grâce à ces conseils, peut mettre en place les “bonnes” règles, adaptées à l’objectif collectif.

Keynes met donc en évidence le rôle pratique des règles. Il intègre l’existence des règles dans le mode de rationalité des agents. Mais il fait supporter à l’Etat la tâche de la régulation. Il ne s’agit donc pas d’une théorie permettant d’expliquer comment les règles émergent à partir du comportement des acteurs, mais d’une théorie de l’action de l’homme d’Etat. Cette conception est problématique, parce qu’elle fait dépendre la raison des acteurs de l’action de l’Etat. Keynes assume la nécessité de la règle. Il n’assume pas son caractère contingent.

C’est sans doute cette prise en compte du caractère négociable, contingent, mouvant, des règles de coordination qui constitue l’avancée principale de *l’économie des conventions*. Ce courant d’analyse récent cherche en effet non seulement à reconnaître le rôle des règles, à intégrer la nécessité de la règle dans la représentation du comportement des individus, mais également à faire de l’acteur le créateur des règles collectives. Elle

opère donc un lien essentiel entre la règle et l'individu qui, jusqu'alors, n'apparaissait pas en économie. L'étude du projet conventionnaliste et des difficultés qu'il rencontre met en évidence la contradiction qui existe entre rationalité instrumentale et prise en compte des règles. Pour autant, les conventionnalistes ne sortent pas tout-à-fait d'une perspective instrumentale. Cette contradiction révèle la profondeur du problème que pose la prise en compte des règles à l'analyse économique.

Cette partie décrit donc un mouvement théorique non standard qui trouve ses origines dans la théorie keynésienne et se poursuit à travers l'*économie des conventions*. La spécificité de ce mouvement non standard est de prendre en compte le caractère "pratique" ou "historique" des règles. Cependant, si cette partie permet de démontrer la pertinence de ce courant de recherche, relativement aux difficultés que rencontre le courant orthodoxe, elle ne conclut pas qu'il existe véritablement une théorie pratique de la règle économique. Une telle théorie se heurte à la difficulté, dans une perspective économique, de s'écarter d'une représentation instrumentale du comportement individuel. En définitive ce courant de recherche donne à voir la contradiction profonde qui existe entre les deux impératifs que sont la prise en compte des règles et le maintien du primat de l'efficacité.

Chapitre 5

*L'analyse keynésienne des
règles*

“La révolution keynésienne ne peut être comprise si l’on ne prend pas entièrement en compte la puissante influence qu’exercent le passé et le futur sur le présent. (...) Keynes avait certainement raison lorsqu’il affirmait que les preuves théoriques issues de modèles dont ces problèmes étaient exclus n’étaient pas recevables” (Arrow et Hahn, 1971, p.369)

L’univers économique décrit par Keynes est profondément différent de l’univers économique néoclassique. Cette différence provient, comme l’indiquent Arrow et Hahn, de la nature temporelle de l’univers keynésien. Keynes accorde en effet une importance capitale au fait que les décisions économiques prennent place dans une forme de temporalité “pratique”, de nature irréversible, qui comprend, comme l’a souligné Shackle, de la nouveauté et donc de l’incertitude.

C’est la raison pour laquelle l’univers économique qu’étudie Keynes ne peut pas être décrit comme un système idéal régi par des lois. L’analyse de la coordination prend donc dans l’analyse keynésienne un sens tout-à-fait différent de celui que lui donne l’analyse walrassienne, centrée autour de l’idée d’un système marchand autorégulateur. Lorsqu’il se pose la question du temps, Keynes n’étudie pas les moyens d’évacuer les problèmes liés à l’incertitude en concevant des institutions permettant de préserver l’autonomie et l’efficacité d’une sphère marchande “naturellement” régie par la loi de l’offre et de la demande. Il cherche au contraire à comprendre dans quelle mesure, étant donné le mode de connaissance et d’action des agents, l’ensemble de leur action peut “faire système” et atteindre une certaine “efficacité”, c’est-à-dire, pour Keynes, assurer le plein emploi. Ce faisant, Keynes se pose d’emblée le problème de la règle, que l’on peut exprimer ainsi : “quels repères collectifs sont nécessaires pour harmoniser des comportements d’agents,

qui, parce qu'ils se déroulent dans le cadre d'une temporalité pratique opaque, sont *a priori* divergents ?". Notre lecture de Keynes vise à démontrer que sa théorie peut servir de fondement à une théorie de la règle économique¹.

Notre étude se déroulera en quatre temps. Nous chercherons tout d'abord à exposer une définition keynésienne "pratique" de la "rationalité", ou de la "raison" individuelle qui est radicalement différente de la rationalité "formelle" néoclassique. Keynes construit dans son *Treatise on Probability*, une théorie de la prise de décision dans l'incertain, qui constitue en fait de façon très générale une théorie de la "raison" ou de la "logique" dans un univers humain et donc contingent. Le point important de cette théorie est de définir la rationalité non pas seulement par des propriétés logiques permettant de définir un ensemble de choix cohérents, mais plus profondément par la mise en évidence de "l'ancrage pratique" du jugement rationnel. Le fait que la rationalité procède chez Keynes d'une connaissance personnelle, et donc forcément relative, sépare sa conception de l'individu de la conception néoclassique. Pour Keynes la "raison" ne peut prendre la forme d'une loi mathématique ou logique précisément parce qu'elle s'exerce dans l'incertain.

Cette conception ne changera plus par la suite, mais il lui faudra résister à deux critiques. La première est formulée par Ramsey, à travers une théorie des probabilités qui critique explicitement celle de Keynes. On cherchera à montrer dans quelle mesure Ramsey ébranle Keynes, sans le convaincre de suivre "sa" conception de la probabilité, "subjective" et sans fondement pratique, qui donnera plus tard lieu à l'axiomatique de Savage. Keynes s'oppose à Ramsey et institue ainsi une ligne de résistance à la théorie de la maximisation de l'utilité espérée que formulera plus tard Savage. La seconde critique est le fait de Keynes lui-même. Keynes constate en effet que le comportement des individus dans l'univers économique diffère de celui que sa théorie "idéale" du comportement rationnel décrit. Dans un essai autobiographique postérieur à la *Théorie Générale*, Keynes semble sourire de cette forme de naïveté de ses premiers écrits. Il explique que, confronté à la réalité des comportements économiques, il a pris conscience du rôle des règles. Keynes ne change pas pour autant de théorie de la "raison". Il constate simplement que les règles collectives peuvent être un moyen de la préserver.

La *Théorie Générale* décrit, précisément, des individus qui sont pris dans un système qui ne leur permet pas d'être "raisonnables", au sens du *Treatise*. Keynes montre que l'anticipation conventionnelle de l'avenir que produisent les marchés financiers est

¹ On reconnaîtra peut-être dans cette présentation l'influence importante de la lecture de Keynes proposée par Olivier Favereau [1985, 1986, 1988]. C'est en effet à partir de cette interprétation qu'il nous a semblé nécessaire d'aller chercher chez Keynes les fondements d'une théorie non-orthodoxe de la règle.

déconnectée de la réalité. Or cette anticipation détermine les conditions de financement des investissements. Son instabilité est donc une des raisons pour lesquelles l'investissement s'établit à un niveau trop faible pour assurer le plein emploi. Pour stabiliser l'investissement à un niveau supérieur il serait nécessaire de mettre en place des règles permettant de remettre le jugement "raisonnable" de l'entrepreneur au centre de l'économie, en lieu et place du jugement conventionnel des marchés financiers.

A travers cette analyse Keynes montre donc que le fonctionnement de l'économie dépend des règles choisies. Cependant, il montre aussi que les "bonnes règles" n'émergent pas forcément spontanément. Le diagnostic keynésien oppose les "conventions" à la "réalité". Nous y voyons la trace d'un "réalisme" keynésien, héritée de Platon via Moore. Ce réalisme est au fondement d'un certain élitisme keynésien. Keynes ne pense pas que les règles économiques puissent être le fruit d'une délibération collective, mais plutôt d'une action habile de l'homme d'Etat. C'est un fondement problématique pour une théorie "individualiste" de la règle.

1- Une rationalité keynésienne

“Le probable est ce qu’il est rationnel de croire au regard des connaissances que nous avons.” (Keynes, 1921, p.339).

Le *Treatise on Probability* est un ouvrage philosophique qui vise à donner un fondement logique au jugement de probabilité, même et surtout lorsqu’il ne prend pas la forme d’un calcul. Le projet de Keynes s’inscrit en cela dans le mouvement des tentatives menées à Cambridge, notamment par Russel, pour donner un fondement logique aux mathématiques². De ce fait, comme le souligne O’Donnel (1989, p.31) le livre de Keynes “aurait tout aussi bien pu être titré *Traité sur la Logique ou Traité sur la Raison plutôt que de recevoir le titre trompeur de Traité sur les Probabilités*”. Keynes désigne en effet par “probabilité” toute la connaissance que l’on n’obtient pas “directement” mais “indirectement”, c’est-à-dire par raisonnement logique. Dans cette optique, son *Treatise* vise à définir les critères inductifs et déductifs permettant de qualifier de “rationnelle” une décision fondée sur des connaissances partielles. Il expose donc une véritable “théorie générale” du comportement rationnel, plutôt qu’une stricte théorie des probabilités.

Pour Keynes, le jugement de probabilité est de nature “pratique” : il repose sur la capacité des individus à percevoir dans la réalité qui les entoure les relations logiques qui unissent les événements. Il dépend donc des connaissances concrètes de l’acteur autant que de ses capacités déductives. C’est la raison pour laquelle les probabilités ne peuvent que rarement prendre une forme mathématique. Or, la prise de décision rationnelle s’appuie nécessairement sur un jugement de probabilité. La rationalité n’est donc pas purement calculatoire. Elle dépend des facultés pratiques de l’acteur et de sa capacité à avoir l’intuition des enchaînements logiques qui peuvent unir les éléments dont il a connaissance. *Non calculatoire, pratique, intuitive* : ces trois grands aspects de la logique keynésienne définissent une forme de “raison” opposée à la “rationalité” standard.

11- Une approche logique non calculatoire.

Bateman [1988] indique que Keynes développe son analyse de la logique des probabilités en réaction à l’utilisation, par Moore, d’une conception “fréquentiste” des

² Keynes dans sa préface, place son *Treatise* sous la double influence de Moore et Russel.

probabilités. Moore, dans ses *Principia Ethica* stipule qu'un individu rationnel *doit* agir en fonction des règles de probabilité indiquant le cas le plus probable, compte tenu des expériences passées. Cette façon de contraindre la réflexion d'un acteur à une sorte de loi causale qui lui serait extérieure était intolérable au jeune Keynes. Pour retrouver une forme de libre arbitre et de libre jugement, il lui fallait déconstruire cette conception, jugée naïve, des probabilités. Keynes cherche donc à définir le jugement rationnel d'un individu souverain en réaction contre les deux prétendues "lois quantitatives" censées régir l'incertain : la loi mathématique de Laplace et la loi fréquentiste de Venn.

Il élabore, en réaction contre cette conception calculatoire des probabilités, une conception épistémique et logique de la relation de probabilité. Le jugement de probabilité n'est plus conçu comme la mesure du degré d'occurrence d'un événement. Il prend la forme d'une relation logique reliant des prémisses certaines à des conséquences possibles. Ces conséquences ne sont cependant pas certaines, elles ne sont que probables. La probabilité correspond donc au degré de croyance que l'acteur raisonnable entretient quant au fait que cette relation logique décrive un enchaînement *réel* qui se produira effectivement.

111- Le rejet des approches quantitatives du jugement dans l'incertain.

"La grande réalisation du dix-huitième siècle a été, aux yeux des premiers penseurs du dix-neuvième, de réconcilier les deux points de vue [de l'induction et de la déduction] et de mesurer les probabilités, qui étaient tirées de l'expérience, par une méthode dont les bases logiques étaient le principe de la raison non-suffisante. (...) Mais il ne fut pas longtemps avant que des personnalités plus sceptiques en viennent à penser que cette théorie prouvait trop de choses. Pour appliquer la méthode de Laplace, il n'était en effet pas nécessaire d'avoir une grande expérience ou des informations particulièrement détaillées. Il fallait simplement que le peu d'information dont on dispose soit parfaitement symétrique." (Keynes, 1921, p.92)

Keynes reproche à Laplace de faire primer le caractère mathématique de la prise de décision rationnelle sur son aspect pratique. Dans cette citation, Keynes fait plus précisément référence à l'importance, dans la théorie Laplacienne de l'axiome d'équiprobabilité des différentes occurrences possibles, qui est la base de l'analyse laplacienne du jugement de probabilité. Laplace s'inspire en effet du principe de la "raison non suffisante" de Bernoulli. Ce principe énonce que "n" événements doivent être dit "équiprobables" lorsqu'il n'existe aucune information permettant de distinguer leur occurrence. On leur affecte alors une probabilité de $(1/n)$. Keynes (1921, p.122)

souligne la difficulté d'application de ce principe dans les cas concrets : il faut, pour qu'il soit valable, que les "n" événements comparés soient des hypothèses "*alternatives*" les unes des autres qui représentent à elles toutes une définition "*exhaustive*" des états du monde possibles. Ces deux conditions définissent une classe très restreinte de phénomènes. Dans la plupart des situations, la complexité du réel ne permet pas de définir des hypothèses absolument alternatives, et encore moins un ensemble "complet" d'états du monde possibles. La méthode de Laplace qui permet de généraliser l'approche mathématique repose donc, selon Keynes, sur une conception trompeuse du jugement dans l'incertain. Elle procède d'une simplification inacceptable du réel visant à ramener les cas "concrets" à des situations expérimentales, du type "tirage au sort dans une urne". Cette méthode mathématique sacrifie donc la connaissance pratique à la déduction logique et, ce faisant, biaise le raisonnement probabiliste.

Par cette critique, Keynes rejoint l'analyse fréquentiste de Ellis et Venn. L'école fréquentiste se constitue en effet contre l'école laplacienne. Elle dénonce la faiblesse des soubassements empiriques de l'approche mathématique et l'accuse d'être une structure "vide". Pour l'école fréquentiste, la seule vérité est en effet d'ordre empirique et naturel : "*l'expérience est notre seul guide*" (Venn, cité par Keynes, 1921, p. 93). La vérité ne peut donc être approchée que par une connaissance empirique de type statistique. Dans cette optique, l'incertitude n'est pas levée par l'utilisation de lois mathématiques abstraites, mais par la mesure des fréquences de l'occurrence réelle d'une classe d'événements.

L'intérêt de la position fréquentiste est évident. L'utilisation de données passées permet de calculer les probabilités sans négliger leur fondement empirique. Les probabilités deviennent en quelque sorte des phénomènes naturels ne devant rien à la subjectivité de l'acteur. Mais les raisons pour lesquelles la théorie fréquentiste ne peut constituer une théorie des probabilités sont, elles aussi, évidentes (Keynes, 1921, chap 8). Tout d'abord, le calcul de la fréquence nécessite de constituer des classes de phénomènes, et cette opération détermine en grande partie le résultat. La subjectivité du jugement n'est donc que dissimulée par la possibilité de mesurer, puisque la mesure repose elle-même sur une construction qui n'est pas naturelle. Toutefois, la véritable limite de l'approche fréquentiste se situe à un niveau plus profond : rien de ce que l'on sait du passé ne nous permet de prédire l'avenir à coup sûr. C'est précisément pour cette raison que le jugement de probabilité est nécessaire. En effet, si l'observation du passé nous permettait à elle seule de connaître parfaitement l'avenir, il n'y aurait plus de problème de probabilité. L'approche fréquentiste repose donc sur une négation du problème de l'incertitude. Elle prétend y répondre en élargissant au maximum la base de

l'ensemble d'informations sans prendre en compte le fait que l'information sur le présent ne peut pas entièrement réduire l'incertitude quant au futur. Une telle attitude consiste à nier purement et simplement le fait que le futur réserve de la nouveauté. Si la connaissance des fréquences passées pouvait nous guider de manière certaine, il n'y aurait en effet pas de problème de probabilité. Nous serions dans le domaine de la certitude. L'approche fréquentiste nie le paradoxe de l'induction, et la nature hypothétique du savoir construit sur des bases empiriques.

Le projet keynésien se construit contre ces deux approches calculatoires des probabilités qui dominent la fin du dix-neuvième siècle. Il s'agit pour Keynes de redonner une certaine valeur au *jugement* de l'individu. Keynes veut montrer que l'acteur raisonnable peut atteindre une certaine forme de vérité sans pour autant être contraint dans son raisonnement par des "lois". Il régénère ainsi la tradition "épistémique" ou "épistémologique" d'approche des probabilités. Ainsi, pour Keynes, le jugement de probabilité ne se résume pas en la connaissance des lois du hasard qui régiraient la nature. C'est un mode logique d'approche de la réalité, prenant pleinement acte de notre incapacité à "tout connaître". Dans la lignée du "pari" de Pascal, Keynes envisage le jugement de probabilité comme un moyen pour l'individu de décrypter un ordre divin qu'il ne peut qu'imparfaitement connaître. *"Il y a une partie de notre connaissance que nous obtenons directement, et une autre partie que nous approchons par le raisonnement. La théorie des probabilités concerne cette forme de connaissance que nous obtenons par le raisonnement."* (Keynes, 1921, p.3). Pour rendre justice à cette forme de savoir, Keynes décrit la rationalité comme une capacité au "jugement" consistant à percevoir, à partir de ce que nous connaissons "directement" de la réalité, les liens "logiques" nous permettant d'approcher la partie "obscur" du réel.

112- La relation de probabilité

Keynes propose donc de concevoir le jugement de probabilité comme un "degré de croyance rationnelle" obtenu par induction à partir d'une connaissance certaine d'une partie de la réalité. Le schéma général que Keynes établit peut être comparé à celui de l'empirisme logique, dont l'élaboration est contemporaine du *Treatise*. Si le choix rationnel se construit par déduction à partir d'un ensemble de prémisses, sa rationalité dépend de la vérité de cet ensemble. En effet, la vérité des prémisses du raisonnement est une condition suffisante à celle des conclusions, puisque la vérité des prémisses est transmise aux conclusions par le biais d'une déduction logique non sujette à l'erreur. Or, la vérité des conditions est nécessaire pour que l'action entreprise sur la base de cette

connaissance soit rationnelle.

Toutefois, le cadre de l'empirisme logique est celui d'une absence de temps. C'est pourquoi la vérité des prémisses est non seulement nécessaire mais *suffisante* pour assurer la validité des conclusions. Keynes, au contraire, s'intéresse au jugement portant sur des périodes futures, pouvant remettre en cause la validité de l'ensemble des prémisses. La relation entre les prémisses informationnelles et les diverses occurrences possibles ne peut donc, selon lui, être certaine : elle est seulement et irrémédiablement probable. C'est pourquoi il définit le jugement de probabilité comme "*le degré de croyance qu'il est raisonnable d'entretenir [en cette relation] dans des conditions données*" (Keynes, 1921, p.4).

Formellement la théorie keynésienne prend la forme suivante: "*Supposons que nos prémisses représentent un ensemble de propositions "h", et que nos conclusions constituent un autre ensemble de propositions "a", alors, si notre connaissance de "h" justifie une croyance rationnelle en "a" de degré α , nous disons qu'il existe une relation de probabilité de degré " α " entre "a" et "h". Cela s'écrit : $a/h=\alpha$.*" (Keynes, 1921, p.4). Sur cette base, Keynes (1921, p.11-13) propose de distinguer deux formes de connaissance : nous avons une connaissance *directe*, par le biais de nos sens, de l'ensemble "h". Nous avons une connaissance *indirecte*, par le biais du raisonnement, de l'ensemble "a". La connaissance, certaine, du lien logique qui unit "h" et "a" est appelé "*proposition secondaire*". Elle est forcément vraie, si notre croyance est rationnelle. En revanche, la proposition "a" est seulement probable, elle peut s'avérer finalement être fausse, même au cas où notre croyance était rationnelle. Keynes la qualifie de "*proposition primaire*". Cette distinction, sur laquelle nous reviendrons, permet d'accrocher la "rationalité" d'une croyance à la "connaissance objective".

12- L'ancrage pratique de la rationalité.

La définition que Keynes propose de la relation de probabilité, fait donc dépendre la rationalité d'un comportement non seulement de la cohérence formelle des choix, mais plus encore de l'adéquation du comportement à la situation "objective" d'action que l'acteur perçoit imparfaitement. C'est, en ce sens, une théorie *pratique* de la rationalité. La rationalité du comportement dépend de la pertinence du regard que l'individu porte sur la réalité qui l'entoure.

121- Le concept de poids de l'argument

O'Donnel [1989, 1991] souligne que le concept de "poids de l'argument" ("weight of argument") met en évidence cette caractéristique de la théorie keynésienne. Ce concept marque la dimension "pratique" ou "pragmatique" de cette théorie. En effet, le "poids de l'argument" représente le poids de la pratique dans un raisonnement logique. Keynes entend ainsi conceptualiser l'idée simple selon laquelle les informations certaines, ou que l'on sait être exhaustives relativement à un problème particulier, pèsent plus lourd dans notre choix que les informations peu fiables ou parcellaires. En quelque sorte, le *poids* de l'argument dépend de la *taille* de l'ensemble d'information³.

Cependant, le "poids" est encore plus difficile à mesurer que la probabilité : il représente d'une part une évaluation subjective de la partie de la réalité que l'on connaît, et d'autre part, suppose une réflexion sur la validité ou la pertinence de ces connaissances relativement au problème traité. Il est donc préférable de saisir le concept de poids dans une perspective non pas absolue mais comparative : ce qui est important, selon Keynes, est de saisir que l'accroissement des informations sécurise le jugement de probabilité. Une probabilité s'appuyant sur des informations nombreuses aura un poids plus important qu'une probabilité reposant sur des informations parcellaires, même si la première probabilité est très faible tandis que l'autre est forte. Or, Keynes affirme qu'entre deux probabilités l'acteur rationnel doit se fonder sur celle qui a le poids le plus important. Cette affirmation est essentielle car elle porte une conception de la prise de décision tout à fait novatrice⁵. En effet, la tradition de l'utilité espérée, puis la tradition benthamienne ont instauré une conception de la rationalité qui fait dépendre la prise de décision de l'intensité des probabilités. Keynes indique au contraire qu'il peut être plus rationnel d'agir sur la base d'une probabilité faible, mais dotée d'un poids important⁶.

Le critère du "poids" permet à Keynes d'opérer une distinction nette entre un événement "*très improbable*" et un événement "*très incertain*" selon la note célèbre de la théorie générale (Keynes, 1936, p.164) L'incertitude est en effet directement liée à la

³ Définir précisément ce qu'est le poids de l'argument relève de la gageure, précisément parce que ce concept désigne ce qui échappe à l'analyse. Carabelli [1988] parle à son sujet de concept indéfinissable. Le concept de poids de l'argument est tout-à-fait symbolique de la tension entre la démarche analytique et le souci de validité pratique qui traverse le *Treatise*.

⁴ Définir précisément ce qu'est le poids de l'argument relève de la gageure, précisément parce que ce concept désigne ce qui échappe à l'analyse. Carabelli [1988] parle à son sujet de concept indéfinissable. Le concept de poids de l'argument est tout-à-fait symbolique de la tension entre la démarche analytique et le souci de validité pratique qui traverse le *Treatise*.

⁵ Nous avons proposé un exemple mettant en jeu le concept de poids de l'argument dans Postel [1999].

⁶ L'afflux d'informations peut faire augmenter, diminuer ou laisser stable la probabilité, alors même qu'elle fait forcément augmenter son poids. La distinction entre poids et probabilité est donc essentielle contrairement à ce que peut laisser croire l'interprétation de Ventelou [1997 a et b].

faiblesse des informations pertinentes dont dispose l'acteur au moment d'agir. Cette distinction ressemble à celle qu'opère Knight [1921] à la même époque, sans toutefois la recouvrir entièrement. Knight distingue en effet les probabilités, calculables, et l'incertitude. Selon cet auteur, l'incertitude régit la plupart des comportements économiques, et se caractérise par l'impossibilité d'un comportement rationnel. Il n'y a pas pour Knight de rationalité dans l'incertain. C'est d'ailleurs là le fondement de sa théorie de l'institution. Keynes ajoute à cette distinction knightienne un échelon intermédiaire, qui transforme cette dichotomie en continuum. La plupart des jugements rationnels sont en effet seulement probables et fondés sur des jugements non quantifiés, sans pour autant être absolument incertains. Il y a donc chez Keynes une forme de jugement rationnel, qui ne s'appuie ni sur la certitude, ni sur une série de probabilités calculables, mais qui tire sa rationalité de l'existence d'une base informationnelle importante.

122- Une incertitude liée à la nature pratique des connaissances

La nature pratique de la rationalité keynésienne l'oppose de manière radicale à la rationalité formelle de la théorie néoclassique. Par l'idée de rationalité, Keynes ne vise pas des facultés de logique "formelle" liée à des capacités déductives, il cherche au contraire à mettre en lumière la rationalité de "l'induction". C'est dans sa capacité à "comprendre" la situation d'action que l'individu fait preuve de sa rationalité. Or, c'est précisément cette attitude inductive qui donne toute son importance à la temporalité de l'action et à l'incertitude. Une relation d'implication purement logique est hors de temps. Mais lorsque cette relation d'implication est censée anticiper un enchaînement causal réel, il ne suffit pas que ce raisonnement logique soit pertinent sur le plan formel, il faut qu'il le soit également sur le plan pratique. La logique qui guide l'acteur doit être identique à celle qui détermine les relations réelles entre les événements⁷. Autrement dit, par "croyance rationnelle" Keynes entend désigner deux propriétés très différentes : la vérité pratique et le sens formel. Il s'appuie donc sur une conception de la rationalité qui a une dimension supplémentaire par rapport à la définition standard : celle qui définit un rapport pertinent au monde réel. Pour mettre l'accent sur l'importance de cette dimension pratique, on utilisera le terme de "raison" pour qualifier la "rationalité selon Keynes".

Cette dimension pratique du jugement de probabilité implique immédiatement un

⁷ Keynes fait se superposer les deux questions que l'empirisme logique distingue. Celle du "sens" logique d'une proposition : c'est une question formelle et hors du temps. Celle de la "vérité" d'une proposition : c'est une question pratique et située dans le temps. Lorsque Keynes définit une relation logique, il entend une relation non seulement sensée, mais aussi actuellement pertinente. C'est bien la raison pour laquelle le problème de l'incertitude se pose. Le monde étant mouvant une relation logique n'est pas toujours vraie dans le long terme.

rapport au temps qui fait de l'incertitude un problème constitutif du choix raisonnable⁸. En effet, les conclusions auxquelles aboutit le jugement de probabilité sont incertaines. L'incertitude qui touche cette "proposition primaire" a trois origines, toutes liées à la nature pratique du jugement. L'ensemble d'informations peut n'être pas exhaustif. Il peut aussi perdre de sa pertinence dans le monde futur. Enfin l'incertitude atteint la relation logique elle-même : elle doit être "perçue" dans le monde réel, et ce pouvoir de perception est relatif à l'acteur. Les individus ont une capacité relative à percevoir dans le monde sensible les connexions logiques unissant certaines causes actuelles à certains effets futurs. La relation de probabilité n'est pas donnée à l'acteur, et sa perception, dans un environnement complexe, est incertaine.

13- Une rationalité intuitive.

Ces trois origines de l'incertitude sont relatives à l'acteur. Elles dépendent de sa capacité à établir des prémisses empiriques puis à y repérer des connexions logiques. Cependant, bien que la relation de probabilité soit liée à l'expérience de l'individu, elle est objective. En effet, les relations de probabilités existent réellement dans l'univers qui est logiquement organisé : la perception intuitive de l'individu est donc un vecteur de la vérité.

131- Une relation dépendant du regard de l'acteur.

La constitution d'un ensemble pertinent d'informations requiert des qualités particulières d'observation et de rapprochement des informations : "*une partie de la connaissance est de manière évidente relative à l'expérience individuelle. Nous ne pouvons pas parler de connaissance dans l'absolu—mais seulement de connaissance d'une personne particulière*" [Keynes, 1921, p.18]. C'est pourquoi le rapprochement de la rationalité et de la connaissance pratique met en lumière l'acteur dans ce qu'il a de plus personnel, c'est-à-dire sa sensibilité, ses facultés de compréhension pratique. La collecte

⁸ La prise en compte du temps est sans doute à l'origine du différend qui oppose Keynes et Wittgenstein sur la question des probabilités. Cette opposition est perceptible dans quelques échanges épistolaires entre Keynes et Wittgenstein. Wittgenstein, dans un brève lettre à Keynes, indique "*Avez-vous progressé dans vos travaux sur les probabilités ? Mon manuscrit comporte quelques lignes qui, je pense, règlent les questions essentielles*" (Wittgenstein, 1919, p.63). Plus tard Keynes lui répond : "*Probability constitue l'achèvement de mes travaux d'avant guerre – Je crains fort qu'il ne vous plaise pas.*" (Keynes, 1924, P.201). En effet, Wittgenstein (1921, aphorisme 5.15 à 5.156) utilise le concept de probabilité pour désigner simplement des rapports logiques d'implication entre des propositions connues et d'autres inconnues. Il situe ces rapports logiques hors du temps.

de l'information s'apparente en grande partie à de la création d'information, et l'acteur apparaît donc comme le véritable créateur des solutions rationnelles qu'il propose. Sa rationalité naît de sa capacité propre à comprendre le monde davantage que de la connaissance d'axiomes mathématiques universels.

La seconde étape du raisonnement probabiliste est la déduction menée à partir des prémisses informationnelles. L'acteur doit établir une relation entre les diverses occurrences entre lesquelles il hésite et les informations dont il a connaissance. C'est sur cette relation que porte la probabilité. La déduction reprend donc ici ses droits : on pourrait s'attendre à ce que l'acteur tire cette fois sa rationalité de l'application de lois "objectives et universelles". Mais Keynes refuse là aussi de considérer que la déduction se résume à l'application d'axiomes mathématiques : "*Nous devons admettre, je pense, qu'elle [la connaissance des axiomes logiques] est également relative à la constitution de l'esprit humain, et que la constitution de l'esprit humain peut différer d'un homme à un autre de plusieurs degrés. Il est évidemment possible que certains aient un plus grand pouvoir d'intuition logique que d'autres*" [Keynes, 1921, p.18]. C'est pourquoi la déduction menée à partir d'un même ensemble d'informations peut différer entre deux individus, bien qu'ils soient tous deux rationnels. L'intuition logique, qui permet d'établir avec certitude une relation entre des informations dont on dispose et les chances de réalisation d'une des occurrences possibles, n'est pas plus universelle que la capacité pratique à comprendre une situation d'action. La relation de probabilité n'est pas donnée à l'individu. Il doit la percevoir, en avoir l'intuition, à partir des informations dont il dispose ; or, "*la perception de certaines relations de probabilité peut être au delà du pouvoir de certains d'entre nous*" [Keynes, 1921, p.19].

132- Une relation objective

Keynes estime pourtant que la relation de probabilité n'est pas subjective : "*Le fait que nous soyons en définitive dépendant d'une intuition ne doit pas nous amener à penser que nos conclusions ne sont plus, dès lors, fondées en raison, ou encore qu'elles sont aussi subjectives, quant à leur validité, que l'est leur origine.*" [Keynes, 1921, p.76]

Ainsi, pour Keynes, la déduction logique menée à partir des prémisses assure l'objectivité des probabilités. La particularité de la théorie keynésienne étant cependant de mettre en avant le caractère "personnel", et donc subjectif, du jugement de probabilité, le fait qu'il atteigne finalement une validité objective est surprenant.

Cette capacité de l'individu à atteindre une forme de vérité objective par le moyen de

l'introspection est un trait spécifique de "l'intuitionnisme" de Keynes⁹. Keynes est en effet très influencé par la théorie intuitionniste de Moore (O'Donnel [1989], Fitzgibbons [1988], Davis [1991], Schionoya [1991]). En quelque sorte, pour Keynes, l'acteur transcende sa propre subjectivité pour parvenir à la vérité. L'intuition permet à l'acteur d'échapper à sa propre subjectivité, de la sublimer, pour atteindre le niveau de la connaissance objective.

Cette conception intuitionniste de la perception du "vrai" suppose l'existence réelle des relations logiques que l'individu repère. La relation logique sur laquelle porte la probabilité existe donc indépendamment de l'acteur. Elle a une existence "objective".

Le *Treatise* définit un comportement rationnel dans des situations où l'individu n'est pas certain des conséquences des connaissances dont il dispose. Cette rationalité s'appuie sur une conception logique du jugement de probabilité. Cependant, la logique à laquelle Keynes fait référence n'est pas seulement la logique formelle que mobilise la conception néoclassique de la rationalité. C'est également une logique pratique qui désigne une capacité à agir de manière cohérente avec la réalité. Pour cette raison, la rationalité keynésienne n'est pas de nature calculatoire. Elle est relative à l'expérience et à la subjectivité de l'acteur. Deux individus ne percevront pas de la même façon des situations pourtant "objectivement" identiques. Cependant, bien que relative, la rationalité est liée à l'intuition de l'acteur : chaque aspect de la réalité que l'individu perçoit et chacune des relations logiques qu'il établit à partir de ses connaissances ont une existence réelle. Or, cette forme d'intuition rationnelle qui permet aux acteurs de se guider dans l'incertain constitue une source "d'inégalité". Non seulement tous les individus ne perçoivent pas la réalité de la même façon, mais certains en perçoivent une part moins importante que d'autres. Il existe donc des acteurs plus "raisonnables" que d'autres : c'est là le fondement de "l'élitisme" de Keynes.

⁹ Keynes adopte, selon Schionoya [1991], les trois principes définissant l'intuitionnisme de Moore : le bien est aussi le vrai et le rationnel ; l'idée du bien est intuitivement comprise de tous, sans être définissable (niveau méta-éthique) ; l'idée du bien s'incarne concrètement dans l'action des acteurs les plus rationnels (niveau éthique et principe de l'excellence).

2- Du *Treatise* à la *Théorie Générale* : fidélité et évolution.

Dans cette seconde section, nous chercherons à préciser deux points importants de la réflexion de Keynes en matière de rationalité. Pour défendre l'idée d'une certaine "continuité" des convictions keynésiennes si précocement acquises (Keynes débute la rédaction du *Treatise* en 1909), il faut en effet expliquer deux textes dans lesquels Keynes semble prendre une certaine distance avec les thèses du *Treatise*.

Le premier texte tient, en fait, en quelques lignes que Keynes consacre à un compte rendu de l'ouvrage de Ramsey sur les probabilités, ouvertement critique de son propre *Treatise*. Keynes semble s'avouer vaincu. Nous montrerons que cette capitulation est extrêmement relative : Keynes, immédiatement après avoir souligné qu'il reconnaissait le bien fondé des remarques de Ramsey, réaffirme les grandes lignes de sa conception de la "rationalité". Ce débat entre Keynes et Ramsey nous permettra de montrer que l'idée de rationalité keynésienne s'oppose directement à ce qui constituera la théorie subjective de l'utilité espérée de Savage.

Le second texte est un essai autobiographique postérieur à la *Théorie Générale*, dans lequel Keynes commente ses premières années à Cambridge. Il se décrit comme un étudiant fasciné par l'idéalisme de Moore et refusant absolument la nécessité de toute contrainte quant à la prise de décision. Il souligne le caractère naïf de ses convictions de jeunesse, et justifie incidemment sa progressive prise de conscience du rôle des règles collectives qui encadrent le choix individuel. Cette conversion de Keynes à la nécessité sociale des règles est une clef de compréhension importante de l'évolution de ses convictions entre le *Treatise* et la *Théorie Générale*.

Pour l'essentiel, les conceptions de Keynes en matière de rationalité ne changent pas. Il les adapte, les élargit, les précise. Il ne les renie pas. Pour Keynes, l'individu rationnel est toujours cet acteur tiraillé entre la nécessité d'asseoir son choix sur une bonne connaissance pratique, et celle d'agir malgré le caractère imparfait de cette connaissance.

21- Le débat Keynes-Ramsey

A l'occasion d'un court compte-rendu posthume de l'ouvrage de Ramsey, dont il était l'ami, Keynes célèbre la profondeur de son étude et "cède" à Ramsey. Cependant, bien qu'il admette que Ramsey avait sans doute raison en remettant en cause le principe de

l'objectivité des propositions secondaires, Keynes ne cède pas sur l'essentiel. Les quelques lignes qui suivent sa "capitulation" suffisent à souligner qu'il réaffirme, contre Ramsey, l'idée selon laquelle la logique qui détermine le jugement de probabilité rationnelle n'est pas seulement formelle. Elle désigne, et cela est essentiel selon lui, une faculté "pratique" que la théorie subjectiviste de Ramsey passe sous silence. C'est précisément cette dimension pratique du jugement de probabilité que Keynes invoque dans son *Treatise on Probability* pour démontrer que la rationalité dans l'incertain ne peut être décrite par le principe de la maximisation de l'utilité espérée.

211- Keynes cède à Ramsey...

"Ramsey affirme, contre la conception que j'ai moi-même défendue antérieurement, que la probabilité ne décrit pas des relations objectives entre des propositions mais (en un sens) des degrés de croyance. Il parvient à montrer que le calcul des probabilités consiste seulement en un ensemble de règles qui assurent que notre système de degrés de croyance forme un système cohérent. Pour cette raison le calcul des probabilités demeure dans le domaine de la logique formelle. Par contre, la base de nos degrés de croyances– ou les probabilités a priori, selon l'expression consacrée – est une part de notre équipement humain, que nous avons peut-être reçu par sélection naturelle, comparable à notre capacité à percevoir et à nous souvenir, plutôt qu'à la logique formelle. Jusqu'à ce point, je cède à Ramsey – Je pense qu'il a raison. " (Keynes, 1931, p.338-339, c'est moi qui souligne).

Ramsey conçoit sa théorie comme une critique de la théorie de Keynes, tout en se situant dans une perspective extrêmement proche. Ainsi Hacking [1975, p.13] souligne que l'approche keynésienne et l'approche de Ramsey constituent les deux branches d'une même "école", qu'il qualifie d'épistémologique. Ce mode d'approche des probabilités, que l'on pourrait également qualifier de "logique" et "d'épistémique", conçoit la probabilité comme une forme de relation logique entre deux propositions, et s'oppose à l'approche "statistique". Pour Keynes comme pour Ramsey, donc, la probabilité décrit un degré de croyance rationnelle en une relation logique mais incertaine. La divergence entre les deux auteurs vient de la manière de concevoir cette relation logique. Chez Keynes, cette relation est, en quelque sorte, duale : il y a d'une part une relation logique "objective" entre l'ensemble d'informations "h" et la connaissance "a" que l'on en déduit. D'autre part cette relation logique est perçue par l'acteur sous la forme d'une croyance rationnelle de degré " α ". La probabilité " α " désigne donc à la fois une relation logique "objective" et un degré de croyance "subjectif". C'est cette articulation complexe que

Ramsey va critiquer¹⁰.

Ramsey montre en effet que le concept central de probabilité non calculable est contradictoire avec le principe du “degré” de croyance. S’il existe un degré, celui-ci est forcément quantifiable. C’est qu’en réalité Keynes, lorsqu’il s’intéresse aux probabilités non calculables, “oublie” cette notion de degré et met l’accent sur le caractère objectif de la relation de probabilité. Keynes affirme effectivement que la “proposition secondaire” (c’est-à-dire la relation logique entre “h” et “a”) est toujours vraie : elle existe en réalité. Or, Ramsey indique : “*il n’existe rien de tel que les relations de probabilités décrites par Keynes. Je ne les perçoit pas, et si l’on doit me persuader qu’elles existent, c’est par argument*” (Ramsey, 1931, p.161).

