

# problèmes économiques

*Problèmes économiques* invite les spécialistes à faire le point

FÉVRIER 2015 NUMÉRO 7

HORS-SÉRIE

comprendre

L'ÉCONOMIE

1. Concepts et mécanismes



# 4 REVUES POUR ÊTRE

# BIEN INFORMÉ

## ÉCONOMIE

Tous les 15 jours,  
le meilleur de la presse et des revues  
pour comprendre l'actualité économique



→ EN VENTE EN KIOSQUE

## SOCIÉTÉ - débat public

Tous les deux mois,  
les grands sujets qui nourrissent le débat public



→ EN VENTE EN KIOSQUE

## EUROPE - INTERNATIONAL

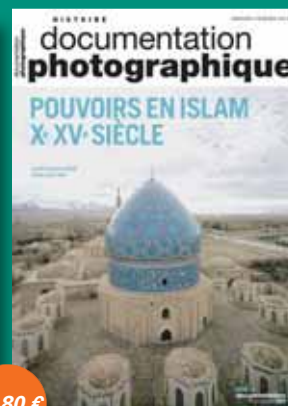
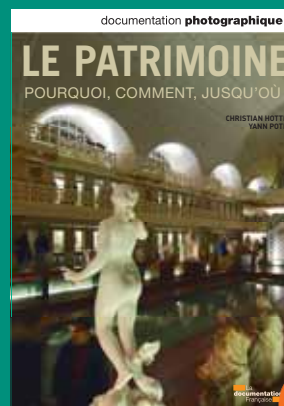
Tous les deux mois,  
vous avez rendez-vous avec le monde



→ EN VENTE EN KIOSQUE

## HISTOIRE - GÉOGRAPHIE

Tous les deux mois,  
l'histoire et la géographie à partir de documents



En vente chez votre libraire,  
sur [www.ladocumentationfrancaise.fr](http://www.ladocumentationfrancaise.fr)  
et par correspondance :  
DILA, 26 rue Desaix, 75727 Paris cedex 15

La  
**documentation**  
Française

# L'ÉCONOMIE AU CŒUR DU DÉBAT PUBLIC

## Direction de l'information légale et administrative

26, rue Desaix  
75015 Paris

### Rédaction

#### de *Problèmes économiques*

Patrice Merlot (rédacteur en chef)  
Olivia Montel  
(rédactrice en chef des hors-série)  
Markus Gabel (rédacteur)  
Stéphanie Gaudron (rédactrice)

### Secrétariat de rédaction

Anne Biet-Coltelloni

### Promotion

Anne-Sophie Château

### Secrétariat

Paul Oury

29, quai Voltaire  
75344 Paris cedex 07  
Tél. : 01 40 15 70 00  
pe@ladocumentationfrancaise.fr  
<http://www.ladocumentationfrancaise.fr/revues-collections/problemes-economiques/index.shtml>  
Abonnez-vous à la newsletter

### Avertissement

Les opinions exprimées  
dans les articles reproduits  
n'engagent que les auteurs

### Crédit photo :

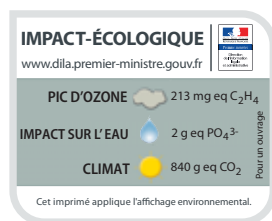
Couverture : Corbis  
© Direction de l'information légale  
et administrative. Paris, 2015

### Conception graphique

Célia Petry  
Nicolas Bessemoulin

En vente en kiosque et en librairie  
(Adresses accessibles en ligne)

#Retrouvez-nous sur  
Facebook et sur Twitter  
@ProbEcoPE



Nous vivons dans un monde où l'économie est omniprésente : pas un jour sans que les prévisions de croissance, les chiffres du chômage ou du déficit public ne fassent l'objet d'un traitement dans les médias. Ces dernières années ont été particulièrement riches en événements, puisque l'économie mondiale, et particulièrement la zone euro, traverse depuis 2008 une ère de turbulences. Clef d'analyse de nombreux problèmes de société, la compréhension des rouages de l'économie constitue un impératif majeur pour tous les citoyens. Il n'est donc pas surprenant que la question de la diffusion des connaissances économiques prenne une place aussi importante dans le débat public.

Avec cette série « Comprendre l'économie », *Problèmes économiques* participe une nouvelle fois à ce projet ambitieux de vulgarisation des savoirs. Le découpage retenu offre une première entrée par les concepts et mécanismes fondamentaux (volume 1) et une seconde par les grands problèmes économiques de notre temps (volume 2).

La réflexion des grands auteurs sur les phénomènes économiques a suscité depuis plusieurs siècles de vives controverses, qui ont progressivement construit la science économique, et dont on retrouve l'héritage dans les savoirs contemporains. C'est pourquoi ce dossier commence par un tour d'horizon des grands courants économiques, en insistant sur les évolutions marquantes de la discipline au cours des trente dernières années. Les trois sous-ensembles qui suivent quittent l'histoire des idées pour se concentrer sur les savoirs fondamentaux : sont ainsi analysés les outils de base des économistes d'aujourd'hui – externalités, imperfections d'information, théorie des jeux, instruments économétriques... –, le comportement – aux niveaux individuel et agrégé – des principaux agents – entreprises, banques, ménages, État –, leurs interdépendances, mais également la dynamique et les mécanismes de régulation de l'économie de marché. Pour finir, l'accent est mis sur le fonctionnement et les dysfonctionnements de certains marchés – marchés financiers, marché du travail.

Les thèmes retenus ne couvrent pas tout le spectre de l'analyse économique. Outre l'impossibilité d'être exhaustif et le manque de place, nous avons fait le choix de nous en tenir aux outils fondamentaux, avec l'ambition que ce numéro ne profite pas seulement aux initiés, mais également aux néophytes. Nombre d'outils classiques mais techniques, tels que le schéma IS-LM-BK, ne sont donc pas abordés ici en tant que tels. Les livraisons récentes de notre revue permettent, pour ceux qui le souhaitent, d'approfondir de nombreux points.

Olivia Montel



# COMPRENDRE L'ÉCONOMIE

## 1. CONCEPTS ET MÉCANISMES

### La science économique et ses évolutions

- P. 5 L'économie : une discipline à haut risque ? *(Pascal Le Merrer)*
- P. 12 Des classiques anglais à *Freakonomics* : l'évolution de la science économique *(Jean-Baptiste Fleury)*
- P. 19 Les grands temps de la macroéconomie *(Bruno Ventelou et Yann Videau)*
- P. 26 Que reste-t-il de l'équilibre général ? *(Sophie Jallais et Ozgur Gun)*

### La boîte à outils de l'économiste

- P. 33 Offre, demande et prix *(Jean-Pierre Biasutti et Laurent Braquet)*
- P. 41 PIB et croissance économique *(Thomas Fabre)*
- P. 50 Biens collectifs et externalités *(Marion Navarro)*
- P. 58 Les asymétries d'information *(Anne Corcos et François Pannequin)*
- P. 65 La théorie des jeux *(Nicolas Eber)*
- P. 73 Les outils économétriques *(Luc Behaghel)*
- P. 79 La monnaie dans les théories économiques *(Alain Beitone)*



## Les acteurs de l'économie

- P. 87 L'entreprise dans la théorie économique *(Olivier Weinstein)*
- P. 94 Comment (dys)fonctionnent les banques ? *(Jézabel Couppey-Soubeyran)*
- P. 102 Consommation/épargne : les choix des ménages *(Nicolas Drouhin)*
- P. 110 L'intervention publique *(Laurent Simula)*

## Fonctionnement et dysfonctionnements des marchés

- P. 118 Marché et marchés : une approche institutionnelle des modalités de l'échange économique *(François Vatin)*
- P. 125 Marchés, concurrence et réglementation *(Frédéric Marty)*
- P. 132 Marchés financiers : des marchés proches de l'idéal concurrentiel ? *(Yves Jégourel)*
- P. 138 Le marché du travail, un marché en mouvement *(François Fontaine)*



L'économie est régulièrement sous le feu des critiques, mais plus encore depuis la crise financière de 2007-2008. Simplement incapable de prévoir les phénomènes économiques pour certains, elle serait d'autant plus dangereuse pour d'autres du fait d'un discours « impérialiste » et de son influence sur les décideurs politiques. Pour d'autres encore, elle serait immorale, légitimant la « marchandisation » du monde et les comportements de l'*homo oeconomicus*. Faisant le point sur ces griefs, Pascal Le Merrer montre que la discipline a su en partie y répondre et évoluer vers plus de réalisme. Loin de rester cantonnés à l'élaboration de modèles abstraits, les économistes orientent leurs recherches vers des sujets qui sont au cœur des problèmes économiques contemporains.

Problèmes économiques

# L'économie : une discipline à haut risque ?

## © PASCAL LE MERRER

Professeur d'économie à l'ENS de Lyon  
Directeur des Journées de l'économie

L'économie est dans une période singulière. D'un côté, les économistes français sont à l'honneur avec le prix Nobel de Jean Tirole, le succès international du livre de Thomas Piketty<sup>1</sup>, la présence de sept Français dans la liste du FMI des vingt-cinq « économistes âgés de moins de 45 ans qui exerceront, au cours des prochaines décennies, le plus d'influence sur notre compréhension de l'économie mondiale<sup>2</sup> ». Dans le même temps, on voit se développer un débat sur le rôle des économistes qui auraient pris la place des philosophes, des sociologues, des historiens, comme le dénonce par exemple Marcel Gauchet (cf. Zoom).

<sup>[1]</sup> Piketty T. (2013), *Le Capital au XXI<sup>e</sup> siècle*, Paris, Le Seuil.

<sup>[2]</sup> FMI (2014), « La génération montante », *Finances et Développement*, septembre.

## Sous le feu de la critique

L'économie est souvent critiquée, mais avec des arguments quelque peu contradictoires. D'un côté, il lui est reproché l'irréalisme des hypothèses et des raisonnements qui soutiennent ses modèles. D'un autre côté, on prête un pouvoir énorme à cette discipline qui légitimerait une marchandisation du monde.

## Des hypothèses et des modèles irréalistes

Steve Keen (2014) illustre la première catégorie de critiques quand il dénonce les erreurs logiques d'une théorie fondée sur le calcul rationnel des agents incapable de décrire la réalité et encore moins d'anticiper des phénomènes comme les crises économiques. C'est l'analyse néoclassique qui est au centre de cette critique, avec le comportement de

## ZOOM

### « LES ÉCONOMISTES SONT-ILS DES IMPOSTEURS ? »

« Que s'est-il passé intellectuellement dans nos sociétés ? La montée de l'économisme correspond, du point de vue des sensibilités, à un changement très profond de la demande sociale d'intelligence. Le désir d'intelligibilité a été supplanté par le souci d'efficacité d'un système conçu comme le seul possible. La question n'est plus de comprendre ce que sont l'homme, la société, l'histoire... La question est juste de savoir comment ça marche et comment faire en sorte que ça marche mieux.

C'est en ce sens que l'expert a pris la relève de l'intellectuel. Le mot est atroce mais parlant : on assiste à une désintellectualisation de nos sociétés. Elle va de pair avec une lecture de l'existence collective réduite au droit, à l'économie et à la technique »\*.

\* Marcel Gauchet, « Les économistes sont-ils des imposteurs ? », *L'Obs*, 23 octobre 2014.

l'offreur et du consommateur, la concurrence pure et parfaite, la rémunération des facteurs de production à leur productivité marginale, l'équilibre général walrasien, les anticipations rationnelles. Pour Steve Keen, c'est le modèle de coordination parfait grâce aux prix flexibles permettant d'équilibrer tous les marchés qui doit être dénoncé. Mais il sous-estime que la théorie néoclassique est compatible avec une grande variété d'arrangements institutionnels : elle ne se limite pas à produire un discours justifiant le libre marché. En revanche, cet auteur dénonce à juste titre le problème de la dichotomie entre les variables réelles et les variables financières dans la théorie néoclassique, ce qui le conduit à proposer « un modèle monétaire du capitalisme ».

## Une discipline immorale

Timothy Taylor (2014) examine l'autre type de critique. Il rappelle que « les économistes préfèrent éluder les questions morales. Ils se plaisent à dire qu'ils étudient les arbitrages, les incitations et les interactions, et qu'ils laissent les jugements de valeur au processus politique et à la société ».

Toutefois, Jean Tirole (2014) précise : « Aux yeux des économistes, le marché est un puissant mécanisme d'allocation des ressources. Cependant, bénéficiaire de ses vertus requiert souvent de s'écarter du laisser-faire. De fait, les économistes ont consacré beaucoup de recherches à l'identification de ses défaillances et à leur correction par la politique publique. Mais les philosophes, les psychologues, les sociologues, les juristes, les politistes, beaucoup de grands acteurs de la société civile et la plupart des leaders religieux ont une vision différente du marché. Ils reprochent aux économistes de ne pas tenir assez compte des problèmes d'éthique, réclament une frontière claire entre domaines marchand et non marchand. » On voudrait penser que les individus sont vertueux mais dès que l'on étudie les comportements addictifs, les actions illégales, les actes individuels qui négligent les effets sociaux et environnementaux, on est confronté à la nécessité de comprendre comment agir pour favoriser des comportements prosociaux. Faut-il réglementer, taxer, créer un marché, jouer sur des incitations non marchandes ? Toutefois, peut-on étendre le raisonnement économique à tous les domaines ? Les débats sur les soldats mercenaires, le commerce du sexe, les droits à polluer sont anciens, mais, aujourd'hui, on voit naître de nouveaux sujets : celui des mères porteuses, celui du commerce des organes et du sang, celui de la brevetabilité du vivant avec le séquençage du génome... Pourquoi alors ne pas aussi faire le commerce des droits d'immigration plutôt que de laisser les passeurs créer un marché parallèle ? Ce serait une erreur d'en déduire que l'économie s'oppose à l'éthique. En fait, sur de nombreux



## ZOOM

### L'ÉCONOMIE, UNE DISCIPLINE CONTRE-INTUITIVE

L'économie est une discipline contre-intuitive depuis ses origines. Cela commence avec Adam Smith dans *La Théorie des sentiments moraux* lorsqu'il insiste sur la vanité et le désir de s'enrichir des hommes. Il montre que l'enrichissement de la société est le produit des comportements égoïstes et vaniteux des individus. Le problème est : comment faire pour que le vice privé devienne une vertu publique ? Il explique que la solution est dans l'instauration d'un régime économique qui donne libre cours aux initiatives privées (propriété privée, concurrence, ouverture des marchés...). Cette approche conduit à un discours qui est très étonnant, puisqu'il propose de se priver de diriger l'économie. On a des unités qui planifient leurs actions et le marché qui coordonne ces actions. Ainsi, le pessimisme concernant la nature humaine conduit à l'optimisme quant aux effets économiques. Et ce discours contre-intuitif va s'appuyer sur une démonstration de la supériorité de l'économie de marché pour coordonner les

intérêts individuels des agents économiques, qui est exposée dans *La Richesse des nations*. Il explique que lorsque la recherche de l'enrichissement illimité prend la forme de l'accumulation du capital, celle-ci perd son caractère immoral car l'accumulation du capital est fondée sur l'épargne qui est une vertu. Cette analyse ne le conduit pas à négliger le rôle de l'État comme le montre la définition qu'il donne de l'économie politique (cf. *infra*). Il consacre d'ailleurs la totalité du livre V de *La Richesse des nations* aux problèmes des dépenses et des recettes de l'État. Pour Adam Smith, « l'économie politique, considérée comme une branche des connaissances du législateur et de l'homme d'État, se propose deux objets distincts : le premier, de procurer au peuple un revenu ou une subsistance abondante, ou pour mieux dire, de le mettre en état de se procurer lui-même ce revenu ou cette subsistance abondante ; le second, de fournir à l'État ou à la communauté un revenu suffisant pour le service public ; elle se propose d'enrichir à la fois le peuple et le souverain »\*.

Pascal Le Merrer

\* Smith A. (1776), *La Richesse des nations*, introduction au volume II, Paris, Garnier Flammarion, 1991, p. 11.

<sup>[3]</sup> Voir le document de travail 14/3 sur [http://www.maxpo.eu/publications\\_DP.asp](http://www.maxpo.eu/publications_DP.asp).

<sup>[4]</sup> Stiglitz J. (2014), *Principes d'économie moderne*, 4<sup>e</sup> édition française, De Boeck, p. 17.

sujets, il faut prendre en compte des questions morales quand on traite de problèmes matériels. C'est vrai pour l'économie comme pour les autres sciences sociales.

Il reste que la crainte non fondée de voir les économistes justifier la marchandisation du monde révèle peut-être une autre inquiétude, celle d'une position dominante des représentants de cette discipline qui conseillent les dirigeants politiques et économiques, produisent des expertises pour les grandes organisations nationales et internationales et enfin s'expriment largement dans les médias. Ce sujet est aussi étudié comme le montre l'article récent de Marion Fourcade, Étienne

Ollion, et Yann Algan : « The Superiority of Economists<sup>3</sup> ».

### L'économie est-elle une science ?

Le manuel de Joseph Stiglitz commence par cette affirmation : « L'économie est une science sociale. Elle étudie le problème des choix dans une société d'un point de vue scientifique, c'est-à-dire à partir d'une exploration systématique qui passe aussi bien par la formulation de théories que par l'examen des données empiriques<sup>4</sup>. »

L'économiste construit des raisonnements logiques qui s'appuient sur des hypothèses

pour élaborer des modèles microéconomiques ou macroéconomiques qui sont confrontés aux données afin d'évaluer leur capacité à expliquer les situations observées et à produire des prévisions. Toutefois, ces théories n'ont pas été capables d'anticiper correctement la crise financière qui s'est déclenchée en 2007 dans une période où les macroéconomistes se félicitaient de la « grande modération<sup>5</sup> », c'est-à-dire de la combinaison d'une réduction de la volatilité du cycle économique et d'une faible inflation.

Alan Kirman (2012) rappelle que les causes identifiées de la crise – « défaillance des réseaux de banques, problèmes de confiance et de contagion à tous les niveaux » – étaient ignorées par les modèles utilisés : on avait d'un côté des modèles macroéconomiques fondés sur l'idée que les marchés s'autorégulent avec des agents qui font des anticipations rationnelles et, de l'autre, un modèle de marchés financiers « fondé sur l'hypothèse des marchés efficaces : toute l'information disponible sur les actifs financiers est contenue dans les prix » ; enfin, le secteur financier était mal intégré dans les modèles macroéconomiques.

Des approches alternatives existent comme le montrent les travaux de Robert Shiller, de Raghuram Rajan, de Roman Frydman ou, du côté francophone, avec Hélène Rey, André Orléan, Gaël Giraud, Michel Aglietta... Mais il reste à faire un effort important pour intégrer ces apports dans l'analyse macroéconomique contemporaine. Joseph Stiglitz est un exemple d'économiste qui plaide pour une reconstruction de la théorie macroéconomique. Dans un article récent<sup>6</sup>, il constate que la macroéconomie *mainstream* a été dominée par deux « églises » : la première raisonnait dans le cadre d'un modèle classique avec des marchés parfaitement flexibles, un agent économique représentatif qui formait des anticipations rationnelles et un secteur bancaire qui ne jouait aucun rôle ; la seconde reconnaissait qu'il pouvait y avoir des rigidités de prix et de salaire, sans concevoir que les défaillances de marché pouvaient aller au-delà de ces rigidités. En particulier, les

anomalies sur les marchés financiers étaient ignorées, laissant ainsi dans l'ombre la croissance anormale de la demande des agents privés par le recours au crédit. Stiglitz propose de fonder la reconstruction de la macroéconomie sur des hypothèses plus réalistes qui intègrent : les asymétries d'information, les distorsions entre les rendements privés et les coûts sociaux, l'incomplétude des contrats négociés qui évaluent mal les prises de risque. Les travaux théoriques qui vont dans cette direction font apparaître qu'il peut y avoir des équilibres multiples avec des agents formant des anticipations différentes. L'économie peut alors se révéler instable avec des ajustements de prix et de salaires qui aggravent la récession et des effets de répartition qui induisent des inégalités de revenu croissantes. L'enjeu de cette reconstruction théorique est important car il conditionne la compréhension des mécanismes qui sous-tendent les choix de politique économique. On le voit clairement avec les débats actuels sur les politiques d'austérité dans la zone euro, sur l'évaluation des multiplicateurs budgétaires, sur les stratégies de désendettement des agents privés et publics, sur l'existence de trappes à liquidité et le risque de déflation...

