

Théorie des conventions

**Contributions de Batifoulier P., de Larquier G., Abecassis P., Biencourt O., Chaserant C., Gannon F., Rebérioux A., Thévenon O., Gabriel P., Merchiers J., Urrutiager D.
Préface de Favereau O., Postface de Eymard-Duvernay F.
Economica, collection Forum.**

Le livre publié par une équipe de chercheurs de Forum (Université ParisX-Nanterre) sous la direction de Philippe Batifoulier qui s'intitule "*Théorie des conventions*" est la hauteur de l'ambition affichée par le titre. En proposant une "théorie", les auteurs se devaient d'en offrir un exposé compréhensif, dans les différents sens de ce terme, ce dont ils s'acquittent brillamment en offrant une synthèse mais aussi en défrichant des questions nouvelles. La notion de convention étant mobilisée par des programmes de recherches différents, voire radicalement opposés dans leur perspective, la tâche n'était pas aisée a priori. L'ouvrage est construit autour de deux grandes visions de la notion de convention, à partir desquelles sont distinguées deux lignes de développement théorique de cette notion. La première vision de la convention (attribuée à Hume) est d'ordre instrumental, la convention est un accord général qui permet de se coordonner sans engagement et n'apparaît que comme un moyen particulier de poursuivre son propre intérêt, elle débouche sur un ordre spontané. Cette vision correspond à l'approche stratégique développée dans le cadre de la théorie des jeux (et dans le prolongement de l'économie standard) pour laquelle la convention est une règle de comportement, résultat d'une interaction stratégique. Dans la seconde vision (attribuée à Weber), le respect de la convention est subordonné à l'existence d'une sanction sociale des comportements autre que le gain immédiat qui en est retiré, ce qui introduit un élément normatif ou prescriptif dans la définition de la convention. Elle est développée par l'approche dite "interprétative", dans laquelle non seulement les règles (ou une partie des règles) sont d'ordre conventionnel, mais aussi les modèles d'évaluation qui permettent d'interpréter toute les règles. Ainsi, au total, non seulement l'ouvrage inclut un exposé systématique et rigoureux de la théorie des conventions développée par la théorie des jeux à la suite des travaux de Schelling et Lewis et avec l'apport récent des jeux évolutionnistes, mais il reconsidère également les rapports entre la conception à la Lewis des conventions, qui est une des inspirations de l'"Economie des conventions" (Revue économique, mars 1989), et celle-ci. Ce travail d'éclaircissement a conduit les auteurs à faire justice de certaines des critiques adressées à cette dernière approche et à développer la place du politique dans celle-ci et même sa dimension politique, qu'affirme Eymard-Duvernay dans sa postface. Cela suscitera de nouveaux débats, c'est en tout cas ce que je souhaite comme succès à ce travail.

L'ouvrage comprend trois parties. La première introduit la convention comme objet théorique en reconnaissant deux pères fondateurs de la théorie économique des conventions : Keynes, qui a le premier introduit la notion de convention comme une représentation commune qui stabilise les anticipations, et Lewis, qui lui a donné ses fondements analytiques, en cherchant à résoudre un paradoxe de la philosophie du langage et en défendant la thèse l'on peut fonder des conventions sans communiquer car la connaissance commune de la rationalité de chacun permet de se mettre d'accord sans recourir au langage. La convention est alors définie comme une régularité de comportement, où chacun se conforme au comportement qu'il croit que l'autre adoptera. En dépit du caractère arbitraire de la convention (une autre solution est toujours possible pour se coordonner) il est rationnel de la suivre. Les deux autres parties développent respectivement l'approche stratégique des conventions, qui modélise l'émergence des conventions mais qui pour l'essentiel ne parvient pas actuellement à en expliquer les dynamiques, et l'approche interprétative, qui s'interroge sur les fondements conventionnels des règles, qui est celle développée par le programme de recherches de l'Economie des conventions (EC).

C'est la relativisation du pouvoir coordinateur du marché qui a constitué la condition nécessaire pour que la notion de convention acquière un statut scientifique dans le champ de l'économie. Cette relativisation s'est faite dans deux directions. Le marché est un moyen (ou un espace) de coordination parmi d'autres, structuré par des règles "constitutives". Les conditions qui définissent le marché de la théorie de l'équilibre général sont particulièrement restrictives (diffusion parfaite de l'information dans le temps et dans l'espace, pas d'interactions stratégiques) et ne constituent qu'un espace particulier de coordination (par les prix). Dans le cas général, les individus se coordonnent sur la base de règles qu'ils élaborent localement, dans des espaces structurés par des règles constitutives, autres que celles du marché. Ces espaces "locaux" sont des cadres d'interactions stratégiques. Les règles de coordination, dans un cadre défini, sont des règles dites "régulatrices" (la distinction entre les deux types de règles est également empruntée à un philosophe, Searle).