Ramsey fait donc voler en éclat la conception keynésienne, et en particulier sa dimension intuitionniste¹¹. Il nie qu’il puisse exister une connaissance du type de celle que représentent les “propositions secondaires” dans la théorie keynésienne. Il nie que le jugement de probabilité puisse être autre chose qu’un jugement purement subjectif ne reposant sur aucun fait réel. Selon Ramsey, la seule chose que l’on puisse connaître, c’est un degré de croyance *subjectif*, portant sur les événements eux-mêmes et non plus sur une relation logique objective. La seule manière que l’on ait de qualifier ce jugement de logique, et de fonder en logique le raisonnement probabiliste, c’est de définir des critères formels que devrait respecter l’ensemble des degrés de croyance d’un individu. La logique des probabilités appartient donc à la logique formelle, et se caractérise par une axiomatique définissant les règles que doivent respecter les degrés de croyance d’une même personne pour former un tout cohérent. La solution astucieuse de Ramsey consiste à “mesurer” le degré de croyance d’un individu comme s’il s’agissait d’un pari. L’intensité de sa croyance dépend de l’enjeu de l’action qu’il serait prêt à tenter sur la base de cette croyance. Cette “astuce” théorique sert de base à la construction de la théorie moderne de la décision dans l’incertain, dont le point d’aboutissement est l’axiomatique de Savage [1950] qui, en fait, a tiré la théorie de Ramsey de l’oubli.

Lorsque Keynes, dans le cadre d’un débat strictement cambridgien, sans écho et sans enjeu, indique qu’il “cède” aux critiques de Ramsey, il ne sait donc pas qu’il cède en fait à l’ensemble de la théorie moderne de la rationalité dans l’incertain qui servira de fondement à l’essor de la théorie économique néoclassique et à la négation du problème “keynésien” de l’incertitude. En effet ce que Ramsey rend possible, c’est une théorie de la rationalité dans l’incertain qui s’affranchisse de son ancrage pratique, si essentiel chez Keynes. En

¹⁰ Sur le débat Keynes-Ramsey, on peut consulter le très bel article de Marion Glatron [1998].

¹¹ En quelque sorte, c’est l’intuitionnisme de Keynes qui se trouve ici mis en cause. Gerrard [1992, p.88] souligne ainsi que Keynes suit Moore, et que Ramsey suit Wittgenstein. Wittgenstein affirme en effet, dans le court passage consacré à la logique probabiliste dans le *Tractatus* : “*La proposition de probabilité n’a pas d’objet particulier propre.*” (Wittgenstein, 1921, p.111)

tranchant le lien entre la logique formelle des degrés de croyance et la logique pratique du jugement de probabilité, Ramsey affranchit l'individu rationnel de la complexité du réel, et lui permet de disposer d'une règle d'action mathématique fondée sur le calcul. Avec Ramsey, la rationalité redevient donc calculatoire. Elle cesse d'être par essence un jugement pratique "personnel" impossible à décrire sous forme de loi générale extérieure à l'acteur. La dimension strictement personnelle, subjective, dans le jugement de probabilité est refoulée hors de la définition de la rationalité et de la logique, alors même qu'elle en était le cœur dans la théorie keynésienne. Mais, est ce bien à tout cela que Keynes cède ?

212- ... Mais pas sur l'essentiel

"Jusqu'à ce point, je cède à Ramsey – Je pense qu'il a raison. Cependant, en cherchant à distinguer des degrés de croyance "rationnels" des croyances en général, il n'a pas encore, je pense, parfaitement réussi. Ce n'est pas pleinement comprendre le principe d'induction que de le désigner simplement comme une habitude mentale utile. Cependant, en tentant de distinguer une logique "humaine" différente de la logique formelle, Ramsey a peut-être identifié un nouveau champ de recherche, qui apparaît une fois que la logique formelle a été mise en bon ordre et que son domaine d'application, restreint, a été correctement défini." (Keynes, 1931, p.338-339, c'est moi qui souligne)

La concession de Keynes à la théorie de Ramsey est marquée d'un "bémol" qui, en réalité, est une réaffirmation de sa propre conception. En effet, Keynes peut accepter la remise en cause par Ramsey du concept de "proposition secondaire" (et plus encore de l'idée obscure d'un degré de croyance qui ne serait pas mesurable) sans pour autant abandonner sa position théorique. Le point essentiel de la théorie keynésienne est de montrer que la logique qui préside au jugement de probabilité s'appuie sur une connaissance pratique non formelle. Or, ce point n'est pas remis en cause par l'idée que la relation de probabilité n'existe pas réellement : elle peut être analysée comme étant uniquement le fruit d'une intuition rationnelle s'appuyant sur des informations empiriques, sans que le cœur de la démarche de Keynes soit touché.

C'est bien en effet à redonner toute son importance à la logique "humaine" de l'induction que vise le *Treatise*¹². Lorsque Keynes indique que Ramsey ne parvient plus à distinguer la croyance "rationnelle" de la croyance en général, il met très précisément l'accent sur ce qui disparaît dans le "tour de prestidigitation" que réalise Ramsey. La

¹² Sur ce point, on peut se reporter à l'analyse que propose O'Donnell [1989, chap.7] du débat Keynes/Ramsey. On peut également consulter l'analyse de Carabelli [1988, chap. 5] qui interprète l'approche keynésienne comme une tentative visant à donner un fondement logique à l'induction.

théorie ne dit plus rien du rapport de l'acteur à la pratique. Le jugement pratique qui fonde les probabilités *a priori* est subsumé sous le pari subjectif régi par la logique formelle. Or l'essentiel pour Keynes était bien de décrire cette faculté qu'a l'acteur d'agir raisonnablement sans être certain. Son *Treatise* vise précisément à rendre justice à la *connaissance* raisonnable mais seulement probable. Il vise à décrire le mécanisme par lequel la subjectivité de l'acteur le guide vers une connaissance finalement objective bien qu'indirecte. Ramsey opère une dichotomie entre logique formelle et logique humaine, et accroche la rationalité du jugement de probabilité à la logique formelle, alors que, précisément, Keynes refuse cette dichotomie et accroche la rationalité des probabilités à une logique de l'action caractérisée par sa dimension pratique et non pas formelle. Ainsi lorsque, dans le *Treatise*, Keynes cherche à distinguer la croyance rationnelle de la croyance en général il indique : "*Il est préférable de reconnaître que la connaissance est l'élément fondamental, et de définir la croyance rationnelle par référence à cette connaissance*" (Keynes, 1921, p.10). C'est bien le fondement cognitif d'une croyance qui la rend rationnelle pour Keynes. Ce n'est pas la cohérence formelle de l'ensemble des croyances d'un acteur qui détermine la rationalité d'une croyance particulière¹³.

213- Le refus du principe de la maximisation de l'utilité espérée

Ce débat entre Keynes et Ramsey permet de mettre en lumière le point central de la position keynésienne. Keynes admet, suite à la critique de Ramsey, le caractère problématique de sa représentation du jugement dans l'incertain. Il admet la contradiction qui existe entre l'idée d'un degré de croyance et le caractère non calculable de certaines probabilités. Il reconnaît sans doute aussi que le concept de "proposition secondaire" est problématique et ne s'appuie sur aucun "fait" vérifiable ou perceptible. Mais il ne cède pas en ce qui concerne sa conception pratique de la logique. Elle est autre chose qu'une capacité calculatoire permettant d'établir un classement formellement cohérent. Le fondement pratique du jugement rationnel est constitutif de sa logique, il n'est pas en amont une simple "*habitude mentale*". Or, en réaffirmant cette conception, Keynes renoue avec ce qui, dans son *Treatise on Probability*, fondait sa critique du principe conséquentialiste de la maximisation de l'utilité espérée.

Dans le chapitre 26 de son *Treatise*, Keynes entreprend de réfléchir au rôle des probabilités dans le choix de l'action adéquate. Il affirme à cette occasion son opposition à

¹³ Cette opposition entre Keynes et Ramsey est fondamentale. Elle trouve ses prolongements dans les critiques contemporaines de la conception standard que nous avons déjà présentées. Ainsi, les réflexions de Suppes [1981] qui souligne que le modèle de la "SEU" ne permet pas de distinguer "*un individu raisonnable d'un fou*", mais également les critiques de Binmore [1987] sur l'opacité de l'origine des "probabilités *a priori*", sont très proches des remarques de Keynes.

la théorie benthamienne, et plus généralement au conséquentialisme, en raison précisément de la nature pratique du jugement de probabilité. Dans l'optique de Bentham, la décision est prise mathématiquement, en pondérant l'utilité attachée à chacune des conséquences possibles d'un acte par les chances de réalisation des états du monde auxquelles ces conséquences sont conditionnées, puis en réalisant la somme des montants ainsi obtenus. Cette somme représente l'utilité espérée de cette décision. Pour que le choix de l'action soit réductible à une somme benthamienne, il faut donc non seulement que l'utilité soit quantifiable, mais aussi que les probabilités le soient. Keynes indique qu'il doute fort de la pertinence de la première condition et qu'il est convaincu du caractère fallacieux de la seconde : *"Elle ignore ce que je nomme le poids de l'argument, c'est-à-dire la quantité des preuves sur laquelle la probabilité est fondée. Elle ignore également le risque et suppose qu'une quantité égale de paradis et d'enfer est équivalente à un certain état de médiocrité"* (Keynes, 1921, p.344).

Autrement dit, cette théorie suppose que les probabilités sont toujours calculables tout en oubliant d'une part que les probabilités sont plus ou moins fiables et que les issues sur lesquelles elles portent sont plus ou moins risquées. Il y a là clairement opposition entre un individu calculeur manipulant des quantités abstraites, et un acteur concret, prudent, agissant en vertu de la fiabilité de ses connaissances. Pour un tel acteur, la possibilité de disposer d'une notion quantitative d'utilité ou de probabilité ne serait d'aucune aide. Son problème n'est pas, en effet, de pouvoir calculer, mais de parvenir à décider, alors même qu'il ne peut être certain de la pertinence de ses prévisions. Pour un tel acteur un pari à la Ramsey donne une illusion de connaissance, et ne constitue donc pas une base sérieuse pour agir. Comme toute notion "quantifiée" de la probabilité opère par une simplification du fondement cognitif et informationnel du jugement de probabilité, Keynes ne croit pas, à la différence de Condorcet et d'Edgeworth, que l'on puisse *"éclairer les Sciences morales et politiques par le flambeau de l'Algèbre"* (Keynes, 1921, p.349). L'économie, que Keynes perçoit comme l'une de ces sciences morales, ne peut donc être mathématique parce qu'elle décrit des acteurs qui sont aux prises avec le réel et qui, pour cette raison, ne peuvent raisonner de façon calculatoire. La théorie de Ramsey ne change pas les convictions de Keynes sur ce point¹⁴.

Ainsi, si Keynes cède aux critiques de Ramsey, il n'endosse pas pour autant sa représentation du comportement rationnel dans l'incertain. Son commentaire de la théorie de Ramsey montre au contraire que Keynes reste convaincu de la nécessité d'éclairer ce qui, dans la logique "humaine" (selon l'expression de Ramsey) est irréductible à la

¹⁴ Keynes réaffirme à plusieurs reprises son opposition à Bentham. Par exemple : *"Notre connaissance ne saurait fournir la base d'une prévision calculée mathématiquement"* (Keynes, 1936, p.168). Voir aussi Keynes [1937].

logique formelle.

22- La découverte des règles.

“Nous n’étions pas conscient que la civilisation est une fine et fragile enveloppe (crust) qui n’existe que grâce à la personnalité et la volonté d’un tout petit nombre de personnes et qui ne se maintient que par le truchement de règles et conventions habilement inculquées et astucieusement préservées” (Keynes, 1938, p. 447).

L’autre évolution de la pensée de Keynes du *Treatise* à la *Théorie Générale* tient à l’apparition du concept de règle, ou plus précisément de convention. En effet, l’individu rationnel que décrit Keynes dans le *Treatise* n’a pas besoin de règles. Il décide seul du choix de ses actions et semble pouvoir atteindre le vrai sans s’appuyer sur aucune règle. L’apparition du concept de règle est contemporaine chez Keynes de la prise en compte de la dimension sociale de la prise de décision raisonnable.

Dans un essai autobiographique postérieur à la *Théorie Générale*, Keynes revient sur cette attitude envers les règles et l’éthique de Moore. Trois caractéristiques de la théorie de Moore animaient les débats cambridgiens, et étaient plus particulièrement au centre des préoccupations du cercle des “apôtres”¹⁵.

La première des caractéristiques de la théorie de Moore est qu’elle repose sur une conception intuitionniste de la vérité. Cette mise en valeur de l’intuition était célébrée par les apôtres comme une théorie libératrice vis-à-vis des conventions sociales ou des règles de décision calculatoires. Keynes et ses collègues de Bloomsberry prônaient ainsi une conception “idéale” de l’individu : un individu parfaitement rationnel, qui d’une part n’avait nul besoin de s’appuyer sur des règles de comportement calculateur à la Bentham pour faire le bien, et qui d’autre part était parfaitement juste, ce qui rendait inutile toute règle sociale. Le choix de l’action était entièrement guidé par une introspection indiquant à l’individu ce qu’il fallait faire : *“Comment savions-nous qu’un avis était le bon ? C’était une affaire d’introspection directe, d’intuition non analysable à propos de laquelle il était inutile et impossible de se justifier.”* (Keynes, 1938, p.438). La détermination de la conduite échappait donc à toute prescription, et même à toute influence “extérieure” à la conscience de l’acteur. La théorie de Moore permettait aux “apôtres” de se libérer à la fois des carcans sociaux et du déterminisme mathématique de la théorie benthamienne. Ce faisant ils prônaient une conduite “intuitive” bien plus que Moore lui-même.

¹⁵ C’est le nom que se donnait à Cambridge le cercle d’admirateurs de Moore dont Keynes faisait partie.

En effet, dans la philosophie de Moore, cette exaltation de l'intuition rationnelle était nuancée par deux éléments importants. Moore s'accordait avec Bentham sur l'idée selon laquelle il était nécessaire, dans l'incertain, de choisir l'action assurant l'utilité espérée la plus élevée. Le *Treatise* de Keynes est précisément une critique de cet aspect de la conception de Moore. Le second élément modérant l'importance de l'intuition rationnelle était la morale. Moore soulignait en effet que les règles coutumières sociales devaient servir de guide à une prise de décision juste et bonne. Or, comme Keynes le souligne, les apôtres n'acceptaient pas ce précepte moorien invitant au respect des normes sociales : *“De Moore, nous conservions la religion, mais rejetions la morale. En fait un des grands avantages de sa religion était à nos yeux de rendre la morale inutile”* (Keynes, 1938, p.436).

L'intuitionnisme des “apôtres” était donc constitutif d'un individualisme profondément rétif à toute “règle” extérieure, qu'elle soit sociale ou mathématique : *“Nous étions parmi les premiers à nous échapper de la tradition benthamienne, (...) mais nous étions aussi les derniers des “Utopistes” du dix-neuvième siècle, (...) qui croyaient en un progrès moral continu grâce auquel la race humaine consistait en des individus fiables, rationnels, décents, guidés par la vérité et l'objectivité, qui pouvaient sans risque être libérés des restrictions extérieures comme les conventions ou la tradition.”* (Keynes, 1938, p.447).

Or, Keynes, sur ce point, affirme avoir profondément évolué. Le refus de la tradition benthamienne, reposant sur une conception “hautement problématique” (*ibid.*, p.445) du calcul des probabilités, constitue, on vient de le voir, une constante dans son oeuvre. Mais ce n'est pas le cas du rejet des règles sociales. Keynes considère en effet que la “naïveté” consistant à supposer que les individus peuvent toujours être rationnels et ne sont guidés que par une quête “illuminée” du bien, ne pouvait résister à l'étude des problèmes sociaux. Une telle position était absolument idéaliste et n'avait pas de pertinence pratique : *“Nous existions dans l'univers des “Dialogues” de Platon, nous n'étions pas encore parvenu jusqu'à la République. (...) L'action sociale avait été rejetée hors de notre univers, et pas seulement l'action sociale, mais l'univers de l'action en général, le pouvoir, la politique, le succès, la richesse, l'ambition. Les motifs économiques et le critère d'économicité étaient encore moins importants dans notre philosophie que dans celle de Saint François d'Assise”* (Keynes, 1938, p.445).

Le *Treatise* adhère à cette perception de l'individu dont une caractéristique importante est l'indépendance absolue. L'individu rationnel que décrit Keynes est, en quelque sorte, libéré de toute contrainte extérieure : les “autres” n'interviennent pas dans sa réflexion. Or, précisément, l'étude des problèmes sociaux, dont font partie les problèmes économiques, est marquée par le fait que les individus sont interdépendants et

poursuivent des objectifs propres et potentiellement contradictoires. Le passage de la réflexion philosophique idéaliste du *Treatise*, dessinant un acteur rationnel devisant seul de la conduite à tenir, à la réflexion économique amène nécessairement Keynes à adapter sa conception du comportement humain. Plus exactement, Keynes se trouve confronté à un problème qui n'a aucune raison d'apparaître dans le *Treatise* : celui de la coordination des actions. C'est à ce moment que l'idée de règle devient essentielle.

L'individu qui cherche à coordonner son action à celles des autres doit parvenir à prévoir leur comportement. Or, à supposer que ses partenaires soient rationnels, l'anticipation de leur comportement dépend encore de leurs connaissances. Keynes a en effet tout de suite montré que l'idée de rationalité n'excluait pas, en raison de l'ancrage pratique du raisonnement, l'idée d'hétérogénéité des comportements. Si l'on se situe dans un univers à information parfaite, l'anticipation du comportement des autres est triviale : le niveau de connaissance étant homogène, tous les individus agissent en fonction d'un processus inductif identique. Mais, précisément, l'information n'étant jamais parfaite, le comportement des individus est difficile à prévoir. L'hypothèse selon laquelle chacun peut décider seul, et indépendamment d'une réflexion sur le comportement des autres, apparaît alors à Keynes comme naïve : *“je peux nous imaginer comme des araignées d'eau, parcourant gracieusement la surface de l'eau, légers et clairvoyants comme l'air, n'entrant jamais en contact avec les turbulences du dessous”* (Keynes, 1938, p.450).

Lorsque Keynes s'attache à comprendre le fonctionnement de l'économie de marché, il est conduit à étoffer sa conception du comportement individuel. En effet, il perçoit que le jugement pratique de l'individu est plus complexe dans un environnement d'interdépendance. Pour se repérer, il faut aux individus plus qu'une rationalité commune, qui n'assure pas à elle seule l'homogénéité de leurs anticipations. Il leur faut des règles : *“habilement inculquées et astucieusement préservées”* (Keynes, 1938, p.447).

3- Raison et convention dans la Théorie Générale

“La théorie orthodoxe nous prête par hypothèse une connaissance du futur tout-à-fait différente de celle que nous avons en réalité. Cette rationalisation fallacieuse reprend les grandes lignes du calcul benthamien. L’hypothèse d’un futur calculable conduit à une interprétation erronée des principes de comportement que la nécessité d’agir nous contraint d’adopter.” (Keynes, 1937, p. 122)

Après avoir étudié quel était, selon Keynes, le fondement d’un comportement raisonnable, nous cherchons maintenant à montrer ce qui dans le système économique de marché “bloque”, selon lui, la possibilité d’agir raisonnablement¹⁶.

La démonstration de Keynes est bien connue, elle a fait l’objet de commentaires précis et complets (Orléan [1986, 1988, 1989], Favereau [1985, 1986, 1988, 1989]). Elle est principalement exposée dans le “livre quatre” de la *Théorie Générale*, et dans un article postérieur (Keynes, 1937) qui recentre le message de la *Théorie Générale* autour de la notion d’incertitude. Le principe démonstratif utilisé par Keynes consiste à montrer que les déterminants du prix des biens capitaux¹⁷ sont conventionnels, et ne sont donc pas fondés en “raison”. Comme le prix des biens capitaux pèse sur la décision d’investir, l’investissement fluctue lui aussi au gré de conventions arbitraires, dont rien n’assure qu’elles s’établissent durablement à un niveau compatible avec le plein emploi.

Nous débuterons notre présentation par un rappel des mécanismes de détermination de l’incitation à investir. Nous montrerons alors que ces déterminants sont conventionnels, puis nous chercherons à montrer que, dans la convention, la rationalité, au sens standard, divorce d’avec la raison, au sens keynésien. Ainsi, une des causes, selon Keynes, de l’instabilité du niveau de l’investissement réside dans le fait que les conventions qui contribuent à le déterminer reposent sur un très petit nombre d’éléments de connaissance “réelle” ou “objective”.

¹⁶ Dans ce chapitre, il est important de conserver à l’esprit que par “raison individuelle” nous désignons très précisément la théorie keynésienne de la rationalité. C’est uniquement à l’aune de cette théorie que l’on peut comprendre pourquoi, selon Keynes, le mécanisme de la convention fait problème.

¹⁷ Le concept de bien capital recouvre dans la théorie keynésienne deux réalités différentes. Il désigne d’une part l’investissement physique réel acheté ou réalisé par l’entrepreneur. Il est alors synonyme de “bien de production”. Mais il désigne aussi la notion d’action, achetée par les investisseurs boursiers opérant sur marché financier. Keynes place en quelque sorte le concept de bien capital à l’interface du monde de l’entrepreneur qui réalise des investissements réels, et du spéculateur qui réalise des investissements financiers. Pour que la distinction entre ces deux conceptions alternatives des biens capitaux soit tout à fait claire, nous utiliserons souvent dans la suite du texte les notions d’*action* et de *bien de production* en lieu et place du terme keynésien de *bien capital*.

31- Les déterminants du prix des biens capitaux.

“Il n'est nullement surprenant qu'ainsi déterminé, le volume de l'investissement soit amené à fluctuer fortement de temps en temps. Il dépend en effet de deux types de jugements sur l'avenir qui s'expriment respectivement à travers la propension à thésauriser et à travers l'estimation des rendements futurs des biens capitaux, deux jugements dont les fondements ne sont ni pertinents, ni solides.” (Keynes, 1937, p.118)

Le livre quatre de la *Théorie Générale*, ainsi que le célèbre article de 1937, dans lequel Keynes présente une version plus concise de ses arguments, permettent de distinguer clairement les deux “canaux” par lesquels un jugement collectif sur l'avenir se répercute sur le niveau de l'investissement, et par là même sur le niveau de l'emploi. Ces deux canaux sont le taux d'intérêt et l'évaluation à long terme du rendement escompté des actions. Or, le financement de l'investissement dépend d'une part du niveau du taux d'intérêt, d'autre part du prix des actions. Il est donc tributaire de l'incertitude des acteurs face à l'avenir.

311- Le financement de l'investissement

“Il est possible en général de produire de nouveaux biens capitaux. L'échelle suivant laquelle ils sont produits est bien sûr déterminée par le rapport entre leur coût de production et le prix que l'on espère leur voir atteindre sur le marché. Par conséquent si la conjonction du niveau des taux d'intérêt et de l'appréciation des rendements futurs fait s'élever le prix des biens capitaux, le volume de l'investissement augmentera, et, à l'inverse, si les déterminants du prix des biens capitaux l'influencent à la baisse, le volume de l'investissement diminuera.” (Keynes, 1937, p.118).

La décision d'investir est suspendue à la réalisation de deux conditions : il faut que l'investissement soit jugé utile en terme d'efficacité productive, et il faut également que le coût de financement de cet investissement n'excède pas les ressources qu'il devrait normalement permettre d'obtenir.

La première condition est déterminée par un calcul réalisé par l'*entrepreneur*. Il évalue l'opportunité d'un investissement nouveau au regard de ce que cet investissement représente comme revenus futurs, tirés de la vente de la production supplémentaire qu'il

permettra de réaliser. C'est ce que Keynes désigne sous l'expression de *rendement escompté de l'investissement*. On se situe ici du côté de la production réelle, et le calcul de l'entrepreneur repose sur une évaluation des "débouchés" possibles des produits que l'investissement permettra de produire. Techniquement, Keynes désigne sous le concept de *l'efficacité marginale du capital*, le rapport entre le rendement escompté du capital et son coût de production.

La seconde condition repose sur des données d'ordre différent. Le financement de l'investissement dépend d'une part du taux d'intérêt en vigueur. L'efficacité marginale du capital doit en effet être au moins égale au taux d'intérêt pour que l'investissement soit rentable. Ainsi, au niveau global, le volume de l'investissement ne peut jamais excéder le niveau à partir duquel l'efficacité marginale du capital de n'importe quel bien capital devient inférieur au taux d'intérêt¹⁸. Mais la décision d'investir dépend également du "prix des biens capitaux" qui est aussi le "prix des actions". Keynes établit en effet un rapport direct entre l'augmentation des capacités productives d'une entreprise, et l'augmentation de son capital boursier. Chaque achat d'un bien de production est analysé comme étant en quelque sorte synonyme de création d'une nouvelle action. Le caractère ambivalent du concept de "bien capital", qui désigne alternativement l'action et le bien de production s'explique par cette forme d'équivalence que Keynes établit entre l'univers de la production et celui de la finance. De ce fait, lorsque le prix des actions est élevé, l'investissement est favorisé. L'entrepreneur-manager parvient plus aisément à convaincre les actionnaires de financer un nouvel investissement.

Or, le prix des actions est déterminé d'une part par leur "rendement escompté"¹⁹, évalué sur le marché financier selon un mécanisme décrit dans le chapitre douze de la *Théorie Générale* ; d'autre part par le niveau du taux d'intérêt. En effet, lorsqu'un investisseur institutionnel a une somme d'argent à placer, il peut choisir de la placer à taux fixe, et toucher ainsi une rémunération garantie dès l'origine de son placement, ou la placer à taux variable, sous forme d'une action dont le rendement n'est pas garanti. De ce fait, le prix des actions évolue jusqu'au niveau pour lequel elles offrent un rendement escompté (qui inclut donc le risque) égal au rendement d'un placement à taux fixe. Lorsque le taux d'intérêt est très élevé, pour que le placement à taux variable demeure intéressant, il est nécessaire que le prix des actions soit très bas. Le prix des actions

¹⁸ Keynes (1936, p.153) définit l'efficacité marginale du capital comme le rapport du rendement escompté sur le coût de production (ou "prix d'offre") d'un bien capital. Il construit une courbe de demande de biens capitaux en reliant l'efficacité marginale globale du capital avec le volume d'investissement. Le volume d'investissement est déterminé par l'intersection de cette courbe avec la courbe représentant le niveau du taux d'intérêt.

¹⁹ Comme pour la notion de bien capital, Keynes utilise l'expression de "rendement escompté" pour désigner alternativement le calcul permettant de déterminer les rendements futurs d'un investissement réel et celui des actions. Or, il faut bien avoir à l'esprit que ces deux formes de rendement escompté ne reposent pas du tout sur les mêmes bases (comme nous le verrons plus loin).

dépend donc d'une part de leur "valeur intrinsèque" (telle que l'évalue le marché financier) et d'autre part du niveau du taux d'intérêt.

En résumé, la décision d'investir dépend de deux types de jugement sur l'avenir de nature radicalement différente. Le jugement de l'entrepreneur est le premier, et en ce sens le plus déterminant. Il porte sur des variables réelles, et consiste surtout à percevoir les débouchés possibles d'une production nouvelle. Le jugement de l'investisseur financier ne repose pas sur les mêmes critères. Il dépend d'une part de l'évolution du prix des actions, d'autre part du niveau du taux d'intérêt²⁰. Or ces deux éléments sont fixés de manière conventionnelle. C'est là le cœur d'une opposition entre le jugement de l'entrepreneur, qui s'appuie sur des données réelles, et le jugement du spéculateur financier qui repose sur un mécanisme conventionnel quasi-indépendant de la "réalité économique"²¹.

312- La détermination du taux d'intérêt

Keynes propose une définition radicalement nouvelle du taux d'intérêt : "*le taux d'intérêt est le prix qui équilibre le désir de détenir de la richesse sous forme de monnaie et la quantité de monnaie disponible*" (Keynes, 1936, p.181). Le taux d'intérêt est donc en quelque sorte le prix de l'argent, et non pas le prix du capital, comme c'est le cas dans la théorie classique. Le taux d'intérêt constitue selon Keynes une "prime" que les individus exigent pour se séparer de leur argent liquide, ce qui leur permet de se protéger contre l'avenir. En effet, les agents économiques possèdent une certaine quantité de monnaie qu'ils n'utilisent pas pour motif de consommation. Cette monnaie épargnée peut, *a priori*, prendre deux formes différentes. Les agents peuvent conserver leur monnaie sous forme liquide, ou choisir de la transformer en placement financier. Lorsqu'ils choisissent le second terme de l'alternative, ils touchent une rémunération qui correspond au taux d'intérêt²². La rémunération qu'ils exigent, en moyenne, dépend étroitement des raisons pour lesquelles les agents souhaitent *a priori* conserver leur richesse sous forme

²⁰ Selon Keynes, le taux d'intérêt intervient donc à deux niveaux en ce qui concerne la détermination de l'investissement. Il intervient directement pour décider de l'alternative placement financier/investissement réel. Il intervient à un second niveau pour déterminer l'alternative action/obligation. Cette seconde alternative intervient dans la détermination du prix des actions, et, par ce biais, dans la détermination du niveau de l'investissement.

²¹ Le terme de "réalité économique" repose sur la distinction investissement réel/placement financier. Il est problématique (puisque la finance est une des réalités économiques). Mais cette difficulté est liée à la conception keynésienne et à son "réalisme platonicien" que nous étudierons plus loin. Nous adoptons donc, temporairement, son point de vue opposant la "réalité" des entrepreneurs à la "virtualité" du monde financier.

²² Quelque soit leur forme (action ou obligation), les placements financiers offrent à l'investisseur marginal, compte tenu des doutes et risques liés au placement sous forme d'action, un rendement à peu près équivalent au taux d'intérêt.

de monnaie.

Keynes (1936, chap.15) énumère les différents motifs de cette “propension à thésauriser” : motif de transaction, motif de précaution, motif de spéculation. Les deux premiers motifs sont relativement stables et liés à l’activité économique réelle. Si ces deux motifs déterminaient seuls le niveau du taux d’intérêt, celui-ci serait indexé sur l’activité économique réelle, montant lors des phases de croissance, et redescendant lors des crises, comme le prévoit l’économie classique. C’est la raison pour laquelle le troisième motif revêt une importance particulière. Il est plus fluctuant, et explique donc davantage les mouvements du taux d’intérêt. Surtout, il n’est pas directement déterminé par l’activité économique réelle mais par les anticipations des agents. Au moment de placer leur monnaie, les agents veulent en effet s’assurer que le taux d’intérêt ne va pas brutalement augmenter, provoquant ainsi une dévalorisation de leur capital. Ils doivent donc anticiper l’avenir.

Or, les agents ne disposent pas d’une série de probabilités associées à une série complète d’états du monde possibles. Ils savent qu’ils ne sont pas suffisamment informés des mécanismes et de la situation économique pour pouvoir décider “en toute connaissance de cause”. Dans les termes du *Treatise on Probability*, on peut dire que leur déduction logique est fondée sur un jugement de probabilité concernant l’avenir qui a un “poids” très faible. De ce fait : “*Les obscures terreurs et les espoirs, tout aussi confus et déraisonnables, ne s’évanouissent jamais tout-à-fait, ils sont, en quelque sorte, toujours tapis sous la surface*” (Keynes, 1937, p.115). Les individus sont donc incertains : ils n’en savent pas assez pour décider raisonnablement. Dans ces conditions, conserver l’argent liquide, c’est-à-dire thésauriser, “*apaise leur angoisse*”. C’est précisément cette “inquiétude” qui détermine le niveau du taux d’intérêt : “*la prime que nous exigeons pour nous séparer [de l’argent liquide] mesure notre degré d’inquiétude*” (*ibid*, p.116).

Ainsi, le taux d’intérêt est un “*baromètre*” des angoisses collectives. Lorsque les agents sont fortement inquiets, leur propension à thésauriser augmente. Si la masse monétaire est fixe, le prix qui équilibre la demande de monnaie pour motif de thésaurisation par rapport à la quantité de monnaie disponible, augmente nécessairement. Ce prix, c’est le taux d’intérêt. La propension à thésauriser joue donc sur le niveau de l’investissement non pas parce qu’elle représente une “fuite” de monnaie qui resterait en dehors du système, mais parce qu’elle influe sur le niveau du taux d’intérêt. Lorsque la propension à thésauriser augmente, la quantité de monnaie effectivement thésaurisée ne varie pas, mais le taux d’intérêt est influencé à la hausse²³. Les variations du taux d’intérêt dépendent donc de manière cruciale des anticipations des agents, et de leur confiance en la valeur de leurs anticipations.

²³ Sauf si le système bancaire intervient et fait augmenter la quantité de monnaie disponible.

313- La détermination du rendement escompté des actions

L'autre canal qui relie la perception de l'avenir à la détermination de l'investissement est l'évaluation de long terme de la rentabilité escomptée des actions, que Keynes décrit dans le chapitre douze de la *Théorie Générale*. Cette rentabilité escomptée est déterminée par les marchés financiers et se traduit par les modifications du cours des actions. Ainsi, au moment d'acquérir des biens capitaux sous forme d'action, les individus sont une fois encore amenés à se prononcer sur l'avenir. Comme dans le cas précédent, Keynes note : "le fait marquant en la matière est l'extrême précarité des bases sur lesquelles nous sommes obligés de former nos évaluations des rendements escomptés" (Keynes, 1936, p.165). C'est donc l'incertitude qui prédomine. Les agents qui opèrent sur le marché boursier ne sont pas des entrepreneurs. Ce sont "des personnes qui ne gèrent pas elles mêmes les affaires qui leur appartiennent et qui n'ont aucune connaissance particulière des circonstances actuelles ou éventuelles qui influent sur elles." (Keynes, 1936, p.169). La séparation entre le management de l'entreprise et sa direction provoque une dégradation des fondements concrets, pratiques, du jugement des individus qui possèdent les actions. Ils doivent évaluer les rendements futurs des biens capitaux sans disposer des connaissances pertinentes qui leur permettraient d'asseoir leur jugement. Ils ne peuvent donc pas être "raisonnables" au sens du *Treatise*. Il leur manque pour cela les facultés pratiques d'observation et de compréhension de la situation économique "réelle" qui leur seraient nécessaires pour juger de la vigueur et des possibilités de développement futur de l'entreprise. De ce fait, leur jugement est profondément influençable. Il n'est pas solidement accroché à des connaissances concrètes certaines. C'est là l'origine du développement de comportements mimétiques : les agents cherchent à travers l'étude du comportement des autres des éléments de connaissance qu'ils n'ont pas.

Dans l'économie keynésienne, le jugement que les agents portent sur l'avenir est donc déterminant. Il détermine le niveau du taux d'intérêt et le rendement escompté des actions, et, par ce biais, influe sur le niveau de l'investissement. En cela, la *Théorie Générale* est dans la lignée du *Treatise on Probability*. Cependant, la situation des individus sur les marchés financiers est marquée par deux éléments qui rendent quasi-impossible la prise de décision raisonnable, telle que Keynes la décrit. Le premier est leur interdépendance, le second est leur faible niveau d'information. Comme ils sont interdépendants, Les individus doivent s'assurer non seulement que leur anticipation est correcte compte tenu de la situation "réelle" de l'économie, mais encore qu'elle correspond bien à celle que font des autres agents. C'est en effet l'anticipation *collective* qui détermine le niveau

effectivement atteint par le taux d'intérêt et le rendement escompté des actions. Chacun cherche donc à savoir ce que pensent les autres. Or, les agents qui détiennent des richesses à placer ont peu de connaissance réelle de l'économie. Aucun élément de connaissance réelle ne vient donc stopper l'observation mutuelle. Le processus d'interrogation spéculaire se déroule donc sur fond d'ignorance généralisée. C'est la conjonction de ces deux éléments qui donne naissance à la "convention".

32- Un jugement conventionnel

"Conscient du peu de valeur de son propre jugement, chacun veille à l'aligner sur le jugement des autres, considérés comme étant sans doute mieux informés. Cela signifie que nous cherchons à nous conformer à l'attitude de la majorité ou de la moyenne. La psychologie d'une société faite d'individus qui, tous, cherchent mutuellement à s'imiter nous conduit à ce qu'il convient d'appeler très précisément un jugement de convention" (Keynes, 1937, p.114)

L'agent qui doit décider, tout d'abord s'il doit conserver sa monnaie ou la placer, puis ensuite s'il doit la placer à taux fixe ou à taux variable, n'a pas de connaissance particulière de l'économie. Keynes le décrit en permanence comme un individu "ignorant". Il ne peut asseoir ses conjectures que sur une faible quantité d'information pertinente. Le jugement qu'il forme seul a donc un faible poids. De plus, la situation qu'il cherche à prévoir est marquée par l'interdépendance de tous les individus. La valeur future du taux d'intérêt, ou le rendement escompté d'une action, dépend en effet des anticipations de l'ensemble des agents. Ayant peu d'information, un agent "ignorant" peut rapidement se convaincre que les autres agents formeront sans doute leurs anticipations sur une base cognitive différente. Dès lors le peu d'information qu'il possède devient inutilisable : l'individu est projeté dans l'ignorance totale. Dans une telle situation, il ne lui reste plus qu'à aligner son jugement sur celui des autres. Or, la plupart des autres sont également "ignorants", et font un raisonnement identique. La conséquence de cette incertitude généralisée est l'émergence d'une logique "mimétique" sur le marché : tous les individus s'épient et cherchent à deviner ce que la majorité d'entre eux va faire. En quelque sorte, l'absence d'une base cognitive commune, provoque un recentrage du marché sur lui-même. Il est alors régi par une logique "auto-centrée", selon l'expression d'Orléan [1988], qui le déconnecte de toute réalité extérieure.

Keynes identifie très clairement le mécanisme par lequel les individus s'extirpent collectivement de cette indétermination : *“En pratique nous sommes tacitement convenus, en règle générale, d'avoir recours à une méthode qui repose à vrai dire sur une pure convention. Cette convention consiste essentiellement – encore bien sûr qu'elle ne joue pas toujours sous une forme aussi simple – dans l'hypothèse que l'état actuel des affaires continuera indéfiniment à moins que l'on ait des raisons définies d'attendre un changement”* (Keynes, 1936, p.167).

Lors de cette première mention de la notion de convention, Keynes rassemble deux idées. La première consiste à souligner que le comportement des individus est conventionnel, en ce qu'il cherche à s'aligner sur celui de la majorité. La seconde consiste à préciser un premier type particulier de convention, qui vise à surévaluer l'importance du présent par rapport au futur. Keynes précise en de multiples endroits que cette convention tient contre l'évidence puisque, comme on le sait, *“le futur ne ressemble jamais au passé”* (Keynes, 1937b, p.124). Cependant cette attitude n'est pas absurde : *“il est assez raisonnable de se laisser guider dans une large mesure par les faits dont on se sent certain”* (Keynes, 1936, p.164, c'est moi qui souligne). En effet, on sait que Keynes est convaincu de la pertinence du raisonnement inductif qui, précisément, repose sur une hypothèse de ce type. Mais, l'induction n'est raisonnable que si l'individu est assez vigilant pour remettre en cause la convention lorsque des faits pertinents le nécessitent. Le caractère “raisonnable” de la convention repose donc sur l'évaluation, par chacun, de ce qu'est une “raison d'attendre un changement”. Tout repose sur la manière dont les agents révisent leur connaissance, ou parviennent, de manière satisfaisante, à identifier les faits pertinents indiquant des changements en cours ou à venir. La convention de départ est donc une position d'attente, un point de départ, un “*a priori*”. Ce qu'il importe de comprendre est la manière dont les individus révisent cette convention.