Enfin, comme le remarque François Bourguignon (2012) : « L'économie n'est pas une science dans le sens où elle ne peut pas fournir des lois universelles et objectives expliquant le fonctionnement des économies en tout point de l'espace et du temps... mais l'économie est peut-être plus scientifique que d'autres sciences humaines dans la mesure où, à partir de quelques hypothèses clairement énoncées, il est possible d'expliquer un certain nombre de phénomènes complexes. » Encore faut-il, toutefois, partir d'hypothèses réalistes.

## Vers des analyses économiques plus réalistes

Les économistes se sont éloignés du modèle abstrait de l'agent économique calculateur, rationnel, en univers certain. Par exemple,

[5] C'est le titre d'un discours célèbre de Ben Bernanke en février 2004.

[6] Stiglitz J. (2014), « Reconstructing Macroeconomic Theory to Manage Economic Policy », *NBER Working paper series*, n° 20517, septembre.

les travaux de Jean Tirole et Roland Bénabou ont fait apparaître que nos comportements répondaient à plusieurs types d'incitations. Il y a évidemment **les motivations extrinsèques** qui sont celles prises en compte par l'analyse économique traditionnelle – les rémunérations ou autres formes de récompenses qui incitent un agent à effectuer une action qui a un coût. Mais, comme le note Jean Tirole : « La plupart des gens entreprennent des tâches ou choisissent des actions ayant des bénéfices privés extrêmement faibles : voter, ne pas polluer même quand on ne peut pas se faire prendre, donner à des organisations caritatives, aider des étrangers, risquer sa vie, etc. Ces comportements ne sont pas toujours faciles à expliquer par la théorie économique classique<sup>7</sup> » car ils relèvent d'un comportement altruiste lié aux valeurs, à la culture de l'agent économique. Ce sont les **motivations intrinsèques**. Enfin, nos actions sont aussi guidées par une troisième catégorie de motivation qui est celle de l'image de soi que l'on veut renvoyer ou simplement avoir de soi-même. C'est ce que Tirole appelle **la motivation réputationnelle**. On sait, en effet, que nos comportements sont influencés par le fait que l'on est observé ou non.

Si nos comportements obéissent à un système complexe de motivations, les incitations à mettre en œuvre pour obtenir un résultat ne se limitent pas à la gratification matérielle. Favoriser la confiance et la collaboration pourra être déterminant. Il reste, comme le montre l'économie expérimentale, que nos comportements souffrent de nombreux biais. Notre rationalité est limitée : nous avons tendance à surévaluer le présent et notre expérience personnelle, nous sommes confrontés à nos identités multiples (nous obéissons à différents processus mentaux : cognitifs, affectifs, instinctifs), nous réagissons différemment selon le contexte... Tout un débat est né, à partir des travaux d'économie comportementale, autour de ce que l'on appelle le paternalisme libéral<sup>8</sup>, c'est-à-dire l'idée que l'État devrait intervenir par des *nudges* (des coups de pouce) afin de corriger les biais de comportement des agents

économiques (les inciter à ne pas fumer, à mieux choisir leur alimentation, à mieux se protéger en prenant une mutuelle...).

Un bon exemple de la prise en compte de la complexité des déterminants qui guident nos comportements économiques est la nouvelle approche de la Banque mondiale concernant la conduite des politiques de développement<sup>9</sup>. La réflexion est centrée sur les modes de prise de décision des individus avec la distinction entre : **la pensée automatique** – celle que l'on mobilise le plus souvent pour les actions répétitives –, **la pensée sociale** – qui est influencée par ce que pensent les autres – et **la pensée par modèles mentaux** – qui sont déterminés par la vision du monde que partagent les humains d'une même société. Ainsi, si on veut inciter les individus à épargner ou à mieux gérer des ressources rares comme l'eau potable, il faut penser de nouvelles formes d'intervention qui s'appuient sur les modèles mentaux, les stéréotypes, les attentes sociales, les schémas de coopération...

L'économie a donc élargi considérablement son champ d'investigation en repensant la rationalité des agents, les problèmes d'incertitude, de défaillance de marché. On pourrait aussi faire référence aux modèles dynamiques qui permettent d'introduire le cadre temporel absent de l'analyse statique et à l'économie géographique qui permet d'introduire l'espace dans le raisonnement économique (problème par exemple de la polarisation des activités).

## Les débats actuels

Pour illustrer l'apport de l'économie dans la compréhension des mutations du monde contemporain, on peut citer quelques exemples de sujets qui mobilisent les réflexions des économistes.

### La stagnation séculaire

La stagnation séculaire est une question récurrente en économie. Pendant les années 1930, certains économistes tels que Alvin

<sup>[7]</sup> Tirole J. « Motivation intrinsèque, incitations et normes sociales », *Revue économique*, 2009/3, vol. 60, p. 578.

<sup>[8]</sup> Voir par exemple Thaler R. et Sunstein C.R. (2003) : « Libertarian Paternalism », *American Economic Review*, vol. 93, n° 2, p. 175 à 179, et la critique d'un « libéral », Gilles Saint-Paul : [http://www.tse-fr.eu/sites/default/files/medias/doc/wp/macros/wp\\_tse\\_339.pdf](http://www.tse-fr.eu/sites/default/files/medias/doc/wp/macros/wp_tse_339.pdf).

<sup>[9]</sup> Le rapport de la Banque mondiale sur le développement dans le monde 2015 a pour sous-titre : « Pensée, société et comportement ».

Hansen, Benjamin Higgins, Paul Sweezy, que l'on a appelés les « stagnationnistes », pensaient qu'il y avait un épuisement de la croissance quand les économies allaient vers la maturité, ce qui leur semblait être le cas dans les années 1930-1940. De nouveau, dans les années 1970, cette thèse est réapparue avec la notion de stagflation, et voilà que Richard Gordon l'a relancée en 2012 en s'interrogeant sur la fin de la croissance économique aux États-Unis. Le sujet mobilise aujourd'hui des économistes réputés comme Richard Baldwin, Larry Summers, Paul Krugman, Olivier Blanchard, Emmanuel Farhi, Barry Eichengreen, qui étudient les manifestations de cette stagnation séculaire, ses causes et les solutions pour la dépasser<sup>10</sup>. Ils pointent du doigt de nombreux facteurs : la démographie, les inégalités, la déflation, l'épuisement du progrès technique, le coût des ressources naturelles, le risque climatique, l'accumulation des dettes. À l'opposé de la thèse de la stagnation séculaire, deux experts du MIT, Erik Brynjolfsson et Andrew McAfee, estiment que nous sommes à l'aube du « deuxième âge de la machine<sup>11</sup> ». Une troisième révolution industrielle serait en train de se produire, mais elle serait encore peu visible dans les statistiques car celles-ci n'arriveraient pas à mesurer correctement les progrès d'une économie de plus en plus immatérielle.

## Les inégalités contre la croissance

Thomas Piketty n'est pas le seul à s'inquiéter du creusement des inégalités dans les pays riches. L'OCDE et le FMI viennent de publier des études qui mettent en évidence leurs effets négatifs sur la croissance économique. Aux États-Unis, les 10 % les plus riches accaparaient environ 33 % du total des revenus dans les années 1970 contre environ 50 % aujourd'hui. Le FMI montre que lorsque les inégalités se creusent, les agents économiques en difficulté s'endettent (c'est une composante importante de la crise financière de 2007-2008). De plus, si les revenus de la classe moyenne stagnent ou régressent, il y

# ZOOM

## L'ÉCONOMIE POLITIQUE DES RÉFORMES

« La conception d'une stratégie de réformes implique aussi d'établir des priorités parmi de nombreuses actions possibles. Les gouvernements essaient rarement de s'engager simultanément sur tous les fronts. Généralement, ils suivent plutôt ce que Dani Rodrik appelle "la liste de courses" (*the laundry list*) et choisissent les réformes qu'ils considèrent politiquement ou socialement faisables. Or, ce n'est pas un moyen efficace pour réformer une économie, car cela ne conduit pas à choisir les réformes les plus utiles ou à tenir compte des complémentarités qui existent entre des décisions relatives à différents domaines.\* »

\* Pisani J., Enderlein H. (2014), « Réformes, investissement et croissance : un agenda pour la France, l'Allemagne et l'Europe ».

[10] [http://www.voxeu.org/sites/default/files/Vox\\_secular\\_stagnation.pdf](http://www.voxeu.org/sites/default/files/Vox_secular_stagnation.pdf).

[11] Brynjolfsson E. et McAfee A. (2014), *The Second Machine Age. Work Progress, and Prosperity in a Time of Brilliant Technologies*, New York, W. W. Norton & Company.

a un effet négatif sur la demande et donc sur l'activité et l'emploi. On observe de surcroît un affaiblissement de la mobilité sociale et une aggravation des formes d'exclusion par l'éducation, l'emploi, le logement, la santé. Les travaux actuels sur la croissance inclusive ou la croissance soutenable traitent des stratégies à mettre en œuvre pour réduire ces inégalités.

## Autres sujets

D'autres sujets sont débattus par les économistes comme **la guerre des monnaies** (y a-t-il une course à la dépréciation monétaire ?), **le ralentissement du commerce international** (est-il le résultat du ralentissement de la croissance économique mondiale ou d'une modification des stratégies de production internationale des grandes

<sup>[12]</sup> Rifkin J. (2014), *La Nouvelle Société du coût marginal zéro*, Paris, Les liens qui libèrent.

<sup>[13]</sup> Un exemple intéressant est le projet : <http://core-econ.org/>.

<sup>[14]</sup> Voir la conférence « Visions d'économistes et projets de société » : <http://www.touteconomie.org/Cycle2010>.

firmer, ce que l'on appelle les chaînes de valeur ?), **l'avenir du capitalisme** (va-t-on vers une économie du quaternaire, une éclipse du capitalisme<sup>12</sup> ?), **l'économie politique des réformes** (cf. Zoom p. 10).

Les réflexions des économistes sont en prise avec les sujets majeurs qui nous interpellent : les défis environnementaux, la recherche d'une croissance « plus juste », la régulation de la finance, la modernisation de l'action publique...

## Conclusion : pour un dialogue fécond

Si nous avons insisté sur des sujets qui font débat entre les économistes, il ne faut pas oublier qu'en fait « les économistes s'accordent sur beaucoup de choses, dont plusieurs sont politiquement controversées. L'économiste de Harvard, Gregory Mankiw, en a énuméré certaines en 2009. Les idées suivantes ont été soutenues par au moins 90 % des économistes :

– les taxes d'importation et les quotas réduisent le bien-être économique général ;

– la réglementation des loyers réduit l'offre de logements ;

– les taux de change flottants offrent un système monétaire international efficace ;

– les États-Unis ne devraient pas empêcher les entreprises de sous-traiter l'emploi à l'étranger ;

– et la politique budgétaire stimule l'économie lorsqu'il y a moins de plein-emploi ».

Et Dani Rodrik de conclure ce billet par : « Les désaccords entre économistes sont une preuve de bonne santé. Ils reflètent le fait que leur discipline se compose d'une collection de modèles variés, et que la correspondance du modèle à la réalité est une science imparfaite qui laisse une grande place à l'erreur. » L'économie est une discipline vivante qui doit être accessible au plus grand nombre et d'abord à ceux qui ont décidé de l'étudier<sup>13</sup>. Reste une question : les économistes ont-ils une idée du type de société vers lequel on devrait converger si on réalisait les réformes qu'ils suggèrent<sup>14</sup> ?

## POUR EN SAVOIR PLUS

❖ **BOURGUIGNON F.** (2012), « L'économie n'est pas une science », *L'Économie*, n° 55.

❖ **KEEN S.** (2014), *L'Imposture économique*, Paris, Les Éditions de l'Atelier.

❖ **KIRMAN A.** (2012), « La théorie économique dans la crise », *Revue économique*, vol. 63.

❖ **RODRIK D.** (2014), « Les dangers du consensus économique », 14 août 2014, [project-syndicate.org](http://project-syndicate.org).

❖ **TAYLOR T.** (2014), « Économie et morale », *Finances et Développement*, FMI, juin.

❖ **TIROLE J.** (2014), « L'éthique face au marché », *Les Échos*, 8 décembre.

❖ Cycle de conférences : comment raisonnent les économistes ? <http://www.touteconomie.org/Cycle2010>.

De l'économie politique des XVIII<sup>e</sup> et XIX<sup>e</sup> siècles à la science économique contemporaine, la discipline a connu de profondes transformations. Celles-ci concernent ses méthodes – l'usage de la modélisation mathématique s'est généralisé –, mais aussi son domaine d'analyse et par suite sa définition même : alors que l'économie politique s'intéressait à la production, la consommation et la distribution des richesses, l'économie contemporaine a étendu ses objets d'étude à l'ensemble des comportements humains et ne tire plus sa spécificité que de ses méthodes et outils. Et, si elle s'est confrontée aux faits, c'est par le recours croissant aux statistiques et à l'expérimentation et non par l'analyse des faits historiques. Des classiques anglais à l'économie empirique contemporaine, en passant par la macroéconomie keynésienne et la constitution de l'hégémonie néoclassique, Jean-Baptiste Fleury retrace les grandes lignes de l'évolution de l'économie depuis la fin du XVIII<sup>e</sup> siècle.

Problèmes économiques

# Des classiques anglais à *Freakonomics* : l'évolution de la science économique

Le terme « science économique » émerge seulement au XX<sup>e</sup> siècle, remplaçant progressivement celui d'« économie politique », lequel accompagnait la discipline depuis la fin du XVIII<sup>e</sup> siècle. Même si les auteurs du XIX<sup>e</sup> affichaient certaines prétentions scientifiques, l'évolution du terme illustre bien les profonds changements qui ont bouleversé la discipline depuis deux siècles. En effet, l'économie a évolué vers plus d'abstraction et de mathématisation, en dépit de critiques et oppositions régulières. Elle semble aujourd'hui stabilisée autour d'un noyau dur issu de la microéconomie néoclassique. Et lorsque les économistes voulurent intégrer dans leur approche des considérations pour le monde « réel » et les faits, ils le firent rarement du point de vue de l'histoire, mais plutôt en référence aux sciences « dures ». En

témoignent l'usage croissant des statistiques durant l'entre-deux-guerres, l'émergence de l'économétrie dans les années 1940 ainsi que de l'économie expérimentale dans les années 1980. Les transformations de la définition de l'économie accompagnent logiquement cette évolution : d'une discipline définie par son champ d'analyse (la production, la consommation et la distribution des richesses), elle finit par ne se distinguer, désormais, des autres sciences sociales que par une méthode scientifique fondée sur la microéconomie et la théorie des jeux.

© JEAN-BAPTISTE FLEURY

Maître de conférences à l'IUT de Cergy-Pontoise  
Chercheur au laboratoire THEMA

## La genèse de la science économique : l'école classique

Il est de coutume de considérer l'ouvrage d'Adam Smith, *La Richesse des nations* (1776), comme fondateur de la théorie économique<sup>1</sup>. Bien que novatrice, la démarche de Smith n'est pas isolée de toute influence intellectuelle : elle s'inspire, par exemple, de la théorisation des motivations humaines élaborée par David Hume, mais également des philosophes politiques tels que Thomas Hobbes ou John Locke. Smith n'est pas non plus le premier à réfléchir sur les mécanismes au cœur de la création de richesse et de sa distribution dans la société : les physiocrates, entre autres, ont développé en France une théorie qui aborde des thèmes et des notions connexes. Néanmoins, si Smith est considéré comme un auteur majeur, c'est parce que son œuvre est à l'origine de l'« école classique », à partir de laquelle se construiront nombre de notions qui restent aujourd'hui diffusées dans le courant dominant de la discipline (qualifié d'approche « néo-classique »).

Cette école classique n'est pas homogène et les sources de discorde entre auteurs sont nombreuses. Cependant, ce qui relie les théories de David Ricardo, Robert Malthus, John Stuart Mill, Jean-Charles de Sismondi ou Jean-Baptiste Say est leur volonté d'écrire en réaction à l'ouvrage de Smith. S'étalant entre la fin du XVIII<sup>e</sup> et le milieu du XIX<sup>e</sup> siècle (l'ouvrage de J.S. Mill, paru en 1848, marque la borne temporelle supérieure), leur contribution à ce qu'on appelle alors « l'économie politique » tente d'apporter une réponse à la question de l'organisation sociale et politique d'une nation. Ainsi, l'économie est « politique » car les principes scientifiques qui gouvernent la production et la distribution des richesses, facteurs essentiels pour la prospérité du Prince et des membres de la société, répondent à des lois qui sont indépendantes du pouvoir politique (ce que les physiocrates considèrent déjà au milieu du XVIII<sup>e</sup> siècle comme un « ordre naturel<sup>2</sup> »).

Par conséquent, la notion de richesse (sa création, l'évolution de son accumulation et de sa distribution dans le temps) est au cœur des préoccupations des classiques. Ils développent tout d'abord une théorie de la valeur fondée sur la notion de travail. C'est en effet la quantité de travail nécessaire à la production d'un bien qui en détermine la valeur, ce qu'ils appellent le « prix naturel ». Les classiques ont bien conscience que la rencontre entre l'offre et la demande *via* le mécanisme de marché fixe également **les prix**. Mais ceux-ci ne peuvent durablement s'écarter du « prix naturel » et ne font, au final, que graviter autour de lui.

À l'instar des physiocrates, la société que les classiques analysent est divisée en classes, chacune jouant un rôle distinct et recevant une rémunération propre. Les travailleurs offrent leur force de travail contre salaire. Les propriétaires fonciers louent la terre contre une rente. Au final, l'accumulation des richesses est assurée par les capitalistes (les entrepreneurs), qui investissent l'épargne pour accroître la division du travail, le progrès technique et la production. Par le jeu du libéralisme économique, mettant en concurrence des individus uniquement mus par leurs intérêts propres, la liberté d'investir là où l'on trouve meilleure rémunération conduit la société à l'enrichissement : c'est l'allégorie de la « main invisible ». Le rôle de l'État dans l'activité économique, bien que minimisé, n'est toutefois pas circonscrit aux seules fonctions régaliennes : chez la plupart des classiques, l'État doit également intervenir dans la création d'infrastructures, de services publics ou la régulation de certaines industries monopolistiques. De plus, la question des crises économiques les divise : alors que certains, suivant Say, pensent que la production d'un bien assure, de par le revenu qu'elle génère et distribue, sa vente (c'est la « loi des débouchés »), d'autres, tels que Malthus ou Sismondi, pensent qu'une crise générale de surproduction peut persister. Enfin, tous ne partagent pas les craintes de Ricardo concernant la baisse du taux de pro-

[1] Cf. Béraud A. et Faccarello G. (2010) [1992]. Les deux premiers tomes de cet ouvrage servent de référence pour les présentations des courants de pensée classiques, de Marx, mais également des marginalistes et premiers néoclassiques.

[2] Voir, par exemple, Say J.-B. (1829), « Histoire abrégée des progrès de l'économie politique » in *Cours complet d'économie politique pratique*, vol. 2, Paris, Economica, 2010.

fit (donc de l'incitation à accumuler), lequel serait comprimé par la hausse de la rente et des salaires du fait d'une croissance démographique forçant l'exploitation de terres toujours moins fertiles.