Pour que la coordination soit envisageable entre individus qui s'adaptent aux comportements des autres (en supposant que des performances cognitives extrêmement élevées soient envisageables) il est nécessaire de supposer un minimum de savoir commun, en l'occurrence le principe de rationalité lui-même ; ce qui permet d'introduire les règles dans le cadre de pensée de l'économie. L'approche stratégique rend ainsi compte des règles qu'elles soient associées à des sanctions explicites (règles juridiques) ou implicites comme c'est le cas de la convention au sens de Weber. Plus précisément sont distinguées parmi les règles régulatrices : la règle-contrat (implicite ou explicite) qui marque un accord de volonté résultant de la maximisation des intérêts individuels sous contraintes et conditionnelle au choix des autres, la règle-convention qui est déjà là et que l'on suit en fin de compte par intérêt, ce qui renforce son pouvoir régulateur, quoique son origine soit arbitraire (plusieurs options étaient, voire restent, envisageables). Mais pour que ces règles puissent être modélisées (et penser par les acteurs) comme le produit d'interactions intéressées, il faut, en amont, des règles, qui sont de l'ordre des représentations, qui constituent l'espace où un calcul d'intérêt est possible et où une formule d'investissement a un sens. Ces règles instituent les pratiques. Ces règles ne sont intrinsèquement présentes dans le principe de rationalité, elles ne sont donc pas préalables à l'interaction. Autrement dit, les règles de l'économie (d'un monde "rationnel") sont à la fois régulatrices (elles permettent de se coordonner ou de coopérer) et constitutives (elles fondent l'espace de coopération). Considérer que les règles sont constitutives, c'est au minimum doter les individus d'une compétence d'interprétation réflexive (de l'environnement comme de leurs propres actions) et plus exactement d'une capacité à former des représentations (collectives) plutôt que d'une capacité d'apprentissage qui conduirait à un savoir presque parfait. C'est à Keynes que l'on doit d'avoir introduit ainsi le rôle des conventions dans la formation des représentations, mais ce versant de son œuvre n'a suscité qu'assez récemment un intérêt (de la part des auteurs conventionnalistes). Ce point de départ est développé dans la troisième partie.

La proposition de Lewis permet d'un point de vue analytique de concilier la rationalité des individus et l'arbitraire de la convention suivie. La convention telle que la définit Lewis rend compte d'une coordination entre individus dotés d'une rationalité parfaite, sans recours à une autorité extérieure. La parenté avec les hypothèses constitutives de la théorie néo-classique est nette. Le chapitre consacré à Lewis fait un compte-rendu détaillé des débats qui ont suivi sa proposition, tant des critiques externes qui portent sur l'hypothèse de *common knowledge*, que des critiques internes qui proposent des éléments de définitions alternatifs et qui aboutissent in fine à réduire le champ des conventions au sens de Lewis.

La seconde partie de l'ouvrage présente différents travaux dans le domaine de la théorie des jeux qui s'intéressent au choix d'une convention dans des contextes variés, à la dynamique des conventions et à la rivalité entre elles ou à leur coexistence (jeux évolutionnistes). L'analyse stratégique des conventions est d'abord rapportée aux différentes classes de jeux non coopératifs : le jeu du "rendez vous" (convention à la Lewis) ; le "dilemme du prisonnier", qui

comme le soulignent les auteurs occupe une place emblématique dans les sciences sociales comme parabole de l'action collective parce qu'il explique comment on peut facilement ne pas choisir la bonne solution mais aussi parce qu'il valorise la coopération (quoique de façon perverse dans le cas des "prisonniers" puisqu'elle permet d'échapper à la règle sanction...) ; le jeu de la "chasse au cerf", moins conflictuel que le précédent, qui oppose une stratégie Pareto optimale (rester à son poste de guet pour rabattre le cerf, coopérer mais sans être sûr du résultat à ce coup ci) et une "stratégie dominante en risque" (Harsanyi et Selten) (quitter son poste pour attraper à coup plus sûr un lièvre et abandonner le projet collectif), que les auteurs préféreraient dans le rôle du précédent, sans doute parce que les deux stratégies en concurrence sont plus aisément justifiables dans un monde commun (plus morales). Lorsque la stratégie Pareto optimale n'est pas un équilibre de Nash (dilemme du prisonnier), on ne peut concevoir que des conventions "externes" qui supportent un équilibre intermédiaire. C'est le cas de la "convention d'effort" proposée par Leibenstein (les salariés ont un intérêt collectif à déployer un certain effort dans leur travail pour maintenir l'entreprise en état de marche, ce niveau est fixé arbitrairement par l'observation de l'effort de l'autre) qui a un caractère assez général dans les situations de coopération dans un contexte économique. D'un point de vue externe à la théorie des jeux, on peut sans doute considérer que c'est ce type de convention (plutôt que celles qui sont représentées par un équilibre de Nash) qui correspond le mieux non seulement à la notion d'arbitraire de la convention mais aussi à l'approche interprétative. En effet, le raisonnement de Leibenstein fait apparaître un groupe de salariés, c'est-à-dire un monde commun autre qui croise le monde commun de l'entreprise. La coordination horizontale des salariés permet la coordination verticale avec l'entreprise (la continuité de l'emploi et plus généralement d'une coopération). La convention qui est "externe" au jeu qui formalise le monde de l'entreprise est "interne" au monde des salariés sous forme d'une "éthique professionnelle", qui cristallise des représentations.