C'est à ce stade qu'il faut distinguer la convention qui fonde le taux d'intérêt de la convention qui fonde le rendement escompté des biens capitaux. La convention qui fonde le taux d'intérêt est stable, parce que, dans la plupart des cas, *“l'opinion dominante croit que le taux d'intérêt s'ajuste automatiquement, de sorte que le niveau établi par convention est considéré comme ancré en des fondements objectifs beaucoup plus résistants qu'une convention”*. Les individus ne perçoivent pas que le taux d'intérêt est une convention. Ils s'illusionnent quant à ses fondements objectifs. Ainsi, comme le note Orléan (1989, p.265), *“la communauté se dote d'une référence commune, produit une représentation collective extériorisée qui fonde les anticipations individuelles”*. La communauté “oublie” que cette représentation commune est un produit de la logique “mimétique”. Elle lui semble être un point de repère extérieur digne de confiance. Par un

étrange jeu de miroir, elle se retrouve face à elle-même, mais croit reconnaître, dans ce qui n'est que son reflet, un signe extérieur digne de confiance. Par ailleurs, le taux d'intérêt est fonction de l'inquiétude des agents. De ce fait, lors des crises, il n'aura pas tendance à diminuer.

Le mécanisme de révision des croyances est différent en ce qui concerne la convention qui gouverne le rendement escompté des actions. Dans ce cas, au contraire, la convention est trop souvent mise en question et dénoncée. En effet, sur le marché financier, les agents peuvent en permanence réviser leur jugement. Or, comme les individus n'ont aucune confiance dans leur évaluation à long terme, ils sont amenés à chercher en permanence à s'assurer de sa validité. Ils sont toujours prêts à céder leurs actions et à modifier leur portefeuille. Les investissements réels, à long terme, dont décident les entrepreneurs, deviennent des actions échangeables à court terme pour les spéculateurs. Le long terme devient donc une succession de courts termes : *“les investissements qui sont fixes pour la communauté sont rendus liquides pour l'individu”* (Keynes, 1936, p.168). Ces révisions successives rendent la convention tout-à-fait instable. Ainsi, à la différence de la convention qui gouverne le taux d'intérêt, celle qui émerge des marchés financiers est “précaire” : *“une évaluation conventionnelle, fruit de la psychologie d'un grand nombre d'individus ignorants est exposée à subir des variations violentes à la suite des revirements soudains que suscitent dans l'opinion certains facteurs dont l'influence sur le rendement escompté est en réalité assez petite.”* (Keynes, 1936, p.168). Cette fois, donc, la fragilité des fondements de la convention est bien perçue par les agents, et ceci les conduit à n'accorder qu'une confiance minimale en son maintien. Autrement dit, toute nouveauté est interprétée comme un fait pertinent susceptible de modifier la conception “moyenne” de l'avenir²⁴.

Keynes indique un contre-poids possible à cette instabilité chronique : *“on pourrait supposer que la concurrence entre les professionnels compétents, doués d'un jugement plus sûr et de connaissances plus étendues que la moyenne des investisseurs privés, corrigerait les fantaisies des individus ignorants livrés à leurs propres lumières”* (Keynes, 1936, p.169,170). Les détenteurs de richesse n'interviennent en effet pas tous directement sur le marché financier. Ils font appel à des intermédiaires financiers qui sont rémunérés pour gérer au mieux les placements de leurs clients. On pourrait donc attendre de ces professionnels, qui connaissent bien l'économie et la situation économique “réelle”, qu'ils raisonnent à long terme et ne cèdent pas aux “excès d'humeur” du marché.

²⁴ Orléan [1989] prend comme exemple les déclarations du président Reagan. La pertinence du jugement économique présidentiel était considérée comme douteuse par la plupart des opérateurs, mais tout le monde pensait que les autres allaient se laisser influencer par ses déclarations. C'est pourquoi tout le monde prenait ses déclarations très au sérieux.

Mais ce n'est pas le cas. En effet, les professionnels des marchés agissent en spéculateurs. Ainsi, ils ne cherchent pas à servir de guide aux individus ignorants, parce qu'il leur est plus facile de chercher à s'enrichir en "*volant le départ*" c'est-à-dire en prévoyant, avant les autres, les retournements de la convention. Le jeu de la spéculation ne consiste donc pas à prévoir la "vraie" valeur des biens capitaux, mais à prévoir comment les "individus de base" vont l'évaluer, et, surtout, à repérer le moment où cette évaluation est susceptible d'être modifiée²⁵. L'action des spéculateurs fragilise donc, plutôt qu'elle ne stabilise, l'évaluation conventionnelle, en accentuant les éventuelles tendances au changement. En définitive, c'est donc le jugement le moins raisonnable, c'est à dire celui qui repose sur les bases cognitives les plus étroites, qui donne le "la". Le marché est guidé par les individus les plus myopes. La convention évolue au gré des changements d'attitude de ces individus ignorants, qui, sachant que leur jugement repose sur un faible "poids", sont enclins à changer fréquemment d'avis.

L'investissement est donc déterminé par deux facteurs, qui sont tous deux *conventionnels*, mais qui n'évoluent pas de la même façon. C'est là l'origine du problème que pointe Keynes : "*le lecteur devrait maintenant voir clairement les difficultés qui s'opposent au maintien d'une demande effective suffisante pour assurer le plein emploi du fait qu'un taux d'intérêt à long terme conventionnel et assez stable se trouve associé à une efficacité marginale du capital capricieuse et fort instable*" (Keynes, 1936, p.214). En effet, le taux d'intérêt est stable et tend à être plutôt surévalué pour calmer les angoisses des individus face au futur. A l'inverse, l'évaluation du rendement escompté des actions est instable, et ne peut se maintenir à un niveau suffisamment haut pour motiver les investisseurs. En situation d'incertitude forte, lors d'une crise, le taux d'intérêt est donc élevé et le rendement escompté est plus bas qu'il ne devrait être. Pour que l'achat des actions soit tout de même intéressant, leur prix doit donc nécessairement être très faible. Les deux déterminants du prix des actions évoluent en effet dans le même sens. Leurs effets se cumulent, ils ne se compensent pas : "*les conditions défavorables à l'un sont en général défavorable à l'autre*" (Keynes, 1937, p.118). De ce fait, lorsque l'incertitude est forte, comme c'est le cas lors des crises, le prix des actions évolue à la baisse, et cela tend à dissuader les actionnaires de financer un nouvel investissement réel. L'investissement risque donc de rester durablement à un niveau inférieur à ce qui serait nécessaire pour assurer le plein emploi.

²⁵ Lorsque l'évaluation est très fluctuante, il y a plus à gagner en prévoyant chacun des retournements qu'en cherchant à déterminer une valeur moyenne.

33- Le jugement "véritable" "de l'entrepreneur"

"Certaines catégories d'investissement sont gouvernées moins par les prévisions véritables des entrepreneurs que par la prévision moyenne des personnes qui opèrent sur le stock exchange.(...) Dans l'évaluation des investissements,(...) l'élément de connaissance réelle a fortement diminué. (...) Le placement fondé sur une véritable prévision à long terme est de nos jours une tâche difficile au point de n'être guère possible" (Keynes, 1936, p. 167-170, c'est moi qui souligne).

L'analyse que propose Keynes des mécanismes complexes déterminant l'incitation à investir met en lumière un problème lié au comportement des agents et à l'organisation du système économique. Keynes montre en effet que les anticipations conventionnelles fonctionnent par une négation des "faits réels". Le comportement spéculatif des agents n'est donc pas raisonnable au sens de Keynes. Le problème est que l'organisation du système économique impose ce jugement spéculatif comme une norme, s'imposant même au jugement "véritable" des entrepreneurs. Il serait donc nécessaire de mettre en place d'autres règles pour replacer le jugement de l'entrepreneur au centre de l'économie.

331- Raison keynésienne versus rationalité néoclassique

Sur le marché financier, les individus n'agissent pas en cherchant à se forger une idée des chances de réussite de tel ou tel bien capital ou des perspectives d'évolution de la masse monétaire, mais en cherchant plutôt à imiter les autres. Ce comportement est rationnel. Il est en effet parfaitement cohérent compte tenu de l'univers dans lequel les individus se situent. Mais il n'est pas raisonnable : il lui manque en effet les fondements pratiques qui déterminent un comportement non seulement rationnel, mais également raisonnable.

Sur les marchés financiers, les individus n'ont pas d'autre choix que de suivre le jugement de la foule. Comme le note Orléan (1986, p.11), *"on ne peut pas avoir raison contre la foule"*. En situation d'interdépendance, il est en effet nécessaire d'accorder son jugement avec celui des autres. La valeur qu'atteindra une action dépend *in fine* uniquement de ce que l'ensemble du marché anticipera. Il est donc parfaitement rationnel d'aligner ses anticipations sur celles des autres. Keynes décrit même un phénomène typique du principe des anticipations rationnelles : lorsque la foule prévoit que le taux d'intérêt va augmenter, il augmente effectivement, lorsque la foule prévoit que le rendement escompté d'un actif va diminuer, il diminue effectivement. Le caractère

autoréalisateur des prophéties du marché assure la rationalité des anticipations conventionnelles. Du point de vue néoclassique, le discours de l'économiste ne peut pas aller plus loin. La rationalité standard est définie de telle manière qu'il n'est absolument pas possible de soulever le problème du manque de fondement informationnel pertinent sur lequel s'appuie le jugement conventionnel. C'est en effet un problème qui se situe hors de la logique formelle, or la logique formelle détermine seule la rationalité néoclassique.

En revanche, il est aisé de souligner que ce comportement n'est pas raisonnable, au sens du *Treatise on Probability*. On a vu en effet que Keynes mettait précisément l'accent sur le caractère essentiel de la dimension pratique du jugement dans l'incertain. Sur cette base, on ne peut que reconnaître qu'il n'est pas raisonnable de changer d'opinion concernant un investissement réel en raison de données n'ayant qu'un rapport lointain avec cet investissement réel. Les jugements que les individus portent sur le rendement des actions ou sur le taux d'intérêt ne sont "*ni pertinents, ni solides*" (Keynes, 1937, p.118). "*Les jugements manquent en effet de racines profondes qui leur permettraient de résister solidement*" (Keynes, 1936, p.168). Des individus peu informés, qui de surcroît abandonnent les quelques d'informations dont ils disposent pour mieux s'aligner sur le choix des autres, ne peuvent être "raisonnables". La logique pratique qui fonde la raison keynésienne ne peut pas être associée au comportement spéculatif.

Un exemple permettra peut-être d'éclairer cette distinction entre un comportement rationnel et un comportement raisonnable. Lorsqu'un feu se déclare dans un lieu public, on sait d'expérience que la cohue et la panique font généralement plus de victimes que le feu lui-même. Chacun, individuellement, sait sans doute cela. Chacun sait aussi que la panique, la "dé-raison", pousse "les autres" à se précipiter vers l'extérieur, de manière désordonnée, et qu'à compter du moment où ce mouvement massif commence, il vaut mieux le suivre... En effet, si l'on attend seul, convaincu de la pertinence de notre point de vue, que la foule soit sortie dans une grande confusion pour nous échapper à notre tour, l'on risque fort d'être carbonisé. La panique de quelques-uns rend impraticable un comportement raisonnable. A compter du moment où tout le monde se précipite en masse pour sortir, on est contraint de faire de même. Chacun se précipite donc vers la sortie, tout en sachant qu'il serait préférable que chacun sorte dans le calme sans paniquer. Un tel comportement est rationnel, mais il n'est pas pour autant raisonnable au sens que Keynes donne à ce mot dans le *Treatise on Probability*.

Cette distinction entre un comportement rationnel et un comportement raisonnable est en rapport avec l'opposition entre la rationalité individuelle et la rationalité collective. On a observé que la rationalité individuelle, au sens standard, pouvait avoir des effets néfastes au niveau collectif. Les réflexions néoinstitutionnalistes menées en théorie des jeux,

permettent d'établir de nombreuses situations paradoxales dans lesquelles un comportement individuellement rationnel mène à un état collectif sous-optimal. Cette opposition n'a pas de sens si on mobilise le concept keynésien de "raison". La raison repose en effet sur une intuition permettant de concevoir la "vérité", même de manière partielle. Or, un comportement éclairé par la vision de la vérité est aussi un comportement moralement bon. Dans l'intuitionnisme de Moore, qui constitue l'assise philosophique de la théorie keynésienne, le vrai et le bon sont synonymes. Ainsi, pour Keynes, la possibilité d'un équilibre de sous-emploi est directement liée à l'absence, à un endroit crucial du système économique, de comportement raisonnable. Le dilemme entre rationalité individuelle et collective découle du divorce entre la rationalité individuelle et la raison, divorce que Keynes repère sur les marchés financiers²⁶.

332- *Le triomphe de la spéculation*

L'opposition entre rationalité et raison trouve donc son point d'orgue avec la victoire, quasi-systématique, de la spéculation sur l'esprit d'entreprise. Or, les causes de ce phénomène sont principalement liées à l'*organisation* des marchés. L'enchaînement des actions des intervenants sur ces marchés est tel qu'il favorise inévitablement le raisonnement de court terme, spéculatif, et rend sans objet la réflexion de l'entrepreneur qui raisonne sur le long terme.

Le "triomphe" de la spéculation sur l'entreprise n'est donc pas lié à l'absence d'individus raisonnables, mais au fonctionnement du marché qui décourage les comportements raisonnables. Keynes indique ainsi qu'il existe "*des esprits sérieux (...) indifférents au passe-temps prédominant, qui persistent à acheter des actifs à la lumière des prévisions véritables à long terme les plus parfaites qu'ils puissent établir.*" Il précise plus loin : "*On a raison de dire que l'entreprise animée par des espoirs s'étendant sur un long avenir bénéficie à la communauté toute entière.*" (Keynes, 1936, p.171 et 176). Mais ce point de vue raisonnable, et donc collectivement bénéfique, ne peut émerger. Les individus sont pour la plupart contraints de faire des bénéfices rapides. De plus, "*il faut plus d'intelligence pour triompher des forces secrètes du temps et de l'ignorance de l'avenir que pour "voler le départ"*" (Keynes, 1936, p.172). La spéculation permet des

²⁶ Cooper et John [1988] ont formalisé le principe keynésien d'un équilibre de sous-emploi sous la forme d'un jeu comprenant deux équilibres de Nash possibles, dont un seul est pareto-optimal. Une fois atteint, l'équilibre sous-optimal est néanmoins stable. Ils identifient ainsi l'analyse keynésienne comme la mise en évidence d'un "défaut de coordination". Il est intéressant de noter (et cela ne remet pas en cause cette formalisation très "parlante") que, du point de vue de ce que Keynes entend par "raison", il ne saurait y avoir d'équilibre sous-optimal atteint du seul fait des agents. S'ils en avaient la possibilité, ils choisiraient forcément l'issue collectivement meilleure, sans même avoir à s'accorder. C'est l'organisation du marché qui les contraint à ne pas agir raisonnablement.

gains plus fréquents (à chaque retournement de la “base conventionnelle d'évaluation) et plus aisés que la prévision sérieuse. Dans ces conditions le raisonnement spéculatif devient dominant. Or, à compter du moment où un marché devient spéculatif, il n'est plus possible d'être raisonnable. Celui qui, à contre-courant de la foule, cherche à fonder sa décision sur une connaissance “réelle” et une prévision raisonnable à long terme, conformément à l'esprit du *Treatise*, ne peut qu'échouer. C'est là une conséquence du mécanisme de l'auto-réalisation, qui rend automatiquement rationnelle toute croyance conventionnelle. Le mécanisme de la convention valide l'attitude spéculative et met en échec le raisonnement raisonnable. La raison, au sens de Keynes, est mise en échec par le fonctionnement des marchés financiers.

Cet échec est directement lié à la séparation de la propriété et de la gestion de l'entreprise. Le marché financier constitue en effet une réponse intéressante, du point de vue individuel à la précarité des fondements cognitifs, sur la base desquels les individus doivent se forger une représentation du futur. Puisque les individus savent que leurs prévisions ont un faible “poids”, selon le concept que Keynes propose dans le *Treatise*, ils manquent de confiance en leurs anticipations, et ne souhaitent pas s'engager pour le long terme. Le marché financier leur permet justement de transformer leurs engagements de long terme, à travers l'achat de biens capitaux, en engagements de court terme, puisque les biens capitaux prennent la forme d'actions constamment revendables. Ce faisant, une disjonction apparaît entre ce qui devrait être au contraire deux faces d'un même objet : l'investissement réel continue de se dérouler dans le long terme, mais la valeur monétaire des actions qui lui correspondent peut varier à court terme pour des raisons indépendantes de la valeur intrinsèque de l'investissement. Le regard de l'entrepreneur et celui du financier divergent. Or, la décision revient, en dernier ressort, à ceux qui détiennent les richesses, plutôt qu'à ceux qui cherchent à en produire. La logique de court terme des actionnaires s'impose à la logique de long terme de l'entrepreneur-manager.

333- *L'intuition raisonnable de l'entrepreneur*

Ainsi, l'organisation du financement de l'investissement contraint et limite l'entrepreneur. Or, l'entrepreneur représente, pour Keynes, l'archétype de l'individu qui, bien qu'il ne soit pas calculateur, est “raisonnable”.

L'entrepreneur est en effet en prise avec l'économie concrète. Il tire sa capacité d'action d'une connaissance concrète d'un secteur particulier de l'économie qui lui permet d'avoir l'*intuition* de ce qu'il convient de faire. Sa connaissance n'est bien sûr pas parfaite. Il n'échappe pas à l'incertitude quant à l'avenir et ne maîtrise qu'une petite partie

du fonctionnement global du système économique. Il n'en sait pas assez pour "calculer" selon le modèle de l'individu maximisateur. Mais son intuition lui permet d'utiliser au mieux le peu d'information dont il dispose, en repérant les déductions logiques que l'on peut en tirer. Il "sent" les secteurs dans lesquels un investissement serait rentable, où existent des besoins insatisfaits. Son "*esprit animal*", à la recherche du profit et du changement, lui permet de déceler, à partir d'une connaissance partielle du réel, les contours de ce que pourrait être le monde futur. Son comportement face à l'incertitude est donc proche du comportement raisonnable tel que Keynes le définit dans le *Treatise*.

Cependant l'entrepreneur n'a pas les moyens d'imposer ses intuitions. Il sait lui-même que son intuition repose sur un faible "poids", puisque les informations dont il dispose sont peu nombreuses. C'est sans doute là l'origine de la scission de la propriété et du management. Ce système permet en effet à l'entrepreneur manager de ne pas supporter seul le risque d'erreur lié à son investissement. Cela permet surtout aux individus qui lui prêtent les ressources nécessaires à la réalisation de ses projets d'investissements, de se garantir contre le caractère incertain de l'analyse de l'entrepreneur. Le système s'organise de manière à éliminer l'incertitude liée à la décision d'investir. Mais par là même, il rend la décision d'investir de l'entrepreneur tributaire de la volonté des actionnaires. Le financement de l'investissement est donc soumis à l'évaluation qu'en font les acteurs des marchés financiers.

Pour l'entrepreneur-manager, les conventions sont donc un phénomène bien "réel" qui détermine le financement des investissements et contrecarre ses "plans". De plus, la logique conventionnelle est étrangère à l'attitude intuitive de l'entrepreneur. Les conventions ne sont pas régies par des préoccupations de long terme, elles ne s'appuient pas sur des données concrètes. De ce fait, les intuitions de l'entrepreneur sont comme "étouffées" par la logique de court terme des marchés financiers. C'est en particulier le cas lors des "crises". Une situation de sous-emploi constitue en effet sur le plan pratique une situation favorable pour l'entrepreneur qui souhaiterait développer sa production ; mais sur le plan des marchés financiers, cette situation incite au contraire à la prudence. L'analyse des mécanismes déterminant le niveau du taux d'intérêt et du prix des actions montre en effet que c'est précisément lorsqu'il faudrait que l'économie redémarre que les possesseurs de richesse sont le moins enclins à placer leur argent dans des projets d'investissement réels fondamentalement incertains.

Keynes déplore cet état de fait, et regrette la disjonction qui s'est opérée entre propriété et management. De nombreux passages de la *Théorie Générale* célèbrent le dynamisme pionnier des premiers entrepreneurs : "*Autrefois l'investissement dépendait d'un recrutement suffisant d'individus de tempérament sanguin et d'esprit constructif qui s'embarquaient dans les affaires sans chercher réellement à s'appuyer sur un calcul précis*

de profit escompté. (...) Si la nature humaine n'avait pas le goût du risque, les seuls investissements suscités par un calcul froidement établi ne prendraient sans doute pas une grande extension" (Keynes, 1936, p.165,166). A travers cette description enthousiaste de l'action intuitive de l'entrepreneur, plus ou moins inspirée selon les compétences spécifiques de chacun, Keynes renoue avec l'esprit du *Treatise*. L'entrepreneur ne calcule pas. Il agit sur la base d'une intuition fondée sur une connaissance "réelle". A rebours de ce que la Science Economique a finalement sanctifié comme "individu rationnel", c'est l'entrepreneur qui, pour Keynes, est vraiment "raisonnable" ! On le mesure encore mieux dans les termes quasi-mystiques employés par Keynes : "du point de vue de l'utilité sociale l'objet des placements éclairés devrait être de vaincre les forces obscures du temps et de percer le mystère qui entoure le futur" (Keynes, 1936, p.170)²⁷. Or, précisément, le système économique moderne ne permet plus au point de vue de l'entrepreneur de s'exprimer. C'est la spéculation qui l'emporte sur l'entreprise : "le risque d'une prédominance de la spéculation grandit à mesure que l'organisation des marchés financiers progresse" (Keynes, 1936, p.173). Ainsi, selon Keynes, l'organisation du système économique empêche l'émergence de la "raison" dans l'économie. Or, l'organisation, c'est un ensemble de règles.

Keynes décrit, dans la partie de son analyse consacrée aux déterminants de l'incitation à investir, un mécanisme organisationnel dont l'effet est de faire disparaître "l'élément de connaissance réelle". Son analyse du fonctionnement des conventions montre en effet une disjonction entre "rationalité" et "raison". Cette disjonction naît du fait que la décision d'investir est finalement suspendue à l'analyse d'une classe d'individus (les ménages qui possèdent des richesses et les spéculateurs) qui ne maîtrisent pas les éléments concrets qui permettraient de juger de la validité à long terme d'un investissement. Ainsi, les problèmes rencontrés par le système marchand pour assurer le plein emploi, sont directement liés au mode de raisonnement pratique que Keynes décrit dans le *Treatise*. L'organisation du financement de l'investissement échoue à maintenir un niveau suffisant d'investissement, précisément parce qu'elle n'est pas adaptée au mode de décision raisonnable. Cette organisation n'est pas satisfaisante parce qu'elle ne permet pas au jugement raisonnable de s'imposer. Ce diagnostic suggère à quel niveau devrait se situer le remède : il faut changer les règles pour replacer la "raison" au fondement de la décision d'investir.

²⁷ On peut noter la constance avec laquelle Keynes associe le caractère raisonnable de l'investissement de long terme fondé sur des données réelles, et son caractère moralement juste et collectivement souhaitable : ce qui est raisonnable est moralement bon.

4- Une théorie “élitiste” des règles

Keynes oppose nettement la détermination conventionnelle du taux d'intérêt et du prix des actions, et la décision intuitive mais raisonnable de l'entrepreneur. Cette opposition est au centre de son analyse de la faiblesse relative du niveau global d'investissement. Or, il s'agit d'un problème de *coordination*, lié à l'absence de *règles* efficaces permettant de financer de manière plus satisfaisante les projets des entrepreneurs.

En effet, la convention est par définition arbitraire. On peut opposer cet arbitraire à l'exigence de justice ou encore d'efficacité. C'est précisément ce que montre Keynes : l'arbitraire de la convention s'oppose à l'objectif du plein emploi. Contre cette conception arbitraire de la valeur des actifs, il faut selon lui rétablir le jugement raisonnable de l'entrepreneur. Il y a donc une opposition entre deux points de vue sur une même réalité. Cette opposition naît d'une différence de conception de l'avenir. Elle est liée à l'existence d'un temps historique laissant place à l'altérité. L'économie est donc conçue comme un jeu entre des acteurs qui ont des conceptions concurrentes de l'avenir, en raison de leur expérience, constitutive des fondements cognitifs de leurs décisions. Dans un tel espace il est nécessaire que des règles viennent stopper le conflit des interprétations, en permettant aux vues de certains acteurs de s'imposer à celles des autres. Ces règles doivent être choisies en fonction de certaines valeurs collectives permettant de déterminer un objectif collectif à atteindre. Keynes propose donc une conception *pragmatique* de l'économie qui met en évidence le rôle des règles.

Toutefois le pragmatisme de Keynes ne débouche pas sur une conception délibérative des règles. Keynes est en effet persuadé que seule la compréhension du *réel* permet de déterminer les bonnes règles. Il est donc attaché à une certaine forme de “réalisme” platonicien, qui prend la forme de l'intuitionnisme. De cette épistémologie “réaliste”, Keynes tire une philosophie politique marquée par l'élitisme intellectuel, à la manière de la *République* de Platon. C'est pourquoi, plutôt que de souligner le caractère contingent des règles d'organisation économiques, il souligne quelles règles particulières *devraient* être mises en place, et attend de l'Etat qu'il fasse office de régulateur.

Ainsi, bien que la théorie de Keynes soit anticonstructiviste et donne une place prépondérante au choix des règles, elle ne contient pas une analyse tout-à-fait convaincante de la dynamique des règles. Cette théorie se clôt en effet sur une vision “salvatrice” de l'Etat qui, sur les conseils du savant, peut “réguler” l'économie. Cette conception est doublement problématique. D'une part, Keynes n'indique pas clairement les raisons pour lesquelles le savant parvient à percevoir, dans un univers contingent, les règles qu'il faut mettre en place. D'autre part, il ne nous montre pas comment l'Etat peut

accéder à une place qui transcende suffisamment le jeu des acteurs pour qu'il lui soit possible d'imposer des règles. En définitive, il existe une forme de contradiction entre le "platonisme" de Keynes, et sa vision pragmatique de l'économie.

41- Une analyse pragmatique ?

411- Une vision pragmatique du système économique

On peut avec Berthoud [1998] reconnaître que Keynes, de par son opposition au constructivisme et à l'idée d'un système autonome et autorégulateur, est "pragmatique". Par "pragmatisme" on entend ici désigner l'idée philosophique selon laquelle toute connaissance est de nature pratique, ancrée dans une situation particulière, et qu'il n'existe pas de possibilité d'accéder à un point de vue abstrait et extérieur qui nous ferait voir l'essence de l'économie. Ainsi, la *Théorie Générale* est dirigée contre le principe d'une "théorie pure" de l'économie. Cela apparaît clairement dans l'analyse que propose Keynes du mécanisme de l'incitation à investir. On constate en effet que l'évaluation des actifs est au confluent de deux regards alternatifs, qui s'expriment dans une organisation économique contingente. Keynes montre que la connaissance pratique de l'entrepreneur est au fondement d'une conception de l'économie plus favorable au plein emploi, et plaide donc pour que le jugement de cet acteur puisse être placé au centre du système.

Cette description d'un conflit concret, limité à un système particulier, est à l'antithèse d'une analyse constructiviste qui permettrait de formuler des lois définitives. Au contraire, Keynes montre que le fonctionnement de l'économie dépend des règles que l'on choisit. Il décrit un univers dans lequel le système ne préexiste pas à l'action concrète des individus mais naît au contraire de cet enchevêtrement d'actions, qui produisent elles-mêmes des conventions. Le système que décrit Keynes est un système de règles et, qui plus est, de règles émergeant de l'interaction entre de nombreux individus.

Cependant, Keynes ne se contente pas de décrire un ensemble d'actions formant un système : "*le système est composé en fonction de son usage possible dans une action. L'action qui se trouve au principe de composition du système est l'action d'une «autorité centrale» ou d'un «gouvernement»*" (Berthoud, 1998, p.273). La théorie de Keynes est une théorie tournée vers l'action. Elle s'adresse à un acteur capable d'intervenir sur le système lui-même : l'Etat. S'il existe un système, c'est d'ailleurs uniquement du point de vue extérieur d'un savant cherchant à comprendre le mécanisme économique de l'investissement, ou d'un homme d'Etat cherchant à le modifier. Keynes éclaire les

fondements pratiques des décisions des acteurs. Il montre leur rapport de force. Il analyse la manière dont la connaissance et les objectifs des uns s'imposent à la connaissance et aux objectifs des autres. Il donne à l'homme d'Etat une représentation organisée de l'économie mettant en évidence les différents rapports de force qui la traverse. Au moment où il rédige la Théorie Générale, Keynes perçoit que ce rapport de force tourne au bénéfice des spéculateurs, et que cette évolution est défavorable au plein emploi. Il montre alors qu'il serait nécessaire de changer les règles du jeu pour assurer le plein emploi. Keynes pense donc que le savant peut conseiller l'homme d'Etat en lui indiquant quelles règles il faudrait mettre en place, compte tenu de l'objectif social à atteindre. Cette conviction, Keynes la tire d'une conception de la "réalité" qui lui permet d'opposer au jugement "conventionnel" la connaissance véritable des entrepreneurs.

412- La connaissance de l'entrepreneur contre l'arbitraire des conventions.

L'analyse que Keynes propose des marchés financiers anticipe, d'une certaine manière sur les développements ultérieurs de Schelling [1960] et de Lewis [1969]. La situation des individus sur les marchés financiers peut être représentée sous la forme d'un jeu de coordination tel que le définit Schelling. La "solution" que Keynes décrit, et qui ne lui convient pas, utilise le concept de convention, que Lewis redéfinira trente années plus tard²⁸.

Dans la théorie de Lewis, la caractéristique importante de la convention est son arbitraire. Lewis insiste sur le fait qu'il est nécessaire qu'une convention soit choisie, mais remarque que la forme particulière qu'elle revêt n'a aucune importance. Il va même plus loin : il intègre dans la définition de la convention l'idée selon laquelle, pour être conventionnelle, elle doit *nécessairement* être arbitraire, c'est-à-dire être choisie alors même qu'une autre aurait tout aussi bien fait l'affaire. C'est en cela que Lewis, on l'a vu, est dans la "pure coordination".

Keynes a pleinement conscience de l'arbitraire de la convention. Il indique clairement que le taux d'intérêt est indéterminé, parce qu'il est conventionnel : "*Le taux de l'intérêt est au plus haut degré un phénomène conventionnel, plutôt que psychologique, car sa valeur effective dépend dans une large mesure de sa valeur future telle que l'opinion la prévoit. N'importe quel niveau du taux d'intérêt dont on est suffisamment convaincu qu'il durera, durera effectivement.*" (Keynes, 1936, p.214). Le problème est précisément que, comme le taux d'intérêt s'établit à un niveau assez arbitraire, il y a en fait assez peu de raison de supposer que ce niveau sera compatible avec le plein emploi. Le même raisonnement peut être tenu à propos du rendement escompté des actions : le problème,

²⁸ Voir chapitre quatre.

pour Keynes, est que le niveau de rentabilité escomptée d'un bien capital sous forme d'action soit à ce point indéterminé.

Ainsi ce qui ne convient pas à Keynes dans le fonctionnement conventionnel des marchés, c'est l'arbitraire des conventions. Contre cet arbitraire, Keynes met en valeur l'importance des données "réelles" et "objectives". Il souhaiterait que l'évaluation des biens capitaux soit fondée sur leur valeur intrinsèque. Il attend de l'autorité monétaire qu'elle mette en place une politique volontariste permettant d'adapter le taux d'intérêt aux objectifs collectifs de la société. Ainsi, de la dénonciation de l'arbitraire des conventions, Keynes passe immédiatement à la dénonciation de leur absence de fondements "réels". Ce faisant, ce n'est plus le choix de la convention particulière qu'il critique, mais le fait même que le choix collectif soit conventionnel.

La convention est en effet par essence opposée à l'idée de réalité naturelle. Il y a précisément raisonnement ou coordination par convention lorsque la réalité n'impose pas un comportement unique. Ainsi, ce n'est pas par convention que l'on décide de chauffer nos habitations durant les périodes d'hiver, mais c'est par convention que l'on choisit de rouler à droite plutôt qu'à gauche. De même, la philosophie "conventionnelle" du langage, au sens le plus large, se dénomme ainsi parce qu'elle suppose que l'origine du langage n'est pas "naturelle". On ne peut donc pas dénoncer le manque de "réalisme" ou "d'objectivité" des conventions. La convention n'est pas justiciable du réel.

Dans son analyse, Keynes dévoile la nature conventionnelle du taux d'intérêt et de l'évaluation du rendement escompté des actions. Il montre également que le comportement conventionnel ne repose pas sur des fondements "véritables" ou "réels". Ce n'est là, finalement, qu'une simple précision de ce qu'est une convention. Mais il procède simultanément à une forme de dénonciation de la forme conventionnelle d'évaluation. Il oppose ainsi à la convention le jugement "véritable" de l'entrepreneur, qui porte sur des préoccupations de long terme et qui est bénéfique à l'ensemble de la communauté. Ici Keynes souligne donc que la nature conventionnelle du comportement dans l'incertain n'est pas une fatalité. Elle est une conséquence du déficit de connaissance des acteurs et de l'absence de règles les guidant de façon adéquate. La convention est donc, selon Keynes, un symptôme indiquant un dysfonctionnement du système. Une économie bien gérée ne devrait pas se laisser guider par des conventions arbitraires. Il ne devrait pas exister de problème de coordination, au sens de Schelling : les éléments concrets de la réalité économique devraient suffisamment guider les acteurs vers une solution satisfaisante. Cette position de Keynes est directement liée à sa conception de la raison. Pour Keynes, en effet, la raison dépend de la connaissance *réelle*, elle porte sur

des relations logiques *objectives* : elle s'oppose précisément au jugement par convention.

C'est sur cette base que Keynes prône la mise en place d'un mode de financement qui servirait mieux les projets de l'entrepreneur. En sécurisant les anticipations de l'entrepreneur, en favorisant le financement de l'investissement, l'Etat permet au jugement raisonnable de refaire surface. Les règles qu'il convient, selon Keynes, de mettre en place pour contrer la mécanique conventionnelle ne sont pas seulement des règles contingentes adaptées à l'objectif particulier qu'est le plein emploi. Elles sont *nécessaires* parce qu'elles permettent de faire émerger la "*vérité*". L'analyse de Keynes, sa défense de l'esprit d'entreprise, son appel à un volontarisme de l'Etat s'appuient donc sur une conviction *réaliste* : pour que l'économie atteigne le plein emploi, il faut redonner sa place à la connaissance réelle. Une économie guidée par des connaissances "*véritables*" est aussi une économie qui assure le bien-être collectif.

Keynes n'est donc pas seulement pragmatique. Son pragmatisme s'accompagne d'une croyance en l'existence d'une vérité de nature transcendante et en l'existence d'un lien étroit unissant le "*vrai*" et le "*bien*". Cette attitude philosophique "*réaliste*" marque profondément les convictions de Keynes. Elle est au fondement d'une conception platonicienne du savant, personnage éclairé capable de percevoir, au-delà de l'enchevêtrement des règles et des actions, les éléments de connaissance véritable qu'il conviendrait de placer au fondement des décisions collectives.

42- *Le réalisme platonicien de Keynes*

Les écrits récents consacrés à la philosophie de Keynes ne manquent pas. Ces dernières années ont en effet vu fleurir, outre-manche, de très nombreuses analyses des premiers écrits, philosophiques, de Keynes²⁹. Ce mouvement a été en partie lié à la redécouverte de textes inédits, conservés à Oxford, qui ont permis de prendre conscience de l'importance que pouvait avoir cette œuvre philosophique. Le *Treatise* est ainsi apparu comme la "*partie émergée*" d'une activité philosophique intense qui a marqué les premières années de formation intellectuelle de Keynes. Comme l'indiquent Bateman et Davis [1991], dans l'introduction d'un ouvrage commun rassemblant des contributions des principaux auteurs de ce courant de recherche, l'idée poursuivie par ces exégètes de la pensée de Keynes est de donner à son "*économie*" des fondations philosophiques, comme cela a pu être fait pour Smith ou pour Marx.

²⁹ On peut citer : Bateman et Davis (eds) [1991], Fitzgibbons [1989], O'Donnell [1989], Carabelli [1988], Skidelski [1983, 1992]. Braithwaite [1975] avait anticipé sur ce mouvement.

Le premier résultat marquant de ces études est la diversité de leurs conclusions, souvent divergentes bien qu'elles soient toutes fondées sur une scrupuleuse analyse des écrits de Keynes. Ainsi, les positions philosophiques de Keynes semblent être ambiguës et sujettes à controverses, ce qui incite surtout à la prudence. Toutefois, certains traits dominants des conceptions de Keynes semblent émerger, et ces traits dominants constituent une assise intéressante pour interpréter les réflexions de Keynes sur les conventions. Nous développerons en particulier deux points, qui font l'objet d'un consensus minimum : Keynes est influencé par Moore, et, à travers lui, par Platon. Keynes est soucieux de la connaissance pratique, ce qui l'oppose aux conceptions calculatoires de la prise de décision. Ces deux aspects dessinent un Keynes attaché à la fois, et de manière un peu paradoxale, au réalisme platonicien et au jugement pratique.

Keynes apparaît donc comme opposé à la fois au *positivisme logique*, qui se construit entre Vienne et Cambridge à cette époque, et au *naturalisme*, qui marque la définition standard de l'économie par Robbins. Contre ces deux écoles enfermant, d'une façon ou d'une autre, le comportement et l'entendement humain dans des lois strictes, logiques ou naturelles, Keynes tente de définir la "raison" sous une forme à la fois plus "pratique" et plus "réaliste".

"Pratique". L'utilisation de ce terme à propos de la philosophie de Keynes nous permet de souligner qu'il s'oppose sans aucun doute au *formalisme* typique, par exemple, du premier Wittgenstein : "*La première impression produite par le travail de Russel fût qu'il étendait énormément le champ de la logique. L'amélioration progressive du traitement formel auquel il contribua lui-même, avec Wittgenstein et Ramsey, a cependant eu pour effet de vider progressivement la logique de tout contenu et de la réduire à une sorte de squelette, excluant non seulement l'expérience, mais aussi tous les principes reconnus en général comme constitutifs d'un mode de pensée logique et raisonnable*" (Keynes, 1931, p.338). Le *Treatise* est entièrement consacré à donner un fondement logique à ces "principes" qu'évoque Keynes, et en particulier à l'induction (Carabelli, 1988). Keynes est donc attaché à une représentation de l'entendement qui inclut un certain "souci du réel". Il ne comprend pas le sens de la révolution qu'opère le formalisme logique à la suite des premiers travaux de Wittgenstein³⁰. Pour Keynes la raison est d'essence pratique, elle désigne non pas des compétences formelles mais une capacité à

³⁰ En fait, la citation ci-dessus montre bien que Keynes ne perçoit pas clairement la distinction qu'opère Wittgenstein entre les notions de *sens* et de *vérité* (que Keynes ne distingue jamais). En cherchant à définir les règles permettant de qualifier un énoncé de "sensé" sur le plan logique, il ne supprime pas pour autant l'importance de la vérité. La définition des énoncés ayant un sens logique est un préalable à la distinction, parmi les énoncés sensés, entre ceux qui sont vrais et de ceux qui ne le sont pas. Wittgenstein n'exclut donc pas le problème de la vérité, il suggère qu'avant de savoir si un énoncé est vrai, il est nécessaire de savoir s'il a un sens.

percevoir la logique qui régit une situation d'action.