## Les critiques des classiques au XIX<sup>e</sup> siècle

### La critique de Marx

Une source importante de critiques des classiques proviendra des écrits de Karl Marx. Il leur emprunte, en particulier à Ricardo, un certain nombre d'idées, mais dans le but de les attaquer dans leur défense du capitalisme. Marx adhère, par exemple, à l'idée que la valeur est dépendante de la quantité de travail nécessaire à produire un bien. Seul le travail crée de la richesse. Mais comme ce sont les propriétaires des moyens de production qui ont acheté la force de travail grâce à leur capital, le salarié se voit déposséder de la « plus-value » que son travail a permis de créer : c'est l'exploitation des travailleurs par le capital. Ce processus de création du profit incite les capitalistes à accumuler, mais cette accumulation mène à la concentration du capital et tend à faire baisser le taux de profit moyen dans une économie. De plus, le rythme d'accumulation n'est pas régulier : le système capitaliste est fondamentalement instable, alternant périodes de prospérité et périodes de surproduction, synonymes de pauvreté. Récusant la loi des débouchés de Say, Marx s'interroge sur l'effondrement potentiel du système. Les contradictions du capitalisme sont alors à placer en lien avec l'idée que celui-ci n'est qu'un moment dans l'histoire des modes de production. À terme, un mode différent pourra se mettre en place, mettant au centre la coopération des travailleurs.

### L'école historique allemande

Marx ne sera pas le seul à considérer que les lois économiques sont historiquement

déterminées. En effet, au milieu du XIX<sup>e</sup> siècle, les principales critiques qui s'élèvent à l'encontre de l'école classique sont le fait de théoriciens appartenant à l'« école historique allemande », par exemple Gustav von Schmoller. Pour ces économistes, la notion de loi naturelle est une chimère. Seule l'analyse historique compte. Cette critique du scientisme des classiques est importante, car on la retrouve sous des formes diverses tout au long du XX<sup>e</sup> siècle, dans l'institutionnalisme américain ou le mouvement radical des années 1960, par exemple. Avec cette critique, venait également une critique de l'individualisme méthodologique. Les historicistes y préférèrent la notion de holisme selon laquelle l'analyse de la société ne peut partir de l'individu, car le tout est plus grand que la somme des parties (une notion au cœur de la future macroéconomie keynésienne).

### Les débuts du marginalisme et l'école autrichienne

Bien que la seconde moitié du XIX<sup>e</sup> siècle soit dominée par l'école historique allemande, une série de contributions, peut-être moins visibles en leur temps, viennent toutefois poser les jalons de la science économique moderne. Ainsi, avec le courant « marginaliste », composé de Carl Menger (Autriche), Stanley Jevons (Angleterre) et Léon Walras (France), un changement radical dans la théorie de la valeur s'opère. En appliquant leur approche au célèbre paradoxe de l'eau et du diamant, les marginalistes récusent l'argument des classiques selon lequel le diamant, quoique beaucoup moins utile que l'eau, a plus de valeur qu'elle, car son extraction demande plus de travail. Déterminée uniquement par le processus de rencontre entre l'offre et la demande sur le marché, la valeur d'un bien est désormais celle de la dernière unité produite et consommée (d'où le qualificatif de « marginal »). Autrement dit, l'eau a moins de valeur que le diamant, car l'utilité (la satisfaction) de la dernière unité d'eau est généralement plus faible que celle de la dernière unité de dia-



mant. En revanche, dans le désert, l'eau peut acquérir plus de valeur que le diamant.

Cette approche peine à s'imposer. Concernant Menger, ses premiers travaux ne sont pas recensés dans la revue de G. Schmoller. Il ne sera reconnu que lorsqu'il s'attaquera directement à la méthodologie de l'école historique, en revendiquant l'individualisme méthodologique et l'importance de la formulation de lois exactes, abstraites et potentiellement anhistoriques. La *Methodenstreit* (« querelle des méthodes ») qu'il déclenche opposera ces économistes jusqu'à la Première Guerre mondiale. Cependant, l'héritage de Menger n'est pas représentatif de l'héritage du marginalisme dans son ensemble. Menger devient plutôt le père fondateur d'une tradition de pensée qui se développera en parallèle du courant dominant tout au long du xx<sup>e</sup> siècle, qualifiée d'« école autrichienne », et dont les contributeurs les plus célèbres seront Eugen Böhm-Bawerk, Ludwig von Mises, Friedrich von Hayek, ou Israel Kirzner. L'école autrichienne adopte notamment une position subjectiviste : l'économiste fait lui aussi partie des agents qu'il analyse, il ne peut s'extraire de la société qu'il observe, comme ce sera le cas dans la théorie néoclassique. Les individus, bien que rationnels, ne sont pas en situation d'information parfaite et complète. Dès Menger, les Autrichiens ne s'intéressent pas à la question de savoir quel sera le prix fixé par le marché et les quantités précisément échangées par les individus – question qui intéresse Walras, Jevons et leurs successeurs – mais plutôt aux mécanismes par lesquels la main invisible coordonne les décisions individuelles et mène à l'émergence d'institutions sociales (comme la monnaie ou certaines lois) par le jeu d'un ordre spontané. Le marché joue dans ces processus un rôle de producteur d'information facilitant la coordination des agents.

## Vers une approche « néoclassique »

Le marginalisme autrichien se distingue de celui de Walras et Jevons, non seulement

dans son objet d'analyse, mais également par la méthode. Menger rejette l'utilisation des mathématiques, tandis que Walras et Jevons les placent au centre de la rigueur scientifique. Malgré leur faible influence immédiate, Walras et Jevons participent d'un mouvement plus général de mathématisation dont les travaux d'Antoine-Augustin Cournot et des ingénieurs français tels que Jules Dupuit sont précurseurs. Ce mouvement se renforce à la fin du xix<sup>e</sup> siècle avec, notamment, les travaux d'Alfred Marshall, qui connurent, eux, un succès immédiat. Marshall et ses contemporains, comme, par exemple, Francis Edgeworth et Vilfredo Pareto, développent le courant qualifié de « néoclassique » en approfondissant les outils du marginalisme autour de thèmes chers aux classiques, mais que ces derniers n'avaient paradoxalement que très peu analysés. Parmi ceux-là, le rôle de la demande et de l'offre, provenant d'agents rationnels au comportement maximisateur, dans la détermination du prix de marché. Ainsi, après Jevons, le cadre analytique et ses hypothèses simplificatrices sont raffinées avec l'introduction du calcul de variation, des multiplicateurs de Lagrange, de la fonction d'utilité. Edgeworth développera aussi les courbes d'indifférence, illustrant bien l'évolution (débattue) d'une théorie qui passe d'une utilité cardinale – c'est-à-dire le fait que la satisfaction d'un individu peut se mesurer dans une unité qui la rendrait directement comparable avec la satisfaction d'un autre individu – à une utilité « ordinale » par laquelle une telle comparaison interpersonnelle est impossible. Avec l'utilité ordinale, la satisfaction individuelle s'analyse comme un ordre de préférences permettant de classer différents biens, et dont la cartographie implique une courbe d'indifférence.

Les néoclassiques reprennent également à leur compte la défense de la concurrence et du capitalisme déjà présente chez les classiques. Ainsi, l'échange marchand est optimal au sens de Pareto, c'est-à-dire que le bien-être d'un individu ne peut être augmenté sans diminution du bien-être d'un autre. Malgré

les ambitions scientifiques des premiers néoclassiques, ce résultat, très important, ne sera rigoureusement démontré que dans les années 1950 par Kenneth Arrow, Gérard Debreu et Lionel McKenzie.

Les transformations de la science économique à la fin du XIX<sup>e</sup> siècle ne sont pas uniquement intellectuelles : l'économie s'institutionnalise. Walras, Menger, Jevons et Marshall furent professeurs à l'université, ce qui n'était pas le cas de leurs prédécesseurs. De plus, la communauté scientifique internationale se développe, ce qui soutient l'influence croissante du néoclassicisme. Il apparaît alors comme plus homogène que ses critiques (réformateurs radicaux, auteurs de l'école historique) qui pensent que l'économie manque de considérations pour l'histoire, la sociologie, ou le socialisme.

Néanmoins, au début du XX<sup>e</sup> siècle, le néoclassicisme qui se développe avec John Bates Clark ou Irving Fisher n'exerce pas la domination observée de nos jours. Jusqu'à la fin de la Seconde Guerre mondiale, il subsistera un important pluralisme au sein de la discipline<sup>3</sup>. D'ailleurs, au tournant des années 1920 aux États-Unis, l'institutionnalisme s'impose rapidement comme un courant alternatif et influent. Les écrits de Thorstein Veblen acquièrent une influence majeure sur ses leaders, à savoir Wesley Mitchell et John R. Commons<sup>4</sup>. Les institutionnalistes proposent une approche scientifique différente et prometteuse, inspirée des sciences naturelles de l'époque, mettant l'accent sur l'empirisme, la collecte d'observations et les tests statistiques. Ils y insèrent également les apports de la psychologie, du droit, de la philosophie et de l'histoire. En rupture avec l'abstraction des néoclassiques, ils étudient l'économie et le comportement des firmes, non pas comme la résultante d'un ordre naturel dont il faudrait dégager les lois, mais comme l'aboutissement d'une construction institutionnelle. Cette approche leur permettra d'aborder des sujets que la théorie néoclassique peine à étudier avec les hypothèses de l'agent rationnel maximisateur, tels que les défaillances de

marché, le rôle du droit et de la réglementation économique (en particulier des services publics), celui des syndicats et des relations de travail. Préoccupés par la réforme sociale et la résolution de problèmes concrets, ces auteurs seront notamment influents au sein du *Brain Trust* de Franklin Roosevelt, lors du New Deal.

## La stabilisation de la science économique après la Seconde Guerre mondiale

Malgré tout, l'institutionnalisme perd pied à partir de la fin des années 1930. En effet, sur le terrain scientifique, l'approche de John Maynard Keynes s'impose progressivement, car elle permet de comprendre les maladies contemporaines de la dépression et du chômage. Keynes raisonne dans un cadre macroéconomique plus large, considérant des grandeurs agrégées. En particulier, la demande globale d'un pays est composée de la consommation des ménages et de l'investissement des entreprises, mais également des commandes publiques à ces mêmes entreprises. Si l'investissement privé venait ainsi à se contracter et, avec lui, la demande et l'offre globales, l'État pourrait se substituer aux entreprises afin de soutenir l'activité pour éviter la dépression et le chômage de masse. Keynes contribue également à développer les instruments statistiques compatibles avec cette approche : la comptabilité nationale émerge.

Force est de constater que la Seconde Guerre mondiale apporte son lot de bouleversements majeurs. En effet, la plupart des économistes sont réquisitionnés pour l'effort de guerre, en particulier aux États-Unis. Ils développent alors leurs compétences théoriques et statistiques pour rationaliser la production militaire, l'efficacité de l'armement, la gestion de la demande globale et des prix. La nouvelle orthodoxie d'après-guerre émerge de ces changements. Les économistes, à l'image des physiciens, construisent désormais des

[3] Cf. Morgan M. et Rutherford M. (1998), *From Inter-War Pluralism to Post-War Neoclassicism*, supplément annuel de *History of Political Economy*, vol. 30, Durham, Duke University Press.

[4] Cf. Rutherford M. (2000), « Institutionalism Between the Wars », *Journal of Economic Issues*, vol. 24.

« modèles ». En particulier, ils approfondissent de manière concomitante le cadre analytique néoclassique walrassien (représentant l'équilibre général issu d'une multitude d'agents rationnels aux comportements interdépendants), les outils de statistique appliquée (l'« économétrie »), mais également le « keynésianisme de la synthèse », qui traduit les principaux apports de Keynes dans un cadre néoclassique moyennant quelques changements d'hypothèses. Les économistes américains et, plus précisément, ceux de la Cowles Commission et du MIT (Kenneth Arrow, Lawrence Klein, ou Paul Samuelson) jouent un rôle central dans ces développements. Dans un contexte de guerre froide, la mathématisation croissante permet aux économistes de faire passer le message keynésien, concernant l'intervention de l'État, sans soulever trop de suspicions. Le keynésianisme de la synthèse gagne une nouvelle fois en respectabilité au cours de la décennie 1960, lorsque ses enseignements seront directement appliqués dans la politique économique de nombreux pays. En plus de faire partie des conseillers de John Kennedy, Samuelson jouera un rôle clé dans la diffusion de cette orthodoxie par le biais de son manuel, *Economics*, qui devient rapidement un best-seller à une période marquée par la massification de l'enseignement supérieur.

À l'orée de la guerre froide, les critères de scientificité ont changé : les sciences sociales doivent pouvoir produire des propositions idéologiquement neutres et vérifiables empiriquement. Le complexe militaro-industriel américain et les fondations philanthropiques soutiennent ainsi une recherche mathématisée qui mène, notamment, au développement de la théorie des jeux, branche des mathématiques développée par John von Neumann et qui étend l'analyse du comportement rationnel aux interactions stratégiques. Cette théorie sera rapidement assimilée par les microéconomistes néoclassiques. La microéconomie s'impose également dans le domaine de l'économie du travail, un des derniers bastions de l'institutionnalisme, en particulier à la suite

des travaux d'économistes de Chicago. L'un d'eux, Gary Becker, fait de sa refonte de l'économie du travail le socle lui permettant d'étudier la discrimination, le mariage ou le crime. Les années 1960 voient ainsi se développer l'« impérialisme de l'économie » menant à la disparition des frontières séparant la science économique des autres sciences sociales. La science économique ne se définit alors plus que par sa méthode : l'offre, la demande et le comportement rationnel peuvent analyser désormais les problèmes de droit, d'anthropologie, de sociologie, de science politique, et même les comportements animaux.

La rationalité des individus fait également son retour en force en macroéconomie au tournant des années 1970. À cette époque, l'échec récurrent des politiques de relance keynésiennes soutient la critique de l'école de Chicago, en particulier celle de Milton Friedman et Robert Lucas. La modélisation de Lucas, notamment, introduit la notion d'« anticipations rationnelles » afin de rendre compte du phénomène de stagflation rencontré dans la plupart des économies occidentales après le choc pétrolier de 1973. Ses fondements microéconomiques définiront la nouvelle orthodoxie académique. Durant la décennie 1970, ce revirement de la pensée accompagne une tendance sociale plus générale conduisant à la remise en cause de l'interventionnisme étatique. Avec les élections de Ronald Reagan et Margaret Thatcher, les idées de dérégulation, de décentralisation et de privatisation marquent le retour en force du libéralisme économique, soutenu, entre autres, par l'école de Chicago (cf. Backhouse et Fontaine, 2010).

## Un tournant empirique

Depuis les années 1970, la discipline semble s'être stabilisée autour d'un noyau dur fait de microéconomie, de théorie des jeux et d'économétrie. Qu'il considère des situations d'asymétries d'information ou qu'il se range au côté des « néokeynésiens », le macroéconomiste a vu sa discipline se renforcer par de

solides fondements microéconomiques. Par ailleurs, l'empirisme gagne du terrain. Tout d'abord parce que l'économétrie devient une méthode très prisée des publications scientifiques : son développement est plus rapide que celui des publications théoriques (Le Merrer, 2008). La popularité, dans les années 1980, de la notion d'« expérience naturelle », qui consacrera, entre autres, Steven Levitt ou David Card, n'en est qu'une illustration. De plus, l'économie se dote d'un volet expérimental après les travaux de Daniel Kahneman, Amos Tversky ou encore Vernon Smith. Ces expérimentations de laboratoire empruntent progressivement les outils de la psychologie et des neurosciences. Ce faisant, les frontières séparant la science économique des autres sciences sociales regagnent en porosité, comme en témoigne également le développement de la sociologie économique. Depuis les années 1970 et 1980, rien ne semble empêcher le développement de la mathématisation de l'économie, déjà accélérée durant les années 1940. À tel point que, de manière récurrente, certains étudiants et chercheurs dénoncent l'isolement des économistes face au monde réel, et, donc, le manque de pertinence et de réalisme de leurs théories et des modèles mathématiques qui les sous-tendent. Ces critiques sont au cœur du mouvement radical des années 1960, inspiré par la pensée marxiste, ou, plus tard, du mouvement attaquant « l'autisme » de l'économie dans les

années 2000. Récemment encore, à l'occasion de la crise des subprimes, nombre de voix se sont élevées contre l'orthodoxie contemporaine. Seul le temps pourra dire si ces critiques parviendront à faire s'éloigner la science économique de ses fondements actuels. En effet, un certain nombre d'économistes orthodoxes ne sont pas prêts à en changer radicalement. En témoigne le développement, depuis les années 1990, d'ouvrages vulgarisant l'approche microéconomique, qui, en s'appuyant sur le phénomène d'impérialisme de l'économie, essaient de démontrer que la science économique est capable d'analyser des problèmes très concrets. La plupart, y compris le best-seller *Freakonomics*, représentent l'économiste comme un naturaliste explorant les paradoxes de la vie quotidienne, analysant les vérités cachées du sport ou des jeux télévisés. Ce faisant, ils promeuvent les récents développements de la discipline marquée par la prévalence de l'empirisme sur fond de microéconomie, ainsi que par l'absence de frontières limitant *a priori* le champ du questionnement économique<sup>5</sup>. D'analyses philosophiques centrées sur la théorie de la valeur au tournant du XIX<sup>e</sup> siècle, le travail de l'économiste politique s'est mué en celui d'un scientifique ancré dans la pop-culture de son temps, agissant aussi bien en qualité de conseiller du gouvernement que d'auteur de livres populaires.

<sup>[5]</sup> Fleury J.-B. (2012), « The Evolving Notion of Relevance », *Journal of Economic Methodology*, vol. 19, n° 3.

## POUR EN SAVOIR PLUS

❖ **BÉRAUD A. et FACCARELLO G.** (2000) [1992] (eds), *Nouvelle histoire de la pensée économique*, Paris, La Découverte.

❖ **BACKHOUSE R. et FONTAINE PH.** (2010) (eds),

*The History of the Social Sciences since 1945*, Cambridge, Cambridge University Press.

❖ **LE MERRER P.** (2008), « Panorama de la pensée économique contemporaine »,

*Cahiers français*, n° 345, Découverte de l'économie, tome 1, Paris, La Documentation française.

Discipline à l'usage des gouvernants, la macroéconomie étudie les relations entre les grands agrégats économiques, à l'inverse de la microéconomie, dont le point de départ est le comportement des individus. Au-delà de cet objectif de base, ses méthodes et ses centres d'intérêt ont évolué au cours du temps. « Science de l'enrichissement national » à l'époque des Classiques, elle s'est constituée en véritable sous-champ de l'économie avec les travaux de J. M. Keynes, en offrant une voie d'analyse et de résolution des crises et des problèmes monétaires. Dans les années 1960, le « keynésianisme de la synthèse » en fait une science de l'interaction entre des marchés en déséquilibre. Enfin, avec la remise en cause de l'hégémonie keynésienne et le succès de la notion d'anticipations rationnelles, le problème central étudié par la macroéconomie devient la « coordination ». Bruno Ventelou et Yann Videau présentent ainsi les grandes évolutions de la macroéconomie, à partir des définitions successives qu'elle s'est attribuées.

Problèmes économiques

# Les grands temps de la macroéconomie

© BRUNO VENTELOU

CNRS Aix-Marseille sciences économiques

© YANN VIDEAU

Maître de conférences, université Paris-Est de Créteil

L'histoire de la macroéconomie peut se relater en évoquant la succession des différentes définitions qu'elle s'est données. La façon dont la macroéconomie se définit – parfois s'autojustifie – par rapport au reste de la science économique est parlante et constitue en soi une bonne introduction aux grands temps du développement du champ d'étude. Ceci invite d'abord à s'interroger sur la période à laquelle remontent les premiers travaux qui traitent de questions macroéconomiques, travaux dont la portée dépassait la simple analyse des phénomènes globaux

pour mettre en lumière des rapports de force et aider le pouvoir en place dans sa prise de décision. Par la suite, le contexte économique marqué par différentes crises, jusqu'à celle, majeure, de 1929, et le changement du rôle accordé aux marchés financiers ont déplacé le curseur vers une macroéconomie vue comme science de la monnaie et des crises. Dans les années 1970, la stagflation et la prise en compte des anticipations en matière de politique économique ont mené à la remise en cause de l'approche keynésienne et fait entrer la macroéconomie dans une ère nouvelle, celle de la science de la coordination. Entre-temps, la macroéconomie en tant que science de l'interaction physico-financière de grands marchés en déséquilibre avait connu son heure de gloire en mettant sur pied des outils d'aide à la décision en termes de politique économique conjoncturelle.