La question de la dynamique des conventions, esquivée par Lewis qui se repose sur la force du précédent, débouche sur la question de l'apprentissage. Si de nombreuses règles de sélection ont été modélisées, la théorie des jeux se heurte au problème de la genèse des conventions, car toute règle de sélection des stratégies est déjà une convention. Pour expliquer l'émergence d'une convention entre des joueurs rationnels, il faut supposer qu'ils suivent déjà certaines conventions. L'idée d'un recours à une méta-règle permettant de choisir la règle à appliquer dans un contexte particulier ne fait que reporter la question. Mais cette impasse est-elle si grave ? Elle ne l'est pas en tout cas dans une approche interprétative puisque celle-ci envisage des individus qui sont effectivement déjà dotés de représentations qui acquièrent, lorsqu'elles sont mises en commun, un sens conventionnel. Il est toutefois intéressant d'imaginer des conjectures dans lesquelles une dynamique des conventions, c'est-à-dire un changement institutionnel, peut intervenir en restant dans un cadre d'interaction stratégique et sans imaginer de grandes capacités réflexives des individus (mais sans grand contrôle non plus sur le sens de ce changement...). La théorie des jeux évolutionnistes répond à cet objectif avec une hypothèse de rationalité limitée qui lui permet de mettre en jeu des individus "myopes" et "naïfs". Son objet est la dynamique d'une convention dans une population qui interagit localement, c'est-à-dire les questions de son existence, de sa rivalité avec d'autres, de sa survie, de son croisement (ou coexistence) avec d'autres. On peut dans ce cadre définir des équilibres. Les travaux les plus récents, présentés dans l'ouvrage, approfondissent l'analyse des différentes structures de population et formes de réseaux. Ces travaux conjecturent des mondes pluriels et structurés et montrent in fine que ni l'ordre ni le changement n'ont besoin pour être conçus en théorie de beaucoup de rationalité et en tout cas bien moins que n'en suppose la version néo-classique de l'économie. Mais tout cela ne nous dit pas où va le monde...

L'approche stratégique, intégrée par la science économique orthodoxe, considère des acteurs (en fait des "joueurs") parfaitement rationnels, maximisateurs d'utilité (persuadés de

l'efficacité du marché) et placés dans un monde où la conception de la justice a un champ réduit et n'est représentée que par un critère unique d'efficacité (Pareto optimalité). Les solutions qu'elle propose pour certains types de problèmes de coordination (toutefois parfois au prix de l'affaiblissement des hypothèses de base comme dans les jeux évolutionnistes) reposent sur un ordre spontané. On le sait, la critique la plus ancienne de cette représentation de l'économie dénonce cet ordre comme un ordre "barbare". L'économie des conventions de son côté avait, semble-t-il, jusqu'à des travaux récents, essentiellement fait porter sa critique sur le caractère arbitraire de l'ordre économique, sur la pluralité des mondes possibles, des modèles d'organisation de l'entreprise et de la cité et sur leurs compositions dans les mondes réels. La troisième partie de l'ouvrage revient sur cette question et la complète en esquissant une économie politique plus radicale. Elle ambitionne à partir de l'incomplétude des règles et de la rationalité procédurale de proposer une théorie de la coordination économique qui va au delà de la l'objet convention.