“Réaliste”. Keynes hérite de Moore sa définition d'une vérité obtenue par intuition. Cette conception du vrai est platonicienne. Davis [1991] démontre clairement comment cette posture platonicienne est au point de départ, pour Keynes, d'une conception de la vérité qui s'oppose à la fois au “subjectivisme” de Ramsey, et au “naturalisme” de Robbins. Keynes, contre Robbins, ne pense pas que la vérité puisse être atteinte à travers une étude brute des “faits”. Dans une lettre à Harrod (Keynes, 1938b), il dénonce les excès de l'économétrie de Timbergen et rejette la définition de Robbins. Contre leurs conceptions, il affirme que l'économie est d'abord une science morale fondée sur l'introspection et non pas une science naturelle qui décrirait des faits objectifs³¹. Mais, pour Keynes, cette introspection est une porte ouverte sur la “réalité” du monde. C'est en cela, on l'a vu, qu'il s'oppose au subjectivisme de Ramsey. Par l'introspection, l'acteur raisonnable atteint le vrai. En scrutant les faits, il a l'intuition de la vérité : “*Nous avons besoin qu'un ensemble de convictions nouvelles naisse d'un examen sincère de nos sentiments intérieurs confrontés aux faits extérieurs*” (Keynes, 1926, p.294). Si l'on s'accorde avec les analyses de O'Donnel [1989], Fitzgibbons [1989] ou encore Skidelski [1983], on doit donc concevoir la présence chez Keynes, d'une conviction platonicienne. Keynes croit sans doute en l'existence d'une réalité transcendante. C'est le sens de la référence de Keynes à la “réalité” dans la *Théorie Générale*.

Cet “intuitionnisme réaliste” est à l'origine de l'opposition de Keynes à toutes les tentatives visant à “systématiser” le raisonnement logique ou rationnel. L'essentialisme de Keynes n'a aucun sens dans la philosophie de Wittgenstein. Il n'en a pas davantage dans la perspective subjectiviste de Ramsey, qui sert de base à la conception économique moderne de l'individu rationnel.

43- L'élitisme de Keynes

La philosophie politique de Keynes découle de sa conception de la connaissance et de la rationalité. Helburn [1991] s'appuie ainsi sur un texte inédit de Keynes consacré à Burke pour défendre l'idée selon laquelle “*Burke a eu une influence majeure sur Keynes*” (Helburn, 1991, p.30). Helburn souligne que Keynes n'accepte pas le conservatisme de Burke, mais qu'il accepte “*la logique de la théorie politique de Burke – une doctrine éthique de l'action et une vue élitiste concernant la pratique du pouvoir*” (*ibid*, p.31). En

³¹“*En second lieu, contre Robbins, l'économie est essentiellement une science morale et non une science naturelle. Cela signifie qu'elle fait usage de l'introspection et des jugements de valeur*” (Keynes, 1938b, p.297).

accord avec cette interprétation, O'Donnel (1989, p.66 et chap.13) souligne l'*élitisme* de Keynes, tandis que Fitzgibbons (1988, p.171-173) le décrit comme ayant une confiance modérée en la démocratie dont l'efficacité est conditionnée à un "*gouvernement des élites*".

Les interprétations convergent donc pour voir en Keynes un auteur marqué par la "*philosophie politique antérieure aux Lumières*" (Fitzgibbons, 1991, p.132). Keynes conçoit sans doute l'art de gouverner comme devant être réservé à certains acteurs particulièrement clairvoyants, ce qui l'oppose en partie à l'idéal démocratique des Lumières. Mais, cette détermination très générale permet aussi, et surtout, de distinguer les conceptions de Keynes, de l'idéal moderne d'une gestion politique "rationalisée" sur le modèle d'un mécanisme. En cela Keynes s'oppose à toute la tradition politique initiée par Hobbes. Carabelli (1988, p.144) souligne que les conceptions exposées par Keynes dans le *Treatise* induisent une conception de la rationalité rigoureusement différente de la rationalité calculatoire sur laquelle Hobbes assoit sa théorie de l'Etat. Pour Hobbes, l'efficacité du gouvernement civil, par opposition à l'état de nature, est liée à l'utilisation d'un langage mathématique non ambigu. La philosophie politique de Hobbes est d'abord une théorie nominaliste du langage qui préfigure par certains aspects le positivisme logique. La théorie de Keynes défend une conception inverse de celle de Hobbes. Selon Keynes, il n'existe pas dans la société de support suffisant sur lequel asseoir un langage mathématique et une organisation mécanique. Le caractère fondamentalement incomplet de notre savoir, dont découle la nécessité de fonder nos actions sur des raisonnements probables, mine à la fois le conséquentialisme et le constructivisme. Ces deux conceptions, liées, de l'organisation de la société reposent en effet sur l'idée selon laquelle il est possible de relier de manière certaine une action à ses conséquences, sur la base d'un savoir certain. Elles ne tolèrent pas le caractère temporaire et seulement probable du savoir, que Keynes décrit dans le *Treatise*.

C'est en ce sens que la philosophie politique de Keynes s'inscrit dans un refus de la "modernité". Keynes opère un retour à une conception de la société dans laquelle même le savoir du "Prince" est incertain³². Keynes est ici proche de Machiavel. Il conçoit le gouvernement comme un art, dans lequel le savant ne peut avoir qu'un rôle de conseiller

³² L'intuitionnisme est associé, chez Keynes, à une conception prudente de la prise de décision fondée sur le fait que personne ne connaît jamais la totalité de la réalité. Le savoir est toujours incomplet. C'est la raison pour laquelle Keynes cite à diverses reprises la phrase de Butler : "*Les probabilités constituent à nos yeux le véritable guide de la vie*" (Keynes, 1921, p. 341). De même, dans la lettre à Harrod déjà citée, Keynes précise : "*L'économie est un savoir penser en termes de modèles doublé de l'art de choisir des modèles applicables au monde contemporain. Elle y est contrainte parce que, à la différence d'une science naturelle typique, l'objet auquel elle s'applique n'est pas homogène dans le temps.*" (Keynes, 1938, p.300).

et non de prescripteur³³. Ainsi, le diagnostic que porte Keynes sur le fonctionnement chaotique du système économique qu'il observe s'appuie sur une conception "réaliste", au sens de Platon, de la connaissance. Ce réalisme philosophique sous-tend une philosophie politique dont les contours sont "élitistes" et "anti-constructivistes".

Ces quelques termes ne permettent pas de définir rigoureusement une conception keynésienne de la connaissance, et encore moins de la philosophie politique. Ils permettent cependant de savoir ce que Keynes n'est sans doute pas. Il n'est pas constructiviste, et, en ce sens, s'oppose au mouvement moderne de la théorie économique. Mais il n'est sans doute pas non plus tout-à-fait en phase avec le mouvement moderne du "pragmatisme" qui définit la rationalité, ou la raison, d'un acteur uniquement en fonction de son objectif propre. Keynes croit en effet à l'existence d'une réalité qui transcende l'acteur, et qu'il faut percevoir. Cette attitude épistémologique l'amène à supposer que certains individus perçoivent mieux la réalité que d'autres, et qu'il serait préférable que la politique publique soit faite sur la base de leurs connaissances. Cette conviction n'est pas entièrement contradictoire avec une conception de type pragmatique de la régulation du système économique. Keynes sait que le système ne peut pas préexister à l'action. Il sait qu'il n'est pas possible de *prescrire* un comportement. L'action est toujours immédiatement présente, elle est en avance sur la règle. Mais il pense qu'il est possible de disposer d'un savoir de type particulier destiné à un acteur de type particulier. Ce savoir est celui du savant, et il s'adresse à l'Etat. Le savoir de l'économiste n'est cependant pas transcendant, et l'action de l'Etat ne se substitue pas à celle des autres acteurs. L'Etat est pris dans un univers où il doit prendre en compte les actions des entrepreneurs et des spéculateurs. Mais il a le pouvoir de trancher le différend entre ces deux catégories d'acteur. Il peut favoriser le point de vue de l'entrepreneur en modifiant les règles du jeu. Il est incité à agir ainsi sur le conseil du savant qui est capable, d'une part d'avoir un point de vue extérieur sur le système, et d'autre part de juger de la validité des connaissances sur lesquelles chacun des acteurs fonde son action. Le savant peut donc percevoir que l'entrepreneur dispose d'un savoir qui est meilleur, plus vrai et donc plus utile à l'ensemble de la communauté, que celui du spéculateur. Son analyse permet donc de mieux organiser la vie économique autour de ce savoir. La connaissance du savant est cependant d'un ordre différent de celle de l'entrepreneur. L'action de l'Etat peut aider celle de l'entrepreneur, mais non s'y substituer. En ce sens Keynes reste pragmatique, il respecte le savoir des acteurs et conçoit leur caractère imprescriptible.

³³ "N'est-ce-pas aux conseillers experts-économistes de dire ce qui devrait être fait, et de laisser aux politiciens dire comment ils le feront ?" (Keynes, cité par O'Donnell, 1989, P.170)-

Cependant, en soulignant que le savant peut disposer d'une connaissance d'un niveau supérieur et que l'Etat peut agir sur le système lui-même au nom de l'intérêt général, il fait l'hypothèse que ces deux acteurs ont un regard et une position surplombante par rapport aux interactions des agents³⁴. A un certain degré son élitisme entre en contradiction avec son pragmatisme. C'est cette contradiction qui rend problématique l'analyse keynésienne des règles.

44- Une théorie "problématique" de la règle.

La dualité de l'attitude philosophique de Keynes rend délicate l'interprétation de sa position en ce qui concerne le rôle et la genèse des règles collectives. L'attitude "pragmatique" de Keynes l'amène immédiatement à reconnaître que le système économique est un ensemble de règles produites par les acteurs. Mais, son attachement "réaliste" à l'existence d'une vérité de nature transcendante l'amène également à penser qu'il existe des règles nécessaires qui devraient s'appliquer contre les conventions en usage. Certes, ces règles sont une réponse ponctuelle dans un système particulier. Keynes ne considère pas que sa théorie énonce des vérités éternelles et transcendantes, mais il suggère que l'économiste sait repérer les règles qu'il faudrait mettre en place, et la nécessité de ces règles est liée au caractère *véritable* du jugement de l'entrepreneur. C'est en définitive cette référence constante à la "réalité" que l'entrepreneur serait capable de percevoir qui pose problème dans le cadre d'une théorie des règles.

Car l'existence d'une "réalité" transcendante légitime la place de "conseiller" que Keynes donne au savant. Le "réalisme" de Keynes est donc une réponse au problème du choix des règles. Il permet en effet de désigner le savant comme conseiller de l'Etat et de supposer que le système peut être régulé "d'en haut". Keynes, parce qu'il suppose que le savant sait à chaque moment quelles sont les règles qu'il convient de mettre en place, vide de sens la question de la genèse des règles. C'est une question qui ne se pose pas dans une optique platonicienne.

Cette articulation du pragmatisme et du réalisme que l'on repère dans l'analyse keynésienne est fragile. En effet, Keynes insiste sur l'importance de la dimension temporelle du processus et des décisions économiques. Il suggère clairement que la

³⁴ Cette capacité du savant à se hisser au dessus des interactions des agents semble être contradictoire avec une position pragmatiste comme celle de Wittgenstein. Celui-ci rappelle ainsi, dès le *Tractatus*, que l'on ne peut précisément pas s'élever au dessus du langage.

décision raisonnable se prend sur des bases cognitives certaines mais nécessairement incomplètes, parce qu'avec le temps elles peuvent s'avérer erronées. Le problème qui se pose au savant et à l'homme d'Etat est donc précisément l'impossibilité de connaître l'avenir avec certitude, et leur tâche est de limiter au maximum cette incertitude. On est là du côté pragmatique de Keynes. Mais Keynes croit aussi qu'il est possible, par intuition, d'atteindre la vérité, même lorsque l'on est incertain. Il pense que le savant est cet acteur inspiré qui peut, en quelque sorte percer les mystères du temps, ou au moins certains d'entre eux. On est cette fois du côté réaliste. La théorie keynésienne des règles se situe à la jointure de ces deux idées. D'une part Keynes perçoit que les agents ont besoin de règles pour harmoniser leurs anticipations, d'autre part il estime que le savant peut fournir à la communauté la connaissance permettant de choisir les "bonnes règles", et que l'Etat constitue le moyen par lequel la société peut mettre en place ces règles adéquates. Cette conception des règles collectives contourne le problème crucial que constitue le mécanisme du choix collectif et de l'émergence des règles, d'une part en faisant reposer la décision collective sur la "vision" du savant (pour reprendre l'expression de Skidelski [1988]) et d'autre part en supposant que l'Etat a la capacité d'accomplir seul l'action collective. Mais on est en droit de se demander d'où le savant tire son don de visionnaire, et comment l'Etat peut à ce point manipuler les règles collectives. Ces deux questions ne sont pas résolues par Keynes, elles font parties de ses *a priori*. C'est cet *a priori* qu'une théorie pratique des règles devrait questionner.

Conclusion

Keynes libère l'analyse économique de l'idée de loi.

Le *Treatise on Probability* expose une conception de la rationalité individuelle qui n'est pas représentable sous la forme d'une loi. L'acteur agit raisonnablement sur la base d'un jugement de probabilité lié à sa capacité à percevoir dans la réalité l'existence de relations logiques. Il ne dispose pas de la base cognitive nécessaire pour représenter ses choix sous forme mathématique. Il sait que l'avenir ne découle pas mécaniquement du passé, et que sa conjecture ne peut être que probable. Lorsque Keynes décrit l'économie, sur la base de cette conception de l'action raisonnable, il ne cherche donc pas à intégrer l'acteur dans un système théorique dans lequel son comportement répondrait à des lois abstraites. Il cherche au contraire à saisir la manière dont ce mode de comportement détermine le fonctionnement de l'ensemble de l'économie. C'est en ce sens que Keynes n'est pas constructiviste.

Keynes montre que les règles collectives sont nécessaires et essentielles.

Dans notre optique, le point crucial de l'économie de Keynes est l'importance qu'il prête au jugement collectif sur l'avenir. Keynes montre que dans la manière dont ce jugement collectif se forme, la part de connaissance personnelle des individus s'efface derrière un processus de massification qui prend la forme de conventions, déconnectées de la réalité. Il souligne clairement que ce processus est lié à l'organisation de l'économie. Autrement dit, il montre que les règles particulières qui régissent l'économie de son époque empêchent l'émergence d'un jugement collectif raisonnable au sens du *Treatise*. Comme ce jugement détermine l'investissement, et donc la demande effective, il n'est pas surprenant que l'on demeure parfois dans la situation "absurde" d'un équilibre de sous-emploi. Le manque de "fondements raisonnables" du jugement collectif empêche que l'on atteigne l'objectif général du plein emploi. Ce que montre Keynes, c'est donc qu'il serait nécessaire de guider le jugement collectif pour lui donner un fondement objectif plus satisfaisant. On peut dire aussi : il est nécessaire d'installer des règles organisationnelles permettant de lier davantage les jugements collectifs à la réalité objective. Keynes pointe donc un problème d'organisation, et un "besoin" de règles collectives évitant le piège de la convention.

Mais Keynes ne conçoit pas la règle de façon délibérative .

Le diagnostic qu'il formule repose sur une opposition des "conventions" qui émergent des interactions entre agents ignorants, à la "réalité" que perçoivent les entrepreneurs.

Cette opposition est révélatrice de l'attitude duale de Keynes vis à vis des règles. Son pragmatisme l'amène à concevoir le rôle des règles, mais son "platonisme" le convainc du fait que les "bonnes règles" ne peuvent émerger directement des interactions entre agents "ignorants", coincés dans la "caverne". C'est pourquoi Keynes attend de l'Etat qu'il mette en place les règles, sur le conseil du savant. Le savant apparaît donc comme un individu éclairé. Néanmoins, on perçoit mal l'origine de cette hauteur de vue qui permet à l'économiste de dominer l'enchevêtrement des règles et actions qui constituent l'économie. On ne comprend pas davantage comment l'Etat, qui est pris lui-même dans cet enchevêtrement, parvient à influencer directement, et de manière souveraine, sur les règles du jeu. A un certain niveau cette extériorité de la connaissance du savant, et cette prééminence de l'action d'Etat semble entrer en contradiction avec l'attitude pragmatiste de Keynes. Keynes peut se passer d'une théorie de la genèse des règles collectives, mais c'est au prix d'une théorie de l'action d'Etat et de la connaissance du savant qui paraît bien idéaliste, peut-être plus encore à l'époque actuelle qu'à celle de la *Théorie Générale*³⁵.

La théorie keynésienne constitue donc un cadre adéquat pour comprendre le rôle des règles. Elle introduit une conception pratique du temps qui permet de bien cerner le problème de la règle. Mais Keynes n'assume pas entièrement le caractère contingent et délibératif des règles collectives, il attend de l'Etat qu'il joue seul le rôle de régulateur. De ce fait, il ne conçoit pas la question de la dynamique des règles autrement que sous la forme d'une action de l'Etat. Après Keynes, il faut donc encore parvenir à formuler une théorie de la genèse des règles qui ne soit pas seulement une théorie de l'action d'un acteur situé au dessus des interactions des agents. La "théorie des conventions" ambitionne précisément à formuler une telle théorie.

³⁵ De fait, Keynes n'est jamais parvenu à faire triompher le principe d'un comité des experts *indépendants*, chargé de gérer les mouvements de change au niveau mondial.

Chapitre 6

L'Economie des Conventions

Le projet théorique de l'économie des conventions s'inscrit dans le droit fil des réflexions de Keynes. Il s'agit de comprendre comment des individus ne disposant pas de ressources cognitives illimitées parviennent à se coordonner en produisant des règles communes. On a cependant vu que Keynes, qui ne pensait sans doute pas que les acteurs pouvait eux-mêmes produire les règles adéquates, ne propose pas de théorie "satisfaisante" de la genèse des règles. L'économie des conventions tente, d'une certaine manière, de combler cette lacune. Elle s'inscrit donc dans la brèche que Keynes ouvre lorsqu'il montre que la coordination économique repose sur des bases collectives *a priori* indéterminées, mais cherche à aller plus loin en construisant une véritable théorie individualiste de la dynamique des règles collectives.

Elle mobilise pour ce faire une hypothèse de rationalité qui s'inspire plus directement des analyses théoriques d'Herbert Simon que de l'analyse keynésienne du *Treatise*. Nous présenterons donc les éléments de la théorie simonienne de la rationalité sur lesquels la théorie des conventions (dans son versant économique) s'appuie explicitement.

La théorie de Simon permet d'envisager le rôle des règles, mais les conçoit plus comme des "ressources" permettant d'améliorer l'efficacité de l'action individuelle, que comme des "produits" de cette action. Elle présuppose, en quelque sorte, l'existence d'une structure organisationnelle produisant les règles. C'est cet aspect de la théorie de Simon que le projet conventionnaliste devrait permettre de dépasser.

Nous énoncerons ce projet en trois points précis. A l'aune de ce projet, on cherchera à mesurer la progression théorique de l'économie des conventions à travers l'étude des propositions théoriques concrètes de deux de ces représentants : Olivier Favereau et Laurent Thévenot. On apercevra alors les obstacles qui entravent la réalisation effective du

projet de recherche conventionnaliste. La réponse "cognitive" que les auteurs tentent d'apporter à ces obstacles n'est en effet pas tout-à-fait satisfaisante. Elle ne lève pas les ambiguïtés de l'hypothèse de rationalité limitée sur laquelle repose cette école. En effet pour Simon, la rationalité, même limitée, est instrumentale. Or, le projet conventionnaliste ne peut se développer qu'en s'appuyant sur une forme de rationalité non exclusivement instrumentale.

1– Une hypothèse de rationalité procédurale

L'économie des conventions s'appuie sur une hypothèse de rationalité "limitée". Ce concept recouvre des conceptions assez différentes de l'idée de rationalité selon que l'on considère qu'il souligne simplement le caractère limité des capacités cognitives d'un individu qui demeure optimisateur, ou, plus profondément, qu'il tend à remettre en cause l'idée d'optimisation. Lorsque l'analyse orthodoxe envisage qu'il peut exister des limites au comportement rationnel, elle n'envisage pas pour autant qu'il faille modifier le concept de rationalité individuelle. Elle tend simplement à réfléchir aux moyens institutionnels capables d'assurer à l'individu un environnement qui lui permette de maximiser correctement. A l'inverse, l'approche de Simon tend à remettre en cause le principe de l'optimisation.

En effet, Simon ne se contente pas de prendre acte des limites "computationnelles" de l'individu, il cherche à montrer que ces limites impliquent de concevoir la rationalité d'une façon différente, en prêtant attention non plus au résultat de la délibération, mais au mécanisme de recension et d'interprétation de l'information qui permet une prise de décision rationnelle. On notera qu'il y a là l'existence d'une parenté de pensée étonnante entre Simon et Keynes¹. Nous étudierons la théorie de Simon en cherchant à mettre en évidence cette parenté. Comme Keynes, Simon montre que l'individu doit être rationnel bien qu'il ne puisse "tout connaître" en cela il est "*limité*". Comme Keynes, Simon montre que pour être rationnel l'individu doit être capable de recenser et d'interpréter des informations pour "comprendre" la situation. En cela, sa rationalité tient au "*processus*" de décision plutôt qu'à la décision elle-même (qui repose toujours sur une connaissance partielle et peut donc s'avérer erronée).

Cette conception de la prise de décision rationnelle constitue une étape décisive dans la prise en compte du rôle des règles. L'individu utilise en effet les ressources cognitives que la société met à sa disposition. On peut concevoir ces ressources comme des "règles" historiquement et socialement déterminées. La règle n'est donc plus uniquement une contrainte, elle constitue une aide à la prise de décision rationnelle. Toutefois, chez Simon, l'archétype de la règle est la règle organisationnelle. L'acteur "hérite" des règles par le biais des organisations collectives. La règle est donc exclusivement envisagée comme une ressource, et non comme un produit, du comportement rationnel des acteurs.

¹ Cette idée est confortée par une remarque récente de Simon : "*Seule la peur de commettre un énorme anachronisme me retient d'affirmer que Keynes est le véritable instigateur de l'économie de la rationalité limitée*" (Simon, 1997, p.16).

11- Une rationalité "limitée"

"Dans les situations complexes, il y a fréquemment un écart important entre l'environnement réel de la décision (tel que le percevrait Dieu ou tout autre observateur omniscient) et l'environnement tel que les acteurs le perçoivent." (Simon, 1978, p.451).

Simon cherche à faire apparaître le problème que constitue pour l'individu l'impossibilité de disposer de toutes les informations nécessaires à une détermination rationnelle, au sens standard, de son choix. Il s'oppose ainsi à l'idée selon laquelle l'incertitude serait liée à un problème externe à l'individu, comme par exemple une défaillance du système de marché. Au contraire, le manque d'information est une caractéristique de la représentation simonienne de la prise de décision de l'individu.

111- Une conception interne de l'incertitude

Pour Simon, comme pour Keynes, l'incertitude n'est pas liée à la défaillance du marché, comme c'est le cas dans l'analyse orthodoxe. Elle est inhérente aux limites cognitives de l'être humain, qui ne peut percevoir qu'une partie de la réalité. La connaissance individuelle est en effet par essence imparfaite. Comme le souligne clairement Simon, "*L'incertitude n'existe pas dans le monde extérieur, mais dans l'esprit de l'individu.*" (Simon, 1976b, p.437).

C'est en ce sens que Simon parle dans ses premières analyses de "bounded rationality", terme traduit en français par l'expression "rationalité limitée"². Les limites, ou plus exactement les "frontières", que pose Simon à la théorie de la rationalité s'opposent à l'idée d'une rationalité "globale" qui caractérise l'approche "standard". Cette opposition a trait à la manière de concevoir le contexte de choix : "*Les conditions et les contraintes (...) peuvent être perçues comme des "caractéristiques objectives" de l'environnement externe de l'organisme qui choisit, ou elles peuvent être des "caractéristiques perçues", ou bien encore des caractéristiques de l'organisme lui-même. La ligne qui sépare la première conception des deux autres est marquée par la distinction entre la rationalité "objective", d'un côté, et la rationalité "subjective" ou "limitée" de l'autre.*" (Simon, 1964, p.405). Dans le premier cas un comportement est dit rationnel s'il est fondé sur les caractéristiques "objectives" de la situation. Dans le second cas, la rationalité d'un comportement est rapportée au niveau de connaissance de l'individu, elle

²La traduction de "bounded rationality" par "rationalité bornée" rendrait peut-être mieux l'idée selon laquelle l'individu n'est rationnel que dans la limite de ses connaissances et de ses capacités cognitives. Toutefois la traduction qui est la plupart du temps retenue est celle de "rationalité limitée". Par convention, c'est celle que nous suivrons ici.

est donc relative à l'acteur. Ainsi, dans le premier cas, on suppose par hypothèse que l'individu sait tout ce qu'il y a à savoir, tandis que dans le second on cherche précisément à rendre compte du caractère subjectif et incomplet de ses connaissances. Le premier cas correspond à une définition normative "idéale" de la rationalité, tandis que le second correspond à une approche plus "positive" qui vise à décrire de manière réaliste le comportement de l'individu.

De même, l'objectif que l'individu cherche à atteindre peut être considéré comme objectivement déterminé par la situation ou subjectivement déterminé par l'individu. Ce peut être le point "*optimum*" qui maximise l'utilité de l'individu rationnel, compte tenu des caractéristiques "objectives" de la situation. Ce peut également être, plus modestement, un niveau de bien-être, qui n'est pas nécessairement optimal, mais qui satisfait l'acteur. C'est la seconde ligne de rupture entre Simon et la théorie standard de la rationalité. Simon oppose en effet à l'attitude "optimisatrice" le principe du seuil de "satisfaction" de l'individu standard. Il y a ici un lien étroit entre la manière dont on considère *l'environnement* de la décision, et la manière dont on conçoit *l'objectif* de l'individu rationnel. L'optimisation est liée à une situation d'information parfaite permettant de percevoir objectivement l'environnement de la décision, tandis que le "satisficing" est lié au regard "subjectif" et "incomplet" que l'individu porte sur la situation. Deux visions s'opposent donc : dans la vision standard, pour être rationnel, un individu doit connaître parfaitement la situation et atteindre l'*optimum* ; dans la vision plus "réaliste" de Simon, la rationalité d'un comportement est rapportée à la manière dont l'individu perçoit l'environnement et ses potentialités. Dans le premier cas, on peut prescrire objectivement un comportement rationnel, dans le second cela est impossible puisque la rationalité est relative aux compétences cognitives de l'individu. Dans le premier cas la rationalité est calculatoire, puisque l'individu raisonne à objectif donné et dans des conditions d'information parfaite, elles aussi données. Dans le second, la rationalité n'est pas uniquement calculatoire et elle n'est plus aussi paramétrique, puisque l'individu doit percevoir *lui-même* l'objectif et construire *lui-même* les conditions de son choix. Il doit donc délibérer avant de pouvoir éventuellement calculer.

112- La tentative alternative de la théorie du "search"

Développée parallèlement aux analyses de Simon, la théorie du "search" peut paraître constituer une tentative de conciliation de ces deux points de vue antagonistes. Mettant l'accent sur le coût de l'information, elle vise, à partir de l'article de Stigler [1961], à intégrer le coût que constitue la recherche d'information dans le programme d'optimisation. La conduite rationnelle optimale est alors celle qui égalise le coût marginal

que représenterait l'obtention d'une information supplémentaire, et le gain marginal que cette information permettrait d'obtenir. Lorsque ce seuil est atteint, il devient rationnel, de façon objective, de stopper la recherche, et de se contenter du point d'aboutissement auquel les recherches antérieures nous ont mené.

Simon perçoit rapidement le risque que représente cette tentative pour sa propre théorie : “Elle vise non seulement à prendre en compte le caractère limité de la rationalité humaine, mais aussi à la maintenir dans le domaine de la rationalité calculable” (Simon, 1978, p.453). Cette tentative tend en effet à réintroduire le principe de rationalité limitée dans le giron de la rationalité optimisatrice et calculatoire, en quantifiant les limites cognitives des individus, à travers une mesure du coût de l'information. Simon s'oppose à cette tentative, en dénonçant son irréalisme qui consiste à prêter aux individus des capacités de calcul identiques sinon supérieures à celle du modélisateur. Il faut en effet que l'individu, avant d'agir, puisse envisager toutes les possibilités dont il dispose pour obtenir de nouvelles informations, avant de décider de stopper ses recherches. De plus, même si l'on accepte cette représentation, il faut supposer que l'individu connaît, dès le départ, le coût de la recherche de n'importe laquelle des informations. Cette hypothèse, bien que très forte, est indispensable. Dans le cas inverse en effet, l'individu devrait rechercher les coûts d'une information supplémentaire, avant de rechercher l'information elle-même. Une telle recherche “préalable” à la recherche d'information est cependant, elle aussi, susceptible d'être évaluée. On se situe donc à un second niveau de réflexion sur le coût de l'information : le coût de la recherche permettant d'évaluer le coût de la recherche d'une information supplémentaire. Ce processus enclenche une régression *infinie* : il faut en effet, avant de rechercher les coûts que constitue cette recherche “préalable” mesurer le coût de la recherche “préalable à cette recherche préalable” (niveau 3), qui devient à son tour un objectif de recherche qu'il faut quantifier (niveau 4)... L'individu rationnel n'entame donc jamais de recherche : il se perd dans une quête infinie du coût de la recherche “initiale”³.

Pour contrer la théorie du “search”, Simon va chercher à préciser le contenu de l'hypothèse de rationalité limitée⁴. Il cherche en particulier à mettre en évidence la dimension temporelle du processus de prise de décision rationnelle, de manière à dénoncer plus explicitement le *principe* de l'optimisation. Dans cette nouvelle optique, la rationalité cesse d'être définie au regard de la rationalité parfaite, et désigne explicitement un *autre* comportement rationnel *irréductible* à l'optimisation mathématique.

³ La mise en forme de cette critique est due à Mongin et Walliser [1987]. Pour une présentation détaillée mais simplifiée, voir Mongin [1997]

⁴ On suit ici l'analyse de Mongin [1984] et Quinet [1990].

12- La rationalité procédurale

“Dans un tel univers [celui des organisations], nous ne devons pas seulement nous intéresser à la rationalité “substantielle” – c’est-à-dire au choix, dans une certaine mesure, d’un comportement approprié. Mais aussi à la rationalité procédurale – c’est-à-dire à la procédure effectivement choisie pour sélectionner les actions, compte tenu des capacités et des limites cognitives des êtres humains” (Simon, 1978, p.452).

Le concept spécifique de rationalité que Simon élabore doit permettre de qualifier le comportement rationnel non plus en fonction du résultat de la délibération rationnelle, mais en fonction de la *forme* de la délibération elle-même. Il oppose ainsi à l’hypothèse traditionnelle de rationalité optimisatrice, qu’il qualifie de *substantielle*, un mode de rationalité de type *procédural*⁵. Cette modification conceptuelle n’est pas une simple précision. Elle désigne en fait un nouveau champ de recherche et implique un *“changement fondamental dans le style scientifique”* (Simon, 1976b, p.442). Simon entend en effet éclairer un nouvel aspect du comportement rationnel : celui de la délibération qui précède le choix effectif, et qui consiste en un recensement, puis une interprétation des informations pertinentes. Cet éclairage constitue donc un glissement dans l’analyse *“depuis l’étude du raisonnement déductif, associé à une axiomatique précise, jusqu’à la mise en évidence de l’exploration empirique détaillée des complexes algorithmes de la pensée”*. (Simon, 1976b, p.442).

En proposant le concept de rationalité procédurale, Simon entend préciser que l’idée de rationalité limitée ne désigne pas une forme faible de rationalité optimisatrice (substantielle), mais une conception différente de la rationalité. Dans un univers transparent, où l’individu peut tout connaître sans coût, comme c’est le cas sur le marché Arrow-Debreu, on peut juger directement de la rationalité d’un acteur à partir de son comportement, compte tenu des caractéristiques objectives de l’environnement. Lorsque l’environnement est plus complexe, que l’individu ne peut pas “tout savoir”, une détermination externe objective n’est plus possible. Il faut alors étudier le raisonnement de l’acteur pour repérer la cohérence qui existe entre les prémisses de son raisonnement, ce qu’il sait, et ce qu’il en déduit. Il faut, de plus, rendre compte de la manière dont l’individu recense les informations qui forment les prémisses de son raisonnement, c’est à

⁵ *“Un comportement est substantiellement rationnel lorsqu’il permet d’atteindre un but donné dans les limites imposées par des conditions et des contraintes données. (...) Un comportement est rationnel de façon procédurale, lorsqu’il est le résultat d’une délibération appropriée.”* (Simon, 1976b, p.130 et 131)

dire comprendre comment l'individu décrypte l'univers qui l'entoure. L'individu fait preuve de rationalité en opérant des rapprochements entre les informations dont il dispose, ou encore en repérant des analogies entre des situations récurrentes, ce qui lui permet d'économiser des ressources cognitives en activant des "routines". L'ensemble des procédures cognitives qui permettent à l'individu de ramener une situation complexe à une alternative décidable constitue le cœur de sa rationalité.

Cette conception de la rationalité est proche de celle que nous avons repérée dans le *Treatise* de Keynes. Simon souligne ainsi : "*la rationalité d'une action implique qu'elle dérive logiquement d'un ensemble de prémisses valides*". (Simon, 1964, p.406). Comme Keynes, Simon met donc l'accent sur la manière dont les individus constituent les prémisses de leur raisonnement, puis déduisent logiquement leur comportement de ces prémisses. Les deux auteurs partagent donc, à cinquante années de distance, une même volonté de faire entrer dans la conception du comportement rationnel les rapports que l'individu entretient avec l'environnement externe. Cet objectif commun les conduit à rompre avec une conception purement calculatoire de la rationalité, et à mettre en évidence l'importance des procédés cognitifs par lesquels l'individu rationnel parvient à se donner une représentation adéquate d'un environnement complexe qu'il ne peut de toute façon pas connaître parfaitement⁶.

Cependant, à la différence de Keynes, Simon entreprend de fonder empiriquement cette conception de la rationalité. Il s'agit alors de décrire les procédés cognitifs *effectivement* mis en œuvre par les individus. L'idée de rationalité procédurale ouvre donc les sciences économiques aux sciences cognitives et à la psychologie⁷. Simon cherche, en quelque sorte, à redonner un fondement empirique à la représentation économique "abstraite" de la rationalité individuelle. Dans un texte récent, il souligne ainsi qu'il est nécessaire de "*construire une théorie adéquate et empiriquement fondée de la rationalité limitée*". (Simon, 1997, p.63). Cette conception empirique de la rationalité est un point de divergence important entre la réflexion keynésienne et la réflexion simonienne, dans la mesure où elle conditionne une conception différente de la science économique elle-même. Keynes considère en effet que : "*l'économie est essentiellement une science morale et non une science naturelle. Cela signifie qu'elle fait usage de l'introspection et des jugements de valeur*". A l'inverse, Simon cherche à en faire une science empirique, sur le modèle des sciences naturelles. La stratégie théorique de Simon, qui vise à "psychologiser" l'économie vise, *in fine*, à la doter d'une représentation

⁶ Ainsi, Simon [1983] met en évidence la nature "intuitive" de la rationalité procédurale. C'est là un point de convergence supplémentaire avec l'approche keynésienne.

⁷ On retrouve là l'apport d'auteurs comme Kahneman et Tversky qui permettent d'étudier empiriquement les procédés précis que les acteurs utilisent en pratique.

empiriquement juste des procédés cognitifs mis en œuvre par les individus.

Simon propose donc une théorie de la rationalité permettant de prendre en compte l'ancrage pratique des décisions des acteurs. Cette tentative vise, entre autres, à forger une représentation empiriquement juste du comportement des individus. Ce savoir empirique n'est pas utile pour une conception abstraite, idéale, de l'économie, mais il est indispensable dans la visée d'une "organisation" optimale des agents. Cet objectif organisationnel est, comme nous allons le voir, un aspect important des conceptions simoniennes.

13- La règle chez Simon.

“La seconde caractéristique saillante de la condition humaine est que nous sommes très dépendants de ceux qui nous entourent. La plus grande partie de ce que nous savons, la plupart de nos savoir-faire, nous l’acquérons grâce à l’environnement social. (...) Nous sommes très réceptifs à l’influence sociale et à la persuasion, je qualifie cette réceptivité par le terme de “docilité”. J’utilise donc le terme docilité dans le sens de capacité à apprendre et à s’éduquer, pas dans l’autre sens possible désignant la “passivité” ou la “douceur”. (...) Il y a une étroite connexion entre la rationalité limitée et la docilité. Si nos capacités computationnelles étaient illimitées, nous n’aurions pas besoin des autres. Plus la complexité du monde réel dépasse nos capacités à connaître et calculer, plus la docilité a de l’importance, car elle nous permet de bénéficier des connaissances collectives et des compétences de notre société.” (Simon, 1997, p.40-41).

La théorie de la rationalité procédurale met l'accent sur la dimension temporelle correspondant au processus effectif du raisonnement⁸. Cet angle de vue permet de mettre en évidence ce qui, dans le raisonnement de l'acteur n'est pas directement lié aux qualités de l'acteur lui-même, mais plutôt à son environnement institutionnel. Chez Simon, l'individu pense dans des règles. Cependant, ces règles sont de nature "organisationnelle" et leur détermination, dictée par un impératif d'efficacité, se situe à un niveau qui n'est pas celui de l'individu mais de l'ingénieur social.

⁸ Lemoigne [1995] dresse l'histoire philosophique de la rationalité limitée et rapproche la rationalité simonienne de l'idée d'ingenium de Vico, contre l'idée de rationalité mécanique de Descartes. Il souligne à cette occasion que la rationalité simonienne est délibérative.