## La macroéconomie comme la science du pouvoir et de l'enrichissement « national »

La macroéconomie a toujours été une discipline à l'usage des gouvernants. S'il l'on veut remonter le temps au-delà même de la fondation de l'économie politique comme discipline autonome, les économistes précurseurs qu'étaient les mercantilistes et les physiocrates ont plutôt été des « macroéconomistes » (mais le terme n'existait pas) qui s'intéressaient d'abord à l'équilibre global du territoire contrôlé par le pouvoir, pouvoir royal à l'époque, et à son enrichissement (autrement dit l'expansion économique). Déjà, à cette époque, la recherche des facteurs déterminant la grandeur ou le déclin des nations est une préoccupation, bien avant Adam Smith et sa *Recherche sur la nature et les causes de la richesse des nations* de 1776. Dès le <sup>xvi</sup><sup>e</sup> siècle, les mercantilistes ont cherché à guider le Prince dans sa volonté de puissance et de richesse en mettant en évidence trois sources potentielles de prospérité (donc de pouvoir) : l'abondance des hommes, main-d'œuvre en puissance (suivant en cela la maxime de Jean Bodin ou encore d'Antoine de Montchrestien : « Il n'est de richesse que d'hommes ») ; l'abondance de métaux précieux (or et argent) ; le commerce avec la recherche d'une balance commerciale excédentaire. Pierre le Pesant de Boisguilbert, ensuite, s'est intéressé aux raisons de l'appauvrissement de la France sous Louis XIV et a avancé comme principal déterminant l'insuffisance de la consommation, anticipant avec deux siècles et demi d'avance l'analyse de John Maynard Keynes. Les économistes classiques, au premier rang desquels Adam Smith, David Ricardo ou encore Thomas Malthus, ont proposé aussi une analyse de la répartition de la richesse entre les différentes catégories d'agents économiques. Ils ont particulièrement mis l'accent sur la dynamique d'accumulation des richesses (du capital pour être précis), préfigurant en cela

les théories de la croissance. Cependant, ces approches s'en écartent fondamentalement dans la mesure où la dynamique de l'économie est caractérisée par une baisse tendancielle du taux de profit devant conduire *in fine* à un état stationnaire.

Dans le même temps, des analyses en termes de circuit économique plus ou moins complexes, mettant en évidence les flux réels et monétaires, sont présentées. On retrouve donc très tôt l'approche macroéconomique selon laquelle une économie fonctionne tel un circuit fermé où la richesse créée circule entre les grands agents économiques – dits « représentatifs ». Mais, les rapports de pouvoir entre les différents groupes sociaux n'ayant pas encore été gommés par les économistes sous couvert de la logique impersonnelle du marché, l'analyse en termes de circuit se double bien souvent d'une description de la hiérarchie sociale qui caractérise cette période. Des auteurs comme Boisguilbert, Cantillon, et un peu plus tard François Quesnay avec son *Tableau économique* (1758), ont fourni une représentation simplifiée du fonctionnement de l'économie d'un pays en retraçant les principaux échanges réels et/ou monétaires entre les différentes classes. Ces travaux annoncent, quelques siècles en avance, ce qui deviendra l'un des instruments les plus importants de la macroéconomie, la comptabilité nationale, dont l'objectif est de fournir un modèle simplifié des interdépendances économiques permettant à la fois la quantification de la richesse créée dans une économie et l'analyse de sa répartition entre différents groupes d'agents économiques homogènes (les « secteurs institutionnels »). C'est aussi l'opérationnalisation des idées de Keynes qui poussera à la création d'outils statistiques sophistiqués de mesure des principaux agrégats (production, consommation, épargne, investissement...) et fera émerger de véritables systèmes nationaux de comptabilité dès les années 1940 (sous l'impulsion de Richard Stone et James Meade en Grande-Bretagne, de Jan Tinbergen aux Pays-Bas...).

## La macroéconomie comme la science de la monnaie et des crises

C'est la crise de 1929 et la remise en cause de l'économie politique classique qui fait apparaître la nécessité de développer une nouvelle section de recherche en économie, section qui serait dédiée tout entière à l'analyse des grandes fluctuations de l'activité d'un pays ou d'une nation et qui sera véritablement intitulée « macroéconomie ». L'économie classique, calquée sur une économie de troc, ne dispose pas d'outils pour penser les crises générales de surproduction : le volume d'échange – à l'intersection de tous les agents de l'économie – doit nécessairement s'équilibrer puisque les dépenses des uns font les ressources des autres (« loi de Say ») et la monnaie n'est qu'un intermédiaire neutre des échanges. J. M. Keynes lance alors la vraie première époque de la macroéconomie, vue comme corpus de réflexion autonome sur le lien entre monnaie, taux d'intérêt et emploi, le taux d'intérêt devenant une variable monétaire là où les Classiques en faisaient une variable réelle sur le marché des fonds prêtables. Il publie en 1930 un premier ouvrage, *Treatise on Money*. Puis dans son livre *Théorie générale de l'emploi, de l'intérêt et de la monnaie* de 1936<sup>1</sup>, Keynes remonte un schéma causal qui va, à rebours du titre de l'ouvrage, du marché de la monnaie vers le niveau du chômage, pour faire du marché monétaire la cause, mais aussi la porte de sortie de la crise économique. Il retrouve d'ailleurs des réflexions antérieures (des bullionistes du XVII<sup>e</sup> siècle à Sismondi au XIX<sup>e</sup>) qui avaient déjà posé le délicat problème de l'intégration de la monnaie dans le schéma de fonctionnement de l'économie.

La macroéconomie s'est donc pendant longtemps identifiée à la question monétaire. Les (op)positions peuvent être séparées en deux camps :

– les keynésiens, pour qui la monnaie est un instrument de politique économique qui va aider à sortir des crises – *a minima* elle doit

convenablement accompagner les politiques de relance budgétaire ;

– les monétaristes (Milton Friedman), eux-mêmes héritiers des « quantitativistes » (Richard Cantillon, Irving Fisher), qui pensent que la monnaie est fondamentalement neutre sur le niveau réel d'activité (et le chômage) et qu'il faut au contraire en éviter toute manipulation. Le débat s'étale sur cinquante ans (de 1936 aux années 1980), et la macroéconomie est tellement associée à la monnaie que dans la nomenclature de l'Association américaine d'économie (AEA), qui propose une partition de l'économie en différentes sous-sections, l'entrée « Macroeconomics and Monetary Economics » résume la production scientifique en macroéconomie. L'AEA propose la définition suivante du champ : « La macroéconomie recouvre les études théoriques et empiriques au sujet des performances agrégées d'une économie : son niveau de production, son niveau d'emploi, ses prix, ses taux d'intérêt et leurs déterminants. »

L'approche keynésienne, en décrivant le fonctionnement d'une économie monétaire de production où les agents économiques sont dotés d'une préférence pour la liquidité et où la demande peut être insuffisante pour conduire l'économie vers le plein emploi, constituera une rupture dans l'histoire des grandes idées économiques. Avec Keynes, c'est un horizon temporel court qui s'impose dans l'agenda. L'étude de la dynamique de long terme de l'économie privilégiée par les économistes classiques laisse la place à une étude des fluctuations de l'activité, à travers notamment la théorie des cycles. Selon l'approche du National Bureau of Economic Research (NBER), ces cycles proposent une séquence de phases telles que l'expansion, la récession, la contraction et la reprise. Cette analyse des cycles repose soit sur l'observation d'un grand nombre de séries statistiques spécifiques concernant le prix des matières premières, les taux d'intérêt..., soit sur l'analyse de l'écart de l'évolution de l'activité économique par rapport à la tendance de long

<sup>[1]</sup> Keynes J. M. (1936) (1988), *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, Paris, Payot.

terme (notion d'*output gap*). L'explication de ces cycles passe par l'analyse de l'investissement, composante la plus volatile du PIB, et de la politique publique. On retrouve notamment l'analyse de Paul Samuelson en termes d'oscillateur (1939<sup>2</sup>), qui combine les approches en termes d'accélérateur simple et de multiplicateur, et celle de Goodwin (1951<sup>3</sup>) en termes d'accélérateur flexible.

## La macroéconomie comme science de l'interaction physico-financière de grands marchés en déséquilibre

L'approche du NBER (et de ses équivalents européens, britannique, français ou hollandais) marque le temps du triomphe de la macroéconomie des ingénieurs économistes et du « modèle IS-LM ». Longtemps enseigné comme le graal de la macroéconomie (quasiement tous les cours de macroéconomie des années 1950 jusqu'aux années 1990 commençaient par là), le modèle IS-LM a été publié en 1937 dans la revue scientifique *Econometrica*. Dans son article, John Hicks, l'auteur, s'efforce de concilier une approche en termes d'équilibre général de marché (la tradition « walrasienne » française) avec une partie du propos de Keynes dans la *Théorie générale*. Chez Hicks, l'accent est mis sur la relation entretenue entre deux types de marché : les marchés monétaires (et financiers), qui déterminent le taux d'intérêt, et le marché des biens et services, qui fixe le niveau d'activité et, partant, le niveau d'emploi. Le taux d'intérêt se situe très exactement à l'interface des comportements risqués, voire « spéculatifs », sur les marchés financiers et des décisions de production sur les marchés de l'économie réelle. Tout en se réclamant d'une interprétation « mathématique » de la pensée keynésienne, le modèle IS-LM n'est rien d'autre que l'étude de l'équilibre simultané de ces deux marchés dans un cadre de prix fixes.

On parle ici de fondements microéconomiques de la macroéconomie keynésienne.

L'intuition générale est que les processus concrets de formation des prix sur les marchés peuvent être défailants : les marchés échouent dans leur rôle d'allocation efficace des ressources et c'est le point d'entrée de la politique économique<sup>4</sup> et de l'intervention de l'État sur les marchés en dehors de ses fonctions régaliennes. Celle-ci s'appuiera notamment sur la comptabilité nationale, qui prend une place grandissante dans l'analyse macroéconomique, à travers le recours aux modèles macroéconométriques popularisés par les travaux de Jan Tinbergen. Le succès d'IS-LM est dû à deux faisceaux de causes bien différentes. D'abord, IS-LM a pu constituer un instrument de clarification du message de la *Théorie générale* pour bon nombre d'économistes parfois peu à l'aise dans la lecture du texte keynésien lui-même. Seconde raison de la réussite de la « synthèse » : le modèle IS-LM et, surtout, ses extensions, ont constitué, de fait, les outils incontournables de l'analyse et de la détermination des politiques économiques conjoncturelles, et cela tout au long des années 1950 et 1960. Les politiques de *stop and go* sont anticipées et dimensionnées à partir de modèles physico-financiers dérivés d'IS-LM. Plus précisément, le modèle offre la détermination du niveau général des prix, permet à l'autorité politique de régler finement l'activité économique par une séquence adaptée de décisions de politique budgétaire : lorsque le plein emploi est atteint et que les tensions inflationnistes paraissent fortes, on décide du « *stop* », avec une réduction des déficits publics ; lorsqu'elles s'éloignent, on « *relance (go)* » pour recréer à nouveau des emplois, cette fois en creusant les déficits. L'intégration de la flexibilité des prix dans le modèle IS-LM avait l'immense avantage de reproduire parfaitement la courbe de Phillips que l'on pouvait alors observer statistiquement : plus d'inflation, moins de chômage<sup>5</sup>. La réduction du chômage s'associe nécessairement à des pressions inflationnistes, car c'est en jouant sur l'illusion monétaire des agents (les agents interprètent mal les signaux de

[2] Samuelson P. A. (1939), « Interactions between the Multiplier Analysis and the Principle of Acceleration », *Review of Economics Statistics*, vol. 21.

[3] Goodwin R.M. (1951), « The Non-linear Accelerator and the Persistence of Business Cycles » *Econometrica*, vol. 19.

[4] Deux auteurs majeurs, Robert Clower et Axel Leijonhufvud, vont reprendre cette synthèse pour développer la « théorie du déséquilibre » : la rigidité des prix devient le point focal de la discussion entre néoclassiques (microéconomistes) et keynésiens (macroéconomistes), sans plus d'oppositions idéologiques. On parle alors de synthèse keynéso-classique.

[5] La courbe de Phillips est une relation statistique. C'est sa version dite « Solow-Samuelson » (1960) qui posera la question en termes de dilemme inflation-chômage.



[6] Lucas R. E., Jr (1976), « Econometric Policy Evaluation: A Critique », *Carnegie Rochester Conference Series on Public Policy*, vol. 1.

[7] Sargent T. J. et Wallace N. (1975), « "Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule », *Journal of Political Economy*, vol. 83, n° 2.

[8] Barro R. J. (1974), « Are Government Bonds Net Wealth? », *Journal of Political Economy*, vol. 82, n° 6.

[9] Diamond P. (1965), « National Debt in a Neoclassical Growth Model », *American Economic Review*, vol. 55, n° 5.

[10] King R., Plosser Ch. et Rebelo S. (1988), « Production, Growth and Business Cycles. I. The Basic Neoclassical Model », et « Production, Growth and Business Cycles. II. New Directions », *Journal of Monetary Economics*.

Kydland F. et Prescott E. (1982), « Time to Build and Aggregate Fluctuations », *Econometrica*, vol. 50.

Long J. et Plosser C. (1983), « Real Business Cycle », *Journal of Political Economy*, vol. 91, n° 1.

[11] Stiglitz J. (1999), *Economics of the Public Sector*, Norton.

[12] Attention, la coupure du droit de propriété des entreprises sur les deux secteurs compte beaucoup dans l'histoire, et il est évident qu'une des solutions au problème posé ci-dessous est la création d'une seule entité décisionnaire, compétente à la fois pour le sucre et le café. Cet exemple est évidemment une parabole simplifiée destinée à faire saisir le problème (y compris le problème de la « propriété »...).

prix) qu'on obtient une relance de l'activité. Ceci jusqu'à un certain point, où les agents réalisent leurs erreurs et ajustent leurs comportements (l'économie redevient alors « classique »). Cette vision va tenir jusqu'à la crise de 1973 et l'entrée des pays de l'OCDE dans la stagflation.

## La macroéconomie comme science de la coordination

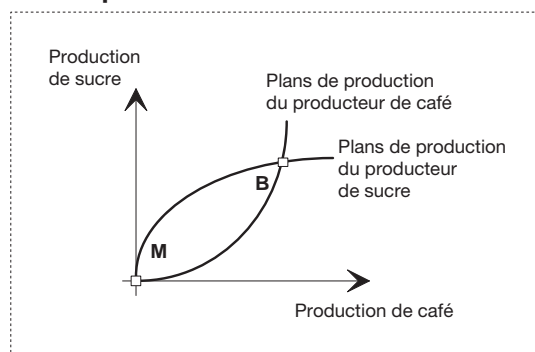
La macroéconomie, en évoluant dans ses réflexions, et face à sa confrontation au réel, a donc été conduite à mettre au premier plan le rôle des anticipations et l'importance des hypothèses faites sur l'horizon temporel des agents. Dans les années 1970, la littérature produite par les économistes et publiée dans les revues scientifiques est édifiante : tous les résultats d'efficacité de la politique économique en viennent à être conditionnés sur la présence, ou non, d'erreur systématique d'estimation des agents (Lucas, 1976<sup>6</sup> ; Sargent et Wallace, 1975<sup>7</sup>) ou sur la présence d'horizons de vie limités (Barro, 1974<sup>8</sup>) : Robert Lucas nous explique que la politique économique est inefficace quand les agents font des « anticipations rationnelles » car toute variation de la demande globale est alors correctement anticipée ; Robert Barro obtient un résultat d'inefficacité de la politique budgétaire lorsque les agents font des choix de consommation calculés sur des horizons de vie dynastique (chose qui ne tient pas, *a contrario*, dans des modèles où les agents ont un horizon d'optimisation fini, comme dans les « modèles à générations imbriquées » – le résultat de suraccumulation de Diamond (1965<sup>9</sup>).

De fait, peu à peu, les questions de niveaux des prix, ou en amont de la monnaie, sont évacuées de la macroéconomie (cela renvoie à la « super-neutralité » de la monnaie avancée par Thomas Sargent et Neil Wallace) et on revient à une réflexion qui s'intéresse uniquement à la coordination des agents au sein d'une économie « réelle » (au sens, d'une

économie « déflatée » du niveau des prix). L'école du « *real business cycle*<sup>10</sup> », née dans les années 1980, en est emblématique : comment les cycles économiques apparaissent-ils lorsqu'ils ne sont pas dus à des manipulations monétaires ? Comment se propagent-ils dans l'économie, à travers le temps et/ou l'espace ? Ce nouveau moment de la macroéconomie récente totalement la recherche sur les mécanismes de formation du capital et de la croissance économique. En la matière, le modèle de référence est celui de Solow (1953). Il est d'ailleurs à noter qu'entre les années 1950 à 1990 et les années 1990-2010, dans les universités et dans les manuels, les cours magistraux de macroéconomie changent de point de départ : dans la première période, tout honnête professeur de macroéconomie passait les cinq premières séances de son cours sur IS-LM ; depuis la fin des années 1990, c'est par le modèle de Solow qu'on commence, en étudiant la dynamique de l'accumulation du capital par tête.

Bien sûr, le modèle de Solow n'intègre pas véritablement de problème de coordination : c'est une « robinsonnade » où l'agent économique, isolé, « joue » contre la nature (hostile, puisque définie par des rendements décroissants). Mais l'histoire, généralement, ne s'arrête pas là. Il suffit d'ajouter un autre agent à la « robinsonnade » (disons « Vendredi », un second entrepreneur) pour que les problèmes surgissent dans l'équilibre économique. Reprenons un exemple de Stiglitz (1999<sup>11</sup>), avec une économie fictive fermée, composée de deux biens, le café et le sucre. Nous supposons que l'on ne consomme le café qu'avec du sucre, et le sucre qu'avec du café : les deux biens sont donc parfaitement complémentaires. Imaginons aussi que, dans cette économie fictive, il existe une seule firme produisant du café, la firme dirigée par Vendredi, et une seule firme produisant du sucre, celle de Robinson. Quelles sont les décisions de production des deux firmes prises isolément<sup>12</sup> ? On peut reconnaître que les deux entreprises font face à un problème de coordination : soit les deux firmes produisent des quanti-

## Marchés et productions complémentaires, un exemple



Ici, au moins deux systèmes de prix d'équilibre sont possibles et deux niveaux de la demande : dans l'un des systèmes, les prix et la production de café et de sucre sont nuls (M sur le graphique) ; dans l'autre non (B). Il y a, en utilisant le registre de la macroéconomie théorique, « multiplicité d'équilibres ».

tés positives des deux biens, tenant compte de la demande jointe *potentielle* du pays pour ce mélange de biens ; soit elles peuvent tout aussi bien se coordonner sur un « équilibre » où aucune d'entre elles ne produit les biens, tablant sur un comportement identique chez l'autre (voir graphique : les plans de production des deux firmes sont imbriqués). Les agents, même parfaitement informés, ne disposent pas d'un critère définitif (précisément, une *valorisation monétaire* – une espérance de profit associé à un prix) pour se coordonner sur une action plutôt que l'autre. Cette dimension de la décision économique s'oppose à la vision développée notamment par Walras d'une coordination complète, réalisée multilatéralement par le système de prix. À l'appui de ces réflexions, mais en oubliant la facilité pédagogique du café et du sucre, on peut par exemple se référer à Cooper et John (1988<sup>13</sup>), s'inspirant eux-mêmes des intuitions de Keynes (1936 et 1937), qui montrent que des défauts de coordination, liés à des complémentarités stratégiques entre agents économiques, débouchent sur des situations d'équilibres macroéconomiques multiples.