La troisième partie débute par une présentation du courant de l'Economie des conventions depuis sa manifestation initiale dans le numéro de mars 1989 de la Revue économique jusqu'à des travaux plus récents. Trois aspects (articulés) sont mis en avant : le recours à l'hypothèse de rationalité procédurale qui reconnaît une double incertitude, une incertitude radicale "externe" qui porte sur l'environnement et une incertitude "interne " ce qui revient à prendre acte des limites des procédures d'optimisation ; une analyse conventionnaliste des principaux objets de la théorie économique, l'entreprise vue comme un espace cognitif et le marché vu comme un espace organisé ; la pluralité des modèles d'interprétation des règles qui est notamment à la base du modèle de la cité (ou des "Economies de la Grandeur") de Boltanski et Thévenot, qui introduit le politique. La capacité d'interprétation des individus ne peut en effet provenir que d'une certaine représentation du collectif (en fait des collectifs) dans lequel (ou lesquels) ils sont insérés : la coopération repose non pas tant sur une régulation des comportements (par des règles conventionnelles) que sur une régulation des représentations, que l'on ne peut se représenter, sauf à admettre la dictature des théorèmes du bien être, que sous une forme conventionnelle (ouverte au changement). L'introduction dans l'analyse du niveau des modèles d'évaluation qui coordonnent les représentations constitue un point de convergence qui apparaît comme le "folk theorem" de ce courant (l'EC). Il me semble toutefois que l'approche cognitive de l'organisation repose également sur une dualité des règles que l'on retrouve dans le couple ressources/compétences (développée notamment à la suite des travaux de Penrose) pour peu que la dimension collective des compétences soit mise en avant. Que serait en effet la ressource travail sans la "convention d'effort", qui est de l'ordre d'une compétence, qui l'active ? Je souligne ce point car il me semble que ce qui en fin de compte est à la source de l'arbitraire des conventions et des institutions (systèmes de conventions) c'est précisément que celles-ci ont simultanément plusieurs déterminants ce qui fait que on ne peut expliquer leur fonctionnalité par leur origine. L'ajustement des comportements (l'interprétation des règles) dans le cours même de l'action, qui passe par la coordination des représentations (sur le collectif) repose sur une capacité propre (ou compétence) du collectif a assuré effectivement cette coordination. C'est sans doute l'analyse de ce type de compétences et de leur dynamique qui forme le domaine de l'économie politique.

L'économie politique de l'EC repose sur deux points. Les représentations formés dans des collectifs portent à la fois sur le bon et le juste. La justice s'incarnant dans différentes conceptions du bien, il n'y pas de conception de la justice qui ait une position surplombante (contrairement à la conception libérale de l'économie). De ce fait les individus s'engagent, simultanément comme des personnes privées et comme des personnes publiques (non seulement dotées de représentations du collectif mais aussi de principes qui expriment ce que devrait être le collectif). Ainsi les entreprises et tout autres sortes de collectifs particuliers se prolongent ainsi dans une action publique et pas seulement dans des transactions privées. Le

passage à l'économie politique me paraît devoir impliquer non seulement la pluralité des règles et des modèles d'évaluation mais aussi la pluralité des collectifs (qui paraît plutôt un point d'entrée de l'approche dite "néo-corporatiste"). Je ne suis pas sûr que ce dernier problème soit complètement cerné par la figure de l'entreprise que propose l'EC. Peut-être parce que les travaux empiriques ont essentiellement porté sur les entreprises (à quelques exceptions près comme les travaux de Thévenot sur les conflits surgissant autour de projets d'aménagement public) négligeant d'autres types de collectif essentiels, comme par exemple les groupes professionnels, qu'il faut sans doute étudier pour enrichir l'analyse de la formation de l'identité collective. C'est dans ce cadre collectif, et non dans un cadre inter-individuel, que l'on peut introduire les conflits comme un élément de dynamique des conventions (de simples conflits interindividuels sont parfaitement saisis par l'analyse évolutionniste en terme de mutation et sélection). Cela pose au fond la question de l'individualisme méthodologique. Suffit-il d'en proposer une version améliorée en faisant des représentations des objets de l'ordre collectif ? En introduisant la notion de "capitalisme" dans leur analyse Boltanski et Chiapello paraissent (en sociologues) se dégager de cette limite. Eymard-Duvernay, dans sa postface fait de même en reprenant aux radicaux américains (en l'occurrence Bowles et Gintis) l'opposition entre la logique des droits de propriété, par lesquels s'exprime la structure autoritaire du capitalisme et la logique démocratique des droits des citoyens entre lesquels se sont construit historiquement des "compromis" institutionnels (et semblant militer comme eux pour une "démocratie post-libérale" qui responsabiliserait démocratiquement l'exercice des droits de propriétés comme de l'autorité publique). La distinction par l'approche interprétative entre les deux niveaux des règles et des modèles d'évaluation, qui enrichi subtilement l'ordre conventionnel d'une substance politique par la prise en compte du collectif, ne suffit pas à constituer une économie politique, du moins si elle en reste à une structuration de type unidimensionnelle. La prise en compte dans l'analyse de l'incomplétude et de la pluralité des règles comme des modèles d'évaluation et des principes, qui sont également d'ordre conventionnel, débouche sur l'idée d'une distribution des représentations qui fonde la multiplicité des collectifs et en fait des sujets de l'histoire.

Gilles Allaire
INRA-ESR