131- Le rôle et la nature de l'environnement institutionnel de l'individu

Le concept de rationalité procédurale permet d'envisager directement l'ancrage de l'individu dans un univers social déterminant largement son "profil cognitif". Dans l'univers standard, contre lequel Simon construit sa propre théorie, le rapport fusionnel entre le mode de rationalité de chaque individu et le "système" marchand permet à l'individu d'exploiter au maximum les potentialités de la situation, sans s'appuyer à aucun moment sur les connaissances des autres. Au contraire, dans l'univers décrit par Simon, l'individu ne peut pas décider seul, il doit mobiliser dans son environnement des éléments collectifs auxquels sont incorporés des éléments de connaissance collective accumulés au cours de l'histoire. Or, à la différence des prix "criés" par le secrétaire de marché walrassien, ces éléments "cognitifs collectifs" sont situés dans le temps et dans l'espace social. L'individu qui les mobilise se trouve donc lui aussi situé dans un environnement institutionnel à un moment historique donné. Et, lorsque cet environnement évolue, il modifie également le comportement des individus. Comme le souligne Simon (1974, p.441), les processus cognitifs mis en œuvre par les individus rationnels évoluent en même temps que la société. Il y a une détermination historique et institutionnelle des "routines et heuristiques" que les individus mettent en place pour décrypter la réalité. L'individu *apprend* à décrypter les situations auxquelles il est confronté. La société dans son ensemble *apprend* elle aussi à adopter une réponse adaptée à certaines situations récurrentes. *"Pour cette raison, les tentatives visant à prédire, et prescrire, les comportements économiques individuels sur la base d'une déduction menée à partir d'une petite série de prémisses inaltérables ne peut qu'échouer, et a effectivement échoué. (...)* La théorie procédurale de la rationalité est compatible avec un monde dans lequel les êtres humains continuent de penser et d'inventer, ce n'est pas la cas d'une théorie substantielle de la rationalité." (Simon, 1976b, p.441)

La théorie de la rationalité procédurale envisage donc le rôle de l'environnement collectif institutionnel dans le processus de décision raisonnable de l'individu. Autrement dit, une telle théorie associe le choix rationnel à l'utilisation de règles collectives. La question qui se pose cependant est de savoir d'où viennent ces règles, comment elles émergent et évoluent, qui les met en place. Or, Simon n'apporte pas sur ce point une réponse plus convaincante que Williamson. Ce n'est qu'à travers le prisme d'une théorie de l'organisation que l'on peut trouver chez Simon une réflexion sur le mécanisme de détermination des règles collectives. Dès lors, s'il existe une "Histoire" de l'environnement institutionnel déterminant les procédés cognitifs mis en œuvre par les individus, c'est celle de la détermination progressive de techniques de plus en plus efficaces. C'est une forme d'histoire naturelle au sens "hayekien" d'un progrès continu

dans la recherche d'efficacité⁹. C'est l'histoire d'une adaptation progressive des structures institutionnelles aux lacunes cognitives des individus. Simon ne cherche cependant pas à forger une représentation temporelle, de type évolutionniste, de la genèse et de la dynamique des règles collectives. Sa théorie de la rationalité peut sans doute justifier une perspective évolutionniste, mais ce n'est pas de cette manière que Simon cherche à poser la question de l'émergence. Il ne cherche pas à décrire un processus effectif, mais plutôt, de manière plus classique, à déterminer quelles règles seraient les plus efficaces socialement, compte tenu du mode de rationalité des agents. La réponse que Simon apporte à la question du choix et de la dynamique des règles est donc une réponse organisationnelle, qui, comme chez Williamson, est marquée par deux *a priori* : le primat de l'efficacité et le rôle prégnant de la hiérarchie¹⁰.

132- Une conception instrumentale des règles

Simon fut un théoricien de l'organisation, et plus précisément de l'administration, avant d'être un théoricien de la rationalité individuelle. C'est à partir de son étude du fonctionnement de l'administration qu'il élabore pour la première fois l'idée de rationalité limitée (Simon [1945]). Sa théorie du "comportement administratif" repose en effet sur le constat des "limites" de l'individu. Prenant acte de ces limites, l'ingénieur social doit parvenir à mettre en place une organisation collective efficace¹¹. La possibilité d'une telle organisation repose sur ce que Simon (1945, p.3-19) qualifie de théorie de "*l'équilibre organisationnel*" : l'organisation permet à l'individu, parce qu'il est limité, d'obtenir plus que ce qu'il aurait seul, et l'individu contribue à l'efficacité globale de l'organisation. Il faut donc que la contribution de chacun à l'efficacité collective soit équivalente aux avantages que l'organisation collective apporte à l'individu.

Cette théorie, que Simon applique à l'administration (Simon [1945]) et qu'il étend ensuite à l'ensemble des organisations (March et Simon, [1958]) est théorisée de façon abstraite à travers l'idée de *complexité*. Simon montre en particulier que la complexité requiert une solution de type *hiérarchique*¹². La réflexion de Simon sur le mode de

⁹ En effet, selon Hayek : "La raison pour laquelle une règle plutôt qu'une autre aura été adoptée et transmise sera que le groupe qui l'a adoptée en sera devenu plus efficace." (Hayek, 1976, t.2, p.5).

¹⁰ Ainsi Simon souligne clairement que la genèse des règles est le fait de l'autorité hiérarchique : "Le rôle principal de l'autorité est de coordonner les comportements en produisant des routines et en édictant les règles du jeu." (Simon, 1991, p.39).

¹¹ "L'individu est intentionnellement rationnel, mais seulement de façon limitée. C'est seulement sur la base de cette limite qu'il existe une place pour une théorie de l'organisation et de l'administration." (Simon, 1957, p.xxiv). On a vu que c'est cet aspect, et uniquement celui-là, que Williamson retient de la théorie simonienne.

¹² Simon [1962] expose sa théorie de la complexité qui découle de son étude des administrations et des organisations, et préfigure son ouvrage sur les sciences de l'artificiel (Simon, 1969). Il souligne que "la

fonctionnement des organisations collectives débouche ainsi sur un ouvrage théorique important exposant l'idée d'une "*science des systèmes*" (Simon [1969]). Il n'est pas utile ici de chercher à résumer, d'une façon qui serait inévitablement réductrice, une réflexion et un ouvrage fondateur dans la thématique aujourd'hui très vaste de "la gestion de la complexité". La stratégie qu'expose Simon vise à résoudre de façon optimale le problème des décisions complexes, en prenant comme hypothèse de travail une représentation de l'individu comme "*un système animé*" (Simon, 1969, p.41). Dans cette optique le travail de l'ingénieur consiste à concevoir les contours d'un "artifice" permettant de gérer "au mieux" cette complexité. Cette "*science de la conception*" s'adresse donc, comme le souligne clairement la postface de Lemoigne [1974] dans l'édition française, à "l'ingénieur" qui cherche à organiser la société. Cette conception du rôle des institutions organisationnelles est d'ailleurs clairement exprimée par Simon [1945, p.91] dès son premier ouvrage : "*en déterminant largement les structures mentales de leurs membres, ces institutions fixent les conditions qui assurent la docilité et par conséquent la rationalité dans la société humaine*".

Il s'agit donc là d'une conception explicitement instrumentale des institutions et des règles. Simon [1945, chap.3] souligne d'ailleurs que l'ensemble de cette théorie s'appuie sur la possibilité de distinguer clairement les "*valeurs*", qui déterminent les fins de l'organisation et les "*faits*". L'étude scientifique de ces faits sert de fondement à une théorie rationnelle des "*moyens*" permettant de satisfaire au mieux ces fins collectives¹³. La rationalité dont parle Simon ne traite donc que des "*moyens*" et se situe dans l'optique d'une recherche d'efficacité maximum. Les règles qui émergent dans un tel cadre ne sont donc pas des règles contingentes, issues d'une délibération collective, elles s'imposent par le biais d'une réflexion rationnelle visant l'efficacité.

S'il existe une théorie simonienne de la règle, c'est à travers une théorie de la conception rationnelle "d'artifices" efficaces conçus par des ingénieurs. Les règles n'émergent pas des interactions, elles sont mises en place par des "savants" chargés de l'efficacité collective. On peut opposer à cette théorie de la règle une partie des remarques que nous avons formulées à l'encontre de l'analyse de Williamson¹⁴. Le cadre

complexité prend fréquemment une forme hiérarchique. La hiérarchie est l'un de schémas centraux que les architectes de la complexité utilisent." (Simon, 1962, p.468). Sa réflexion vise à montrer que le mode hiérarchique est le mode le plus efficace de résolution de la plupart des problèmes complexes, sur le modèle des algorithmes mathématiques.

¹³ On retrouve ici la position épistémologique de Simon, qui est à l'opposé de celle de Keynes. C'est une position popperienne que revendique Simon [1945,1976].

¹⁴ Alain Boyer [1994] souligne que la conception simonienne de l'agent "*tend à en faire un automate programmé, instruit pour ainsi dire de l'extérieur par des règles routinières. On pourrait risquer la thèse selon laquelle plus un comportement constitue une action authentique, moins le modèle "behavioriste" de Simon sera pertinent pour en rendre compte.*" Roger Frydman développe également une brève critique de Simon remarquant que l'accent mis sur le caractère systématique du comportement des individus présente

organisationnel de la théorie de la règle de Simon conditionne celle-ci à un objectif technique d'efficacité et en fait le produit de l'action d'une hiérarchie régulatrice. Elle n'opère donc pas véritablement de lien entre le comportement de l'individu et les règles qu'il utilise, qu'il respecte, mais qu'il ne produit pas lui-même. Il est possible que les individus innovent en découvrant des solutions efficaces, généralisées ensuite à l'ensemble de l'organisation sous forme de règles. Mais, l'innovation est étroitement circonscrite aux objectifs de l'organisation. Simon envisage donc clairement le rôle des repères collectifs, mais n'envisage pas qu'ils puissent être véritablement choisis et produits par les individus.

En s'appuyant sur la théorie de Simon, l'économie des conventions mobilise une théorie de la rationalité qui permet d'envisager le rôle des règles. Pour autant, elle ne peut simplement être une exégèse de la pensée simonienne. Simon n'envisage en effet la règle qu'à travers l'idée d'organisation rationnelle, ce qui, comme chez Williamson, laisse encore ouverte la question de l'émergence de cette "organisation rationnelle". Chez Simon, l'individu est situé "historiquement", dans une structure institutionnelle donnée. Mais, cette situation est le fruit, et évolue en fonction, des "calculs" de l'ingénieur social qui conçoit au mieux les règles collectives. La règle n'émerge donc pas à partir des actions des individus. Son origine se situe à un niveau qui lui est inaccessible et qui se situe, en fait, hors du temps. L'économie des conventions cherche précisément à dépasser cet aspect de la théorie de Simon. Elle tente de comprendre, à partir d'une hypothèse de rationalité procédurale, comment les enchaînements des actions des individus peuvent donner naissance à des règles, dans un processus véritablement historique.

2– Le projet conventionnaliste

La constitution¹⁵ d'une économie, ou d'une école, des conventions s'est concrétisée par la parution successive de trois ouvrages collectifs en 1986, 1989, et 1994¹⁶. Elle s'est accompagnée d'une notoriété croissante pour cette école qui, à partir de cercles restreints (le C.E.E, l'I.N.S.E.E., le C.R.E.A.), est parvenue à rassembler un collectif de chercheurs importants et prestigieux, et à assurer une large diffusion de ses thèses, en France comme à l'étranger (comme en témoigne l'article de Wagner[1994]). Cette montée en puissance du courant conventionnaliste n'est pas le fait du hasard. Ces auteurs¹⁷ se sont en effet efforcés de proposer un message, le principe du dépassement de l'opposition entre holisme et individualisme, qui soit clair et percutant. Ils ont incarné ce message en se constituant en école de pensée et en se choisissant un titre. Ils ont enfin et surtout choisi un habile positionnement stratégique vis-à-vis du paradigme économique dominant, proposant une critique sans rupture qui leur a permis d'être critiques sans être immédiatement mis au ban de la communauté scientifique. Favereau rappelle le principe de cette posture critique : *“la poursuite du programme de recherches traditionnel de l'individualisme méthodologique en économie (c'est à dire la question de la coordination des activités économiques), mais avec une hypothèse de rationalité limitée”* (Favereau, 1993, p.2).

Le succès des thèses de l'économie des conventions s'explique cependant surtout par la force novatrice de son projet. L'apparition de la problématique conventionnaliste a suscité un grand espoir chez de nombreux chercheurs qui ont vu en elle la promesse d'un renouvellement de l'analyse de la coordination économique passant par une prise en compte accrue des facultés de l'acteur. Plutôt que d'expliquer l'action économique en la réduisant à l'application de règles et de normes préexistantes, elle semblait devoir s'attacher à analyser l'émergence de ces règles et conventions économiques en les appréhendant comme autant de repères créés et modifiés en permanence dans et par l'action. C'est à l'aune de cette promesse dont le projet conventionnaliste était porteur que

¹⁵Cette section est parue sous la forme d'un article dans la *Revue Economique* (Postel [1998]).

¹⁶ On pourrait ajouter à cette liste deux autres ouvrages communs : Dupuy, Livet, Reynaud [1997] et Chatel, Rivaud-Danset, Salais [1998]. Cependant, à la différence des trois premiers, ces deux ouvrages ne sont pas spécifiquement consacrés à l'économie des conventions. Ils s'inscrivent dans la mouvance de l'économie des conventions, plus qu'ils ne participent de sa définition.

¹⁷Nous désignons sous ce terme les auteurs qui constituent les piliers de “l'économie des conventions”. Il s'agit notamment de : JP Dupuy, F. Eymard Duvernay, O. Favereau, A. Orléan, R. Salais, L. Thévenot, B. Reynaud. Il conviendrait d'ajouter à ce noyau dur une liste de plusieurs “électrons libres” qui participent plus ou moins régulièrement et depuis plus ou moins longtemps à l'économie des conventions. Une telle liste serait cependant pléthorique et forcément partielle, car la liste des collaborateurs à l'économie des conventions est en perpétuelle évolution et extension.

nous tentons, une dizaine d'année après son apparition, de mesurer les avancées réelles que le cadre analytique des conventions a rendues possibles.

Pour cela, nous chercherons à montrer, dans ce premier temps, que le projet conventionnaliste se compose à l'origine de trois exigences: (1) reconnaître le caractère construit du marché ; (2) étudier des modes de coordination différents du mode marchand ; (3) affirmer que le fondement des règles et institutions nécessaires à la réussite de la coordination n'est pas naturel mais historique ou négocié.

Il faut, pour satisfaire cette troisième exigence, doter les acteurs d'une capacité à créer ou à faire évoluer les règles et institutions qui leur permettent de se coordonner, or la prise en compte de cette capacité implique d'élargir l'hypothèse standard de rationalité individuelle. Nous montrerons, dans un second temps, que cet élargissement représente un obstacle théorique que l'économie des conventions a eu tendance à contourner en privilégiant les deux premiers points de son projet, au détriment du troisième. La question de la dynamique des règles n'est principalement envisagée qu'à travers le prisme des sciences cognitives, et reçoit une formalisation qui s'inscrit dans le cadre de la théorie des jeux. Cette évolution des propositions de l'économie des conventions en ce qui concerne la genèse des institutions et la dynamique des règles tend à recentrer son discours sur les aspects techniques et stratégiques de la coordination. Ce recentrage menace de gommer peu à peu ce qui faisait la force de son projet : l'analyse des rapports entre règle et action économique.

21- Le projet conventionnaliste

Le programme de recherche de l'économie des conventions est présenté à deux reprises. On en trouve une ébauche sous la plume de Thévenot [1986], puis une présentation plus précise et plus ramassée dans l'introduction commune de la *RE*. Ce projet s'inscrit dans la continuité des critiques formulées par Keynes (Favereau[1988]) et Simon (Favereau [1998]) à l'encontre de l'analyse classique de la rationalité et de la coordination. Il a cependant une spécificité importante qui tient à son rejet explicite des deux grandes formes de réductionnisme que sont l'individualisme et le holisme méthodologique. Ce rejet signe la volonté de produire une théorie de la règle accrochée à une véritable théorie de l'action.

211- La relativisation du marché

Dans sa présentation de l'économie des conventions, Olivier Favereau insiste sur le concept de rationalité limitée, jusqu'à en faire l'emblème du courant conventionnaliste¹⁸. Or ce concept recouvre, on l'a vu, des hypothèses très variées sur le comportement prêté aux acteurs économiques. Il apparaît en particulier que le fait de reconnaître que l'acteur est limité dans sa maîtrise de l'information, parce qu'il ne parvient pas à se la procurer ou parce qu'il ne parvient pas à la traiter, n'implique pas nécessairement de considérer que l'acteur poursuit un but plus large que la stricte satisfaction de ses intérêts (même si cela nécessite un effort de coordination). En d'autres termes, la rationalité limitée peut être une rationalité exclusivement instrumentale.

Que serait le projet conventionnaliste s'il s'attachait exclusivement à une telle conception de la rationalité limitée ? Il retrouverait les analyses développées notamment par les néo-institutionnalistes américains, et en particulier par Oliver Williamson. Ce courant défend l'idée qu'à côté du marché existe un autre mode de coordination, l'entreprise, qui peut s'avérer plus efficace notamment sous l'hypothèse de rationalité limitée. On peut le caractériser par deux exigences qui s'appliquent aussi au projet conventionnaliste :

- 1- Faire apparaître le caractère non naturel et non universel du marché walrassien.
- 2- Étudier le fonctionnement d'autres modes de coordination alternatifs aux marchés.

Ces deux principes de recherche dessinent, un projet critique vis-à-vis du cadre marchand de la théorie standard mais non vis-à-vis de ses racines utilitaristes. Ils sont constitutifs de la "Théorie Standard Etendue", selon la terminologie de Favereau [1989]. La "TSE" partage avec la théorie standard une même conception de l'individu, qui n'est perçu qu'à travers la maximisation de son intérêt. Elle intègre cependant une dimension institutionnelle au cadre marchand de la théorie standard. L'organisation est ainsi présentée comme étant une alternative efficace à la coordination marchande, notamment en situation d'incertitude.

Notre présentation de l'analyse de Williamson nous a permis de cerner quelques limites de cette tentative. Elle rencontre des problèmes qui ne sont pas sans rapport avec le problème déjà soulevé par Hobbes. Pour être efficace, la régulation contractuelle doit s'appuyer sur une structure institutionnelle plus large, qui spécifie le droit, réduit l'incertitude et limite les interprétations possibles du contrat. Le contrat est efficace

¹⁸Il est juste de préciser que Favereau qualifie ainsi l'entrée "économique" dans le projet conventionnaliste, et qu'il identifie aussi une entrée sociologique fondée sur "*les exigences d'une méthode compréhensive*" (Favereau, 1995, p.104). C'est, de façon naturelle, l'entrée "économique" que nous suivons ici.

lorsqu'il est protégé de la contingence inhérente aux affaires humaines par une structure hiérarchique. Celle-ci règle de fait les principaux problèmes posés par la coordination des activités. L'intégration de l'institution organisationnelle au paradigme individualiste nécessiterait donc d'endogénéiser cette structure hiérarchique sans pour cela invoquer de mécanismes holistes.

C'est précisément ce que tentent les conventionnalistes qui cherchent à dépasser le diptyque du contrat et de la contrainte en s'appuyant sur l'action individuelle pour expliquer la constitution des cadres collectifs de coordination.

212- La critique de la figure réductrice de l'*homo œconomicus*

Cette volonté conduit au rejet simultané des deux méthodes traditionnelles d'explication en sciences sociales que sont l'individualisme méthodologique et le holisme méthodologique¹⁹. Les tenants d'une méthode individualiste reprochent à la méthode holiste de ne pas prendre en compte l'action individuelle, qui est pourtant le seul niveau d'étude pertinent. Le principe de la démarche conventionnaliste est de montrer en retour que l'individualisme méthodologique, dans sa version walrassienne, conduit à un réductionnisme similaire : le refus de prendre en compte le rôle des structures collectives conduit les tenants de cette méthode à se rabattre sur une perception réductrice de l'action individuelle.

Cette conviction se nourrit de plusieurs travaux qui constituent en quelque sorte un préalable à l'analyse conventionnaliste. C'est le cas de l'analyse de Berthoud [1988] qui a montré que l'approche walrassienne reposait en fait sur un système holiste : les individus se déchargent de toute responsabilité sur le crieur, et peuvent ainsi jouir de la liberté "d'être quitte de tous les autres". Dans la même optique, Frydman[1992] rappelle que le crieur constitue "une aporie, propre à tout modèle de socialisation qui entend faire surgir l'accord global du choix libre d'agents" (Frydman, 1992, p.10). L'*homo-œconomicus* néoclassique a la particularité d'être en phase avec l'ordre marchand auquel il participe. De ce fait l'ensemble des "individus" correspondent au même moule, ils n'existent qu'en étant à l'image du marché. L'approche qui consiste à étudier le comportement de tels individus n'a d'individualiste que le nom : derrière l'individu, c'est la machine marchande qui est étudiée. L'individu n'est animé que par les forces de cette machine : il

¹⁹Le rejet de deux hypothèses classiques permet de se démarquer de l'individualisme néoclassique : "La première hypothèse rejetée stipule que les relations marchandes et les contrats d'échange entre les personnes les affranchissent de toute référence extérieure dépassant le cadre de leur volonté." (RE, p.142). "La seconde hypothèse [rejetée] consiste à considérer que le cadre néoclassique peut s'étendre continûment à l'analyse de relations non strictement marchandes." (RE, p.143). La reconnaissance du principe de Popper-Agassi ("selon lequel on ne peut prêter à des objets collectifs ni objectifs ni intérêts") institue le refus du recours à une approche holiste en remplacement du cadre utilitariste.

ne s'appartient pas.

A l'inverse, les conventionnalistes cherchent à faire de l'individu un précédent à l'ordre marchand. Le marché, loin d'être naturel, est perçu comme une construction collective et permanente des individus qui y participent. Ce refus de la référence à un ordre transcendant les conduit à renouer le lien entre l'analyse de l'action individuelle et l'existence d'un ordre collectif. Le renouvellement de la réflexion sur les compétences de l'acteur économique devient alors nécessaire puisque l'on cherche à comprendre ou à expliquer comment, à partir de l'individu, se constituent des mouvements collectifs qui se réifient en norme, règle, institution qui sont autant de formes de conventions²⁰. L'économie des conventions apparaît ainsi comme une tentative pour constituer une grammaire générale de la régulation de nos économies dans une perspective dynamique fondée sur l'action. Plutôt que d'étudier un nouveau mode de coordination qui serait la convention, il s'agit de proposer pour toutes les modalités de coordination un formalisme, une grammaire commune, qui pourrait constituer un paradigme novateur à même de fédérer des chercheurs d'horizons différents²¹. Le pivot de cette démarche c'est l'acteur, et sa capacité à créer de nouvelles façons d'échanger et de produire à plusieurs.

Cette réflexion sur la figure de l'acteur explique la posture individualiste revendiquée par les auteurs dans leur entreprise de dépassement de l'opposition entre holisme et individualisme. Il s'agit toujours d'individualisme, mais ce n'est plus le même individu : l'économie des conventions se construit aussi contre la figure walrassienne de l'individu, en prônant une prise en compte accrue des facultés de l'acteur²². La relativisation du marché et l'étude de modes de coordination alternatifs ne couvrent donc qu'une partie du programme conventionnaliste, et il convient de compléter ces deux premiers objectifs par un troisième :

3- Affirmer que le fondement de ces conventions n'est pas naturel mais historique ou négocié : qu'il est l'objet d'une délibération humaine et n'a aucun caractère de nécessité logique. Ceci implique de rompre avec le positivisme et le naturalisme pour renouer avec

²⁰ C'est ce que souligne la définition commune de la convention, dans l'introduction de la RE, p. 143 : *"Le terme de convention qui désigne le dispositif constituant un accord de volonté tout comme son produit, doté d'une force normative obligatoire, doit être appréhendé à la fois comme le résultat d'actions individuelles et comme un cadre contraignant les sujets."*

²¹ Le discours de l'économie des conventions est avant tout méthodologique comme l'annonce Thévenot lorsqu'il présente, pour la première fois un programme de recherche conventionnaliste : *"ce programme devrait conduire à l'élaboration d'un cadre théorique original ne relevant d'aucune de ces disciplines et permettant de comprendre dans une formulation commune la composition de ces divers univers de référence . (...)"* (Thévenot, 1986, p.198). Favereau (1994b, p. 135) le rappelle avec insistance : *"L'économie des conventions se caractérise par une méthode générale d'approche des coordinations organisationnelles, plutôt que par l'analyse d'une classe particulière de tels phénomènes"*.

²² Les auteurs soulignent dans l'introduction de la RE qu'ils porteront *"attention à la compréhension et aux attentes des acteurs"*(p.142). Ce point est notamment développé par Laurent Thévenot dans le premier chapitre du livre qu'il a écrit en commun avec Boltanski (Thévenot et Boltanski [1991]). Ils refusent de considérer que *"la règle est une loi scientifique qui s'applique aux personnes comme aux choses"*. Thévenot [1990] propose au contraire de faire *"une théorie de l'action"*.

un certain pragmatisme. Dès que l'on se situe hors d'un modèle systémique objectivé, l'acte de se coordonner ne peut plus être considéré comme un phénomène mécanique soutenu par une rationalité instrumentale illimitée.

Le projet conventionnaliste est donc ambitieux et radicalement critique vis-à-vis de l'économie standard. Il nécessite en particulier d'élargir l'hypothèse de rationalité instrumentale. Cette radicalité est toutefois en contradiction avec la prudence dont font preuve les auteurs dans leur positionnement théorique. Cette contradiction est révélatrice d'un problème de fond : l'économie des conventions, de par son projet, ne s'intègre dans aucun corpus théorique déjà constitué. Or l'ambition légitime des conventionnalistes est de proposer une théorie rigoureuse, qui puisse être confrontée au formalisme standard. Cette volonté implique de reformuler plusieurs concepts de manière originale, à commencer par ceux qui nous semblent être cruciaux : les concepts de rationalité et de convention.

22- Brève histoire de l'économie des conventions

Le premier souci des auteurs conventionnalistes a été d'harmoniser leurs analyses et de proposer des concepts de convention cohérents entre eux. Devant les difficultés qu'ils rencontrent, ils tendent à se retrouver sur des champs théoriques déjà formalisés, comme les sciences cognitives ou la théorie des jeux. Or ces champs théoriques ne leur permettent pas de développer le troisième point de leur projet, qui marquait pourtant l'originalité de leur démarche. Ainsi, pour parvenir à s'organiser et à disposer d'un formalisme commun, les conventionnalistes tendent à affaiblir leur propos. Nous suivrons ces difficultés et cette brève histoire de l'économie des conventions à travers la trajectoire de deux auteurs emblématiques, Olivier Favereau et Laurent Thévenot.

221- Olivier Favereau : la convention comme instrument entrepreneurial

Olivier Favereau propose dès les premiers textes conventionnalistes, une formalisation rigoureuse de l'idée de convention. Il propose, dans sa contribution au premier ouvrage commun (Favereau [1986]), une approche constituée de deux types de conventions : convention 1 (C1) et convention 2 (C2). C1 s'inspire de l'approche proposée par Lewis[1969]. La convention que Keynes repère sur les marchés financiers et selon laquelle "l'état présent des affaires persistera" en constitue l'exemple typique. Ce type de

convention émerge spontanément. Ce n'est pas le cas de C2 qui désigne le type de règle que l'on rencontre par exemple dans les marchés internes du travail formalisés par Piore et Doeringer [1971]. Ce type de convention est une construction hiérarchique, fondée sur la "relation d'autorité". Favereau précise qu'elle "*correspond (...) à la part obligatoire de chaque rôle qui se retrouve dans tous les mondes possibles*" (...). "*Il s'agit cette fois de situations où les individus se définissent par ce qu'ils font avant de se définir par ce qu'ils pensent. Les comportements sont régulés de l'extérieur (...).*" (Favereau, 1986, p.62,64). C2 correspond donc clairement à l'approche instrumentale, et fonctionnaliste, de la règle ou de la convention.

L'analyse proposée dans le numéro spécial de la Revue économique [1989] revient sur cette construction binaire entre règles interprétatives harmonisant les anticipations et règles fonctionnelles harmonisant directement les comportements. En proposant un concept unique, le "Dispositif Collectif Cognitif" (désormais : DCC), Favereau privilégie l'aspect C2 sur l'aspect C1. Ainsi le DCC s'appuie sur une relation d'autorité (p.297) et répond à un "savoir comment" sans disposer du "savoir pourquoi"²³. L'aspect cognitif souligne que le DCC vient remplir les lacunes d'acteurs à rationalité limitée, l'aspect collectif souligne que le savoir qu'il concentre est un savoir collectif, qui n'appartient pas à un individu omniscient mais à tous les individus qui font l'organisation. Un aspect du DCC limite cependant sa collégialité : bien qu'il émerge des interactions de tous les salariés, sa mise en place définitive demeure la prérogative des postes élevés de la hiérarchie²⁴. La constitution de la convention échappe donc en partie à l'acteur, comme dans l'analyse de Simon.

La principale difficulté de l'analyse conventionnaliste est, on l'a vu, de faire le lien entre acteur et ordre. La difficulté est donc ici en quelque sorte contournée : la question de l'ordre continue d'être réglée à l'avance (l'organisation est toujours déjà constituée hiérarchiquement et l'étude de cette constitution est clairement hors sujet), l'acteur suit peu ou prou la conduite impulsée par les tenants de cet ordre. Ceci pose un problème de cohérence : comment expliquer, si l'on rejette les explications holistes et naturalistes, l'existence d'un ordre qui ne dépend pas de l'acteur ? En fait ce que l'on découvre sous cet aspect du DCC, c'est que cette construction est encore empreinte de fonctionnalisme.

²³Cette approche de la convention nous semble s'approcher de celle de Robert Salais : "*une convention est un système d'attente réciproque sur les compétences et les comportements conçu pour aller de soi et comme allant de soi*" (Salais, 1989, p.213).

²⁴Favereau (1989, p. 298) : "*Le premier cas polaire correspond aux postes inférieurs de la hiérarchie. L'Apprentissage individuel du salarié se réduit au «comment» d'un ensemble de tâches entre lesquelles l'employeur se réserve la faculté de choisir (...). Le salaire intervient alors comme dédommagement de la renonciation à l'intentionnalité.*

Le second cas polaire correspond aux postes supérieurs de la hiérarchie. Cette fois l'Apprentissage individuel est susceptible de déboucher sur une extension du savoir collectif de l'organisation après cristallisation de cet apprentissage individuel dans de nouvelles règles de fonctionnement interne. (...). Le salaire intervient cette fois comme encouragement."

La convention est justifiée par la fonction qu'elle remplit, et détermine elle-même la fonction de chacun des individus. La rationalité des acteurs est cognitivement limitée mais n'en n'est pas moins instrumentale. De ce fait, le problème de la constitution de l'ordre, ou de l'équilibre, se résout toujours par une aporie : ici celle de la hiérarchie. Si l'on se réfère aux trois principes du projet conventionnaliste, on peut conclure qu'Olivier Favereau développe brillamment les deux premiers, mais laisse le troisième en friche.

222- Les conventions comme "formes" de coordination.

La recherche que Laurent Thévenot a menée en commun avec Luc Boltanski et qui a donnée lieu à la parution d'un ouvrage commun (Thévenot et Boltanski [1987 et 1991]), s'appuie sur la notion de forme ou de cité. Nous ne pouvons ici présenter même de manière très résumée un ouvrage important et complexe²⁵, et nous nous contenterons de mettre en avant le type de convention que Thévenot étudie et les problèmes inhérents à son interprétation du projet conventionnaliste.

Le modèle des économies de la grandeur repose sur six principes supérieurs représentant les six conceptions du "juste" faisant l'objet d'un consensus dans la société occidentale actuelle²⁶. Ces principes de justice se sont constitués historiquement et ont acquis une extériorité par rapport aux acteurs, ce qui leur permet de ne plus être soumis à discussion. Ils sont étayés par des objets qui les incarnent et les rendent manipulables par les acteurs. L'accord entre acteurs, condition *sine qua non* de la coordination, est obtenu grâce à la justification par chacun de sa propre action à l'aune d'un principe supérieur commun dont la sollicitation est rendue possible par la présence d'objets qui l'incarnent. C'est le moment de "l'épreuve" où chacun est amené à rendre compte de sa conduite et qui peut aboutir à un différend si chacun sollicite un principe différent pour caractériser la situation qui les réunit²⁷. Les "objets" permettent alors de déterminer "objectivement" qui a "raison" dans la situation présente.

La structure du modèle de Thévenot et Boltanski est donc double. Les conventions de niveau supérieur, c'est-à-dire les principes supérieurs communs, ne sont directement invoquées par les acteurs que lorsque la discussion atteint un haut niveau de généralité. Dans les situations les plus courantes, ces "principes" se manifestent sous une forme plus simple et plus concrète : à travers les "objets" conventionnels qui les incarnent. Dans ce

²⁵ Le lecteur intéressé par cet ouvrage peut consulter, outre la synthèse qu'en propose Thévenot [1989], la présentation très stimulante de N. Dodier [1991] ou encore l'excellente synthèse proposée par Philippe Corcuff [1995].

²⁶ Il s'agit des principes : "marchand", "industriel", "civique", "de l'opinion", "domestique", "inspiré".

²⁷ La légitimité de l'opinion permet d'engager des réformes, mais donne quelquefois lieu à des fraudes... Le fraudeur pourra s'appuyer sur sa cote de popularité, ses détracteurs feront valoir que sa conduite entre en contradiction avec le principe civique.

modèle, c'est donc uniquement à travers l'argumentation, ou la justification qui suit une accusation d'injustice, qu'apparaît la nécessité de se référer à des conventions communes. Le comportement individuel n'est en aucun cas directement régi par des conventions, mais l'impératif de justification est potentiellement toujours présent. Le moment de l'épreuve agit comme un révélateur des principes qui ont guidé l'action passée. La coordination par convention s'appuie donc sur les paroles par lesquelles les acteurs qualifient et justifient leur action. C'est pourquoi les auteurs parlent de grammaire de l'accord social.

Dans ce modèle, l'acteur est doté de compétences excédant le domaine traditionnel de la rationalité instrumentale et approchant le concept traditionnel de "prudence". Il s'agit cependant d'une forme particulière de prudence, "encastrée" dans les formes conventionnelles²⁸. Le fait même que les compétences de l'acteur ne puissent se décliner que selon une liste précise de six principes, nous incite à étudier l'autonomie de l'acteur vis-à-vis de ce cadre conventionnel. La question centrale semble être : l'acteur est-il à l'origine de la convention ? Elle pose d'une manière générale le problème de la dynamique du cadre conventionnel proposé par L. Thévenot. Or, on se heurte sur ce point à un principe de construction du modèle. Les principes de justice tirent en effet leur légitimité de leur extériorité par rapport aux interactions humaines. Ils représentent le "bien commun". Mais cette notion est équivoque : elle revêt un aspect moral sans pour autant faire référence à une validité transcendante. De plus les acteurs n'ont pas de visée critique en ce qui concerne ces principes qu'ils considèrent, à tort, comme naturels. Il est donc difficile d'imaginer qu'un ou plusieurs acteurs puissent soudain accéder à une lucidité qui leur ferait voir la nature comme un "construit", et leur permettrait de se réclamer du bien commun pour fonder un nouveau principe de justice. Le moment de la légitimation d'un principe de justice appartient par construction à un passé mythique : le moment de la genèse est forcément derrière nous. Le sens de la justice reste ainsi désespérément inaccessible à l'homme, à partir du moment où l'on pose le modèle. Par leur extériorité, les principes de justice tendent à devenir, de fait, quasi-transcendants par rapports aux acteurs.

Cet aspect du modèle limite sa portée : l'analyse de la rationalité continue de se faire en référence à un cadre constitué antérieurement. Le lien entre l'ordre et l'action n'est pas renoué. Les principes de justice apparaissent finalement comme une concession,

²⁸ "Cette hypothèse [le cadre des économies de la grandeur] conduit à replacer la rationalité intéressée au sein d'une pluralité de justifications possibles de l'action : c'est pourquoi nous préférons le terme de justifiable ou de raisonnable au terme de rationnel qui s'accorde plus difficilement à cette multiplicité" (RE, p. 148). "Nous ne pouvons pas limiter la notion de rationalité à une capacité au calcul, mais devons l'inscrire dans le cadre plus général du jugement sur l'action". (Thévenot [1993], p. 30). "L'exigence d'épreuve modifie l'optique de notre examen qui, de l'étude des constructions de la philosophie politique passe à celle de raison pratique ou, pour remonter plus avant dans la tradition, à celle de prudence". (Thévenot et Boltanski, 1991, p. 59.).

minimale, au holisme méthodologique²⁹. Bien que l'approche de Thévenot prenne en compte les trois points du projet conventionnaliste, elle ne parvient pas à proposer une analyse satisfaisante du troisième. Dans cette approche comme dans la précédente, la convention échappe aux acteurs. Son historicité est fictive : une fois passé le moment originel de la constitution des six principes supérieurs communs, le temps suspend son vol...

223- Des problèmes symétriques

La confrontation de ces deux approches de la convention est en elle-même signifiante. Elle l'est tout d'abord parce qu'elle permet de constater que les deux auteurs se heurtent au même problème : celui de la constitution ou de la genèse des conventions. Mais elle permet aussi de souligner que face à ce problème les auteurs empruntent des chemins diamétralement opposés, l'analyse de Favereau reprenant par certains aspects une méthodologie individualiste de type néoclassique alors que le modèle de Thévenot repose sur une structure holiste.

La démarche proposée par Olivier Favereau [1989] vise à étudier et styliser le mode de coordination organisationnel pour pouvoir ensuite l'intégrer et l'imbriquer dans le mode de coordination marchand, mode de coordination exclusif de l'économie standard. Cette présentation met en lumière le mode de régulation des organisations productives. Mais il faut garder présent à l'esprit que le projet conventionnaliste est essentiellement méthodologique. Il s'agit de proposer une nouvelle façon de concevoir la constitution des institutions, en partant de l'individu. Or, sur ce point précis, l'approche en terme de DCC apporte peu de nouveauté par rapport à la démarche walrassienne. L'individu entre en contact avec une totalité, ici l'organisation, qui détermine la façon dont il doit se comporter, en fonction des intérêts "collectifs" de l'organisation. Sa façon d'être rationnel est de rechercher l'efficacité. Or rien n'est dit sur les raisons de l'adhésion de l'individu aux objectifs de l'organisation et au critère d'efficacité qui préside à sa régulation³⁰. Tout

²⁹ Le qualificatif "holiste" désigne ici une démarche supposant l'existence d'une structure hiérarchique s'imposant aux individus sans que ceux-ci aient aucun moyen de la modifier et sans qu'existe une théorie associant la genèse de cette structure à une série d'actes individuels. Pour reprendre le vocabulaire d'Agassi (cité par Defalvard, 1992, p.129) : "*la structure sociale influence les comportements individuels sans que cette structure soit modifiable par les individus*". Nous utilisons donc ici le terme "holiste" dans un sens affaibli par rapport au sens que lui donne Agassi [1973], au risque de lui donner un sens proche du terme de "contrainte". Mais ceci nous permet de souligner en quoi le modèle de Thévenot entre ici en contradiction avec le projet de dépassement de l'opposition entre holisme et individualisme. Nous dirons que le holisme fonctionnel dont a besoin l'économie des conventions est ici peu ou pas relayé par un individualisme morphogénétique (selon les termes de Defalvard [1992], p.128-129). Dans la suite du texte le qualificatif holiste est toujours utilisé dans ce sens.

³⁰ Alors même qu'un des grands apports de Thévenot et Boltanski [1991] est d'avoir présenté l'entreprise comme étant par excellence un univers composite, dans lequel coexistent tous les principes de justice (chap IV). La confusion entre entreprise et monde industriel, qui revient à faire d'un monde abstrait un lieu

se passe comme s'il existait une hypothèse non écrite selon laquelle l'individu suit la loi de l'organisation avant même d'entrer en contact avec elle, comme pour l'individu walrassien : *"la loi de l'agent est celle du système"* (Frydman, 1992, p. 10). Du strict point de vue du traitement méthodologique du comportement individuel, on assiste à une translation du cadre marchand vers le cadre entrepreneurial³¹. L'hypothèse de rationalité demeure instrumentale, même si elle est limitée. L'analyse des conventions reprend donc ici en partie le principe de l'individualisme néoclassique qu'elle voulait dépasser.

A l'inverse, la démarche de Laurent Thévenot, lorsqu'elle se heurte au problème de l'ordre ou de l'équilibre, suit un type de solution holiste. C'est en tout cas ce que semble révéler l'impossibilité de proposer une analyse dynamique des "principes supérieurs" fondée sur l'action individuelle ou collective.