Fondamentalement, l'incertitude économique subie par les entrepreneurs au moment de la détermination des niveaux de production doit être reliée à l'absence de marchés d'as-

surance contre les méventes : c'est parce que les entrepreneurs ne peuvent pas couvrir le risque d'une surproduction lors de la mise en vente de leur production qu'ils sont dépendants l'un de l'autre et que l'incertitude demeure<sup>14</sup>, et ce malgré des anticipations rationnelles « justes » en moyenne.

Ce problème définit selon nous le champ actuel occupé par la macroéconomie, à savoir la question des équilibres multiples et des prophéties autoréalisatrices<sup>15</sup>. Les économistes travaillant sur ce sujet retiennent notamment de Keynes que l'équilibre macroéconomique résulte d'un processus de coordination entre des décideurs placés en information incomplète (parfois en « incertitude radicale »). Éclairée par les concepts de la théorie des jeux, l'analyse économique montre alors que les solutions de coordination entre un petit nombre d'agents-investisseurs sont multiples : au premier titre, la croissance ou la récession sont toutes les deux possibles *a priori*<sup>16</sup>. Face à cette indétermination, « l'état de confiance » (les « *animal spirits* ») est un déterminant essentiel de l'investissement en situation d'incertitude, comme Keynes l'avait déjà souligné<sup>17</sup>. La détermination d'un équilibre peut se faire en fonction de croyances apparemment arbitraires (déclenchées par des aléas purement extrinsèques, comme des taches solaires), mais ces croyances collectives ont une dimension autoréalisatrice<sup>18</sup>. À l'équilibre, l'économie est interprétée dans le cadre de récits qui viennent valider et donner une apparence de rationalité à nos actes. Notons que les systèmes théoriques eux-mêmes, classique et keynésien, peuvent jouer ce rôle de « récits coordonnateurs<sup>19</sup> ». Les rêves deviennent réalité... C'est précisément la conclusion de la *Théorie générale*, qui présente les « chimères » – le système théorique construit par Keynes lui-même – comme susceptible de se réaliser.

\* \* \*

Cette très brève présentation de la macroéconomie, surtout cette dernière section,

[13] Cooper R., John A. (1988), « Coordinating Coordination Failures in Keynesian Models », *Quarterly Journal of Economics*, vol. 103, n° 3.

[14] Grandmont J.-M. (1976), « Théorie de l'équilibre temporaire général », *Revue économique*, vol. 27, n° 5, septembre.

[15] Cass D. et Shell K. (1983), « Do Sunspots Matter? », *Journal of Political Economy*, n° 91 ; Azariadis C. et Guesnerie R. (1986), « Sunspots and Cycles », *Review of Economic Studies*, n° 53 ; Guesnerie R. et Woodford M. (1992), « Endogenous Fluctuations » in Laffont J.J., *Advances in Economic Theory*, Cambridge MA., Cambridge University Press, vol. 2.

[16] Cooper R. et John A. (1988), *op. cit.* ; Evans G.S. Honkapohja *et al.* (1998), « Growth Cycles », *American Economic Review*, vol. 88.

[17] Le processus de coordination est déjà précisément dépeint par Keynes dans la *Théorie générale* sous le terme de « demande effective » au chapitre 5. On y trouvera un détail du processus, d'abord mental puis effectif, par lequel les entrepreneurs fixent le volume de l'activité par un « équilibre en anticipations rationnelles croisées » (comme pour Robinson et Vendredi, mais sans l'idée de biens complémentaires).

[18] Azariadis C. (1981), « Self-Fulfilling Prophecies », *Journal of Economic Theory*, vol. 25. Azariadis C. et Guesnerie R. (1982), « Prophéties créatrices et persistance des théories », *Revue économique*, vol. 33. Une littérature étudie les trajectoires multiples, le chaos et les cycles endogènes dans les modèles dynamiques agrégés ; voir par exemple Grandmont J.-M. (1985), « On Endogenous Competitive Business Cycles », *Econometrica*, vol. 53 ; Benhabib J. et Farmer R. (1994), « Indeterminacy and Increasing Returns », *Journal of Economic Theory*, vol. 63.

[19] Ventelou B. (2001), *Au-delà de la rareté, la croissance économique comme construction sociale*, Paris, Albin Michel ; Ventelou B. (2004), *Millennial Keynes: The Origins, Development and Future of Keynesian Economics*, New York, M.E. Sharpe Publisher.

[20] Boyer R. (2014), « L'économie peut-elle (re)devenir une science sociale ? » À propos des relations entre économie et histoire, *Revue française de socio-économie*, n° 13.

[21] Cartapanis A. (2011), « La crise financière et la responsabilité des économistes », *Tracés. Revue de sciences humaines* [En ligne].

[22] Dans une interview donnée au journal *Les Échos* le 25 juin 2013.

pourrait sembler faire de celle-ci un simple complément de la science économique. La macroéconomie se serait constituée avec un objectif de *correction* des insuffisances de la microéconomie, correction parfois très accessible, parfois très critique, mais toujours avec l'idée de simple « patch » correctif. Mais la macroéconomie, c'est aussi, pour certains économistes, la volonté de poser une opposition, un « *no bridge* », entre ce qui se passe au niveau des agents et au niveau de la société tout entière. On revient au premier temps de la macroéconomie, comme la science du pouvoir, à ne pas interpréter comme la science *au service du pouvoir* (en tant qu'institution gouvernementale), mais la science qui *instille la relation de pouvoir* (synonyme de « domination »), dans la microéconomie simpliste de l'individu représentatif.

Aujourd'hui, en effet, on peut reprocher à la macroéconomie, non seulement de s'être laissé ré-enfermer dans un cadre microéconomique, où le primat donné à l'individualisme méthodologique et à la recherche de l'équilibre stationnaire pourrait empêcher de penser les interactions complexes à l'échelle d'une nation (Boyer<sup>20</sup>, 2014), mais aussi de

s'être détournée de ses sujets fondamentaux : les sujets du moment, notamment des questions relatives à l'instabilité financière<sup>21</sup> ; et les sujets fondamentaux intemporels, à savoir les relations de domination sociale qui structurent l'échange économique entre des acteurs de puissance inégale. Il serait alors nécessaire de « rebâtir » (selon les termes d'Olivier Blanchard<sup>22</sup>) une macroéconomie trop aveuglée par ses certitudes. Le nouvel agenda de recherche s'intéresserait par exemple aux interactions entre la dynamique des systèmes financiers et celle du système économique réel, dans le but de mieux comprendre et prévoir le caractère endogène de certaines crises. Il pourrait aussi retourner à une pluralité de modèles d'équilibre partiel qui, réassemblés, prendraient peut-être mieux en compte la complexité d'une économie globalisée (un « meta-NBER », coordonnateur de politiques publiques mondiales ?), ou encore se rapprocher d'autres sciences sociales comme l'histoire, la sociologie ou la psychologie qui permettraient de mieux comprendre l'économie, insérée dans la société, et en interaction avec elle.

## POUR EN SAVOIR PLUS

❖ **BLANCHARD O. et COHEN D.** (2010), *Macroéconomie*, Paris, Pearson, 5<sup>e</sup> éd.

❖ **DEVROEY M. et MALGRANGE P.** (2007), « La théorie et la modélisation macroéconomiques, d'hier à aujourd'hui », *Revue française d'économie*, vol. 21, n° 3.

❖ **ROMER D.** (1997), *Macroéconomie approfondie*, Edi-science.

Cœur de la théorie néoclassique, le modèle de l'équilibre général a longtemps été le modèle « standard » de la microéconomie à partir duquel on pouvait, en relâchant progressivement les hypothèses les plus éloignées de la réalité, comprendre les imperfections de marché. La théorie de l'équilibre général a toutefois buté sur de sérieuses difficultés : ni l'unicité de l'équilibre ni la forme décroissante de la demande agrégée n'ont pu être démontrées. Si ces résultats décevants ont poussé les microéconomistes à raisonner plutôt en termes d'équilibre partiel, les macroéconomistes, soucieux de donner des fondements microéconomiques à leurs modèles, se sont en revanche saisis à leur tour de la théorie de l'équilibre général, ce qui a donné naissance aux modèles « DSGE » – équilibre général dynamique et stochastique. Selon Ozgur Gun et Sophie Jallais, des problèmes sérieux demeurent dans ces modèles, notamment celui de la non-prise en compte de l'hétérogénéité des agents et de leurs interactions, éléments essentiels, par exemple, pour comprendre la crise récente.

Problèmes économiques

# Que reste-t-il de l'équilibre général ?

En économie, le terme « équilibre général » désigne l'approche qui part des goûts, des techniques de production et du cadre institutionnel pour étudier la coordination des décisions économiques des agents (ménages ou entreprises) en tenant compte de leurs interdépendances et de toutes leurs conséquences directes et indirectes. Les décisions dont il s'agit ici portant exclusivement sur les offres et demandes de biens, dire qu'elles se coordonnent, qu'elles sont compatibles, c'est dire que la quantité demandée de chaque bien est égale à la quantité qui en est offerte. Et c'est cette dernière situation que l'on désigne par l'expression « équilibre général ».

Si l'approche par l'équilibre général – qui constitue le cœur de la théorie économique néoclassique – a longtemps été considérée comme relevant de la seule microéconomie, depuis quelques années, pourtant, elle est surtout revendiquée par les macroéconomistes, dans le cadre d'une recherche de « fondements microéconomiques ». Ceci est

© **OZGUR GUN**

REGARDS – Université de Reims Champagne-Ardenne

© **SOPHIE JALLAIS**

PHARE – Université Paris-I Panthéon-Sorbonne

paradoxal dans la mesure où la macroéconomie traite d'entités globales (les agrégats), dont les interactions sont par nature bien plus limitées que dans le cas général. Cette évolution s'explique en partie par les écueils qu'a rencontrés la théorie de l'équilibre général dans les années 1970-1980.

## Donner un contenu à la « loi de l'offre et de la demande »

La théorie de l'équilibre général est née du besoin de fonder théoriquement l'idée très répandue selon laquelle, si les prix sont soumis à la « loi de l'offre et de la demande »,

autrement dit s'ils augmentent lorsque la demande est supérieure à l'offre et diminuent dans le cas contraire, alors ils convergent vers un équilibre général – une situation où, pour chaque bien, l'offre est égale à la demande, et où, en conséquence, chacun est satisfait, ses objectifs étant atteints.

Pour que cette question de la convergence vers (ou stabilité) des prix d'équilibre se pose, il faut bien évidemment que des prix d'équilibre existent (en théorie). Préalablement au questionnement sur sa stabilité, les théoriciens devaient donc démontrer qu'un tel équilibre général existait. Pour ce faire, ils imaginèrent un monde dans lequel les offres et les demandes des agents résultent de la maximisation de leur fonction objectif (fonction d'utilité pour les ménages et de profit pour les entreprises) et sont confrontées globalement. Ils s'attachèrent ensuite à prouver que, sous certaines conditions supplémentaires, il existe des prix auxquels ces offres et demandes sont égales. Plus précisément, supposant que les agents formulent leurs demandes et leurs offres, notées respectivement  $d(P)$  et  $o(P)$ , à des prix  $P = (p_1, \dots, p_n)$  donnés (un prix pour chacun des  $n$  biens de l'économie), leur question fut alors de savoir si, parmi tous les prix possibles, il en existe certains – les prix d'équilibre général –, notés  $P_e$ , tels que  $d(P_e) = o(P_e)$  : l'offre est égale à la demande pour chaque bien.

Au milieu des années 1950, Kenneth Arrow et Gérard Debreu apportèrent une réponse positive à cette question en faisant appel à des mathématiques relativement compliquées et à des hypothèses concernant le cadre institutionnel et les « fondamentaux » de l'économie, c'est-à-dire les goûts et les techniques de production. Ces hypothèses sont donc autant de conditions d'existence de prix d'équilibre  $P_e$  au sens (mathématique) d'existence d'une solution au système d'équations  $d(P_e) = o(P_e)$ . L'ensemble de ces conditions et la démonstration du théorème d'existence, appelé « modèle Arrow-Debreu », est la version la plus élaborée du **modèle d'équilibre général**.

Après ce résultat remarquable, on pensait qu'il ne restait plus qu'à prouver que l'équilibre était unique et que, grâce à la « loi de l'offre et de la demande », il était stable (les prix convergeaient vers l'équilibre). La démonstration n'arriva cependant jamais et les théoriciens s'accordèrent pour dire qu'elle était impossible, et que l'unicité et la stabilité de l'équilibre étaient l'exception plutôt que la règle : en raison de l'interdépendance des offres et des demandes des différents biens, une hausse du prix d'un bien peut, en effet, avoir un impact globalement positif sur sa demande, remettant ainsi en cause la forme décroissante habituellement attribuée aux fonctions de demande<sup>1</sup>. Il s'ensuit que les courbes d'offre et de demande du modèle peuvent avoir des formes telles que, même à supposer que la loi de l'offre et de la demande soit vérifiée, les prix ne convergent pas vers les prix d'équilibre général.

Il s'ensuit également qu'il est impossible de prouver, dans un cadre suffisamment général, l'existence de « lois » ou de régularités concernant, par exemple, les effets sur les prix d'une variation dans les fondamentaux de l'économie. L'habitude a été prise de parler de « théorème de Sonnenschein-Mantel-Debreu (SMD) » à propos de ces résultats négatifs, qui portèrent un coup fatal au programme de recherche sur l'équilibre général, aussi bien en concurrence parfaite qu'en concurrence imparfaite.

## Des problèmes qui persistent en concurrence imparfaite

Le modèle d'équilibre général évoqué dans la partie précédente est construit autour de l'hypothèse selon laquelle les agents considèrent les prix comme donnés : ils sont « preneurs de prix » – condition pour que les expressions  $d(P)$  et  $o(P)$  aient un sens. Tout le monde étant de ce point de vue sur un pied d'égalité, on dit que l'on est en « concurrence parfaite ». C'est d'ailleurs la raison pour laquelle le modèle d'équilibre général de Arrow et Debreu est

<sup>[1]</sup> Du moins quand on raisonne en « équilibre partiel », c'est-à-dire sur le prix et la demande d'un seul bien en faisant comme si les autres biens n'existaient pas.

également appelé **modèle de concurrence parfaite**. Cela dit, sortir de ce cadre, idéal mais difficile à justifier, ne permet pas d'éviter l'écueil décrit plus haut.

Lorsque l'on suppose que certains agents prennent l'initiative de proposer des prix, on parle de « **concurrence imparfaite** » par opposition au modèle précédent. Le cas le plus simple est celui où un bien est produit par une unique entreprise, ses éventuels clients se comportant en « preneurs de prix ». C'est le cas dans le modèle dit du **monopole**, qui est relativement facile à traiter... du moins en équilibre partiel, où l'on ne tient compte que du bien en question, à l'exclusion de tous les autres, le modélisateur « se donnant » alors une courbe de demande du bien décroissante. La situation est autrement plus ardue en équilibre général, puisque l'entreprise doit déterminer la demande qui s'adresse à elle pour les divers prix qu'elle peut proposer, mais en tenant compte, cette fois, de ce que cette demande dépend, entre autres, de la rémunération de ses salariés et des profits qu'elle distribue à ses actionnaires. On se retrouve en fait, côté demande, dans la même situation que dans le cas de la concurrence parfaite, où s'applique le théorème de SMD : la forme de la fonction de demande peut être quelconque. Il se peut donc que le problème du monopole n'ait pas de solution et, en conséquence, que le modèle n'ait pas d'équilibre. Or, sans équilibre, un théoricien néoclassique n'a plus de point de repère. C'est la raison pour laquelle, loin de l'idéal d'une théorie unifiée que la théorie de l'équilibre général semblait promettre, les microéconomistes durent se replier sur une approche d'équilibre partiel, selon eux bien moins rigoureuse, et qui, surtout, se traduit par une multitude de petits modèles qui dépendent de choix *a priori* du théoricien.

## Équilibre général et macroéconomie

Alors que les microéconomistes constataient que la théorie de l'équilibre général était dans l'impasse, des macroéconomistes commencèrent à brandir son étendard au nom de la

rigueur et de la cohérence. Selon eux, elle devait notamment permettre de donner des « fondements microéconomiques » à leur discipline – la macroéconomie étant alors caractérisée par des modèles superposant des identités comptables, des relations de comportement observées globalement et des équations en équilibre partiel, le tout n'étant pas forcément cohérent (Kydland et Prescott, 1990).

Parler d'équilibre général en macroéconomie ne va toutefois pas de soi. À la différence de la microéconomie, la macroéconomie a, en effet, pour vocation de construire des modèles susceptibles d'être testés empiriquement puis utilisés dans le cadre de la mise en œuvre des politiques économiques. Elle opère, en outre, sur des agrégats tels que le PIB, la consommation, l'investissement, le niveau des prix, dont le calcul mêle le plus souvent des quantités et des prix. Les agents, enfin, y sont eux aussi considérés de façon globale – du moins par grandes catégories. L'approche diffère donc de celle de l'équilibre général décrite plus haut, où prix et quantités sont soigneusement séparés et où une attention toute particulière est accordée aux interactions entre les décisions des agents considérés séparément.

Deux grands types de modèles macroéconomiques se réclament pourtant aujourd'hui de la théorie de l'équilibre général : les modèles d'« équilibre général calculable » et les modèles « dynamiques et stochastiques d'équilibre général ». Les premiers accordent une place essentielle aux interactions entre secteurs et catégories d'agents. Les seconds mettent l'accent sur la maximisation de la fonction objectif d'un agent qualifié de représentatif car ses choix sont censés représenter ceux de l'ensemble de l'économie.

## Du calcul de l'équilibre général à l'équilibre général calculable

Après que l'existence d'un système de prix en équilibre général a été établie sur le plan théorique, certains économistes se

lancèrent dans la recherche d'algorithmes permettant le calcul (théorique) de ces prix. Leur première idée fut d'appliquer la « loi de l'offre et de la demande » en partant de prix quelconques et de fonctions d'offre et de demande ayant les propriétés du modèle Arrow-Debreu. Elle fut toutefois vite abandonnée, la convergence s'avérant trop lente, voire inexistante (confirmant le théorème de SMD). D'autres techniques mathématiques furent alors proposées, notamment par Herbert Scarf (1973). Contre-intuitives sur le plan économique, et en particulier fort éloignées de l'idée de « loi de l'offre et de la demande », elles demandent de longs calculs dès que les biens et les agents sont en nombre important.

Au même moment, la collecte et le traitement statistique des données connaissent un essor considérable, dans le cadre de la comptabilité nationale, dont la logique « emploi-ressources » présente des similitudes avec celle de l'équilibre général. L'économiste norvégien Leif Johansen construisit sur cette base des « matrices de comptabilité sociale » pouvant notamment être utilisées par des planificateurs (Johansen, 1960). Inspirées par les tableaux *input/output* de Vassily Leontief, ces matrices rendaient compte à la fois des interdépendances entre les divers secteurs productifs de l'économie, des flux de revenus qui y sont engendrés, de leur répartition entre les diverses catégories d'agents et de l'affectation de leurs dépenses entre les divers produits. À la différence des modèles d'équilibre général « purs » étudiés par Scarf et ses étudiants, ces modèles, dont le but était de tester diverses politiques économiques, raisonnaient directement sur des agrégats calculés en valeur, à prix donnés.

La différence entre les conceptions de Scarf et de Johansen apparaît nettement au niveau des mathématiques utilisées : alors que Scarf faisait appel à des techniques relativement compliquées consistant à approcher l'équilibre par itérations successives, Johansen n'avait qu'à inverser la matrice qui liait

les variables exogènes (celles que le théoricien considère comme données) aux variables endogènes (déterminées dans le modèle).