A travers ces deux auteurs, on retrouve donc, à propos de l'articulation entre comportement individuel et régulation collective, l'opposition entre l'individualisme strict de l'analyse néoclassique et la méthodologie holiste. Il leur manque la théorie de la genèse ou de la dynamique des conventions qui leur permettrait d'harmoniser leurs analyses et de s'affranchir de ces deux types de réductionnisme.

Ces difficultés à appréhender de manière dynamique la coordination par convention est révélatrice de l'ambiguïté profonde d'une démarche conventionnaliste qui se cantonnerait aux deux premières exigences de son projet, au détriment de la troisième. La dynamique des conventions repose en effet sur la capacité de l'acteur à s'approprier les règles collectives auxquelles il participe pour les faire évoluer. Cela implique que l'acteur ne perçoive pas les conventions comme naturelles ou transcendantes, qu'il ne s'illusionne pas quant à leur nature contingente et historique. Cette lucidité de l'acteur est la condition nécessaire de l'évolution dynamique des conventions. Faute de développer suffisamment cet aspect de la coordination par convention, les premières analyses conventionnalistes sont condamnées à demeurer statiques et d'une certaine manière divergentes.

concret, va à l'encontre de cette présentation, en posant comme acquis l'unicité du principe de justice régissant l'entreprise, à savoir le principe industriel d'efficacité.

³¹ La réflexion d'Olivier Favereau vise à montrer que même le cadre marchand est organisé. Ceci remet en question la naturalité du cadre marchand et le naturalisme profond de l'analyse économique néoclassique. Mais, avec l'affirmation d'un principe d'organisation unique, l'approche de Favereau ne rompt pas avec le positivisme.

23-De la convention à la cognition

La réponse de l'économie des conventions aux problèmes rencontrés dans l'analyse de la dynamique des conventions tient en un mot : la cognition. L'analyse cognitive devient le point central d'une démarche visant désormais à assouplir le concept de convention en mettant l'accent sur sa faible généralité et son ancrage dans les particularités de la situation. L'analyse des procédés cognitifs des agents devrait ainsi permettre d'étudier la structure et la dynamique des conventions.

Mais que faut-il entendre par cognition ? Ce terme recouvre au minimum trois démarches très différentes dans le champ des sciences sociales³². Les tenants d'un cognitivisme fort cherchent à naturaliser les sciences sociales en remplaçant l'approche classique en terme de "raison d'action" par une approche neuro-psychologique en terme de "cause". Une version plus faible de la démarche cognitiviste consiste à étudier l'activité intellectuelle rationnelle et à mieux la comprendre pour dévoiler, sous des comportements apparemment normés, des démarches individuellement rationnelles. La troisième démarche cognitiviste, plus spécifiquement économique, consiste, dans la lignée des travaux d'Herbert Simon, à abandonner l'idée d'une rationalité parfaite et substantielle pour s'attacher à comprendre et formaliser toutes les étapes d'une réflexion rationnelle et à faire alors porter la question de la rationalité sur le raisonnement lui-même et non plus sur le résultat de ce raisonnement. Le point commun de ces démarches axées sur la cognition est de mettre l'accent sur la rationalité individuelle en cherchant à remplacer le point de vue traditionnellement extérieur que portent les sciences sociales sur l'individu, par un point de vue interne cherchant à expliquer son comportement par son mode de raisonnement rationnel. Le présupposé de ces démarches est donc explicitement rationaliste, ce qui les conduit à s'écarter également de l'herméneutique qui, bien qu'elle favorise aussi le point de vue interne de l'acteur, n'en privilégie pas un traitement analytique. Cela se traduit d'une part par le glissement du concept herméneutique d'interprétation vers celui de cognition qui laisse prise à l'analyse et l'explication, et d'autre part par le remplacement de l'hypothèse de rationalité instrumentale parfaite et substantielle par celui de rationalité limitée et procédurale, mais néanmoins toujours instrumentale³³.

Les approches cognitives peuvent permettre à l'économie des conventions de préciser son propos et de le rendre plus analytique. En revanche l'apport des approches cognitives

³²Les deux premières démarches sont présentées dans l'ouvrage dirigé par Boudon, Bouvier et Chazel [1997].

³³ Sur la complémentarité des notions de rationalité cognitive et instrumentale, on peut consulter Walliser [1995].

en ce qui concerne la dynamique des règles et l'analyse de la constitution historique des institutions paraît limité. L'unique explication cognitiviste de l'émergence des règles serait une réponse naturaliste qui concevrait les règles comme des réponses aux causes neuro-psychologiques des imperfections de coordination. Le cognitivisme à la Simon présuppose l'existence d'un cadre hiérarchique reconnu et accepté, de manière à ce que les problèmes à résoudre soient dépourvus d'aspects éthiques. D'une manière générale, le glissement du concept d'interprétation au concept de cognition va à l'inverse du mouvement de l'économie des conventions qui cherchait précisément à ouvrir la question de l'émergence des règles sur ces aspects historiques et éthiques, et non pas purement techniques³⁴.

Nous étudierons cette tentative d'ouverture dynamique des approches conventionnalistes par l'étude des procédés cognitifs des agents à travers les contributions de Laurent Thévenot et Olivier Favereau.

231- La "cognition située" : l'accent mis sur les objets

L'ouverture dynamique du cadre d'analyse des *Economies de la Grandeur* peut être envisagée sous trois angles différents. La première forme de dynamique est attachée à la notion d'épreuve. Lors des "épreuves", la place de chacun est réévaluée, ce qui assure que personne ne soit attaché à un rang de manière permanente. L'ordre hiérarchique est donc dynamique, mais cette première modalité de dynamique est exclusivement interne et ne concerne pas les principes de l'ordonnement.

La seconde modalité vise à l'inverse à étudier la formation de nouveaux principes. La confrontation permanente entre les divers principes de justice assure une mise en question permanente de l'ordre établi et permet d'envisager des conflits entre principes. Lorsque rien ne permet de trancher entre la référence à deux principes différents, la coordination est dans l'impasse. Les acteurs sortent de cette impasse en frayant des compromis. Le moment du compromis est typiquement économique³⁵ : la nécessité de maintenir une certaine efficacité dans la coordination prend le pas sur le souci de justice des acteurs. Cette suspension du sens du juste n'est que temporairement justifiée par la nécessité de poursuivre l'action. L'action hors du juste n'est qu'une parenthèse dans l'ensemble des interactions réunissant le souci de justice et l'impératif d'efficacité. L'action est comme

³⁴ Comme le suggère cette remarque de Thévenot [1993] : "*La réflexion de Lewis [en terme de CK] nous éloigne d'une réflexion strictement cognitive et nous ramène à la question des repères du jugement*" (P.22). La démarche des conventionnalistes semble en effet devoir rapprocher la question de la coordination des actions à celle de l'implication éthique des acteurs qui ont à se coordonner, ce qui excède la stricte question de la cognition.

³⁵ Il est d'ailleurs tout à fait révélateur que le "compromis" soit l'un des concepts les plus souvent convoqués par les économistes lorsqu'ils utilisent le modèle des *économies de la grandeur*.

suspendue à une réflexion a posteriori sur ce qu'il convient de faire pour réunir à nouveau le juste et l'efficace. Or, paradoxalement, c'est cette période transitoire de divorce entre les deux dimensions de la coordination que les auteurs présentent comme le moment de genèse de nouveaux principes de justice. Cette étrange ouverture dynamique semble aller à l'encontre de la philosophie de ce modèle analytique. Comment, en effet, les principes du juste pourraient naître du basculement de l'impératif de justice derrière celui d'efficacité ? Cette première modalité de dynamique du cadre des économies de la grandeur le fait implorer³⁶.

L'autre piste de recherche permettant de comprendre comment se forment les principes de justice est fondée sur le concept d'objet. Les objets qui étaient à l'origine de simples "représentants matériels" des principes abstraits de justice acquièrent ainsi le statut de convention autonome. Cette piste est esquissée par Laurent Thévenot (Thévenot [1990] et postface de Thévenot et Boltanski [1991]) et poursuivie au travers de recherches menées en commun avec Pierre Livet (Livet et Thévenot [1994 et 1997], Livet [1992]). L'idée centrale de cette nouvelle orientation est que les situations de coordination ordinaires ne nécessitent pas de recourir à une justification aussi générale que celle qui passe par la référence à des principes de justice et que, le plus souvent, la justification est restreinte à quelques personnes dans un espace limité. Elle s'appuie alors sur des objets dont le domaine de définition n'excède pas les limites du groupe concerné.

Cette étude de la dimension locale de la coordination vise à établir une architecture générale des principes supérieurs jusque dans leurs appuis les moins généraux. Le concept d'objet permet de souligner la manière dont les principes se manifestent concrètement dans le cours de l'action. En sollicitant des objets pour qualifier la situation qui les réunit, les acteurs participent à l'édification des principes généraux. Le modèle apparaît alors comme une succession de niveaux, à l'image des couches sédimentaires, allant du plus particulier au plus général. L'économiste ou le sociologue joue le rôle du géologue et dévoile les différents registres et niveaux d'action dissimulés sous des principes de justice en apparence indépendants de nos actes. Le but est ainsi de boucler une perception pragmatique de la coordination par convention.

Les auteurs se concentrent toutefois sur un aspect particulier de la coordination : la dimension cognitive. C'est par l'étude de la cognition qu'ils tentent d'éclaircir les mécanismes de construction du juste. La coordination est essentiellement perçue comme un problème quasi-technique de traduction d'une action particulière en termes assez généraux pour qu'ils soient compréhensibles de tous les partenaires. L'acteur, qui a une

³⁶Cette conception de la genèse pose enfin un problème de pure logique : le principe du compromis étant de relier des principes déjà existants, ce concept ne peut pas permettre d'expliquer la genèse de tous les principes. Aucun des six principes du modèle n'est d'ailleurs présenté comme provenant originellement d'un compromis.

intention précise, la traduit aux autres en s'appuyant sur des objets³⁷. Il considère donc la sémantique constituée par ces objets comme neutres. Il pense pouvoir caractériser sa propre action en s'appuyant sur les "objets" sans pour autant modifier le sens même de cette action. Pourtant, ces objets ne sont pas neutres, ils portent par définition un sens du juste puisqu'ils caractérisent un des six principes supérieurs de justice. Cette double facette des objets conventionnels est problématique. Par leur biais, les acteurs contribuent involontairement à consolider les principes de justice. Ceux-ci ne disparaissent pas de la situation mais uniquement de l'esprit de l'acteur qui perpétue ainsi inconsciemment, en pliant son action à des objets qu'il considère à tort comme neutres, une régulation qui lui échappe. L'illusion de l'acteur devient alors un vecteur efficace de stabilisation des principes conventionnels de justice, mais la question de l'origine de ces principes n'en devient que plus délicate. Cette nouvelle piste de recherche assouplit considérablement le modèle des économies de la grandeur, mais l'acteur conventionnaliste, libéré en apparence du carcan des principes supérieurs, ne se trouve-t-il pas dans le même temps dessaisi de la justice³⁸ ?

232- Une interprétation cantonnée aux limites de l'entreprise.

Olivier Favereau [1994a, 1994b, 1995, 1996, 1998] aboutit lui aussi à une analyse axée sur la cognition. Il insiste sur le fait qu'une théorie conventionnaliste des salaires doit traiter les salaires comme règles, et non comme prix, et doit insister sur le caractère interprétatif de la règle. Le sens de la règle n'est pas disjoint de son interprétation par les salariés. L'analyse initiale s'enrichit ainsi d'une prise en compte des difficultés de compréhension et de négociation entre la hiérarchie d'entreprise et les salariés.

³⁷"Agir c'est d'abord déclencher des mouvements qui ne nous semblent pas en discordance avec un état ou un but, et contrôler ces mouvements en corrigeant ceux qui se trouveraient disconvenir au type d'action engagée." Cette définition de Livet ([1992], p. 293) souligne l'aspect instrumental de l'action. Une telle action se coule dans un monde déjà existant, déjà balisé par des "objets", elle n'induit pas de changement. Bien que les auteurs s'appuient sur les travaux de Ricœur pour énoncer leur théorie de l'action, il semble qu'ils aient une perception de l'action assez fondamentalement différente de cet auteur pour qui : "Agir signifie avant tout opérer un changement dans le monde" (Ricœur, 1986, p.172).

³⁸L'ensemble de cette critique des travaux de Thévenot peut sembler prendre à rebours la démarche de l'auteur en cela qu'elle met l'accent sur la nécessité d'intégrer le rôle créatif de l'action dans le modèle, alors même que le modèle cherche à montrer que la coordination ne porte pas directement sur les comportements mais sur des énoncés de justification qui sont "à resituer dans un cours d'action qui, en amont et en aval du moment du jugement se déroule en dehors de contraintes fortes de réflexion et de justification" (Thévenot et Boltanski [1991], cités par Corcuff [1995] p.111). C'est en fait précisément le problème que nous cherchons à poser, il semble qu'à aucun moment l'action et son caractère imprévisible ne parvienne à tordre la sémantique instituée historiquement et qui se fait jour lors des moments de justification. Tout se passe comme si la grammaire de l'accord n'était jamais influencée par la pratique des agents. Il nous semble que la rétroaction de l'action sur le langage n'apparaît pas assez dans cette analyse de la coordination fondée sur la justification. Les principes de justice semblent en définitive naturellement ancrés dans l'esprit des acteurs : on retrouve une dimension naturaliste dans une approche fondée sur la pragmatique (en particulier dans Thévenot et Livet [1997]).

Favereau [1994a] propose une analyse de la règle fondée sur l'étude des règles salariales³⁹. Elle sont perçues comme des compromis entre salariés et entreprise visant à compenser les exigences de l'entreprise⁴⁰. L'accent mis sur ce type de règle, et sur la forme de négociation entre salarié et entrepreneur, rabat la dimension éthique de l'interprétation sur une interaction de type stratégique. On retrouve ici un glissement caractéristique du passage du concept d'interprétation à celui de cognition⁴¹. Ce glissement est d'une certaine manière inévitable dès lors que l'on situe le débat à l'intérieur des frontières de l'entreprise, espace dans lequel le primat de l'efficacité ne peut pas être débattu.

Favereau [1994b] affirme précisément, "*qu'il convient d'aller aux règles par une théorie de l'organisation*"(p.131), et "*d'aller d'une théorie de l'entreprise à une théorie de l'organisation plutôt que l'inverse*"(p.127). L'entreprise est donc confirmée comme étant l'objet d'étude fondamental. Or, mettre l'idée d'entreprise avant celle de règle revient à affirmer clairement ce qui est tacite dans l'approche cognitive telle qu'on la trouve chez Simon : on considère que cette structure hiérarchique est donnée et non problématique. Ce postulat limite singulièrement l'analyse de la dynamique des règles. En effet, la démarche cognitive conditionne une perception particulière de "l'interprétation", beaucoup plus restreinte que celle que défend le courant herméneutique. Cette restriction est sensible dans la façon dont est appréhendée la question de la genèse : tout se passe comme si seul le contenu technique des conventions était amené à évoluer tandis que leur contenu éthique était fixé une fois pour toutes suivant une procédure non explicitée.

Le choix de l'approche cognitive suppose que l'on se situe d'emblée dans un espace où les choix sont neutres du point de vue éthique⁴². Dans le domaine des relations productives, cela revient à considérer que l'espace organisationnel est épuré de tout problème non technique, c'est-à-dire que les questions éthiques y ont déjà été tranchées,

³⁹ Il est à noter que le concept de règle a remplacé celui de convention. La règle est "*une prescription à laquelle il est possible de se conformer et qui indique quel comportement est requis ou préféré ou prohibé, dans des contextes déterminés*" (Shimanoff[1980], p.57; cité par Favereau [1994a], p.51).

⁴⁰ Favereau (1994a, p.55). "*Le salaire est, du point de vue (descendant) des salariés, la clef de répartition des efforts collectifs, et du point de vue (ascendant) de l'entrepreneur le schéma de rémunération qui incite le mieux les salariés à contribuer au résultat collectif*".

⁴¹ "*La recherche philosophique (...) conclut à l'extrême importance des problèmes d'interprétation des règles, ou pour reprendre un vocabulaire familier aux économistes, des problèmes cognitifs dans la coordination par les règles*" (Favereau, 1994a, p.51)

⁴² Ce point est essentiel. On retrouve ici l'épistémologie popperienne. Popper considère qu'il est toujours possible d'isoler des fins éthiques, ouvertes à la délibération. Une fois ces fins fixées, il est possible de déterminer techniquement le meilleur moyen pour les atteindre. Cette posture épistémologique s'oppose à celle d'Habermas qui nie la possibilité de distinguer la fin des moyens dans un environnement humain. Pour cet auteur, aucune réflexion sur les moyens ou les règles sociales ne peut pas être purement technique, elle se fait toujours sur la base de présupposés éthiques. (Sur cette controverse : Adorno, Habermas, Popper [1969]).

que la place et le rôle de chacun ainsi que les buts de l'organisation sont clairement fixés et ne sont pas sujets à discussion. Or cette hypothèse est très forte, non seulement sur le plan empirique, mais dans l'optique même des conventionnalistes qui cherchaient à faire de ces choix à la fois éthiques et techniques l'objet de conventions toujours ouvertes à délibération. L'accent mis sur la cognition amène à considérer les conventions comme des normes échappant en grande partie à la discussion des acteurs.

24- La force d'attraction de la théorie des jeux

L'étude et la confrontation de ces deux approches de la convention nous permettent de tirer des enseignements plus généraux sur les problèmes inhérents à la démarche conventionnaliste et sur la manière dont ils tendent à être contournés. Olivier Favereau et Laurent Thévenot étudient deux manières opposées de suivre leur projet commun. L'un et l'autre butent sur un même problème : celui du moment de la genèse, celui de l'action non institutionnalisée. Ils évoluent l'un et l'autre vers une approche marquée par les sciences cognitives. Cette mise en avant de la cognition ramène le débat sur le terrain de la rationalité instrumentale et permet de lancer des ponts en direction de la théorie des jeux.

En effet, les relations entre acteurs peuvent, dans un cadre cognitif, être analysées selon le modèle de l'interaction stratégique, ce qui est précisément l'objet de la théorie des jeux. Cet espace théorique représente dès lors un terrain d'investigation de plus en plus utilisé par l'économie des conventions, alors même que le problème de l'émergence des règles constitue justement un objet théorique particulièrement difficile à traiter dans le cadre de la théorie des jeux.

Cette évolution est perceptible dans la manière dont les conventionnalistes se situent par rapport à Lewis [1969]. Il représente dès l'origine une référence essentielle, qui marque une approche individualiste de la convention et traduit l'accroche "orthodoxe" du pôle conventionnaliste. Les auteurs se situent alors dans une perspective de dépassement par rapport à l'approche lewissienne. La critique de Dupuy [1989] vise ainsi à montrer que la spontanéité du Common Knowledge repose en fait sur l'existence de repères communs préexistants qui s'actualisent lors de la coordination. Il est alors nécessaire, parce que "*le paradigme de la rationalité est radicalement incomplet*", de redécouvrir "*ce que l'on nomme d'ordinaire l'histoire*" (Dupuy[1989], p.398), c'est-à-dire de s'intéresser à ce passé commun qui permet la coordination. L'espace de la convention commence là où s'arrête la capacité explicative du concept de Common Knowledge. C'est sa critique qui ouvre les portes d'une réflexion sur la genèse de l'accord collectif qui préside à la

convention économique. Or, aujourd'hui, pour les conventionnalistes, l'article de Dupuy [1989], bien qu'il constitue une référence importante, ne représente plus la trame d'une réflexion critique vis-à-vis de Lewis.

Il semble en effet qu'une partie seulement de cette critique ait été réellement intégrée, celle qui concerne l'absence de théorie lewissienne de la genèse. Le principe de Common Knowledge devient ainsi la référence principale, même si elle est "sceptique" (Orléan, 1994, p.26)⁴³. Pourtant, au-delà de la question de la genèse, la nature historique de la convention rejaillit sur la forme de rationalité que mettent en œuvre les acteurs. Le paradoxe de la théorie lewissienne n'est pas tant qu'elle ne comporte pas de théorie positive de la genèse, mais bien que la convention de Lewis permet un oubli de l'Autre et de l'histoire commune qui lie les individus et devrait constituer le fondement de la convention. Un concept de convention qui fonctionne comme celui de Lewis occulte l'interprétation créative permanente à laquelle procèdent les individus, et replace la question de la coordination dans le cadre de la théorie des jeux. Dès lors celle-ci tend à envahir l'espace conventionnaliste, et à constituer une des matrices principales de la théorie des conventions, au moins dans sa dimension économique.

L'économie des conventions avait pour objet de départ de comprendre comment, de l'action, naissent des conventions et comment celles-ci régulaient par la suite les actions collectives. L'attention s'est ensuite progressivement centrée sur les aspects techniques de la coordination et a amené les conventionnalistes à des préoccupations proches des sciences cognitives. Cette restriction du champ d'étude aux aspects exclusivement cognitifs de la coordination a eu pour effet de mettre de côté les interrogations sur le sens de l'action et sa dimension créative. Paradoxalement ce rétrécissement de point de vue a ainsi permis à l'économie des conventions de s'ouvrir au vaste domaine de la théorie des jeux, les interrogations sur les limites d'une perception instrumentale de la rationalité n'étant plus d'actualité. Ce faisant, le projet conventionnaliste a été réduit à ses deux premiers principes de recherche. Or, c'est selon nous le troisième principe qui faisait son originalité et sa force. Puisque le terme de convention se comprend comme ce qui pallie les manques de la rationalité instrumentale, il faut sortir de ce paradigme, ou cesser d'en parler.

⁴³ Ainsi, "une coordination réussie nécessite la présence d'une convention proche de la définition lewissienne, à ceci près que la connaissance commune y est remplacée par une forme affaiblie de savoir collectif" (Orléan [1994, p.26]).

conclusion

L'impact des thèses de l'économie des conventions sur les orientations de la recherche économique française de ces dernières années est sensible, et ce en raison de la pertinence de son projet théorique. Ce projet met en effet l'accent sur un problème essentiel de la micro économie standard : son déficit d'analyse des règles comme supports essentiels de la coordination entre individus, sur les marchés comme dans les organisations productives. Toutefois, le projet conventionnaliste ne se borne pas à ajouter un étage intermédiaire entre l'individu et le marché standard. Il tire toutes les conséquences de l'idée relativement simple selon laquelle les individus ont besoin de repères communs pour coordonner leurs actions, en particulier en soulignant que ce simple fait remet en cause la dialectique entre un individu parfaitement rationnel et un marché auto-régulé. Si l'on admet que les acteurs doivent s'appuyer sur des règles pour se coordonner, on doit également admettre qu'une structure collective, produit d'un accord institutionnel non marchand, est nécessaire au fonctionnement de l'économie. Deux concepts suivent donc immédiatement celui de convention, le concept d'action collective, et celui d'accord non marchand. Ces trois concepts ont permis à l'économie des conventions de proposer une analyse renouvelée du fonctionnement du marché et du fonctionnement des organisations, au point de pouvoir proposer un modèle commun à ces deux modes de coordination antagonistes.

Mais le projet conventionnaliste va plus loin. L'accord collectif entre acteurs qu'il analyse, reflète, bien qu'il soit contraignant, une volonté consciente des acteurs en opposition avec l'idée d'une norme transcendante les orientant malgré eux. Cette lucidité dont fait preuve l'acteur conventionnaliste lorsqu'il cherche à s'accorder avec les autres, fait apparaître le lien qui existe entre problème de coordination et problème de justice. La règle ou la convention dont parle l'économie des conventions soutient une vision du monde qui intègre ses aspects éthiques. C'est là le second apport de l'économie des conventions.

Il reste cependant encore un pas à franchir aux conventionnalistes pour mener à bien leur projet. Les deux premiers pas franchis, l'intégration des notions de marché et d'organisation et le lien établi entre coordination et équité, portent essentiellement sur la nature et le fonctionnement des conventions, pas sur leur dynamique. Or il n'y a pas d'économie des conventions sans théorie de la dynamique des règles. Cette épineuse question peut donner lieu à deux sortes de réponses. La première est celle qui semble se profiler de plus en plus nettement dans les analyses conventionnalistes. Elle est analytique et s'appuie essentiellement sur les sciences cognitives et sur la théorie des jeux de

coordination et des jeux évolutionnaires. Cette façon d'analyser l'émergence des conventions dans un cadre marqué par la rationalité instrumentale des acteurs n'est pas satisfaisante parce qu'elle remet en cause les deux premières avancées. En effet réinterpréter les notions de règles et d'action collective dans un cadre marqué par l'interaction stratégique constitue un recul dans la perspective que contenait le projet conventionnaliste d'une analyse renouvelée de la coordination économique. Ce renouvellement avait précisément pour intérêt d'ouvrir de nouvelles pistes d'analyse de la genèse des règles, question qui ne reçoit que des réponses très partielles en théorie des jeux⁴⁴.

Une seconde piste est envisageable, et serait selon nous plus conforme au projet de départ des conventionnalistes. Elle s'appuie en particulier sur la notion d'interprétation, au sens large que ce terme revêt dans la tradition herméneutique⁴⁵. Mais la prise en compte de cette notion nécessite d'élargir l'hypothèse de rationalité instrumentale, en intégrant à l'idée de rationalité, ou de raison, des réflexions éthiques portant sur le choix de l'action, indépendamment de son efficacité relative à la fin poursuivie. Cette ouverture est l'un des objets principaux de l'œuvre de Karl Polanyi [1944], écarté d'emblée par les auteurs conventionnalistes en raison de sa tendance à "socialiser l'économie"⁴⁶. Ce rejet empressé nous semble pourtant devoir s'adresser davantage à Granovetter qu'à Polanyi, dont on peut faire une lecture en termes de rationalité. Polanyi cherche à mettre en évidence l'indépassable lien qui existe entre la capacité des acteurs de l'économie à faire preuve d'une forme de raison irréductible à la rationalité instrumentale et l'existence de règles nécessaires au bon déroulement du processus économique de production et d'échange. Pour Polanyi l'existence d'une économie exclusivement axée sur la rationalité instrumentale marchande est une utopie dangereuse⁴⁷, parce que la sélection d'une seule

⁴⁴ La théorie des jeux ne parvient finalement à expliquer l'émergence des normes... "que si elles existent déjà en pratique" (Cordonnier [1994], p.227).

⁴⁵ Avec la notion d'interprétation, on s'inscrit dans la longue tradition de la raison pratique aristotélicienne. Les concepts de prudence et de raison pratique d'Aristote (tels qu'ils sont présentés par exemple par Aubenque [1976]) ont été repris par Arendt [1960] en termes d'action, par Gadamer [1960] en terme d'interprétation et par Habermas [1975, 1987] en terme de communication. Ces trois auteurs forment une sorte de tradition reliant l'interprétation à l'action, l'action à la communication, et la communication à une forme non exclusivement instrumentale de rationalité.

⁴⁶ Ce rejet est exprimé dès l'introduction de la RE : "Plutôt que d'aller chercher dans les institutions sociales non marchandes le soutien de relations marchandes qui y seraient immergées (Polanyi, Arensberg Pearsons [1965], Granovetter [1985]), il faut s'attacher à élucider le présupposé commun propre à l'échange marchand, et montrer qu'il constitue une forme de collectif." (Revue économique [1989], p.143). Ce rejet est motivé par une crainte que traduit bien la remarque de Quéré : "Il s'agit d'abord d'écarter une stratégie possible qui consisterait à faire dépendre la capacité autorégulatrice du marché de formes sociales non marchandes (Polanyi [1944]), bref à socialiser l'économie" (Quéré [1993], p.22).

⁴⁷ Polanyi [1944] critique la mise en place d'un marché autorégulateur au cours du dix-huitième siècle, à travers notamment la marchandisation de la terre, du travail et de la monnaie. Pour cet auteur, le marché autorégulateur s'est révélé être une utopie lorsqu'il s'est effondré dans les années trente, projetant la société occidentale dans le chaos du fascisme puis de la guerre.

des dimensions de la raison humaine a pour effet de briser le fondement des règles nécessaires non seulement à l'équilibre social, mais aussi au fonctionnement de l'économie⁴⁸. Polanyi exhume la relation directe qui lie l'existence de règles ou de conventions économiques à la capacité des acteurs à faire preuve d'une rationalité débordant la stricte instrumentalité.

L'économie des conventions est parvenue à imposer une méthode originale et pertinente d'analyse des problèmes de coordination. Le développement de ses thèses passe aujourd'hui par sa capacité à analyser la dynamique des règles sans remettre en cause son projet théorique. Cela nécessite la formulation d'une hypothèse de rationalité plus large que la stricte rationalité instrumentale. La convention n'existe qu'à travers la faculté des individus à l'interpréter, c'est-à-dire à la créer. La prise en compte de cette faculté dans une hypothèse suffisamment large de rationalité individuelle rendrait inutile la formulation d'une théorie positive de la genèse des règles. C'est à la mise au point d'une telle hypothèse que nous conviait le projet conventionnaliste. Elle reste aujourd'hui à formuler⁴⁹.

⁴⁸Polanyi estime ainsi que : "*La véritable critique que l'on peut faire à la société de marché n'est pas qu'elle était fondée sur l'économie – en un sens toute société quelle qu'elle soit doit être fondée sur lui-mais que son économie était fondée sur l'intérêt personnel.*" (Polanyi [1983], p.320, c'est moi qui souligne)

⁴⁹Pour une interprétation convergente du projet conventionnaliste et ses rapports avec le concept de raison pratique on peut se reporter à Kandyl [1998].

Conclusion générale

“Créer des règles de relation c’est donner un sens à l’espace social” (J.D. Reynaud, 1989, p.280).

Le choix d’une règle collective contraint les individus. Penser la règle revient donc à éclairer le mécanisme par lequel une collectivité d’acteurs cherche à se construire un horizon commun. C’est pourquoi la règle crée doublement un “sens” : elle désigne une *direction* collective et dote l’action collective d’une unité et d’une *signification* qui la justifie.

C’est cette “création” de sens qui échappe inévitablement à l’analyse économique. Elle se situe en effet dans une perspective temporelle non prévisible et non réversible. Elle met donc en échec un comportement instrumental fondé sur une parfaite prévisibilité de l’avenir. Or, la science économique moderne s’appuie exclusivement sur la modélisation d’un comportement instrumentalement rationnel. Elle peut donc repérer le *rôle* des règles, mais ne peut penser leur *émergence* et leur *dynamique*. La règle est pour elle un concept *indispensable* mais *insaisissable*.

Nécessité et opacité des règles

Notre recherche fait apparaître un paradoxe : *“l’économie a besoin du concept de règle, mais le concept de règle ne peut pas être pensé dans le cadre économique standard”*. Le cadre économique standard peut en effet être décrit à partir de trois postulats de base que la prise en compte des règles remet inévitablement en cause.

Trois postulats de base

La science économique moderne, de type néoclassique, repose à l’origine sur trois hypothèses “fortes” qui lui permettent de bâtir une représentation scientifique de

l'économie :

- 1- L'espace économique est de nature marchande.
- 2- Dans cet espace le temps prend une forme prévisible et réversible.
- 3- Les individus sont instrumentalement rationnels.

Le premier chapitre nous a permis de mettre en évidence la cohérence de ces trois postulats et leur efficacité analytique. Ils permettent en effet de construire une représentation purement mathématique de l'économie et de démontrer ainsi l'efficacité du mode de coordination marchand. Toutefois, cette démonstration mathématique n'est pas en contact direct avec la réalité. Elle permet d'exprimer dans toute sa clarté la *logique* d'un mécanisme, mais ne nous dit rien de la *pratique* économique concrète. Pour avoir un "sens pratique", il serait en effet nécessaire que la démonstration de l'existence *formelle* de l'équilibre marchand soit complétée par la démonstration de son existence *pratique*, c'est-à-dire par une théorie satisfaisante de la "stabilité" du processus marchand. Ce n'est pas le cas et ce ne peut sans doute pas être le cas : la question de l'existence se situe sur un plan différent de celle de la stabilité et, si l'univers formel Arrow-Debreu est bien adapté pour résoudre la première, il ne l'est pas du tout pour résoudre la seconde. Il apparaît donc nécessaire de se doter d'une représentation *pratique* du processus marchand. C'est, en partie, l'objectif du programme "critique" d'Arrow.

Arrow tente en effet de déterminer les conditions pratiques qui permettraient à un marché de type "Arrow-Debreu" de fonctionner. Il développe pour ce faire une critique "interne" à l'univers néoclassique qui relativise le premier postulat standard mais tente de préserver les deux autres.

La critique d'Arrow ne remet pas en cause l'existence du mécanisme de marché mais son autonomie. L'existence d'une sphère marchande autonome se heurte à l'hétérogénéité des acteurs et au caractère divergent de leurs anticipations du futur. Le marché ne peut fonctionner que si les rapports de causalité entre les différentes actions des agents sont logiquement nécessaires et donc hors du temps. Or, pour que des rapports de type logique puissent être établis entre les individus, il faut nécessairement que l'espace des interactions soit ordonné par des règles collectives non marchandes qui permettent une "suppression" du temps pratique, de la contingence, de l'incertitude. Arrow montre ainsi que l'efficacité du mode de coordination marchand dépend de deux conditions précises : l'existence d'un accord moral collectif définissant un état social à atteindre, et l'existence d'une parfaite uniformité des profils cognitifs des agents. Or, pour que ces deux "conditions" soient remplies, il faut qu'il existe des règles collectives. Arrow relativise donc le mode de coordination marchand. Le marché ne se suffit pas à lui-même. Il

s'insère nécessairement dans un ensemble de règles collectives.

La critique d'Arrow a cependant un statut spécifique : elle conçoit la nécessité des règles à partir du modèle abstrait de marché. Elle ne vise pas à sortir de ce modèle pour décrire le *fonctionnement* et la *forme* de ces règles collectives. Elle montre simplement leur *rôle*.

La remise en cause du premier postulat

L'analyse néoinstitutionnaliste franchit un pas supplémentaire. Elle prend au sérieux la question du *choix* des règles collectives. Toutefois, elle hérite du mode de questionnement d'Arrow : elle se pose donc le problème de la règle à *partir de la représentation marchande de l'économie*. Pour cette raison, on peut résumer le projet néoinstitutionnaliste comme suit : "*parvenir à élargir l'univers économique, au-delà du marché, sans modifier les postulats 2 et 3*". L'analyse néoinstitutionnaliste cherche en quelque sorte à établir une condition institutionnelle nécessaire et suffisante pour que les deux derniers postulats puissent être maintenus.

Pour O. Williamson, cette institution est l'*organisation hiérarchique*. La hiérarchie résout les problèmes de coordination qui apparaissent dans le cadre d'une temporalité pratique en imposant aux membres de l'organisation un certain nombre de règles qui réduisent la diversité et assurent une forte cohérence des comportements des individus. A l'intérieur des organisations, l'enchaînement des actes des individus redevient nécessaire. Les individus peuvent être instrumentalement rationnels.

Cependant, l'intégration d'une structure institutionnelle dans l'univers néoclassique pose un problème de taille : il faut parvenir à rendre compte de la manière dont cette structure a pu être produite par les acteurs. Or, la démonstration de Williamson repose sur une mise en opposition des niveaux de rationalité individuel et collectif. Elle fait apparaître la profonde *altérité* qui sépare la logique de l'organisation hiérarchique de celle de chacun de ses membres. Cette altérité empêche que l'organisation hiérarchique *émerge* comme le résultat des actions des agents : il faudrait supposer que ceux-ci sont capables par eux-mêmes de pallier leurs propres défauts de rationalité... L'organisation hiérarchique apparaît donc comme une aporie du modèle économique. Elle est un élément irréductible aux actions des agents qui la composent. Il faut supposer qu'elle est le fruit d'une action d'ordre supérieur, d'une force sociale capable d'imposer des règles aux individus. En définitive, Williamson fait resurgir le problème de Hobbes au cœur de l'économie politique. Il "sauve" les deux derniers postulats néoclassiques (la temporalité réversible et le mode de rationalité des agents) en réintroduisant la *coercition* comme élément fondateur

de la coordination.

L'analyse stratégique des institutions s'inscrit dans l'optique de Williamson mais cherche à montrer que les règles collectives peuvent émerger *spontanément* comme des solutions collectives dans le cadre d'interactions individuelles. Ainsi, dans cette optique, la règle ne s'oppose pas aux agents, elle apparaît au contraire comme une solution individuellement rationnelle.

Pour défendre cette idée, la théorie des jeux, dans sa version néoinstitutionnaliste, mobilise la notion de convention proposée par Lewis [1969]. La convention permet en effet de penser l'existence d'un "ordre" collectif sans pour autant recourir au principe de la coercition. Cependant le concept de convention n'a aucun sens pratique chez Lewis : il désigne de manière habile une forme "arbitraire" de "nécessité". C'est pourquoi Lewis ne traite pas de la question du choix des conventions.

En revanche cette question devient essentielle dans le cadre d'une théorie de l'émergence. Pour expliquer le moment de la genèse des règles, il faut donner un "sens" à leur "arbitraire" et éclairer les raisons qui président à leur choix. Or, comme l'avait déjà perçu Hume, le sens de la convention dépend du cadre social et historique des interactions. Pour saisir le processus de l'émergence il faut donc introduire une dimension temporelle "réelle" : ce qui revient à remettre en cause le second "postulat de base". L'analyse stratégique est donc face à un dilemme : rendre compte de l'émergence des règles et sortir du cadre standard, ou rester à l'intérieur du cadre standard sans rendre compte de l'émergence des règles. Aucune des deux solutions n'est satisfaisante puisque cette théorie cherche à concilier ces deux objectifs : respect du cadre standard et description du mécanisme de l'émergence.

C'est pourquoi, la théorie des jeux ne parvient vraiment qu'à expliquer la manière dont une norme peut se *transformer* en institution. Pour aller plus loin et rendre compte de l'*émergence* des conventions, il faudrait parvenir à étudier la transformation progressive des situations d'interaction au fur et à mesure que le jeu se répète et que les joueurs "apprennent" à bien se comporter. C'est l'objectif de la théorie des jeux évolutionnistes. Cette théorie réintroduit donc l'Histoire. Mais elle le fait de manière bien spécifique, puisqu'elle ne peut rien dire de l'action effective des agents dans un cadre temporel non réversible sans remettre en cause le second postulat. Elle fait donc jouer un rôle central à la notion de "hasard". Or, l'Histoire des règles institutionnelles ne peut pas exclusivement être représentée comme un processus autonome, doté d'une dynamique propre et indépendant de l'action des agents. Pour *expliquer* l'émergence des règles, il faut rendre compte des raisons pour lesquelles les agents choisissent collectivement telle règle plutôt que telle autre. Dire qu'une règle perdure parce qu'elle est stable est insuffisant voire

tautologique : ce qu'il faudrait parvenir à faire apparaître, ce sont les raisons pour lesquelles elle est stable. Cela passe par la remise en cause du "second postulat" standard.

La remise en cause du second postulat

La remise en cause du caractère prévisible et réversible de la temporalité des actes économiques nous fait sortir du cadre standard, pour entrer dans un cadre "keynésien".

C'est en effet Keynes qui a mis le premier l'accent sur le caractère opaque de l'horizon des décisions économiques des agents. Dans la représentation théorique qu'il construit, les agents ne peuvent connaître avec certitude l'avenir. Ils ne disposent pas d'une représentation quantifiée des différentes chances de réalisation d'une série de "mondes possibles". Ils font donc appel à leurs capacités personnelles d'observation, à leur "sens pratique", pour prévoir l'avenir. Or, dès lors que le comportement individuel dépend sur l'expérience propre de chacun des acteurs, se pose un problème de coordination. L'uniformité cognitive des agents n'est plus une donnée.

Keynes montre en particulier que deux conceptions s'affrontent : celle des entrepreneurs et celle des spéculateurs. Ces deux types d'acteurs sont situés dans deux univers différents. Les spéculateurs agissent sur le marché, dans lequel les engagements sont pris à court terme et sont réversibles. Les entrepreneurs sont situés dans l'univers de la production, celui du long terme et de l'irréversibilité. Or c'est de l'organisation globale de l'économie que dépend l'issue du "combat" entre ces deux regards alternatifs sur l'économie. Keynes montre donc que les règles sont nécessaires pour résoudre les conflits d'interprétation du réel.

Mais il n'envisage pas du tout que le choix des règles puisse être contingent. Au contraire, selon lui, c'est précisément le caractère contingent des conventions qui est source d'instabilité. Keynes adopte une posture pragmatique : les règles doivent être choisies en fonction des désirs collectifs de la collectivité incarnée par l'Etat. Il réintroduit par ce biais une certaine idée de nécessité : l'objectif de l'Etat justifie et impose le choix des règles. Chez Keynes, la règle naît donc de la "vision" du savant et de l'action de l'Etat. Le savant est capable de percevoir quelles règles seraient nécessaires pour atteindre un objectif donné. L'Etat est supposé pouvoir mettre en place les règles adaptées à l'objectif qu'il poursuit (chez Keynes : le plein emploi). C'est en définitive l'Etat qui donne aux *individus* les moyens d'être raisonnables. Dans une perspective individualiste, cette position est difficile à accepter : il faudrait en effet supposer que l'Etat est un organisme collectif agissant en dehors des individus.

L'*Economie des Conventions* se nourrit des réflexions keynésiennes, tout en cherchant à dépasser ce problème. Elle cherche en effet à rendre compte de la genèse des règles à partir d'une conception renouvelée de l'action économique individuelle. Son projet théorique vise à développer une théorie permettant non seulement de relativiser l'importance du mode de régulation marchand, de mettre en évidence l'efficacité de modes de coordination alternatifs, mais encore d'éclairer la dynamique des repères collectifs qui fondent ces mécanismes de coordination non-marchands.

Ce programme de recherche remet en cause le premier postulat néoclassique. Il remet également en cause le second : dans le cadre de l'économie des conventions l'individu est en effet situé dans un espace temporel incertain et irréversible, qui nécessite l'existence de règles. Mais le projet conventionnaliste permet également de montrer que, dans ce cadre temporel, l'accord des acteurs ne repose pas exclusivement sur un calcul efficace. Il met en jeu des dimensions sociales, culturelles, éthiques. Une théorie de la genèse devrait donc articuler ces dimensions multiples du choix collectif. Pour cette raison, le projet conventionnaliste fait apparaître un troisième niveau critique : la remise en cause du postulat de la rationalité instrumentale. Pour rendre compte de la manière dont les acteurs économiques s'accordent sur des règles, il est nécessaire de les doter d'un mode de rationalité qui ne soit pas exclusivement caractérisé par la recherche de la plus grande efficacité possible. Les conventionnalistes hésitent à franchir cet ultime pas. Il risque en effet de les faire sortir de la science économique moderne.

L'étude du projet conventionnaliste nous amène ainsi au terme de notre analyse historique de la prise en compte des règles dans la théorie économique moderne. La reconnaissance du rôle des règles implique de s'écarter du *premier* postulat néoclassique. L'étude de la dynamique des règles implique de s'écarter du *second*. L'étude de l'émergence des règles à partir des actions individuelles implique de s'écarter du *troisième*. Or, c'est ce dernier pas qui est le plus problématique. Il conduit en effet à prendre des distances par rapport à une conception purement "technique" de l'efficacité.

La règle est utile. Elle permet à une collectivité d'agir "efficacement". Mais, l'efficacité sociale et pratique des règles n'est pas de même nature que l'efficacité mécanique de la rationalité instrumentale. La réussite d'une action collective ne prend pas la forme d'une relation mécanique de cause à effet. Elle naît d'un enchaînement d'actions que la règle favorise, accompagne, mais qui n'est pas mécaniquement prévisible. La règle ne peut donc être le produit d'un *calcul*. Elle est le fruit d'un *accord* dont les conséquences sont nécessairement incertaines. C'est sans doute là l'origine des nombreuses difficultés théoriques que l'économiste rencontre lorsqu'il tente de représenter l'*Histoire* de l'émergence des règles. Ces difficultés sont directement liées à la tension qui existe entre

temporalité “réelle” et rationalité instrumentale.

Rationalité instrumentale et prise en compte du temps

Si la théorie économique a besoin de la notion de règle mais ne parvient pas à la penser, c'est parce qu'elle est construite sur le concept de rationalité instrumentale qui, précisément, requiert *a priori* l'existence de repères collectifs. En effet, le temps de la rationalité instrumentale est nécessairement un temps *prévisible* et *réversible*, tandis que le temps de l'émergence des règles est celui qui correspond à l'enchaînement des “actions” individuelles et collectives. La principale caractéristique de cette forme de temps est de faire apparaître de la *nouveauté* et d'être par essence *imprévisible* et *irréversible*. Ces caractéristiques mettent en échec la rationalité instrumentale.

La temporalité mécanique de la rationalité instrumentale

Pour pouvoir asseoir un raisonnement théorique fondé sur une hypothèse de rationalité instrumentale, il est nécessaire de supposer que les individus peuvent, au moment où ils décident d'agir, connaître exactement les conséquences possibles de leurs actions. En effet un mode de rationalité de type instrumental repose sur un principe “conséquentialiste” : on évalue l'action en fonction de ses conséquences.

L'hypothèse de prévisibilité des conséquences repose sur une conception du temps qui est de type “logique” ou mécanique”. Ces deux types de temporalité reposent sur le principe d'un enchaînement “prévisible”, “nécessaire”, des actions de l'ensemble des agents. La relation qui unit le comportement des individus marchands aux mouvements de prix est de type logique. Elle se situe à un niveau de généralité qui exclut le temps. La relation qui unit le comportement de l'ensemble des acteurs dans une organisation est de type mécanique. Elle est rendue parfaitement prévisible par l'existence d'une organisation hiérarchique qui installe les individus dans un rapport mécanique en “ordonnant leurs interactions”. Dans un univers logique ou mécanique, l'agent peut prévoir de quelle manière sa propre action va s'enchaîner avec celles des autres. Il connaît les relations causales qui unissent nécessairement certaines actions à certaines conséquences. Il peut donc décider du choix de ses actions en contemplant au moment de les initier, les conséquences que chacune d'elles aura dans le futur.

Cependant, si l'individu peut agir ainsi de manière conséquentialiste, c'est parce qu'il

existe une *structure collective* qui l'isole de ses semblables. C'est, à l'origine, le marché qui joue ce rôle. Mais, la prise en compte des problèmes d'information et d'incertitude a mis en évidence le rôle des organisations centralisées et hiérarchiques.

La démarche néoinstitutionnaliste consiste ainsi à substituer un espace *organisationnel* à l'espace *marchand*. Cette substitution ne requiert pas de modifier la représentation instrumentale du comportement individuel. Mais elle met en évidence une contrepartie nécessaire à cette attitude instrumentale : la coercition. Dans le cadre organisationnel, la rationalité des individus dépend de l'existence d'une structure transcendante qui n'est plus naturelle (comme le marché est supposé l'être) mais explicitement *artificielle*. L'organisation est en effet construite sur le modèle du contrat social hobbién.

Le glissement du marché à l'organisation fait donc apparaître un paradoxe : la rationalité et la liberté individuelle dépendent du respect d'une structure collective. Pour lever ce paradoxe, il serait nécessaire d'éclairer le moment où cette structure collective se constitue. Or, cela est impossible : des individus instrumentalement rationnels ne peuvent pas s'accorder sur des règles communes. Ils ne peuvent pas "s'entendre" puisqu'ils agissent de manière paramétrique et instrumentale : l'individu rationnel n'est pas conçu pour entrer en contact avec les autres. On retrouve ici un enseignement du "théorème d'impossibilité" démontré par Arrow : la diversité des préférences "individualistes" empêche l'émergence spontanée d'un choix collectif.

La théorie économique se trouve de ce fait confrontée à une aporie interne. Une optique individualiste repoussant le principe d'une "autonomie" des structures collectives, il lui faut simultanément affirmer que :

- 1- Il est *nécessaire* qu'il y ait des règles, compte tenu du fait que les individus sont instrumentalement rationnels.
- 2- Il est *impossible* qu'il y ait des règles, compte tenu du fait que les individus sont instrumentalement rationnels.

Un individu instrumentalement rationnel a besoin de règles mais il ne peut pas les produire.

L'invention collective de la nécessité

Ce paradoxe ne se dénoue que dans un cadre temporel non-mécanique. En effet, s'il existe un paradoxe, c'est en raison du lien étroit qui unit les deux postulats néoclassiques : temporalité formelle et rationalité instrumentale sont simultanément déterminées. Un individu ne peut être instrumentalement rationnel que si le temps est réversible et prévisible. Le temps n'est réversible et prévisible que lorsque les actions des individus sont dans un rapport de causalité nécessaire ne permettant pas l'apparition de

nouveauté. C'est pourquoi la reconnaissance du rôle des règles et l'étude de leur émergence nous fait d'emblée sortir d'une perspective instrumentale *et* d'un cadre temporel "formel".

Car, pour pouvoir être instrumentalement rationnel, il faut se situer dans un espace dans lequel une cause produit *nécessairement* un effet prévisible. Lorsque les actes ne sont reliés entre eux par aucune forme de nécessité logique, les individus rationnels sont incertains, leurs actions ne se coordonnent pas mécaniquement. Or, c'est précisément la règle collective qui *produit* cette nécessité : la règle marchande et la règle organisationnelle établissent entre les individus des rapports logiques ou mécaniques qui excluent l'incertitude. En quelque sorte, concevoir des règles collectives, c'est "inventer" une forme de nécessité.

Le terme "d'invention" souligne bien que la question de l'émergence des règles est directement liée à la création d'une *nouveauté*. Elle se situe en amont de l'univers standard dans lequel les conséquences collectives des actes individuels sont parfaitement prévisibles. Lorsque les acteurs s'entendent sur des règles, ils *créent* un élément qui n'était pas présent à l'origine de l'action. L'émergence de la règle n'est donc pas représentable sur le modèle de l'analyse logique, puisqu'elle constitue un moment où un élément radicalement nouveau naît de la confrontation entre acteurs¹.

Pour éclairer le moment de l'accord collectif, il faut donc se situer dans un cadre temporel qui permette d'envisager l'apparition de cette "nouveauté". La temporalité de la règle n'est pas la temporalité mécanique qui rythme un comportement instrumentalement rationnel, mais la temporalité réelle de l'action, dont le propre est d'opérer un changement dans le monde². La question de l'émergence des règles échappe ainsi à une analyse fondée sur l'étude du comportement d'un individu instrumentalement rationnel.

La science économique moderne s'est construite autour de l'hypothèse de rationalité instrumentale. En cherchant à comprendre à quoi servent les règles collectives, et comment elles émergent, elle perçoit donc ses propres limites. Pour penser la règle, il lui faut en quelque sorte faire un retour critique sur elle-même. L'économie a besoin des règles, mais elle ne peut pas les penser.

¹ L'idée de nouveauté s'oppose au principe de la prévisibilité parfaite mais aussi au principe d'une incertitude probabilisable. L'hypothèse selon laquelle il existe une liste complète d'états du monde possibles est en effet contradictoire avec le principe de "l'invention".

² "Agir signifie au sens le plus général prendre une initiative, entreprendre, mettre en mouvement. (...) Il est dans la nature du commencement que débute quelque chose de neuf auquel on ne pouvait s'attendre." (Arendt, 1961, p.234).

Bibliographie

- ADORNO** Théodore W., **HABERMAS** Jürgen, **POPPER** Karl R., 1969, *Der Positivismusstreit in der Deutschen Soziologie*. Traduction Française : (J. Dewitte et ali.) *De Vienne à Francfort. La querelle allemande des sciences sociales*, Bruxelles, Complexes, 1979.
- AKERLOFF** George A, 1970, "The Market for Lemons : Qualitative Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, vol 84, p.488-500.
- ALCHIAN** Armen et **Harold DEMSETZ**, 1972, "Production, Information, Costs, and Economic Organisation", *American Economic Review*, vol 62, n°5, p.77-795.
- ALCHIAN** Armen et **Harold DEMSETZ**, 1973, "The property right paradigm", *Journal of economic theory*, vol.33, p.16-27.
- ALCHIAN** Armen et **Susan WOODWARD**, 1988, "The Firm is Dead, Long Live the Firm", *Journal of Economic Literature*, vol 26, n°1, p.65-79.
- ALLAIS** Maurice, 1953, "*Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine*", *Econométrie*, XL, Colloques internationaux du CNRS, Paris.
- ARENDT** Hannah, 1961, *The Human Condition*, traduction française : *La Condition de l'Homme Moderne*, Calmann Levy. Paris. Réédité chez Presse Pocket, Agora, 1988.
- ARROW** Kenneth.J. et **Franck HAHN**, 1971, *General competitive Equilibrium*, Holden Day, San Francisco and Oliver et Boyd, Edimburgh.
- ARROW** Kenneth J. et **Gérard DEBREU**, 1954, "Existence of an equilibrium for a competitive economy", *Econometrica*, vol.22, p.265-290.
- ARROW** Kenneth J., 1950, "A difficulty in the concept of social welfare", *Journal of Political Economy*, 58, p.328-346. Republié dans *The Collected Papers*, vol.1, Basil Blackwell, 1984, p.1-29.
- ARROW** Kenneth J., 1950, "A difficulty in the concept of social welfare", *Journal of Political Economy*, 58, p.328-346. Republié dans *The Collected Writings*, vol.1, Basil Blackwell, 1984, p.1-29.
- ARROW** Kenneth J., 1951a, "An extension of the basic theorem of classical welfare economics", in, J Neyman (ed), *Proceedings of the Second Bekerley Symposium on Mathematical Statistics and Probability*, p.507-532.
- ARROW** Kenneth J., 1951b, *Social Choice and Individual Values*, Cowles Foundation for Research, Yale university.

- ARROW Kenneth J.**, 1951c, "Alternative approach to the theory of choice in risk-taking situations", *Econometrica*, vol. 19, p.404-437. Republié dans *The Collected Papers*, vol. 3, Basil Blackwell, 1984, p.5-41.
- ARROW Kenneth J.**, 1952, Le principe de rationalité dans les décisions collectives, *Economie Appliquée*, 5, p.469-484. Republié en anglais dans *The Collected Writings*, vol.1, Basil Blackwell, 1984, p.45-58.
- ARROW Kenneth J.**, 1953, "Le rôle des valeurs boursières dans la répartition la meilleure des risques", *Econométrie*, Paris, CNRS, p.41-48.
- ARROW Kenneth J.**, 1958, "Utilities, attitudes, choices : a review note", *Operations Research*, vol. 5, p.765-774. Republié dans *The Collected Papers*, vol. 3, p.55-84.
- ARROW Kenneth J.**, 1963a, *Social Choice and Individual Values*, New Haven, Yale University Press. Seconde édition d'Arrow [1951b] avec une annexe supplémentaire.
- ARROW Kenneth J.**, 1963b, "Utility and Expectation in Economic Behavior", in *Psychologie : A Study of a Science*, Koch (ed), New York, Mac Graw-Hill, vol 6, p.724-752. Republié dans *The Collected Papers*, vol 3, p.117-143.
- ARROW Kenneth J.**, 1963c, "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review*, 53, p.941-973. Republié dans *The Collected Papers*, Vol. 6, The Belknap Press of Harvard University Press, 1985, p.15-50.
- ARROW Kenneth J.**, 1965, "Insurance, risk and resource allocation", in *Aspects of the Theory of Risk-Bearing*. Republié dans *The Collected papers*, Basil Blackwell, 1984, vol. 4, p.77-86.
- ARROW Kenneth J.**, 1967, "Values and collective decision making", in Laslett and Ruciman (eds) *Philosophy, Politics and Society*, Third series, Oxford, Basil Blackwell. Republié dans *The Collected Papers*, vol.1, p.59-77.
- ARROW Kenneth J.**, 1967, "Values and collective decision making", in Laslett and Ruciman (eds) *Philosophy, Politics and Society*, Third series, Oxford, Basil Blackwell. Republié en anglais dans *The Collected Writings*, vol.1, Basil Blackwell, 1984, p.59-77.
- ARROW Kenneth J.**, 1968a, "Economic Equilibrium", *International Encyclopedia of the Social Sciences*, D.L.Sills (ed), Macmillan. Republié dans *The Collected Papers*, Basil Blackwell, 1983, vol 2.
- ARROW Kenneth J.**, 1968b, "The economic of moral hazard : further comment", *American Economic Review*, 58, p.537-539. Republié dans *The Collected Papers*, vol. 4, p.103-105.

- ARROW Kenneth J.**, 1971a, "Exposition of the theory of Choice under uncertainty", in : C. Mc Guire and R. Radner, *Decision and Organisation*, North Holland, p.19-55. Republié dans *The Collected Papers*, vol. 3, p.172-208.
- ARROW Kenneth J.**, 1971b, The firm in the general equilibrium theory, R.Marris et A. Wood (ed) *The corporate Economy : Growth, Competition, and Innovative Potential.*, Cambridge, Mass. Republié dans *The Collected Papers*, vol 2, p.156-197.
- ARROW Kenneth J.**, 1973a, "General economic equilibrium : purpose, analytic techniques, collective choice", in *Les prix nobel en 1972*, Stockholm, Nobel fondation. Republié dans *The Collected Papers*, vol. 2, p.199-226.
- ARROW Kenneth J.**, 1973b, "Information and Economic Behavior", lecture to the Federation of Swedish Industries. Publié dans *The Collected Papers*, vol 4, p.136-152.
- ARROW Kenneth J.**, 1974a, *Choix collectif et préférence individuelle*, traduction française de Arrow [1963], Paris, Calmann-Levy.
- ARROW Kenneth J.**, 1974b, "Limited knowledge and economic analysis", *American Economic Review*, vol 64., n°4, p.153-166. Republié dans *The Collected Papers*, vol 4, p.153-166.
- ARROW Kenneth J.**, 1974c, *The Limits of Organisations*, New York, Norton. Ed. Française : 1976, *Les Limites de l'Organisation*, Paris, PUF.
- ARROW Kenneth J.**, 1978, "Risk, allocation and information : some recent theoretical developpement", *The Geneva papers on Risk and Insurance*, n°8. Republié dans *The Collected Papers*, vol 4, p.197-215.
- ARROW Kenneth J.**, 1981, "Grandeurs réelles et grandeur nominale en économie", in D. Bell et I. Kristol, *The Crisis in Economic Theory*, Basil Book, New York. Traduction Française : *Crise et Renouveau de la Théorie Economique*, Publisud-Bonneil, 1985, p.233-249.
- ARROW Kenneth J.**, 1982, "Risk Perception in Psychologie and Economics", *Economic Inquiry*, n°20, p.1-9. Republié dans *The Collected Papers*, vol 4, p.261-270.
- ARROW Kenneth J.**, 1985a, "The economic of the agency", in Pratt J. et R. Zeckhauser (ed), *Principal and Agent : The Structure of the Business*, Harvard Business School Press, p.37-54.
- ARROW Kenneth J.**, 1985b, Arrow's Theorem, in J. Eatwell (ed) *The New Palgrave*, p.124-126.

- ARROW Kenneth J.**, 1986, "Economic Theory and the Hypothesis of Rationality", *Journal of Business*, vol 59, n°2, pt 4. Republié dans J. Eatwell (ed) *The New Palgrave*, p.69-75.
- ARROW Kenneth J., BLOCK H., URWICZ Léonid**, 1959, "On the stability of the competitive equilibrium II", *Econometrica*, vol.27, p.82-109.
- ASCHWORTH T.**, 1980, *Trench Warfare, 1914-1918 : The Live and Let Live System*, New York, Holmes and Meir.
- AUBENQUE Pierre**, 1976, *La prudence chez Aristote*, Paris, PUF.
- AUMANN Robert J.**, 1964, "Markets with a continuum of traders", *Econometrica*, vol.32, p.39-50.
- AUMANN Robert J.**, 1989, "Game Theory", in J. Eatwell, M.Milgate, P. Newman (eds), *The New Palgrave, Game Theory*. London, Macmillan, p.1-54.
- AUMANN Robert J.**, 1976, "Agreeing to disagree", *Annals of Statistics*,4, p.1236-239.
- AUMANN Robert J.**, 1987, "Correlated Equilibrium as an Expression of Bayesian rationality", *Econometrica*, 55, p.1-18.
- AUMANN Robert J.**, 1994, "L'irrationalité dans la théorie des jeux", in ORLEAN André (dir) *Analyse Economique des Conventions*, Paris, PUF, p.43-61.
- AXELROD Robert**, 1984, *The Evolution of Coopération*, BasicBooks.
- AXELROD Robert**, 1986, "An evolutionary approach to norms", *American Political Science Review*, vol.80, p.1096-1111.
- BATEMAN Bradley**, 1988, "G.E. Moore and J.M. Keynes, a missing chapter in the History of the Expected utility Model", *American Economic Review*, vol 5, Dec., p.1098-1106.
- BATEMAN Bradely et John B. DAVIS**, 1991, *Keynes and Philosophy. Essays on the origin of Keynes's thought*, Londre, Edward Elgar.
- BAUDRY Bernard**, 1999, "L'apport de la théorie des organisations à la conception néo-institutionnelle de la firme", *Revue Economique*, vol 50, n°1, p.45-69.
- BENETTI Carlo et Jean CARTELIER**, 1994, L'intérêt limité de l'hypothèse de rationalité individuelle, *Cahiers d'Economie Politique*, n°24-25, p. 21-35.
- BENETTI Carlo et Jean CARTELIER**, 1980, *Marchands, Salarier et Capitalistes*. Maspéro, Paris.
- BERNOULLI Daniel.**, 1738, "Specimen Theoriae Novae de Mensura Sortis", *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, p.175-192.
- BERTHOUD Arnaud**, 1988, "Economie politique et morale chez Walras", *Æconomia*, série PE de la revue *Economie et Société*, n°9, p.65-93.

- BERTHOUD Arnaud**, 1992, "Remarques sur l'histoire de la pensée économique", *Cles*, n°20, 2ème semestre, p.129-138.
- BERTHOUD Arnaud**, 1994, "Remarques sur la rationalité instrumentale", *Cahiers d'Economie Politique*, n°24-25, p.105-124.
- BERTHOUD Arnaud**, 1998a, "L'épistémologie du modèle Arrow-Debreu", *Economie et Société. Economia, Série pe*, n°26, octobre, p.3-18.
- BERTHOUD Arnaud**, 1998b, "Economie et action politique dans la Théorie Générale", *Cahiers d'économie politique*, n°30-31, p.265-280.
- BICCHIERI Christina**, 1997, "Learning to cooperate", in Bicchieri, Jeffrey, Skyrms (eds), *The Dynamics of Norms*, Cambridge Mass., Cambridge U.P. p.17-46.
- BINMORE Ken et Larry SAMUELSON**, 1994, "An economist's perspective on the evolution of norms", *Journal of Institutional and Theoretical Economics*, p.45-63.
- BINMORE Ken et Pasta DASGUPTA**, 1986, *Economic Organisation as Games*, New York, Basil Blackwell.
- BINMORE Ken et Pasta DASGUPTA**, 1987, *The Economics of Bargaining*, New York, Basil Blackwell.
- BINMORE Ken**, 1987, *The Foudation of Game Theory*, Cambridge, Basil Blackwell.
- BLAUG Marc**, 1992, *The methodologie of Economics* (2nd édition), Traduction française (A. et C. Alcouffe) : *la Méthodologie Economique*, Economica, 1994.
- BOLTANSKI Luc et Laurent THEVENOT**, 1987 *Les économies de la grandeur*, cahiers du CEE, série Protée, n°31, C.E.E.PUF.
- BOLTANSKI Luc et Laurent THEVENOT**, 1991, *De la justification. Les économies de la grandeur*, Paris, Gallimard, coll Essais.
- BOUDON Raymond, BOUVIER Alain, CHAZEL François**, 1997, *Cognition et Sciences Sociales*, PUF (coll. sociologies).
- BOUVERESSE Jacques**, 1987, *La Force de la Règle*, Paris, Les éditions de minuit.
- BOYER Alain**, 1994, "La rationalité simonienne est elle satisfaisante ?", *Cahiers d'Economie Politique*, n°24-25, p.161-166.
- BOYER Robert et André ORLÉAN**, 1991, "Les transformations des conventions salariales entre théorie et histoire. D'Henry Ford au fordisme", *Revue Economique*, vol 42, n°2, p.233-272.

- BOYER Robert et André ORLÉAN**, 1994, "Persistence et changement des conventions", in A. ORLEAN (dir) *Analyse Economique des Conventions*, Paris, Puf, p.219-248.
- BOYER Robert et André ORLEAN**, 1997, "Comment émerge la coopération ? Quelques enseignements des jeux évolutionnistes". in Dupuy et Reynaud (eds) , *Les limites de la rationalité*, tome 2, Paris, La découverte.p.19-45.
- BRAITHWAITE Richard B.**, 1973, "Editorial Forewords", in Keynes, 1921.
- BRAITHWAITE Richard B.**, 1975, "Keynes as a Philosopher", in Keynes M.(ed), *Essays on J.M. Keynes*, Cambridge, Cambridge U.P., p.
- BROUSSEAU Eric**, 1989, "L'approche néo-institutionnelle des coûts de transaction", *Revue Française d'Economie*, vol IV, n°4, p.123-166.
- BROUSSEAU Eric**, 1993, *L'Economie des Contrats*, Paris, PUF, col "Economie en liberté".
- CARABELLI Anna**,1988, *On Keynes 's Method*, London. Macmillan.
- CARTELIER Jean**, 1995, "Le mode d'accord marchand : monnaie versus équilibre", in Jacob A. et Vérin H.(eds), *L'inscription sociale du marché*, Paris, L'Harmattan.
- CHATEL Elisabeth, RIVAUD-DANSET Dorothee, SALAIS Robert** (eds), 1998, *Institutions et Conventions, Raison Pratique* n°9, Paris, Ed. de L'EHESS.
- CHOI Young B.**, 1994, *Paradigms and Conventions*. The University of Michigan Press.
- COASE Ronald**, 1937, "The Nature of the Firm", *Economica*, n°4, p.386-405
- COASE Ronald**, 1960, "The problem of social cost", *Journal of Law and Economics*, 3, oct., p.1-44.
- COASE Ronald**, 1972 , "Industrial Organisation : a proposal for research", In V.R. Fuchs (ed), *Policy Issues and Research Opportunities in Industrial Organisation.*, New York. National Bureau of Economic Research, p.59-73.
- Conlisk J.**, 1996, "Why bounded rationality ?", *Journal of Economic Literature*, vol XXXIV, p.669-700.
- COOPER R. et JOHN A.**, 1988, "Coordinating Coordination Failures in Keynesian Models", *Quarterly Journal of Economics*, vol. 103, August, p.441-463.
- CORCUFF Philippe**, 1996, *Les nouvelles sociologies*, Paris, Nathan, Col 128.
- CORDONNIER Laurent**, 1993, *Coopération et Réciprocité*, Thèse pour le doctorat ès Sciences Economiques, Université de Paris X.
- CORDONNIER Laurent**, 1994, "Rationalité stratégique et émergence des normes", *Cahiers d'Economie Politique*, n°24,25. p.227-244.

- CORDONNIER Laurent**, 1997, *Coopération et Réciprocité*, PUF.
- DAVIS John B.**, "Keynes view of Economics as a Moral Science", in Bateman and Davis, 1991, pp 89-103.
- DE BANDT Jacques et Jean GADREY**, 1994, *Relations de Service, Marché de Service*, Paris, CNRS Editions.
- DE FINETTI Bruno**, 1937, "La prévision : ses lois logiques, ses sources subjectives", *Annales de l'Institut Henri Poincaré*. Paris.
- DE MONTBRIAL Thierry**, 1988, *La Science Economique*, PUF, Paris.
- DEBREU Gérard**, 1974, "Excess Demand Functions", *Journal of Mathematical Economics*, 1, p.15-21.
- DEBREU Gérard**, 1959, *Theory of Value: an Axiomatic Analysis of Economic Equilibrium*, New York, Willey. Edition française : *Théorie de la valeur*, 1963, Dunod.
- DEFALVARD Hervé**, 1990a, "La main invisible, Mythe et réalité du marché spontané", *Revue d'Economie Politique*, 100(6), Nov-Dec.
- DEFALVARD Hervé**, 1992, "Critique de l'individualisme méthodologique revu par l'économie des conventions", *Revue économique*, 1, p.127-142.
- DODIER Nicolas**, 1991, Agir dans plusieurs mondes, *Critique*, 529-530, juin, p.427-458.
- DOERINGER Peter et Michael PIORE**, 1971, *Internal labor market and manpower analysis*, Heath, Lexington, Mass.
- DOW Gregory**, 1987, "The function of Authority in Transaction Cost Economics", *Journal of Economic Behavior and Organisation*, vol 8, p.13-38.
- DUPUY Jean Pierre**, 1989, "Convention et Common Knowledge", *Revue Economique*, n°2, Mars, p.361-398.
- DUPUY Jean Pierre**, 1992, *Introduction aux sciences sociales*, Paris, Ellipses.
- DUPUY Jean Pierre, LIVET Pierre., REYNAUD Bénédicte.** (eds), 1997, *Les limites de la rationalité*, Paris, La découverte, col recherche.
- ENGEL Pascal**, 1994, "Les post-quiniens : réalisme et anti-réalisme", in Meyer (dir), *La philosophie anglo-saxonne*. Paris. PUF. p.348-367.
- EYMARD DUVERNAY François**, 1989, "Conventions de qualité et formes de coordination", *Revue économique*, 2, mars, p.329-360
- EYMARD DUVERNAY François**, 1986, "La qualification des produits", in SALAIS Robert, THEVENOT Laurent (eds), *Le travail - marché, règles conventions*. Paris, Insee-Economica, p.239-248.
- FAVEREAU Olivier**, 1985a, "L'incertain dans la révolution keynésienne. L'hypothèse Wittgenstein". *Æconomia*, série PE, 3, Mars, p.29-72

- FAVEREAU Olivier**, 1985b, "Evolution récente des modèles et des représentations théoriques du fonctionnement du marché du travail", *Problèmes économiques*, n°1955, janvier.
- FAVEREAU Olivier**, 1986, "La formalisation du rôle des conventions dans l'allocation des ressources", in SALAIS Robert, THEVENOT Laurent (eds), *Le travail – marché, règles conventions*. Paris, Insee-Economica, p. 250-265.
- FAVEREAU Olivier**, 1988, "La théorie générale : de l'économie conventionnelle à l'économie des conventions". *Cahiers d'économie politique*, janv, 14-15, p.197-220.
- FAVEREAU Olivier**, 1989, "Marchés internes , marchés externes", *Revue économique*, 2, mars, p.273-328
- FAVEREAU Olivier**, 1993, "Théorie de la régulation et économie des conventions : canevas pour une confrontation", *Lettre de la régulation*. N°7, Mai, p.1-3.
- FAVEREAU Olivier**, 1994a, "Suggestions pour reconstruire une théorie des salaires sur une théorie des règles", séminaire interdisciplinaire Technologie et sciences de l'homme, U.T.C. de Compiègne, 24-28 Janvier, Doc Ronéo.
- FAVEREAU Olivier**, 1994b, "Règles, organisation et apprentissage collectif : un paradigme non standard pour trois théories hétérodoxes", in ORLEAN André (dir.), *Analyse Economique Des Conventions*, Paris, PUF, p.113-135.
- FAVEREAU Olivier**, 1995, "Apprentissage collectif et coordination par les règles : application à la théorie des salaires", in Lazaric N et Monnier JM., *Coordination Economique et Apprentissage des firmes*, Economica, Paris, p. 23-38.
- FAVEREAU Olivier**, 1997, "L'incomplétude n'est pas le problème, c'est la solution", in Dupuy, Livet, Reynaud (eds), *Les limites de la rationalité*, Paris, La découverte.
- FAVEREAU Olivier**, 1998, "Note sur la théorie de l'information à laquelle pourrait conduire la théorie des conventions", in Petit P. (ed) *Economie de l'information*, Paris, La découverte, coll. recherche, p.195-252.
- FEIWEL George**, 1987a, *Arrow and the Foundations of the Theory of Economic Policy*, New York, Mac Millan Press.
- FEIWEL George**, 1987b, *Arrow and the Ascent of Modern Economic Theory*, New York, Mac Millan Press
- FIELD John A.**, 1984, "Microeconomics, Norms and Rationality", *Economic development and Cultural Change*, p.683-711.
- FISHER Francklin**, 1983, *Disequilibrium Foundations of Equilibrium economics*, CUP.

- FISHER Franklin**, 1990, "La formation des grandeurs économiques : déséquilibre et instabilité", in Cartelier (ed), *La Formation des Grandeurs Economiques.*, Nouvelle encyclopédie Diderot, PUF, P.19-55.
- FITZGIBBONS Athol**, 1989, *Keynes's Vision*, Oxford. Clarendon Press.
- FRYDMAN Roger**, 1992, "Le territoire de l'économiste, marché et société marchande", *Revue Economique*, 1, janvier, p.5-29.
- FRYDMAN Roger**, 1994, "Sur l'opposition de la rationalité substantive et de la rationalité procédurale", *Cahiers d'Economie Politique*, n°24-25, p.167-177.
- FUDENBERG Drew et Erik MASKIN**, 1986, "The folk theorem in repeated games with discounting or with incomplete information", *Econometrica*, vol.54, n°3, p.533-554.
- FUDENBERG Drew et Jean TIROLE**, 1991, *Game Theory*, Cambridge Mass., MIT press.
- GABRIER Hubert et Jean Louis JACQUIER**, 1994, *La théorie moderne de l'entreprise ; l'approche institutionnelle*, Paris, Economica.
- GADAMER Hans G.**, 1960, *Wahrheit und Methode*, Tübingen, Mohn, traduction Française (E. Sacre et P. Ricœur) : *Vérité et méthode*, Paris, Seuil, 1976.
- GADREY Jean**, 1990, "Rapports sociaux de service : une nouvelle régulation", *Revue Economique*, vol 41, n.1, p.49-69.
- GADREY Jean**, 1992, *L'économie des services*, Paris, Seuil, La découverte.
- GADREY Jean**, 1994, "Les relations de service et l'analyse du travail des agents", *Sociologie du Travail*, 3(94), p.381-389.
- GADREY Jean**, 1995, "Socio-économie de la relation de service", in Jacob A. et Vérin H. (eds), *L'inscription sociale du marché*, Paris, L'Harmattan. p. 145-167.
- GADREY Jean**, 1996, *Services : La Productivité en Question*, Paris, Desclée de Brouwer.
- GEANAKOPOLOS John**, 1985, "Arrow-Debreu model of general equilibrium", in Eatwell and ali.(eds) *The New Palgrave*, p.116-123
- GÉRARD-VARET Louis André**, 1995, "Quelques problèmes de rationalité posés par la théorie des jeux", in J.C. Passeron et L.A. Gérard Varet (eds), *Le modèle et l'enquête*, Paris, Editions de l'EHESS, p.479-526.
- GERRARD B.**, 1992, "From A Treatise on Probability to the General Theory : Continuity or Change in Keynes's Thought ?", in B. Gerrard et J. Hillard (eds), *The Philosophy and Economics of J.M. Keynes*, Aldershot, Hants, Elgar Pub. Co., p.80-95.

- GLATRON Marion**, 1998, "Keynes, Ramsey et les probabilités : un débat fondateur pour la théorie de l'action économique", in: Salais, Chatel, Rivaud-Danset (eds) *Institutions et conventions, Raison Pratique*, n°9, Paris, Ed de l'EHESS, p.51-72.
- GOCHET Paul**, 1978, Préface de l'édition française de Quine [1960].
- GOCHET Paul**, 1994, "L'empirisme relatif de Quine", in Meyer (dir), *La philosophie anglo-saxonne*. Paris. PUF. p.318-347.
- GOLDBERG Victor**, 1985, "Production Functions, Transactions Costs and the New Institutionalism". In Feiwel G.(ed), *Issues in Contemporary Microeconomic and Welfare*, London, Macmillan Press, p.395-404.
- GOYARD FABRE Simone**, 1994, "La philosophie morale et politique : entre le contractualisme et l'utilitarisme", in Meyer (dir), *La philosophie anglo-saxonne*. Paris. PUF. p.35-164
- GRANOVETTER Marc**, 1985, "Economic action and social structure : the problem of embeddedness, *American journal of sociology*, 91 (3), p. 481-510.
- GUERRIEN Bernard**, 1990, *La Théorie Néoclassique. Bilan et perspectives de modèles d'équilibre général*. Paris. Economica.
- GUERRIEN Bernard**, 1993, *La théorie des jeux*, Paris, Economica.
- GUESNERIE Roger et Jean Yves JAFFRAY**, 1971, "Probabilités subjectives et utilité de Von Neumann Morgenstern", *Bulletin de Mathématiques Economiques*, n°6.
- GUESNERIE Roger**, 1973, Actes du colloque "Valeur et prix de production", *Annales de la faculté de droit et des sciences économiques*. Université de Lyon II, n°2.
- GÜTH W., MARCHAND N., RULLIÈRE J.L.**, 1998, "Equilibration et dépendance du contexte", *Revue Economique*, vol.49, n°3, p.785,794.
- HABERMAS Jürgen**, 1975, *Théorie et pratique* (Trad. G. Raulet), Paris, Payot.
- HABERMAS Jürgen**, 1987, *Théorie de l'agir communicationnel*, (Trad de JM Ferry), Paris, Fayard.
- HACKING Ian**, 1975, *The Emergence of Probability*, Cambridge, Cambridge University Press.
- HAHN Franck**, 1980, "Stability", in Arrow et Intriligator (eds), *The Handbook of Mathematical Economics*, volume 2, Amsterdam, North Holland. p.745-793.
- HAHN Franck**, 1981, "Théorie de l'équilibre général", in D. Bell et I. Kristol, *The Crisis in Economic Theory*, Basil Book, New York. Traduction Française : *Crise et Renouveau de la Théorie Economique*, Publisud-Bonneil, 1985,

- HAHN Franck**, 1982, Réflexions sur la "main invisible", *Loyds Bank Review* (Londres), Avril. Traduit dans *Problèmes Economiques*, N°1797, nov., p.16-22
- HAHN Franck et Thomas NEGISHI**, 1962b, "A theorem on non-tâtonnement stability", *Econometrica*, vol.30, p.463-469.
- HAHN Franck**, 1962, "On the stability of pure exchange equilibrium", *International Economic review*, vol.3, p.206-213.
- HARGREAVES-HEAP Sean P. et Yanis VAROUFAKIS**, 1995, *Game Theory, a critical Introduction*, London, Routledge.
- HARSANYI John et SELTEN Reinhard**, 1988, *A General Theory of equilibrium selection in games*, Cambridge Mass, MIT Press.
- HART Oliver**, 1990, "Is bounded rationality an important element of a theory of institutions ?", *Journal of Institutional and Theoretical Economics*, n°146, p.696-702.
- HAYEK Friedrich**, 1976, *Law, Legislation and Liberty*, traduction française : *Droit, legislation et liberté*, Paris, PUF, 1986.
- HELBURN Suzan**, "Burke and Keynes", in Bateman et Davis, 1991, p. 30-53.
- HELWIG Martin F.**, 1993, "The challenge of monetary theory", *European Economic Review*, vol 37, p.215-242.
- HICKS John R.**, 1939, *Value and Capital*. Oxford. Clarendon Press.
- HOBBS Thomas**, 1651, *Léviathan*, Paris, Sirey, 1971
- HUME David**, 1740, *Traité de la Nature Humaine*, Tome 3 : La morale. GF, Paris, Flammarion, 1993.
- JENSEN M. et W.H. MECKLING**, 1976, "Theory of the firm, managerial behavior, agency cost, and ownership structure", *Journal of Financial Economics*, (3), p.304-360.
- KAHNEMAN Daniel et Amos TVERSKY**, 1979, "Prospect Theory : an analysis of Decision under Risk", *Econometrica*, 47, p.263-291.
- KAHNEMAN Daniel, SLOVIC Paul, TVERSKY Amos**, 1982, *Judgement under Uncertainty : Heuristics and Biases*, New York, Cambridge University Press.
- KANDYL Ferial**, 1998, "De la rationalité à la raison pratique dans les actes économiques", in Chatel, Rivaud-Danset, Salais (eds), p.145-172.
- KARPIK Lucien**, 1998, "La confiance : réalité ou illusion ?", *Revue Economique*, Vol 49 (4), p.1043-1056.
- KEYNES John Maynard**, 1921, *A Treatise on Probability*, reprinted in *The Collected Writings*, vol 8.