Malgré leur grande différence – dans leurs références théoriques, leurs objectifs, le choix des variables exogènes et endogènes, les techniques utilisées –, les approches par l'équilibre général et par les matrices de comptabilité sociale furent « combinées » dans les années 1980, sous l'impulsion de la Banque mondiale, pour créer les modèles dits d'« équilibre général calculable » (EGC). Ceux-ci sont ainsi constitués par un modèle d'« équilibre général » agrégé, avec des agents et des biens « représentatifs », les prix d'équilibre étant obtenus par égalisation des offres et des demandes globales calculées à partir des données fournies par les matrices de comptabilité sociale ; ce qui pose un problème de cohérence : les éléments de la matrice de comptabilité sociale, mesurés en valeur, dépendent des prix, alors que les prix font partie des « inconnues » du modèle d'équilibre général (ses variables endogènes) – problème récurrent, qui se pose chaque fois que des agrégats sont utilisés dans une démarche de type microéconomique (Guerrien et Gün, 2014). Une autre difficulté logique rencontrée par les modèles d'EGC provient de ce qu'ils assimilent les identités comptables de la matrice de comptabilité sociale, toujours vérifiées par définition, à des conditions d'équilibre, qui ne sont vérifiées qu'exceptionnellement. Enfin, il découle de la façon dont est construite la matrice de comptabilité sociale que le nombre de variables est en général supérieur à celui des équations, ce qui oblige à considérer certaines de ses variables comme exogènes (problème de « bouclage » du modèle). Souvent, c'est l'offre de « facteurs » (travail, capital) ou l'investissement qui joue ce rôle, alors que, pour les théoriciens de l'équilibre général, seuls les fondamentaux de l'économie sont exogènes.

Toutes ces limites expliquent que les théoriciens soient très réservés à l'égard de ces modèles<sup>2</sup>, qui sont pourtant utilisés par les praticiens – surtout des pays en développe-

<sup>[2]</sup> Aucun « prix Nobel » d'économie n'a par exemple été décerné à des travaux sur l'EGC.

ment – auxquels ils fournissent une boîte à outils d'utilisation relativement facile, dans des contextes variés.

## L'équilibre général « dynamique et stochastique » de la macroéconomie

Aujourd'hui, la voie « noble » d'équilibre général en macroéconomie – qui a déjà eu droit à plusieurs « prix Nobel » – est celle des modèles dits d'« équilibre général dynamique stochastique » (DSGE, en anglais) : « dynamique » car ces modèles traitent de l'évolution dans le temps d'un certain nombre d'agrégats ; « stochastique » en raison des chocs aléatoires, exogènes, qui perturbent celle-ci ; « équilibre général » car cette évolution résulte de la maximisation d'une fonction d'utilité intertemporelle d'un agent qualifié de « représentatif » (sous-entendu de l'ensemble des agents de l'économie, eux-mêmes supposés identiques sans quoi leurs choix ne pourraient être agrégés). Le modèle n'étudie donc pas les interdépendances entre les choix d'agents différents comme dans le modèle d'équilibre général, mais celles existant entre les diverses composantes dans le temps du choix d'un unique agent : consommer moins « aujourd'hui » – et donc investir plus – a une incidence sur la consommation future ; augmenter son temps de loisir restreint la production et donc les possibilités de consommation ou d'investissement dans le futur.

Les modèles DSGE comportent, dans leur version de base, deux principaux ingrédients : en premier lieu, une **fonction d'utilité intertemporelle** qui décrit les préférences de l'agent représentatif, d'une part, entre consommation présente et consommation future, et, d'autre part, entre temps consacré au travail et temps consacré au loisir ; en second lieu, une **fonction de production** dont les variables sont le travail de l'agent et le « capital » accumulé.

Ils se présentent donc comme des programmes au sens mathématique, avec une fonction objectif (l'utilité de l'agent repré-

sentatif) à maximiser, dont les variables sont soumises à des contraintes techniques (résumées par la fonction de production de l'économie). La solution de ces programmes (production, consommation, emploi, etc., présents et futurs, de l'agent représentatif) est comparée aux évolutions observées dans les pays étudiés. La qualité de la concordance entre la solution théorique et les observations dépend notamment de la valeur donnée aux paramètres qui caractérisent les fonctions d'utilité et de production. Il revient au modélisateur de donner des valeurs à ces paramètres. Une procédure courante consiste ici à effectuer un calibrage, en choisissant ces valeurs pour que l'évolution théorique reproduise le mieux possible l'évolution observée.

Il manque toutefois à ces modèles les éléments essentiels de l'approche par l'équilibre général : des offres, des demandes et les prix censés les rendre compatibles (c'est-à-dire les prix d'équilibre général). Les promoteurs des modèles DSGE pensent combler cette lacune en évoquant un « grand nombre d'agents identiques », le « grand nombre d'agents » étant un argument souvent invoqué pour tenter de justifier que les agents soient preneurs de prix dans le modèle de concurrence parfaite. Ils associent alors, à la trajectoire-solution du modèle, des trajectoires de prix, puis d'offres et de demandes à ces prix – ce qui permettrait, selon eux, de parler de « marchés ». Mais ce sont là des leurres : des agents identiques n'ont, en effet, aucune raison de faire des échanges. Quant aux prix (en fait, le salaire et le taux de rendement du capital), ils sont rajoutés *a posteriori*, après que l'agent a fait son choix. Le temps que celui-ci a décidé de consacrer au travail est ensuite présenté comme une « offre » de sa part à ces prix, et comme une « demande » de la part de l'entreprise dont il est l'unique propriétaire. Mais ces offres et demandes ne sont en fait que des quantités d'« équilibre », puisque, par construction, le temps offert est égal au temps demandé. Comme il en va de même pour la quantité de



biens produite et pour le « capital », l'équilibre est qualifié de « général ». Quant à l'affectation des ressources, elle est « optimale », puisque le choix de l'agent l'est, par définition<sup>3</sup>.

<sup>[3]</sup> Il est courant chez les auteurs de ces modèles de décrire la situation comme résultant du choix d'un « planificateur » (*social planner*) (King et Rebelo, 2001, p. 17 ; Christiano et Eichenbaum, 1992, p. 433), tout en l'assimilant à une sorte de marché idéal (une fois les prix introduits).

À l'origine, les modèles DSGE avaient pour objectif de rendre compte d'un certain nombre de « faits stylisés » en adoptant une démarche « fondée microéconomiquement », contrairement à celles des modèles existants, obtenus par superposition de relations de comportement ou d'équilibre partiel. Puis des économistes formés à la macroéconomie à agent représentatif furent recrutés par les grandes administrations économiques, notamment les banques centrales, où ils cherchèrent à adapter ce qu'ils avaient appris pour répondre aux questions qui leur étaient posées. Ils furent ainsi conduits à introduire la monnaie comme variable dans la fonction d'utilité de l'agent représentatif – ce que les microéconomistes refusent de faire, la monnaie n'étant pas, en soi, source d'utilité. Ce nouveau besoin, qui se traduit par une demande de monnaie, ne peut être satisfait que si quelqu'un en offre, sous certaines conditions. D'où la nécessité de supposer l'existence d'un nouvel agent, appelé « banque centrale » et caractérisé par une règle concernant l'offre de monnaie. On peut ensuite ajouter un agent supplémentaire – l'« État » – qui émet des « bons » porteurs d'intérêt, contrairement à la monnaie à

laquelle ils peuvent en partie se substituer. L'État utilise l'argent emprunté pour ses dépenses et affecte ainsi le choix, et l'utilité, de l'agent représentatif. Il le fait aussi lorsqu'il lui fait payer des taxes, ne serait-ce que pour pouvoir lui rembourser les bons qu'il lui a vendus.

Au nom du « réalisme », les modèles DSGE ont été ainsi progressivement élargis en y introduisant, outre la monnaie et l'État, un temps de travail réglementé, des « frictions » dans la production, des décalages dans la mise en route des investissements ou dans le versement du salaire par l'entreprise, une perception limitée ou avec retard des « chocs », etc. (Smets et Wouters, 2002). Toutes ces entraves font que le choix de l'agent représentatif n'est plus optimal comparé au cas sans entraves, d'où la nécessité de « réformes structurelles » pour les éliminer et se rapprocher ainsi de l'optimalité. Si cela n'est pas possible, ou en attendant, il revient aux politiques monétaires et budgétaires d'atténuer les effets de ces « imperfections ».

La crise récente a porté un coup très rude aux modèles DSGE. Elle ne peut, en effet, être comprise qu'en considérant un ensemble hétérogène d'agents dont les décisions peuvent, contre leur gré, les précipiter dans un gouffre.

—

## POUR EN SAVOIR PLUS

---

❖ **CHRISTIANO L. J. et EICHENBAUM M.** (1992), « Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations », *The American Economic Review*, 82(3), p. 430 à 450.

❖ **GUERRIEN B. et GÜN O.** (2014), « En finir pour toujours avec la fonction de production agrégée ? » *Revue de la régulation*, n° 15.

❖ **JOHANSEN L.** (1960), *A Multi-Sectoral Study of Economic Growth*, Amsterdam, North-Holland.

❖ **KING R. G. et REBELO S.T.** (1999), « Resuscitating Real

Business Cycles », in Taylor J.B. et M. Woodford (ed.), *Handbook of Macroeconomics, Volume 1*, Elsevier, p. 927 à 1007.

❖ **KYDLAND F. et PRESCOTT E.** (1990) « The Econometrics of the General Equilibrium Approach to Business Cycles », *Research Department Staff Report 130, Federal Reserve Bank of Minneapolis*.

❖ **MITRAH-KAHN B.** (2008), « Debunking the Myths of Computable General Equilibrium Models », *Schwartz Center Economic Policy Analysis*.

❖ **SCARF H.** (1973), *The Computation of Economic Equilibria*, New Haven et Londres, Yale University Press.

❖ **SMETS F. et WOUTERS R.** (2002), « An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area » *ECB Working Paper Series*, n° 171.

❖ **SUWA A.** (1991), « Les modèles d'équilibre général calculable », *Analyse et prévision*, n° 97.

L'équilibre de l'offre et de la demande et le processus d'ajustement par les prix sont des mécanismes fondamentaux des économies de marché. Implicites chez Adam Smith dans la métaphore de la « main invisible », ils ont été théorisés dans les années 1950 sous le modèle de l'équilibre général. Pourtant, comme nous le rappellent Jean-Pierre Biasutti et Laurent Braquet, le mécanisme d'égalisation de l'offre et de la demande, simple en apparence, est en fait complexe et incertain. Le passage d'une analyse en termes d'équilibre partiel à une analyse en termes d'équilibre général proposé par la microéconomie pose notamment des difficultés non résolues liées aux incertitudes pesant sur les courbes de demandes agrégées. L'analyse macroéconomique des fonctions d'offre et de demande, présentée ici sous la forme du modèle offre globale/demande globale, permet toutefois de comprendre la détermination et les fluctuations des principales grandeurs macroéconomiques telles que le PIB et le niveau d'emploi.

Problèmes économiques

# Offre, demande et prix

## © JEAN-PIERRE BIASUTTI

Professeur en classe préparatoire ECE au lycée M. Montaigne de Bordeaux

## © LAURENT BRAQUET

Professeur de SES au lycée G. Flaubert de Rouen et en classe préparatoire IEP

On prête au célèbre essayiste écossais du milieu du XIX<sup>e</sup> siècle, Thomas Carlyle, la boutade suivante : « Apprenez à un perroquet les mots "offre" et "demande" et vous aurez un économiste. » C'est dire l'importance de ces concepts dans le langage courant de cette profession, qui évoque régulièrement la « loi de l'offre et de la demande » comme principe régulateur des économies de marché généralisé au travers de la formation d'un système de prix d'équilibre. John Stuart Mill le formule ainsi dès le milieu du XIX<sup>e</sup> siècle : « La hausse ou la baisse (du prix) ont lieu jusqu'à

ce que l'offre et la demande soient exactement égales l'une à l'autre ; et la valeur (c'est-à-dire le prix) à laquelle une marchandise s'enlève sur le marché n'est autre que celle qui, sur ce marché, détermine une demande suffisante pour absorber toutes les quantités offertes ou attendues »<sup>1</sup>.

Cette loi de l'offre et de la demande, dont on affirme souvent l'évidence empirique (du marché des fraises à celui du pétrole) et qui semble régir le monde économique comme la loi de l'attraction universelle le monde physique, est en fait une construction bien plus complexe que ne le laisse entendre l'évocation commune. Elle impose donc de préciser ce qu'on entend par offre et demande, ainsi que la forme du processus dans lequel elles interviennent.

Les concepts d'offre et de demande opèrent aussi à plusieurs niveaux de l'analyse économique. La similitude des termes employés

<sup>[1]</sup> John Stuart Mill (1848), *Principes d'économie politique*, livre III, chap. 2.

ne doit pas tromper car elle masque des phénomènes très disparates et de plus en plus vagues lorsqu'on passe du niveau microéconomique au niveau macroéconomique.

## Offre et demande comme résultat d'un comportement individuel de maximisation

L'étude du comportement rationnel du consommateur et du producteur permet de déterminer les fonctions d'offre et de demande individuelles. Par sommation de ces dernières, on détermine alors l'offre et la demande sur un marché donné dont on analyse les conditions d'équilibre. Le raisonnement est dit alors en « équilibre partiel ».

### Calcul économique du producteur et courbe d'offre

Intuitivement, on comprend que le producteur a d'autant plus intérêt à offrir une quantité importante de son bien que le prix qu'il en retire est élevé. Ce comportement s'appuie cependant sur un calcul économique plus précis. Connaissant sa fonction de coût, le producteur détermine l'offre qui va maximiser son profit. Cette fonction de coût suppose que le producteur a préalablement déterminé la combinaison de facteurs de production qui minimise le coût de production pour une quantité donnée. Comme la concurrence est pure et parfaite, elle fait du producteur un « *price taker* » en lui imposant le prix du marché. Quand ce prix est égal au coût marginal (celui de la dernière unité produite), le profit total du producteur n'augmente plus ; il est donc maximal. En faisant l'hypothèse que les rendements des facteurs sont décroissants, le coût marginal va augmenter à partir d'un certain niveau de production : il faut chaque fois plus de facteurs pour obtenir une unité supplémentaire de produit. Seule une augmentation du prix de vente peut alors permettre d'augmenter la quantité mise sur le marché,

ce qui introduit une relation positive entre prix de marché et quantité offerte : c'est la **fonction d'offre  $O(p)$** . L'offre totale d'un bien sur le marché sera ensuite obtenue en additionnant les offres individuelles pour chaque niveau de prix.

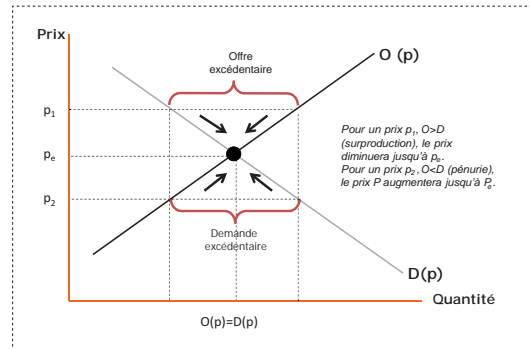
Sur le graphique 1, la courbe d'offre associée à chaque niveau de prix la quantité offerte d'un bien. La courbe d'offre est croissante : plus le prix d'un bien est élevé, plus les producteurs présents augmentent la quantité offerte. D'autre part, sur un plus long terme, l'entrée éventuelle de nouveaux producteurs sur le marché ou le changement de l'échelle de production des offreurs déjà présents, déplace la courbe d'offre vers la droite. C'est aussi le cas lorsque les coûts diminuent sous l'effet du progrès technique, de la baisse du prix des intrants ou d'une réglementation moins contraignante. Dans le cas contraire, elle se déplace vers la gauche.

### Calcul économique du consommateur et courbe de demande

La demande individuelle correspond à la quantité du produit qu'un consommateur est disposé à acheter, compte tenu du prix du marché. La demande a de nombreux déterminants, parmi lesquels les goûts et les préférences, le revenu du demandeur, ou le prix des biens substituables, et, naturellement, le prix du bien considéré. Pour isoler l'effet de ce dernier, le raisonnement se fait « toutes choses égales par ailleurs » (tous les autres facteurs sont constants). Dans la grande majorité des cas, on montre alors que la demande individuelle est croissante lorsque le prix du marché baisse. Le comportement économique du consommateur permet cependant de préciser cette idée intuitive. Face à deux biens substituables A et B, lorsque le revenu du consommateur et le prix du bien B sont donnés, la baisse du prix du bien A provoque la hausse de sa consommation parce que deux effets se composent le plus souvent :

– la baisse du prix de A rend le produit B moins attractif et le consommateur remplace

## 1. Courbes d'offre et de demande



ce dernier par des unités supplémentaires du bien A (**effet de substitution**) ;

– la baisse du prix de A augmente le pouvoir d'achat du revenu, ce qui permet au consommateur d'acheter une quantité plus élevée du produit (**effet de revenu**).

Cette relation décroissante entre le prix et la quantité demandée est la **fonction de demande  $D(p)$** . La demande totale est alors obtenue en additionnant les demandes individuelles pour chaque prix. Sur le graphique 1, la courbe de demande associe dès lors, à chaque niveau de prix, la quantité demandée d'un bien. La courbe de demande est décroissante : plus un bien est cher, moins nombreuses sont les personnes qui veulent l'acheter et moins élevées sont les quantités demandées.

## La loi de l'offre et de la demande : une « main invisible » ?

### L'équilibre partiel par la flexibilité des prix

La forme des courbes « normales » d'offre et de demande garantit qu'elles se croisent en un seul point. À ce prix d'équilibre ( $P_e$ ), les transactions des uns et des autres correspondent à leurs souhaits. Il n'y a, ainsi, ni rationnement de l'offre ni rationnement de la demande. Même si, très souvent, de nombreux acheteurs potentiels sont exclus par le

## ZOOM

### LA NOTION D'ÉLASTICITÉ-PRIX

L'élasticité-prix de la demande (ou de l'offre) d'un bien permet de mesurer la sensibilité de la demande (ou de l'offre) d'un bien par rapport à son prix : ainsi, l'élasticité-prix de la demande (offre) du bien X se mesure en calculant le rapport entre le taux de variation de la demande (offre) du bien X et le taux de variation du prix de ce bien.

prix, leurs ressources ne leur permettant pas d'entrer sur le marché. Il existe donc toujours une combinaison prix-quantités qui permet d'égaliser l'offre et la demande. En cas de déséquilibre sur le marché, la variation du prix permet d'atteindre l'équilibre (voir graphique 1). Une demande supérieure à l'offre pousse le prix à la hausse, ce qui conduit certains demandeurs à renoncer et incite les offreurs à développer leur offre. Le processus se poursuit jusqu'à ce que l'équilibre soit atteint. Un mécanisme inverse se produit lorsque l'offre est supérieure à la demande : la baisse des prix attire de nouveaux demandeurs, tandis que certains offreurs sont moins incités. Ainsi, quand un choc quelconque modifie la demande ou l'offre sur un marché, se met en œuvre automatiquement un processus d'ajustement des prix qui garantit le retour à l'équilibre.