- KEYNES John Maynard**, 1924, Lettre à Wittgenstein, in *Ludwig Wittgenstein, Cambridge Letters*, p. 201
- KEYNES John Maynard**, 1926, "La fin du Laissez faire", *The Collected Writings*, vol.IX, p.272-295.
- KEYNES John Maynard**, 1931, "Franck Ramsey", *The Collected Writings*, vol X, p. 335-346.
- KEYNES John Maynard**, 1936, *La Théorie Générale de l'Emploi, de l'Intérêt et de la Monnaie* (trad De Largentaye), Paris, Payot. (La pagination renvoie à l'édition de 1990)
- KEYNES John Maynard**, 1937a, "La théorie générale de l'emploi", in *The Quaterly Journal of Economics*, Fevrier. Republié dans *The Collected Writings*, vol XIV, p.109-123.
- KEYNES John Maynard**, 1937b, "Some economic conséquences of declining population", *The Collected Writings*, vol XIV, p.124-138
- KEYNES John Maynard**,1938a, "My Early Beliefs", in *The collected Writings*, Vol X, p.433-451.
- KEYNES John Maynard**, 1938b, "Lettre à Harrod", *The Collected Writings*, vol XIV, p.297-299.
- KEYNES John Maynard**, 1973-79, *The collected Writings of J.M. Keynes*, London, Macmillan.
- KIRMAN Alan P.**, 1985, "Organisation et communication dans les marchés", *Economie Appliquée*, tome XXXVII n°3, p.597-609.
- KIRMAN Alan P.**, 1991, "Information et communication dans les marchés", *Economie Appliquée*, tome XLIV, n°1, p. 105-124..
- KNIGHT Franck**, 1921, *Risk, Uncertainty and Profit*, Boston, Houghton, Mifflin Company.
- KOLM Serge Christophe**, 1986, *Philosophie de l'économie*, Paris, Seuil.
- KREPS David**, 1984, "Corporate Culture and Economic Theory", in Alt et Shepsle (eds) *Perspectives on Political Economy*, Cambridge Mass., Cambridge U.P., Chap.4, p.90-143.
- KREPS David**, 1990a, *A Course In Microeconomic Theory*, traduction française : *Leçons de Théorie Microéconomique*, PUF, coll. Finance, 1997.
- KREPS David**, 1990b, *Game Theory and Economic Modelling*, New York, Oxford University Press.
- LAFFONT Jean Jacques**, 1984, *Economie de l'Incertain et de l'Information*, Paris. Economica.

- LANGLOIS Richard, 1990, "Bounded rationality and behavioralism : a clarification and critique", *Journal of Institutional and Theoretical Economics*, n°146, p.691-695.
- LEMOIGNE Jean Louis, 1995, "Sur la capacité de la raison à discerner : rationalité substantive et rationalité procédurale d'Aristote à Simon par Descartes et Vico", in J.C. Passeron et L.A. Gérard Varet (eds), *Le modèle et l'enquête*, Paris, Editions de l'EHESS, p.245-277.
- LEWIS David K., 1969, *Convention. A philosophical study*. Cambridge Mass., Harvard University Press
- LEWIS David K., 1983, "Langages and Languages", in *Philosophical Papers*. New York. Oxford U.P.
- LEWIS David K., 1986, *On the Plurality of Worlds*, New York, Basil Blackwell.
- LIVET Pierre et THEVENOT Laurent, 1994, "L'action collective", in Orléan (dir.), *Analyse économique des conventions*, Paris, PUF.
- LIVET Pierre et THEVENOT Laurent, 1996, "Modes d'action collective et construction éthique", in Dupuy, Livet et Reynaud (eds), *Les limites de la rationalité*, Paris, La Découverte.
- LIVET Pierre, 1992, "Théorie de l'action et conventions", *Raison Pratique* 2, Editions de L'EHESS.
- LUCE Duncan et Howard RAIFFA, 1957, *Games and Decision*, New York, Willey.
- MAC GUINNESS Brian et Georg H. VON WRIGHT., 1995, *Ludwig Wittgenstein : Cambridge Letters*, Oxford, Basil Blackwell.
- MANTEL R., 1974, "On the characterisation of Aggregate Excess Demand". *Journal of Economic Theory*, vol 7, n°3, p.348-353.
- MAS-COLELL A., 1989, "Cooperative equilibrium", in J. Eatwell, M.Milgate, P. Newman (eds), *The New Palgrave, Game Theory*, London Macmillan, p.95-103.
- MARCH James et Herbert A. SIMON, 1958, *Les Organisations*, traduction française : Rouchy et Prunier, 1991, Paris, Bordas.
- MATHIOT Jean, 1990, *Adam Smith, Philosophie et Economie*, Paris, PUF.
- MAYNARD-SMITH J. et PRICE G.R., 1973, "The Logic of Animal Conflict", *Nature*, n°246, p.15-18.
- MENARD Claude, 1990, *L'Economie des Organisations*, Paris, La découverte.
- MENARD Claude, 1994a, "Organisations as coordinating devices", *Metroeconomica*, vol 45, n°3, p.224-247.

- MENARD Claude**, 1994b, "Comportement rationnel et coopération, le dilemme organisationnel", *Cahiers d'Economie Politique*, n°24-25, p.163-186.
- MENARD Claude**, 1995, "Markets as institutions versus organisations as markets ? Disentangling some fundamental concepts". *Journal of Economic Behavior and Organisation*, Vol 28, n°2, p.161-182.
- MENGER Karl**, 1934, "Das unsicherheitsmoment in der Wertlehre", *Zeitschrift für Nationalökonomie.*, vol.5, n°4, p.459-485.
- MEYER Michel (ed)** , 1994, *La philosophie Anglo-saxonne*, Paris, PUF.
- MONGIN Philippe**, "Modèle rationnel ou modèle économique de la rationalité ?", *Revue Economique*, vol.35, n°1, p.9-63.
- MORGENSTERN Oscar**, 1974, "Treize points critiques de la théorie économique contemporaine". *Economie Appliquée*, tome XXVII, n°2-3.
- MOULIN Hervé**, 1981, *Théorie des jeux pour l'économie et la politique*, Paris, Hermann.
- MOULIN Hervé**, 1995, "An appraisal of cooperative game theory", *Revue d'Economie Politique*, 105, n°4, p.617-632.
- MYERSON Roger B.**, 1991, *Game Theory, Analysis of Conflict*, Harvard, Harvard University Press.
- NACHBAR John**, 1992, "Evolution in the finitely repeated prisoner's dilemma", *Journal of Economic Behavior and Organisation*, vol.19, p.307-326.
- NASH John F.**, 1950a, "Equilibrium in n-person games", *Proceedings of the National Academy of Sciences*, vol.36, p.48-49.
- NASH John F.**, 1950b, "The bargaining problem", *Econometrica*, 18, p.155-162.
- NASH John F.**, 1953, "Two person cooperative person games", *Econometrica*, 21, p.128-140.
- NEISSER H.**, 1932, "Lohnhöhe und beschäftigungsgrad im Marktgleichgewicht", *Weltwirtschaftliches Archiv*, vol.36, p.415-455.
- NORTH Douglas**, 1986, "The new institutionnal economics", *Journal of Institutional and Theoretical Economics*, vol 142, n°1, p.230-237.
- NORTH Douglas**, 1991, "Institutions", *Journal of Economic Perspectives*, vol.5, n°1, p.97-112.
- O'DONNEL Rob**, 1989, *Keynes : Philosophy, Economics and Politics*, London, Macmillan.
- O'DONNEL Rob**, 1991, "Keynes Weight of Argument and it's Bearing on Rationality and Uncertainty", in Bateman and Davis, p.69-88.
- ORLEAN André**, 1986, "Le rôle des conventions dans la logique monétaire", in Salais et Thévenot (eds). PUF, p. 217-238.

- ORLEAN André, 1988, "L'auto référence dans la théorie keynésienne de la spéculation", *Cahiers d'économie politique*, 14-15, janv, p.229-242
- ORLEAN André, 1989, "Pour une approche cognitive des conventions économiques", *Revue Economique*, Mars, p.241-270
- ORLEAN André (dir.), 1994, *Analyse économique des conventions*, Paris, PUF.
- ORLEAN André, 1994b, "Vers un modèle général de la coordination économique par les conventions", in Orléan (dir) [1994], p.9-38.
- ORLEAN André, 1994c, "Sur le rôle respectif de la confiance et de l'intérêt dans la constitution de l'ordre marchand", *Cahiers du Mauss, Recherche*, vol 4, p.17-36.
- ORLEAN André, 1997, "Jeux évolutionnistes et normes sociales", *Economie Appliquée*, tome L, n°3, p.177-198.
- OSTROY J. et R.M. STARR, 1974, "Money and the decentralisation of exchange", *Econometrica*, vol.42, n°6, Novembre, p.1093-1113.
- PARETO Wilfredo, 1909, *Manuel d'Economie Politique*, Paris, Girard et Brière.
- PETTIT Philip, 1993, "Normes et choix rationnel", *Réseaux*, n°62, p.87-111.
- PHILIPSON Henri, 1995, *L'Economie contre la Nature*, Lille, Ester.
- POLANYI Karl, 1944, *The Great Transformation*, traduction française (C. Malamoud) : *La grande transformation. Aux origines politiques et économiques de notre temps*, Paris, Gallimard, 1983.
- POLANYI Karl, Arensberg Conrad, PEARSON Harry, 1965, *Trade and Market in the Early Empires. Economies in History and Theory*, traduction française (C. et A. Rivière) : *Les systèmes économiques dans l'histoire et dans la théorie*, Paris, Larousse, 1975.
- POLANYI Michael, 1962, *Personal Knowledge : Toward a Post-Critical Philosophy*, Chicago, Chicago University Press.
- PONSSARD Jean Pierre, 1995, "A propos du rôle des singularités dans l'analyse économique", in J.C. Passeron et L.A. Gérard Varet (eds), *Le modèle et l'enquête*, Paris, Editions de l'EHESS, p.527-547.
- POSTEL Nicolas, 1998, "L'économie des conventions : une approche instrumentale de la rationalité individuelle ?", *Revue Economique*, n°6, vol.49, p.1473-1495.
- POSTEL Nicolas, 1999, "Règle et coordination dans la théorie keynésienne", CLES (à paraître).
- PRATT John et Richard ZECKHAUSER (ed), 1985, *Principal and Agent : The Structure of the Business*, Harvard Business School Press.
- QUÉRÉ Louis, "A-t-on vraiment besoin de la notion de convention ?", dans *Réseaux*, 62, Octobre, p. 21-41.

- QUINE Williard V.O.**, 1936, "Truth By Convention", in *Philosophical Essays for A.N. Whitehead*, O.H. Lee (ed), New York, Longmans.
- QUINE Williard V.O.**, 1951, "Two dogmas of empiricism", *Philosophical Review*, vol 60. Traduction française : "Les deux dogmes de l'empirisme", in *De Vienne à Cambridge*, P. Jacob (ed), Paris, Gallimard, 1980.
- QUINE Williard V.O.**, 1960, "Le Mot et la Chose", Traduction française : Gochet et Dopp, Paris, Flammarion, 1978.
- QUINET Catherine**, 1994, "Herbert Simon et la rationalité", *Revue Française d'Economie*, vol IX, n° 1, p.133-182.
- QUIRK James et Rubin SAPOSNIK**, 1968, *Introduction to General Equilibrium Theory and Welfare Economics*, Mc Graw Hill, New York. Traduction Française de J.J. Laffont et G. Laroque : *Théorie de l'équilibre général, et économie du bien être*. 1974, Paris. PUF.
- RALLET Alain**, 1993, "La théorie des conventions chez les économistes", *Réseaux*, 62, Octobre, p. 43-59.
- RAMSEY Franck P.**, 1926, "Truth and Probability", in Ramsey, 1931, *The Foundations of Mathematics and Other Logical Essays*. P. Trench, Trumber and co. London.
- REVUE ECONOMIQUE**, 1989, "l'économie des conventions", 2, Mars.
- REYNAUD Jean Daniel**, 1989, *Les règles du jeu, L'action collective et la régulation sociale*, Paris, Armand Colin.
- RICŒUR Paul**, 1986, "Expliquer et comprendre. Sur quelques connexions remarquables entre la théorie du texte, la théorie de l'action et la théorie de l'histoire.", dans *Du texte à l'action. Essais d'herméneutique II*, Paris, Seuil-Esprit, p 161-182.
- SALAIIS Robert et Laurent THEVENOT (eds.)**, 1986, *Le travail – marché, règles, conventions*. Paris, Insee-Economica.
- SALAIIS Robert**, 1992, "Flexibilité et conventions du travail", *Problèmes économiques*, 2277, mai.
- SALAIIS Robert et Michael STORPER**, 1994, *Les mondes de production, enquête sur l'identité économique de la France*. Paris, Ed. de l'EHESS
- SAMUELSON Paul A.**, 1957, "Wages and interest : a modern dissection of the marxian economic models", *American Economic Review*, vol.47, p.884-912.
- SAMUELSON Paul A.**, 1947, *Foundations of Economic Analysis*, Harvard University Press, Cambridge Mass.
- SAMUELSON Paul A.**, 1959, "A modern treatment of the ricardian economy", *Quarterly Journal of Economics*, vol 73, p.1-35.

- SAVAGE Leonard**, 1954, *The Foundations Of Statistics*. New York. Willey.
- SCHELLING Thomas C.**, 1960, *The Strategy of Conflict*, Cambridge Mass., Harvard University Press, 1980.
- SCHIONOYA Y.**, 1991, "Sigwick, Moore and Keynes : a Philosophical Analysis of Keynes's "My Early Beliefs", in Bateman and Davis, p. 6-30.
- SCHLICHT W.**, 1990, "Rationality, bounded or not, and institutional analysis", *Journal of Institutional and Theoretical Economics*, n°146, P.703-719.
- SCHOTTER Andrew**, 1981, *The Economic Theory of Social Institutions*. Cambridge. Cambridge University Press.
- SCHOTTER Andrew**, 1983, "Why take a game theoretical approach to Economics ? Institutions, economics and game theory". *Economie Appliquée*, XXXVI, n°4, p.673-695.
- SELTEN Reinhard**, 1990, "Bounded rationaliy", *Journal of Institutional and Theoretical Economics*, vol 146, p.649-658.
- SHUBIK Martin**, 1982, *Game Theory and Social Science*, Cambridge, MIT Press. Traduction française (B. Guerrien) : *Théorie des Jeux et Sciences Sociales*, Paris, Economica, 1991.
- SIMON Herbert A.**, 1951, "A formal Theory of Employment Relationship", *Econometrica*, vol.19, p.293-305.
- SIMON Herbert A.**, 1962, "The architecture of complexity", *Proceedings of the American Philosophical Society*, 106, p.467-482.
- SIMON Herbert A.**, 1964, Rationality, in Gould et Kolb (eds), *Dictionary of Social Science*, Glencoe, The Free Press.
- Simon, H.A.**, 1969, *The Science of Artificial*, Cambridge MIT Press. Traduction française (J.L. Lemoigne) : *Sciences des Systèmes, Science de l'Artificiel*, , EPI, Paris, 1974.
- SIMON Herbert A.**, 1972, "Theories of bounded rationality", in Mc Guire and Radner, (ed), *Decision and Organisation*, Amsterdam, North Holand Publishing Company.
- SIMON Herbert A.**, 1978, "Rationality as a process and a product of thought", Richard T.Ely lecture, *American Economic Review*, 68(2), p.1-16.
- SIMON Herbert A.**, 1976a, *Administrative Behavior*, The Free Press, New York.
- SIMON Herbert A.**, 1976b, "From Substantive to procedural Rationality", in Latsis (ed), *Method and Appraisal in Economics*, Cambridge, Cambridge University Press.
- SIMON Herbert A.**, 1983, *Reason in Human Affairs*, Stanford, Stanford University Press.

- SIMON Herbert A.**, 1991, "Organisations and markets", *Journal of Economic Perspectives*, vol 5, n°2, p.25-44.
- SIMON Herbert A.**, 1997, *An Empirically Based Microeconomics*, Cambridge, Cambridge University Press.
- SMITH Adam**, 1776, *Essais sur les causes de la richesse des nations*, Gallimard, coll. idées, 1976.
- SONNENSCHNEIN H.**, 1973, "Do Walrass identity and continuity characterise the class of Community Excess Demand Function ?", *Journal of Economic Theory*, vol 6, n°4, août, p.345-354.
- SONNENSCHNEIN H.**, 1972, "Market Excess Demand Functions", *Econometrica*, 40, n°3, mai, p.549-563.
- STIGLER George J.**, 1968, "The economics of information", *Journal of Political Economy*, vol.69, p.213-225.
- SUGDEN Robert**, 1989, "Spontaneous order", *Journal of Economic Perspectives*, 3, p.85-97.
- SUGDEN Robert**, 1995, "The Coexistence Of Conventions", *Journal of Economic Behavior and Organisation*, vol 28, p.241-256.
- SUGDEN Robert**, 1986, *The Economics of Rights, Coopération and Welfare*. Oxford, Basil Blackwell.
- SUPPES Patrick**, 1981, *Logique du Probable. Démarche bayésienne et rationalité*. Paris, Flammarion.
- TALLON Jean Marc**, 1997, *Equilibre Général : une introduction*, Paris, Vuibert.
- THEVENOT Laurent**, 1986, "Economie et formes conventionnelles", in SALAIS Robert et THEVENOT Laurent (eds.) *Le travail-marché, règles, conventions*, Paris, Insee-Economica, p.195-215.
- THEVENOT Laurent**, 1989, "Equilibre et rationalité dans un univers complexe", *Revue Economique*, 2, mars .
- THEVENOT Laurent**, 1990, "L'action qui convient", dans Pharo P., Quéré L.(eds), *Les formes de l'action* . Paris, Editions de L'EHESS, p.39-66.
- THEVENOT Laurent**, 1993, "Norme ou rationalité : une opposition dépassée ?", dans Gérard-Varet L.-A., Passeron J.-C., *Le modèle et l'enquête. Les usages du principe de rationalité dans les sciences sociales*, Paris, Ed. EHESS, p.2-30.
- TRICOU Fabrice**, 1994a, *Constitution de la Société Marchande. Règles d'Organisation et Règles de Comportement*, Thèse de Doctorat. Paris X-Nanterre.
- TRICOU Fabrice**, 1994b, "Rationalité du comportement des marchands et règles d'organisation des marchés", *Cahiers d'Economie Politique*, n°24-25, p. 38-67.

- ULLMANN-MARGALIT Edna., 1977, *The Emergence of Norms*, Clarendon Press, Oxford.
- VENTELOU Bruno, 1997a, "L'incertitude keynésienne et le projet pragmatique : l'hypothèse de l'instrumentalisme scientifique", *Economie et Société*, *Æconomia*, Histoire de la pensée économique, Série P.E., n°26, p.115-134.
- VENTELOU Bruno, 1997b, *Lire Keynes et le comprendre*, Paris, Vuibert.
- VON NEUMAN John et Oscar MORGENSTERN, 1944, *Theory of Games and Economic Behavior*, Princeton University Press.Princeton.
- VON STACKELBERG H., 1932, "Zwei Kritische bemerkungen zur preis-theorie Gustav Cassels", *Zeitschrift für Nationalökonomi*", vol 4, p.456-472.
- WAGNER P., 1994, " Action, coordination and institution in recent french debates", *The journal of political phylosophy*, 3, p.270-289.
- WALD Abraham, 1933, "über die eindeutige positive Lösbarkeit der neuen produktion-gleichungen", *Ergebnisse eines mathematischen kolloquiums*, vol.6, p.12-20.
- WALLISER Bernard, 1989, "Théorie des jeux et genèse des institutions", *Recherches économiques de Louvain*, vol 55, n°4, p.339-364.
- WALLISER Bernard, 1995, "Rationalité instrumentale et rationalité cognitive", *in* J.C. Passeron et L.A. Gérard Varet (eds), *Le modèle et l'enquête*, Paris, Editions de l'EHESS, p.309-322
- WALLISER Bernard,1996 , "A spectrum of equilibration process in game theory", *Actes du colloque "évolutionnisme : fondements et perspectives"*, Paris, Mai.
- WALLISER Bernard et Jean Louis RULLIÈRE, 1995, "De la spécularité à la temporalité en théorie des jeux", *Revue d'économie politique*, vol 105, n°4, p.601-615.
- WHITAKER John, 1987, *The Limits of Organisation Revisited*, *in* Feiwel G, *Arrow and the Ascent of Modern Economic Theory*, New York, Mac Millan Press
- WILLIAMSON Oliver E. et Sydney G. Winter, 1991, *The Nature of the Firm*, New-York, Oxford University Press.
- WILLIAMSON Oliver E., 1975, *Market and Hierarchies. Analysis and Antitrust Implications*, New York, The Free Press.
- WILLIAMSON Oliver E., 1979, "Transaction-cost economics : the governance of contractual relations", *Journal of Law and Economics*, vol.22, n°2, p.233-261.
- WILLIAMSON Oliver E., 1981, "The modern corporation : origins, evolution, attributes", *Journal of Economic Literature*, vol XIX, p.1537-1568.

- WILLIAMSON Oliver E., 1984, "The economics of governance : framework and implications", *Journal of Institutional and Theoretical Economics*, vol 140, p.195-223.
- WILLIAMSON Oliver E., 1985, *The Economic Institutions of Capitalism*, New York, The Free Press.
- WILLIAMSON Oliver E., 1987a, "Kenneth Arrow and the new institutional economics", in Feiwel G.(ed), *Issues in Contemporary Microeconomic and Welfare*, London, Macmillan Press, p.584-599.
- WILLIAMSON Oliver E., 1987b, "Transaction cost economics", *Journal of Economic Behavior and Organisation*, vol.8, p.617-625.
- WILLIAMSON Oliver E., 1989, "Transaction cost economics", in Schmalensee R. et R.D. Willig, *Handbook of Industrial Organisation*, vol 1, p.135-182.
- WILLIAMSON Oliver E., 1990, "Introduction", in Aoki, Gustafsson, Williamson (ed), *The Firm as a Nexus of Treaties*, London, Sage, p.1-25.
- WILLIAMSON Oliver E., 1993a, "The evolving science of organisation", *Journal of Institutional and Theoretical Economics*, vol 149/1, p.36-93.
- WILLIAMSON Oliver E., 1993b, "Calculativeness, trust and economic organisation", *Journal of Law and Economics*, vol.XXXVI, p.453-486.
- WITTGENSTEIN Ludwig, 1921, *Tractatus logico Philosophicus.*, traduction française (G.G. Granger) Paris, Gallimard, NRF, 1993.
- WITTGENSTEIN Ludwig, 1933-1945, *Philosophische Untersuchungen*, traduction française (P. Klossowski) : *Investigations Philosophiques*, Paris, Gallimard, NRF, 1961.
- WITTGENSTEIN Ludwig, 1919, "Lettre to J.M. Keynes", in Mac Guinness et Von Wright. [1995], p.114.
- WITTGENSTEIN Ludwig, 1937-1944, *Bemerkungen über die Grundlagen der Mathematik*, reviedierte und erweiterte Ausgabe, Francfort, Suhrkamp Verlag, 1974.
- YOUNG Harvey P., 1993, "The evolution of conventions", *Econometrica*, vol.61, n°1, p.57-84.
- YOUNG Harvey P., 1996, "The economics of convention", *Journal of Economic Perspectives*, Vol 10, n°2, Spring, P.105-122.
- ZEUTHEN F., 1932, "Das prinzip der knappheit, technische konbination und ökonomische qualitat", *Zeitschrift für Nationalökonomie.*, vol.4, p.1-24.

Table des Matières

| | |
|--|----|
| Introduction générale | 1 |
| Première partie : La modalité marchande | 17 |
| Chapitre 1 : Le modèle Arrow-Debreu : la formulation d'une loi ? | 21 |
| 1-De la main invisible au modèle d'équilibre général Arrow-Debreu | 24 |
| 11- La référence à Smith | 25 |
| 12- L'héritage walrassien | 28 |
| 13- Une question mathématique | 31 |
| 2- La théorie de l'équilibre général : le cadre d'analyse | 35 |
| 21-L'espace | 36 |
| 211- <i>Les biens Arrow-Debreu</i> | 36 |
| 212- <i>Un système complet de marché</i> | 37 |
| 213- <i>La temporalité de l'univers Arrow-Debreu</i> | 39 |
| 22- La loi de comportement de l'individu rationnel | 41 |
| 221- <i>Une rationalité instrumentale</i> | 42 |
| 222- <i>Une rationalité paramétrique</i> | 44 |
| 223- <i>Une rationalité calculatoire</i> | 44 |
| 224- <i>La loi de la demande</i> | 48 |
| 3- Une loi qui n'existe qu'à l'équilibre... | 50 |
| 31- Existence et optimalité de l'équilibre concurrentiel | 50 |
| 311- <i>L'optimalité de l'équilibre concurrentiel</i> | 51 |
| 312- <i>Le théorème d'existence</i> | 52 |
| 32- Comment atteindre l'équilibre ? | 54 |
| 321- <i>Le commissaire priseur</i> | 54 |
| 322- <i>Le théorème de Sonnenschein</i> | 57 |
| 33- Une économie bloquée à l'équilibre | 59 |
| Conclusion | 62 |

| | |
|--|-----------|
| Chapitre 2 : La critique d'Arrow | 65 |
| 1- Le théorème d'impossibilité | 69 |
| 11- Énoncé du problème | 71 |
| 111- <i>Platon versus Bentham</i> | 71 |
| 112- <i>Un classement portant sur des "états sociaux"</i> | 72 |
| 113- <i>La rationalité du choix collectif</i> | 73 |
| 12- L'axiomatique | 75 |
| 121- <i>Exposé de l'axiomatique</i> | 75 |
| 122- <i>Commentaire de l'axiomatique</i> | 76 |
| 13- Le théorème d'impossibilité | 79 |
| 131- <i>Démonstration</i> | 79 |
| 132- <i>Énoncé</i> | 80 |
| 14- Une reformulation du problème du choix collectif | 80 |
| 141- <i>Un changement dans l'axiomatique</i> | 81 |
| 142- <i>Un individu complètement rationnel ?</i> | 82 |
| 143- <i>Le rejet du marché</i> | 83 |
| 15- La "solution" d'Arrow | 84 |
| 151- <i>Un consensus d'origine "biologique et culturelle"</i> | 86 |
| 152- <i>L'obstacle de l'incertitude</i> | 86 |
| 153- <i>Un rapprochement entre Arrow et Hobbes</i> | 90 |
| 2- L'analyse de l'incertitude | 94 |
| 21- L'incomplétude du système de marché | 95 |
| 211- <i>L'absence de marchés futurs</i> | 95 |
| 214- <i>Le principe des marchés contingents</i> | 96 |
| 215- <i>L'échec des marchés contingents</i> | 98 |
| 22- La représentation standard du comportement rationnel dans l'incertain | 100 |
| 221- <i>La maximisation de l'utilité espérée</i> | 101 |
| 222- <i>Des probabilités subjectives</i> | 103 |
| 23- La conception "bayésienne" d'Arrow | 105 |
| 231- <i>le jugement subjectif est un jugement de probabilité</i> | 106 |
| 222- <i>Le fondement informationnel des probabilités</i> | 106 |
| 233- <i>Arrow : l'accent sur la dimension bayésienne de la prise de décision</i> | 107 |
| 24- L'information | 109 |
| 241- <i>L'information n'est pas une marchandise</i> | 110 |
| 242- <i>L'information fait resurgir l'altérité</i> | 111 |

| | |
|--|----------------|
| 3- Vers l'organisation... | 113 |
| 31- Un besoin d'unité | 114 |
| 311- <i>Le rôle des codes éthiques</i> | 114 |
| 312- <i>La gestion collective de l'information</i> | 116 |
| 32- Le rôle des organisations. | 118 |
| 321- <i>La justification des organisations</i> | 119 |
| 322- <i>Autorité versus consensus</i> | 120 |
| 323- <i>Une description de l'autorité</i> | 122 |
| 323- <i>La centralisation de l'information</i> | 122 |
| 33- La structure hobbienne des organisations | 123 |
| Conclusion | 127 |
| Seconde partie : la modalité organisationnelle | 129 |
| Chapitre 3 : L'analyse de Williamson | 133 |
| 1-La synthèse williamsonienne | 136 |
| 11- Deux révélateurs : les lois antitrust et l'article de Coase. | 136 |
| 12- Le retour, limité, aux institutionnalistes. | 138 |
| 13- La filiation Arrow-Williamson. | 140 |
| 2- Contractual man | 144 |
| 21- Un individu à rationalité limitée ? | 145 |
| 211- <i>Une rationalité maximisatrice mais limitée ?</i> | 145 |
| 212- <i>Arrow-Stigler ou Simon ?</i> | 146 |
| 22- Un individu opportuniste | 148 |
| 221- <i>Dans la lignée d'Arrow</i> | 148 |
| 222- <i>Une généralisation</i> | 149 |
| 23- La spécificité des actifs | 149 |
| 231- <i>La notion d'actif spécifique</i> | 151 |
| 232- <i>La théorie de la transformation fondamentale</i> | 152 |

| | |
|---|------------|
| 24- La nécessité des règles | 154 |
| 3- La règle hiérarchique | 157 |
| 31- Un seul critère de choix : l'efficacité. | 158 |
| 32- Informer, inciter, surveiller | 159 |
| 33- La "démonstration" de la supériorité du mode hiérarchique | 161 |
| 331- <i>La supériorité "historique" de la hiérarchie</i> | 161 |
| 332- <i>La supériorité empirique de la hiérarchie</i> | 162 |
| 34- Les raisons de l'efficacité de la hiérarchie | 164 |
| 341- <i>La nécessité d'une structure de gouvernance</i> | 164 |
| 342- <i>Langage et coercition</i> | 165 |
| 35- Le pouvoir aux propriétaires. | 167 |
| 351- <i>Une hiérarchie naturelle</i> | 167 |
| 352 - <i>Le pouvoir décisionnel aux actionnaires</i> | 168 |
| 353- Une justification des structures capitalistes ? | 169 |
| 4- Force analytique et faiblesse pratique du projet Williamsonien | 171 |
| 41- Le refus de l'éthique : l'exemple de la confiance | 171 |
| 411- <i>L'absence de fiabilité de "contractual man."</i> | 172 |
| 412- <i>Une confiance institutionnalisée</i> | 173 |
| 42- L'organisation est-elle la seule institution ? | 175 |
| 421- <i>Un modèle abstrait ...</i> | 175 |
| 422- <i>...dans la suite logique de la théorie de l'équilibre général</i> | 176 |
| 43- Un problème pratique. | 177 |
| 431- <i>Le problème de l'engagement</i> | 178 |
| 432- <i>L'impossible mesure des coûts de transaction</i> | 179 |
| 433- <i>L'obscur origine des organisations</i> | 180 |
| 5- Une théorie typiquement hobbienne | 183 |
| Conclusion | 186 |

| | |
|---|------------|
| Chapitre 4 : L'analyse stratégique de l'émergence des règles | 188 |
| 1- L'analyse de l'institution en théorie des jeux | 191 |
| 11- Le cadre d'analyse | 192 |
| 111- <i>Une approche non coopérative</i> | 192 |
| 112- <i>L'analyse de Schelling</i> | 194 |
| 12- Les situations paradigmatiques qui nécessitent une règle | 198 |
| 121- <i>Le problème de la coordination: l'exemple du "jeu du rendez-vous"</i> | 198 |
| 122- <i>Le cas de la coopération : l'exemple du dilemme du prisonnier</i> | 199 |
| 123- <i>Un troisième cas paradigmatique "intermédiaire" : la guerre des sexes</i> | 202 |
| 13- La typologie d'Edna Ullmann-Margalit | 203 |
| 131- <i>Un projet structuraliste et fonctionnaliste</i> | 203 |
| 132- <i>Trois types de normes</i> | 204 |
| 133- <i>Coopération et coercition</i> | 206 |
| 14- La question de la genèse des institutions | 208 |
| 141- <i>Répétition versus coercition</i> | 208 |
| 142- <i>Les règles collectives comme support des croyances</i> | 209 |
| 2- L'analyse de Lewis | 212 |
| 21- Un problème logique quant à la constitution du langage | 213 |
| 211- <i>Les deux dogmes de l'empirisme logique</i> | 213 |
| 212- <i>Le radicalisme de Quine</i> | 215 |
| 213- <i>La critique de Lewis</i> | 217 |
| 22- La définition de la convention chez Lewis | 219 |
| 221- <i>"Putting ourselves in the other fellow's shoes"</i> | 220 |
| 222- <i>Un exemple de coordination réussie</i> | 220 |
| 223- <i>La solution de Lewis</i> | 221 |
| 23- Les caractéristiques de la convention lewissienne | 223 |
| 231- <i>La convention : avant les mots</i> | 224 |
| 232- <i>La convention n'est pas contractuelle</i> | 225 |
| 233- <i>La coordination sans obligation</i> | 225 |
| 24- De Hume à Lewis : la disparition du temps historique. | 227 |
| 241- <i>L'héritage de Hume</i> | 227 |
| 242- <i>Une théorie logique de la convention</i> | 230 |

| | |
|---|------------|
| 243- <i>Une convention sans dimension "pratique"</i> | 232 |
| 3- La convention comme fondement des institutions : une impasse ? | 234 |
| 31- L'essai de synthèse néoinstitutionnelle de Schotter | 235 |
| 311- <i>Le projet de Schotter</i> | 235 |
| 312- <i>Une "convention hobbienne" ?</i> | 236 |
| 313- <i>Un modèle évolutionniste-bayésien</i> | 238 |
| 314- <i>Une explication partielle</i> | 241 |
| 32- La théorie "humienne" de Young Back Choi. | 243 |
| 33- L'analyse d'Axelrod | 246 |
| 331- <i>La théorie des jeux évolutionnistes</i> | 246 |
| 332- <i>L'analyse d'Axelrod</i> | 247 |
| 333- <i>Critique du "modèle" d'Axelrod</i> | 251 |
| 34- L'évolutionnisme conventionnaliste | 256 |
| Conclusion | 263 |
| | |
| <i>Troisième partie : vers une théorie pratique des règles ?</i> | 265 |
| | |
| Chapitre 5 : Information, coordination et règles chez Keynes | 270 |
| | |
| 1- Une rationalité keynésienne | 274 |
| | |
| 11- Une approche logique non calculatoire | 274 |
| 111- <i>Le rejet des approches quantitatives du jugement dans l'incertain</i> | 275 |
| 112- <i>La relation de probabilité</i> | 277 |
| 12- L'ancrage pratique de la rationalité | 278 |
| 121- <i>Le concept de poids de l'argument</i> | 279 |
| 122- <i>Une incertitude liée à la nature pratique des connaissances</i> | 280 |
| 14- Une rationalité intuitive. | 281 |
| 141- <i>Une relation relative à l'acteur</i> | 281 |
| 142- <i>Une relation objective</i> | 282 |

| | |
|---|------------|
| 2- Du Treatise à la Théorie Générale : fidélité et évolution | 284 |
| 21- Le débat Keynes-Ramsey | 284 |
| 211- <i>Keynes cède à Ramsey...</i> | 285 |
| 212- <i>...Mais pas sur l'essentiel</i> | 287 |
| 213- <i>Le refus du principe de la maximisation de l'utilité espérée</i> | 288 |
| 22- La découverte des règles. | 290 |
| | |
| 3- Raison et convention dans la Théorie Générale | 293 |
| 31- Les déterminants du prix des biens capitaux. | 294 |
| 311- <i>Le financement de l'investissement</i> | 294 |
| 312- <i>La détermination du taux d'intérêt</i> | 296 |
| 313- <i>La détermination du rendement escomptée des actions</i> | 298 |
| 32- Un jugement conventionnel | 299 |
| 33- Le jugement "véritable" de l'entrepreneur | 303 |
| 331- <i>raison keynésienne versus rationalité néoclassique</i> | 303 |
| 332- <i>Le triomphe de la spéculation</i> | 305 |
| 333- <i>L'intuition raisonnable de l'entrepreneur</i> | 306 |
| | |
| 4- Une théorie "élitiste" des règles | 309 |
| 41- Une analyse pragmatique ? | 310 |
| 42- La connaissance de l'entrepreneur contre l'arbitraire des conventions | 311 |
| 43- Le réalisme platonicien de Keynes | 313 |
| 44- L'élitisme de Keynes | 315 |
| 45- Une théorie "problématique" de la règle. | 318 |
| | |
| Conclusion | 320 |

| | |
|--|-----|
| Chapitre 6 : l'Économie des Conventions | 322 |
| 1- Une hypothèse de rationalité procédurale | 325 |
| 11- Une rationalité "limitée" | 326 |
| 111- <i>Une conception interne de l'incertitude</i> | 326 |
| 112- <i>La tentative alternative de la théorie du "search"</i> | 327 |
| 12- La rationalité procédurale | 329 |
| 13- La règle chez Simon. | 331 |
| 131- <i>Le rôle et la nature de l'environnement institutionnel de l'individu</i> | 332 |
| 132- <i>Une conception instrumentale des règles</i> | 333 |
| 2- Le projet conventionnaliste | 336 |
| 21- Le projet conventionnaliste | 337 |
| 211- <i>La relativisation du marché</i> | 338 |
| 212- <i>La critique de la figure réductrice de l'homo œconomicus</i> | 339 |
| 22- Brève histoire de l'économie des conventions | 341 |
| 221- <i>Olivier Favereau : la convention comme instrument entrepreneurial</i> | 341 |
| 222- <i>Les conventions comme "formes" de coordination</i> | 343 |
| 223- <i>Des problèmes symétriques</i> | 345 |
| 23- De la convention à la cognition | 347 |
| 231- <i>La "cognition située" : l'accent mis sur les objets</i> | 348 |
| 232- <i>Une interprétation cantonnée aux limites de l'entreprise</i> | 350 |
| 24- La force d'attraction de la théorie des jeux | 352 |
| Conclusion | 354 |
| Conclusion générale | 357 |
| Bibliographie | 366 |