Pour avoir un équilibre sur le marché, il faut donc que les prix soient flexibles. En égalisant l'offre à la demande, ces prix guident les ressources vers les productions pour lesquelles la demande est forte et assurent une allocation optimale des ressources. Les situations de sur ou sous-production sont donc éliminées. À l'équilibre, les quantités offertes et les demandes sont optimales puisque les consommateurs maximisent leur utilité tan-

## ZOOM

### LA CONCURRENCE IMPARFAITE ET LES « FAISEURS DE PRIX »

L'irréalisme de l'hypothèse des « preneurs de prix » (qui suppose un processus hyper-centralisé) a conduit au développement des modèles de concurrence imparfaite qui se rapprochent de l'évidence empirique : ce sont les entreprises qui font les prix. Dans ces situations modélisées, les agents, du fait de leur faible nombre ou de leurs interactions stratégiques, ont un pouvoir sur le marché qui s'exprime dans leur capacité à fixer ou à influencer les prix ; ils sont « *price makers* ». Le cas le plus extrême est celui du **monopole** dans lequel un seul producteur satisfait à toute la demande du marché. Il peut alors fixer le prix qui détermine la quantité vendue maximisant son profit. Cette situation peut être approchée par des entreprises qui, par l'innovation ou le marquage publicitaire, réduisent la substituabilité des produits pour le consommateur et peuvent alors fixer un prix plus élevé pour leur produit ainsi différencié (**concurrence monopolistique**). Les modèles d'oligopole montrent que la « compétition entre quelques-uns » conduit à une manipulation des prix mais ont un résultat plus indéterminé sur l'équilibre final.

Jean-Pierre Biasutti et Laurent Braquet

dis que les producteurs maximisent leur profit. Les gaspillages sont éliminés.

Cependant, un certain nombre de conditions doivent être respectées pour que les mécanismes du marché fonctionnent ainsi. Il est d'abord nécessaire que la concurrence joue normalement sur ce marché ; en particulier, aucun intervenant ne doit pouvoir imposer ses

prix. Par ailleurs, l'offre et la demande doivent être fortement élastiques (*cf.* Zoom p. 35).

### Tâtonnement ou marchandage ?

Le modèle précédent pose un problème de fond. Si les hypothèses de la concurrence pure et parfaite sont respectées, tous les acteurs sont « preneurs de prix », ce qui signifie qu'aucun d'entre eux ne peut agir sur ces prix : comment expliquer alors que ceux-ci fluctuent ? L'analyse économique recourt à une solution coûteuse pour son réalisme mais indispensable pour sa cohérence. Elle suppose en effet que le tâtonnement qui conduit à l'équilibre est le fait d'un agent fictif qui se charge d'annoncer les prix en fonction des écarts qu'il constate entre les offres et les demandes pour chaque bien. Ce « secrétaire de marché » ou « commissaire-priseur », seul agent non intéressé du modèle, doit agir d'une façon bien spécifique : si la demande nette (demande-offre) est positive, il augmente le prix ; si elle est négative, il le baisse jusqu'à arriver au prix qui l'annule (offre = demande). Dans l'attente, les agents ne procèdent pas aux échanges et ne les mettent en œuvre qu'une fois le prix d'équilibre établi. Ce processus implique donc une organisation très centralisée peu congruente avec l'idée que l'on se fait spontanément de la « main invisible » du marché.

En revanche, l'analyse d'un processus plus décentralisé, comme l'échange entre deux agents en face-à-face, montre que le choix du prix est indéterminé à l'intérieur d'une plage qui rend l'échange mutuellement avantageux. Domine alors un marchandage dont le résultat dépend de variables étrangères à l'économie (rapport de forces, us et coutumes).

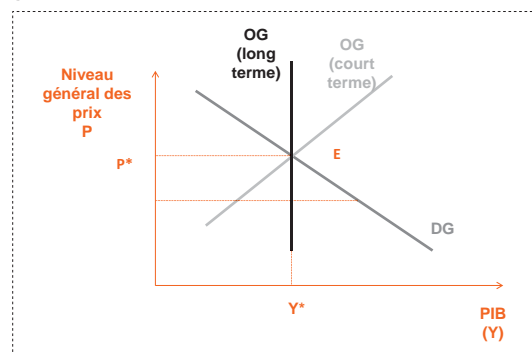
Dans la pratique, force est de constater que les prix ne sont pas déterminés ainsi. Ils sont dans l'ensemble très rigides, ce qui rend les mécanismes d'ajustement supposés entre l'offre et la demande très lents et générateurs de déséquilibres importants dans l'entre-deux. Beaucoup sont fixés par le producteur et déterminés par application

d'un coefficient de marge à ses coûts (« *mark up* »). Pour se rapprocher de cette situation, il faut accepter de sortir de la concurrence pure et parfaite en considérant que certains agents ont un pouvoir de marché et qu'ils ne subissent plus la « loi de l'offre et de la demande » comme une contrainte comparable à la gravitation universelle. On entre alors dans un monde d'interactions stratégiques dans lequel chacun essaie d'être le plus fort mais pour lequel les efforts de modélisation, appelant aux techniques complexes de la « théorie des jeux », montrent qu'il débouche sur des équilibres multiples, donc toujours aussi indéterminés.

### Les incertitudes de l'agrégation des demandes

La facilité lorsque l'on étudie un marché est de supposer que les offres et les demandes peuvent être agrégées en partant des comportements individuels. Mais cette approche en équilibre partiel (un seul marché observé « toutes choses étant égales par ailleurs ») ne prend pas en compte, par hypothèse, les interdépendances entre les marchés. En le faisant, il devient alors impossible de séparer les effets de revenu et de substitution car les variations de prix sur un marché induisent des variations de revenu qui affectent les autres (le prix d'un bien qui baisse sur un marché affecte le revenu de celui qui l'offre ; la demande de cet offreur sur un autre marché en sera donc modifiée, ce qui agira en retour sur le premier marché). La « loi de la demande », qui fait décroître celle-ci lorsque le prix augmente, peut en conséquence ne pas s'appliquer à la courbe de demande agrégée. Ce résultat n'a rien à voir avec un quelconque effet de snobisme (signalement de la qualité par le prix), il a été démontré pour n'importe quel bien par des économistes néo-classiques au début des années 1970 sous le nom de « théorème Sonnenschein-Mantel-Debreu » et stipule que la courbe de demande agrégée peut prendre à peu près n'importe quelle forme. De fait, si

## 2. Courbes d'offre globale (OG) et de demande globale (DG)



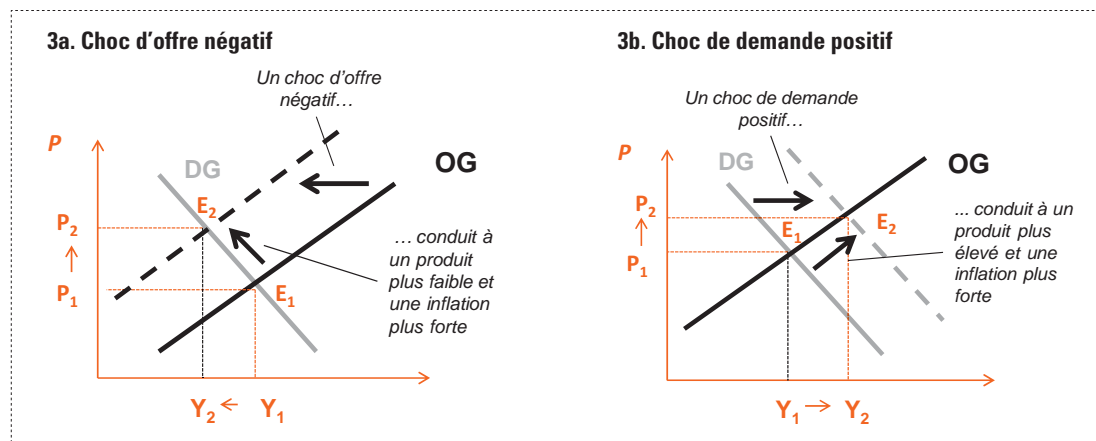
la convergence vers un prix d'équilibre est démontrable en équilibre partiel, elle ne l'est plus pour l'ensemble des marchés.

## L'offre et la demande comme fonctions macroéconomiques

C'est donc dans une autre logique que les concepts d'offre et de demande sont traités au niveau des grandeurs globales d'une économie. On les définit dans le cadre d'un modèle dit « offre globale/demande globale » (OG/DG). Celui-ci est l'outil de base pour étudier les fluctuations à court terme de l'économie (PIB, prix, emploi). Il répond, en le complétant, aux insuffisances du premier modèle de la macroéconomie, le modèle dit IS/LM<sup>2</sup>. L'hypothèse de rigidité des prix de ce dernier n'a pas résisté, en effet, aux chocs inflationnistes des années 1960-1970 qui ont rendu nécessaire de comprendre les interactions entre inflation, emploi et activité économique. Dans ce nouveau modèle, les variations de la demande n'ont pas un effet nécessairement exclusif sur les volumes produits mais peuvent aussi affecter les prix. Le modèle permet aussi d'analyser comment les effets des politiques monétaires et budgétaires se partagent entre la croissance (et le reflux du chômage) et l'inflation. Il peut enfin justifier le passage de politiques centrées sur la demande à des politiques plus orientées vers l'offre.

<sup>[2]</sup> Le modèle introduit par John Hicks en 1937 se polarise sur la demande ; l'hypothèse sous-jacente étant qu'en situation de sous-emploi l'offre globale s'ajuste à la demande globale sans variation des prix. L'offre est donc « endogène » à la demande et infiniment élastique par rapport au prix (sous-emploi et capacités de production inutilisées). Il exprime la double condition d'équilibre de la demande et de l'offre sur le marché des biens et services ( $I = S$ ) et du marché de la monnaie ( $L = M$ ).

### 3. Chocs d'offre et de demande



#### Les courbes de quasi-offre et de quasi-demande

Au carrefour du modèle classique qui ne regarde que l'offre et du modèle keynésien qui ne concerne que la demande, le modèle OG/DG introduit la flexibilité des prix pour cerner l'impact des chocs qui peuvent affecter l'économie à court terme (chocs d'offre comme les chocs pétroliers, ou de demande comme la « Grande Récession » de 2008-2009). Pour ce faire, il va « ajouter » au modèle IS/LM un « bloc offre ». Celui-ci est constitué d'un marché du travail et d'une fonction de production  $Y = f(L)$  et suppose que, à court terme, seule la quantité de travail  $L$  détermine le volume de production. De fait, le modèle ne concerne que les variations de l'offre à court terme.

La représentation graphique du modèle OG/DG se fait dans un plan (PIB  $Y$ , niveau général des prix  $P$ ). La courbe DG se déduit du modèle IS/LM et représente les couples activité/prix qui correspondent à un équilibre sur le marché des biens et sur le marché de la monnaie. Elle est décroissante car la demande sera d'autant plus forte que le niveau des prix est faible. En effet, une baisse des prix stimule la demande des ménages car le pouvoir d'achat de leurs encaisses monétaires et d'une partie de leurs actifs financiers augmente (effet de richesse positif). En économie ouverte, la

baisse des prix stimule en plus la compétitivité-prix des exportations.

La courbe OG décrit, pour chaque niveau de prix, la production que les entreprises sont, dans leur ensemble, disposées à mettre sur le marché après prise en compte du fonctionnement du marché du travail. Comme les entreprises ont tendance à augmenter leur offre lorsque les prix s'élèvent, cette courbe est croissante. En effet, lorsque les prix augmentent, le chiffre d'affaires des entreprises est plus grand. Il suffit alors que le coût du travail progresse moins vite pour que leur profit s'accroisse. En faisant l'hypothèse d'une rigidité des salaires nominaux (souvent fixés pour des périodes assez longues), la hausse des prix conduit à une baisse des salaires réels qui stimule la demande de travail des entreprises et leur permet d'augmenter la production. S'ajoute l'hypothèse que les chefs d'entreprise connaissent mieux l'évolution des prix (ce sont eux qui les fixent en grande partie) que les salariés : l'information est donc asymétrique. Les salariés sont victimes d'une « illusion monétaire » qui leur fait confondre l'augmentation de leur salaire nominal (« fiche de paie ») et celle de leur salaire réel. C'est essentiellement cette différence entre la vitesse d'ajustement de salaires et des prix, liée à la présence des rigidités ou d'erreurs d'anticipation, qui engendre une courbe d'offre croissante.



Dans le cas contraire (prix flexibles sans erreurs d'anticipation), la courbe d'offre globale est une droite verticale qui retrouve les prédictions de l'analyse néoclassique. Le marché du travail est à l'équilibre de plein emploi (tous les facteurs sont utilisés), une augmentation des prix ne peut avoir d'effet sur l'offre. Ce niveau de production correspond au PIB potentiel (ou offre globale de long terme) qui augmente en général au fil du temps sous l'effet de l'accroissement de la population active, des équipements et du progrès technique. Son augmentation n'est plus liée alors à celle du niveau général des prix.

Le point d'équilibre est donné par l'intersection des deux courbes. Pour autant, il n'y a pas de mécanisme automatique qui permettrait un retour à l'équilibre en affichant des prix et en attendant la réponse des offreurs et demandeurs. Qu'une courbe ou l'autre se déplace et ce sont les niveaux de prix et de l'activité qui se modifient. Malgré la similitude formelle avec les courbes associées aux fonctions microéconomiques (*cf. supra*), elles ne procèdent pas d'un calcul maximisateur de la part des agents économiques. D'autre part, alors que, dans le modèle microéconomique, le prix est celui d'un bien donné, dans le modèle OG/DG le prix représente le niveau général des prix de l'économie. Enfin, les courbes d'offre et de demande sont construites conditionnellement à ce qui se passe sur d'autres marchés (marché de la monnaie, marché du travail). Pour cette raison, on qualifie ces courbes de « quasi-offre » et de « quasi demande ».

### Les chocs et les déplacements des courbes d'offre et de demande globale

Les déplacements des courbes d'offre et de demande dans le plan résultent soit de chocs exogènes affectant l'économie de manière spontanée, soit des politiques économiques menées. Les chocs d'offre (exemple : législation environnementale restrictive, revendications salariales, prix du pétrole modifié par un cartel de producteurs) modifient la compéti-

tivité des entreprises et déplacent la courbe OG ; les chocs de demande (exemple : pessimisme des consommateurs, politique de relance) affectent la dépense et déplacent la courbe DG. Le modèle OG/DG permet d'analyser les conséquences de ces chocs sur l'activité économique et sur le niveau général des prix (graphique 3).

Un choc d'offre négatif (graphique 3a) va par exemple déplacer la courbe OG vers la gauche. C'est le cas d'un choc pétrolier qui, parce qu'il réduit la rentabilité des entreprises en élevant leurs coûts de production, les conduit à réduire leur offre globale pour un niveau de prix plus élevé, c'est-à-dire une situation de « stagflation ». En revanche, un choc positif (exemple : contre-choc pétrolier, exploitation du pétrole bitumeux) déplace la courbe vers la droite. Dans les deux cas, on note que les mouvements de la production et des prix sont inversés.

Un choc de demande positif (graphique 3b) déplace la courbe de DG vers la droite. Il entraîne, à la différence d'un choc d'offre, à la fois une inflation et une hausse de la production. C'est le cas d'une politique budgétaire expansionniste mais ce choc pourrait découler aussi d'un relèvement des anticipations de croissance des entreprises ou des ménages. Symétriquement, un choc de demande négatif (exemple : politique budgétaire restrictive à partir de 2010 dans la zone euro) entraîne une baisse simultanée des prix et de la production, c'est-à-dire une « déflation ».

Le graphique permet aussi de comprendre les dilemmes des politiques économiques face à ces chocs. Par exemple, dans le cas d'un choc d'offre négatif, les autorités peuvent choisir de maintenir inchangée la demande globale en misant sur le fait que la baisse de l'activité finira par peser sur les coûts de production avec la montée du chômage et poussera les entreprises à augmenter leur offre. Elles peuvent aussi choisir de stimuler la demande pour relancer l'activité au prix d'une inflation plus importante.

## POUR EN SAVOIR PLUS

---

❖ **ASENSIO A.** (2008), *Le fonctionnement des économies de marché*, Bruxelles, De Boeck.

❖ **BEITONEA., BUISSON E. et DOLLO CH.** (2012), *Économie – Aide-mémoire*, Paris, Sirey, 5<sup>e</sup> éd.

❖ **GÉNÉREUX J.** (2014), *Jacques Généreux explique l'économie à tout le monde*, Paris, Seuil.

❖ **GUESNERIE R.** (2013), *L'économie de marché*, Paris, Le Pommier.

❖ **MANKIW G.** (2010), *Macro-économie*, Bruxelles, De Boeck, 5<sup>e</sup> éd., chap. 9.

❖ **MONTOUSSÉ M. et WACQUET I.** (2013), *100 Fiches pour comprendre la microéconomie et la macroéconomie*, Paris, Bréal.

❖ **WASMER É.** (2014), *Principes de microéconomie*, Paris, Pearson, 2<sup>e</sup> éd., chap.12.

Au cours des Trente Glorieuses, la croissance du PIB s'est imposée comme l'objectif premier des politiques économiques, le plein-emploi, la réduction des inégalités ou l'augmentation du bien-être collectif étant supposés en découler directement. Si ce que certains dénoncent comme un véritable « dogme » est toujours d'actualité, les critiques – formulées dès les années 1970 – à l'égard du productivisme, de l'impact écologique de la croissance et de l'utilisation du PIB comme indicateur principal des performances économiques d'une nation sont de plus en plus nombreuses et ont donné lieu à des réflexions sur l'élaboration d'autres indicateurs de richesse. Que mesure réellement le PIB et quelles sont ses limites ? Thomas Fabre fait le point sur ces questions, avant de présenter les conceptions théoriques de la croissance.

Problèmes économiques

# PIB et croissance économique

© THOMAS FABRE

Professeur en CPGE au lycée Faidherbe, Lille

C'est après la Seconde Guerre mondiale que la croissance économique, perçue et présentée comme une condition nécessaire au plein-emploi et au développement des sociétés, devient un objectif politique prioritaire. Depuis une vingtaine d'années cependant, en raison de la multiplication des crises financières et économiques, des dégâts avérés du productivisme et de l'inaptitude du produit intérieur brut (PIB) à évaluer le bien-être social et écologique, ce que certains économistes appellent le « dogme de la croissance » affronte des critiques de plus en plus vives et partagées, annonçant peut-être une transformation du référentiel dominant des politiques publiques.

Après avoir rappelé la signification et les limites du PIB, nous présenterons les principales conceptions savantes de la croissance, quant à ses origines, à sa forme et à son devenir.

## Signification et limites du PIB

### Qu'est-ce que la croissance ?

La croissance économique désigne, pour un territoire donné, l'augmentation de la production de biens et services sur une longue période. Il y a donc « croissance » lorsque, d'une année sur l'autre et de façon répétée, on constate un accroissement d'un flux de produits (biens de consommation, biens de production, services) dont l'élaboration a donné lieu à une distribution de revenus monétaires, dans le cadre d'une activité déclarée et légale. Pour un territoire donné, ce flux de production est mesuré par le produit intérieur brut (PIB).

D'après l'INSEE, « le PIB est égal à la somme des valeurs ajoutées brutes des différents secteurs institutionnels ou des différentes branches d'activité, augmentée des impôts moins les subventions sur les produits ». Considéré comme un indice de la puissance et de la vitalité économiques d'un territoire

pris dans son ensemble, c'est au nom de l'intérêt général que la plupart des gouvernements font de son augmentation un axe majeur des politiques économiques, comme moyen indirect d'accroître le niveau de vie de la population, de faciliter le financement des services collectifs et de la protection sociale, ou encore de lutter contre le chômage – à condition toutefois que la croissance du PIB soit supérieure à celle de la productivité. Mais le PIB permet-il pour autant d'évaluer la « véritable » richesse d'une société ? Est-il un indice de bien-être social ? En somme, que mesure-t-il vraiment ?

## Que mesure réellement le PIB ?

### Le PIB mesure les activités marchandes

Tout d'abord, le « contenu » en biens et services d'un PIB se transforme au cours du temps. Si la croissance économique est un phénomène historiquement récent – selon Angus Maddison, elle débute réellement au XIX<sup>e</sup> siècle pour « exploser » dans le second XX<sup>e</sup> siècle –, les comparaisons intertemporelles ne considèrent que l'aspect quantitatif de la croissance : on produit beaucoup plus aujourd'hui qu'il y a cent ans. Elles masquent ainsi de nombreux aspects qualitatifs : on produit, en proportion, de plus en plus de services, dont certains n'existaient pratiquement pas il y a cent ans, dans les domaines du transport, de la santé ou de la finance par exemple. En outre, certains biens ou services (on pense en particulier au vaste ensemble des « services à la personne »), au fur et à mesure qu'ils deviennent des marchandises, contribuent à alimenter la croissance. Cette dernière a donc une dimension strictement *sociale*, dans la mesure où elle reflète l'extension de rapports ou d'échanges sociaux faisant l'objet d'une contrepartie monétaire.

Par ailleurs, le PIB ne comptabilise que les biens et services marchands ou non marchands obtenus à l'aide de facteurs de pro-

duction payants. Il ne peut donc pas rendre compte de certaines activités productives qui échappent encore au marché, relevant par exemple des sphères domestique, amicale, ou associative. Encore que cette lacune soit aussi affaire de conventions : au début des années 1970, l'INSEE a par exemple intégré au PIB les services non marchands fournis par les administrations publiques, en les évaluant, faute de prix de vente, au coût de production ; plus récemment, en juin 2014, la France, à la différence de l'Espagne ou de la Belgique, a refusé de se plier à une recommandation européenne invitant à comptabiliser les activités liées à la drogue et à la prostitution dans la richesse nationale.

### Le PIB mesure un flux de production

Ensuite, le PIB ne mesure pas un stock de richesses, mais un flux : il est plus juste de dire qu'il mesure une « richesse nouvellement créée » en une année que simplement la richesse d'un territoire. C'est donc, au sens strict, plus un indicateur d'activités donnant lieu à rémunération que de quantités produites et accumulées – on parle d'approche « non patrimoniale » –, même si le volume de ces activités est nécessairement fonction d'un « capital » (matériel, technique, humain, naturel, etc.) mis en œuvre pour produire. Une des critiques principales adressée au PIB est justement sa non-prise en compte de l'enrichissement (ou de l'appauvrissement) patrimonial d'une société, notamment lorsqu'on inclut la notion de capital naturel ou de capital humain (dont la santé des individus). Une catastrophe nucléaire comme celle de Fukushima, en 2011, constitue un exemple parlant : la mobilisation des habitants et du gouvernement a permis au PIB japonais de repartir à la hausse en 2013, sans que cela donne la moindre idée de l'ampleur et de l'irréversibilité des pertes liées à la pollution des sols, de l'air et de l'océan, ni des dégâts sanitaires et psychiques ayant affecté le bien-être de la population (Fitoussi et Stiglitz, 2011).

## ZOOM

### LE PIB : APPROCHES ET MODES DE CALCUL

Il existe trois modes de calcul du PIB aux « prix du marché » d'un pays ou d'une région.

**Premier cas : l'approche par la production**  
**PIB = somme des valeurs ajoutées + impôts sur les produits – subventions sur les produits.** En sommant des « valeurs ajoutées », on fournit une estimation des flux de production annuels. Les valeurs ajoutées étant évaluées sans les impôts sur les produits comme la TVA, on ajoute ces derniers pour obtenir le PIB aux « prix du marché », puis on soustrait les subventions sur les produits (dont le but est d'influencer les niveaux de production, les prix ou la rémunération des facteurs de production).

**Deuxième cas : l'approche par la demande**  
**PIB = dépenses de consommation finale + FBCF + exportations – importations.** On considère ici le PIB sous l'angle de la « demande globale », c'est-à-dire de la consommation et de l'investissement

(mesuré par la formation brute de capital fixe), en ajoutant le solde de la balance commerciale (cf. Complément).

**Troisième cas : l'approche par les revenus**  
**PIB = rémunération des salariés + EBE + revenus mixtes + impôts sur la production et les importations – subventions.** On additionne ici les revenus primaires (du travail et du capital). Puis, pour obtenir le PIB au « prix du marché », on ajoute le supplément payé du fait des impôts (nets des subventions) sur la production et les importations.

On obtient le PIB « réel » (ou PIB « en volume ») en retranchant du PIB à prix courants (ou PIB « en valeur ») les effets de l'inflation. Si un PIB en valeur croît de 3 % et que l'inflation est de 3 %, la quantité supplémentaire de marchandises produites est globalement nulle : on dit alors que le PIB réel stagne.

Pour mesurer le niveau de vie moyen de la population, on utilise le PIB par habitant qui, par définition, ne croît que si la croissance du PIB est supérieure à celle de la population.

Thomas Fabre

### Le PIB n'est pas un indicateur de bien-être social

Le PIB n'a en effet pas vocation à être un indicateur de bien-être social (ou de « développement »). Nous avons vu qu'il laissait de côté un ensemble d'activités non marchandes nécessaires à la cohésion sociale. De plus, il évalue des « *outputs* », simples enregistrements monétaires de quantités produites, ne disant rien des effets de l'activité productive sur le bien-être (les « *outcomes* »). Les exemples sont désormais fameux : « Une société où il y a beaucoup d'accidents de la route, qui vont exiger des soins médicaux, des réparations de véhicules, des services d'urgence, etc., aura tendance, toutes choses égales par ailleurs, à avoir un PIB plus important qu'une société où les gens conduisent prudemment »,

de même que « la destruction organisée de la forêt amazonienne est une activité qui fait progresser le PIB mondial » (Gadrey et Jany-Catrice, 2005). Les mêmes auteurs, dont les travaux font référence, ajoutent que « la mesure du PIB est tout aussi indifférente à la répartition des richesses comptabilisées, aux inégalités, à la pauvreté, à la sécurité économique, etc., qui sont pourtant presque unanimement considérées comme des dimensions du bien-être à l'échelle d'une société ».

### Un indicateur de richesse partiel et partial

En somme, le PIB est un indicateur partiel et partial de la « richesse » d'un territoire ; partant, sa croissance (c'est-à-dire « la »

croissance) peut recouvrir des situations qualitativement fort disparates, que ce soit en termes d'inégalités sociales, de sécurité, de libertés des individus ou de dégradations de l'environnement. Ces lacunes expliquent la constitution, depuis les années 1980, de nouveaux indicateurs de richesse, mieux à même de rendre compte du bien-être social et de son évolution, tels l'indice de développement humain (IDH), l'épargne nette ajustée, le « PIB vert » ou encore l'indice de santé sociale (cf. Gadrey et Jany-Catrice, 2005). Mais ceux-ci peinent à détrôner le PIB, et plus largement la croissance comme référents premiers des politiques et discours économiques. G.W. Bush, alors président des États-Unis, affirmait en 2002 : « La croissance n'est pas le problème ; c'est la solution. » De nombreux gouvernants continuent de se tourner vers les économistes, afin de trouver les ressorts permettant de continuer à « libérer la croissance » (selon l'expression de J. Attali), ceux-ci n'étant pas toujours d'accord ni sur les moyens ni sur la possibilité même de continuer à croître.

## Origine, forme et devenir de la croissance économique

Que l'on prête ou non allégeance à la croissance du PIB comme objectif premier des politiques publiques, le phénomène historique demeure : depuis les débuts du capitalisme contemporain, soit depuis deux siècles, nos sociétés s'organisent autour de la production toujours plus importante de marchandises. Selon l'exclamation de Karl Marx dans le premier tome du *Capital* (1867), le mot d'ordre en serait : « Accumulez, accumulez ! C'est la loi et les prophètes ! ». Comment, alors, rendre compte de la perpétuation dans le temps de cette accumulation sans cesse renouvelée ? Est-elle vouée à se poursuivre indéfiniment ? De façon régulière ? Telles sont les questions posées par les différentes théories de la croissance (voir tableau).

## Les sources de la croissance

Pour produire, il faut des ressources naturelles (parfois appelées « capital naturel »), du capital physique ou technique (soit, pour aller vite, des machines, des terrains et des bâtiments), et du travail (parfois assimilé à la mise en œuvre d'un « capital humain »). La croissance requiert par conséquent l'utilisation d'une quantité accrue et/ou l'amélioration de l'efficacité productive (ou productivité) de ces « facteurs de production ». Les théories de la croissance se rejoignent d'ailleurs pour voir son origine dans l'augmentation des capitaux productifs et/ou dans le « progrès technique ». La diversité des raffinements et sophistications théoriques est en partie imputable à la priorité accordée à tel ou tel facteur, à tel ou tel acteur économique à même d'en favoriser l'essor.

Les historiens de l'économie, et notamment l'école de la régulation, considèrent la croissance du premier capitalisme comme principalement extensive, c'est-à-dire fondée sur la mobilisation toujours plus importante de capital et de travail, la diffusion du progrès technique jouant au XIX<sup>e</sup> siècle un rôle de déclencheur et d'accompagnateur. Mais la plupart des explications de la croissance ont été développées au XX<sup>e</sup> siècle, lorsque celle-ci s'accélère pour atteindre des taux exceptionnels pendant les Trente Glorieuses (autour de 5 % par an dans les pays occidentaux) : c'est alors le progrès technique – qui comprend au sens large les gains de productivité permis par l'organisation scientifique du travail – qui est mis en avant. Reste à en identifier les sources : le développement scientifique, l'innovation impulsée par les « entrepreneurs » (J.-A. Schumpeter), les rendements d'échelle croissants permis par l'accumulation de capitaux humains et publics (croissance endogène) ou encore l'efficacité des « institutions » formelles et informelles (D. North). Notons que la plupart des théories récentes, même lorsqu'elles s'inscrivent dans la tradition du libéralisme économique, soulignent le rôle primordial des pouvoirs publics, *via* la

## Diverses conceptions de la croissance économique

| Auteurs/théories   | Sources principales de la croissance  | Facteurs limitant la croissance  | Place et rôle des crises économiques   |
|--|---|--|--|
| A. Smith   | Division du travail, progrès technique  | Croissance potentiellement infinie   | Absence de crises économiques majeures, autorégulation supposée des marchés (tradition libérale)               |
| Modèle de Solow  | Progrès technique « exogène »   | Croissance de la population, rythme du progrès technique ; croissance potentiellement infinie  |  |
| Théories de la « croissance endogène »   | Accumulation de capitaux (physique, humain, public), innovations et externalités positives                                | Degré d'internalisation des externalités positives de productivité par l'État ; croissance potentiellement infinie : hypothèse de « croissance durable » |  |
| D. Ricardo   | Libéralisation du commerce international, progrès technique   | Rendements décroissants de la terre ; croissance finie, butant sur un « état stationnaire »  |  |
| « Nouvelle histoire économique » (R. Fogel, D. North)                          | Institutions (contraintes formelles et informelles, allant des lois aux normes et usages), changements techniques         | Inefficacité des institutions d'encadrement du marché ; question de la finitude de la croissance non posée   | Analyse centrée sur les différentiels de croissance à long terme, non sur la question des crises               |
| J.-A. Schumpeter   | « Grappes d'innovations » autour d'un facteur clé, concurrence et progrès technique                                       | Ralentissement des gains de productivité et des taux de profit en fin de cycle long ; question de la finitude de la croissance non posée                 | Alternance de crises et périodes de croissance économique (cycles longs)                                       |
| K. Marx  | Concurrence, accumulation de capital, gains de productivité (via les gains de « plus-value relative »)                    | Baisse tendancielle du taux de profit ; croissance probablement finie via l'épuisement de la terre et des travailleurs                                   | Crises économiques répétées, non-viabilité du système capitaliste à long terme                                 |
| Théorie de la régulation (M. Aglietta, R. Boyer)                               | Accumulation extensive et/ou intensive du capital ; compromis sociaux permettant de trouver des débouchés à la production | Croissance sinon finie du moins « non durable » en raison de l'incompatibilité entre dynamique capitaliste et préservation des ressources naturelles     | Crises régulières des « modes de régulation » (épuisement des gains de productivité, conflits de répartition)  |
| Club de Rome (rapport Meadows), écologie politique, courant de la décroissance | Ressources naturelles transformées en énergie mécanique par le biais d'innovations techniques                             | Croissance finie en raison du caractère fini des ressources naturelles (fossiles en particulier) qui sont à la base de la croissance capitaliste         | « Crise » économique pensée sous l'angle écologique, s'accroissant avec l'épuisement des ressources naturelles |

construction d'infrastructures de transport et de communication, l'éducation, les investissements dans la recherche et le développement, la sécurisation des échanges, etc. La croissance économique apparaît en ce sens redevable d'un cadre institutionnel adéquat<sup>1</sup>. Est-ce cependant suffisant pour assurer une accumulation infinie de marchandises ?

## La croissance peut-elle se poursuivre indéfiniment ?

Si l'on excepte les penseurs qui ne posent pas explicitement la question, le caractère potentiellement fini ou infini de la croissance donne lieu à deux grandes postures, à la fois prédictives et normatives.

### « Durabilité faible » contre « durabilité forte »

D'une part, la plupart des économistes libéraux<sup>2</sup> font l'hypothèse que la croissance et le développement peuvent être « durables », autrement dit se poursuivre sans entamer les capacités de bien-être des générations futures, misant sur les progrès techniques et les gains d'efficacité énergétique à venir. Cela implique de considérer capital naturel et capital artificiel comme substituables (hypothèse dite de la « durabilité faible », formulée par J.-M. Hartwick en 1977).

Au contraire, les partisans de l'écologie politique, de la « décroissance » ou les théoriciens de la régulation pensent que la croissance propre aux sociétés capitalistes est en train de buter sur des limites naturelles et humaines, sur l'épuisement de ressources non renouvelables auquel les avancées techniques ne sauraient trouver de substituts suffisants ; et ce d'autant que les ressources naturelles renouvelables et l'environnement en général connaissent des dégradations en partie irréversibles, qui altèrent les capacités de production futures<sup>3</sup>. Autrement dit, le développement durable (soit la poursuite de la croissance moyennant un « authentique »

respect de la nature et des hommes) serait incompatible avec le capitalisme comme système fondé sur la recherche incessante de nouveaux moyens d'accumulation du profit<sup>4</sup>.

### « Durabilité » et crises du capitalisme

On ne sera alors pas forcément surpris de constater que les économistes libéraux partisans de la « durabilité faible » ne prêtent pas une grande attention aux crises majeures : ils postulent non seulement la toute-puissance de l'innovation technologique, mais aussi l'autorégulation des marchés.

Opposées aux conceptions libérales, les théories issues de la tradition de critique du capitalisme et du productivisme affirment en revanche que les crises majeures sont « normales » et ont partie liée, depuis les chocs pétroliers des années 1970, avec la raréfaction des ressources naturelles. Leur donner raison conduit à souligner derechef les carences du PIB mondial comme *lanceur d'alerte*, puisqu'il n'a depuis lors pratiquement pas cessé de croître, même à un rythme toujours faiblissant. Nombreux sont d'ailleurs ceux qui, aujourd'hui, estiment que le PIB a, au plan de la légitimité scientifique ou politique, largement fait son temps.

## Conclusion : « Au-delà du PIB »<sup>5</sup> ?

Toute comptabilité nationale repose sur des conventions sociopolitiques : il s'agit de définir des agrégats (ce qui ne va déjà pas de soi), de choisir des instruments et des unités de mesure, des critères de démarcation, etc. Lorsqu'il s'agit de définir la « richesse nationale », on devine la multiplicité des voies à emprunter. Étudiant le cas français, François Fourquet montre en 1980 comment les « comptes de la puissance », au sortir de la Seconde Guerre mondiale, préjugent de la supériorité de systèmes économiques fondés sur la « juste et libre concurrence », la production industrielle de masse et l'extension des rapports sociaux marchands<sup>6</sup>. Le PIB,

[1] Sans compter ses aspects « culturels », déjà mis en avant par Max Weber dans *L'Éthique protestante et l'esprit du capitalisme* (1905).

[2] À l'exception notable de David Ricardo qui anticipait l'avènement d'un « état stationnaire » en raison de la fertilité décroissante des terres cultivées.

[3] H. Daly définit la « durabilité forte » comme le développement n'altérant pas le stock existant de capital naturel. Pour des auteurs comme Dennis et Donella Meadows, tout porte à croire que cette notion n'a de sens que dans une société « postcroissance ».

[4] Dans le documentaire de Marie-Monique Robin « Sacrée croissance ! » (Arte, 2014), Jean Gadrey rappelle la célèbre boutade de Kenneth Boulding : « Celui qui pense qu'une croissance exponentielle infinie est possible dans un monde fini est soit un fou, soit un économiste. » Voir aussi Rousseau S. et Zuideau B. (2007), « Théorie de la régulation et développement durable », *Revue de la régulation*, n° 1.

[5] En référence à l'ouvrage éponyme de Dominique Méda (2008).

[6] Fourquet F. (1980), *Les Comptes de la puissance. Histoire de la comptabilité nationale et du plan*, Paris, Éditions Recherches.



comme indice de richesse, et sa croissance, comme priorité des politiques économiques, sont alors au faite de leur légitimité. Dans un contexte de reconstruction, puis de guerre froide, alors que l'État-providence s'est généralisé, interroger les insuffisances et la légitimité du PIB ou dénoncer les nuisances environnementales accompagnant la croissance ne fait pas recette, si l'on peut dire.

Avertissements et remises en cause ne manquent pourtant pas, dès les années 1970<sup>7</sup>, mais leur audience demeure limitée et éphémère. C'est que définir le bien-être collectif et ses enveloppes langagières est aussi un enjeu de luttes sociales et politiques<sup>8</sup>. Nous l'avons suggéré, ne parler que de croissance – et implicitement, ne faire référence qu'au PIB – peut être un moyen de ne rien dire des

contenus et des externalités économiques associés au phénomène, dont on oublie ou ignore parfois le caractère statistiquement construit. Sans compter l'inertie certaine des représentations économiques : « À court et moyen termes, [la croissance] est donc appréciée positivement par de larges fractions de la population, d'autant plus larges d'ailleurs qu'on ne leur communique qu'un chiffre final (le taux de croissance réalisé ou espéré), en esquivant la question de son contenu (qu'est-ce qui s'est amélioré ?) ou de sa répartition (le partage de la valeur ajoutée) » (Gadrey et Jany-Catrice, 2005). Aujourd'hui, l'idée perdue mais chancelle selon laquelle, sans la croissance, point de salut.

<sup>[7]</sup> Par exemple, le rapport du Club de Rome, (ou « rapport Meadows »), *Halte à la croissance !*, est publié en 1972.

<sup>[8]</sup> Sur ces questions, voir Pessis C., Topçu S. et Bonneuil C. (dir.) (2013), *Une autre histoire des « Trente Glorieuses »*. Modernisation, contestations et pollutions dans la France d'après-guerre, Paris, La Découverte.

## POUR EN SAVOIR PLUS

❖ **FABRE T.** (2012), *L'Essentiel pour comprendre l'économie*, chapitre « Mesurer la richesse », Paris, Ellipses.

❖ **FITOUSSI J.-P. et STIGLITZ J.** (2011), « Nouvelles réflexions sur la mesure du progrès social et du bien-être », in *Économie du développement soutenable*, *Revue de l'OFCE*, n° 120.

❖ **GADREY J., et JANY-CATRICE F.** (2005), *Les Nouveaux Indicateurs de richesse*, Paris, La Découverte, coll. « Repères ».

❖ **MÉDA D.** (2008), *Au-delà du PIB. Pour une autre mesure de la richesse*, Paris, Flammarion, coll. « Champs actuel ».

❖ **STIGLITZ J., SEN A. et FITOUSSI J.-P.** (2009), *Rapport de la Commission sur la mesure des performances économiques et du progrès social*.

❖ [Http://www.stiglitz-sen-fitoussi.fr/documents/rapport\\_francais.pdf](http://www.stiglitz-sen-fitoussi.fr/documents/rapport_francais.pdf)

## ↓ COMPLÉMENT

### La balance des paiements

La balance des paiements est un document statistique (élaboré, en France, par la Banque de France) qui retrace sous une forme comptable l'ensemble des flux d'actifs réels, financiers et monétaires entre les résidents d'une économie et les non-résidents au cours d'une période déterminée. Les flux

économiques et financiers à l'origine de ces opérations sont répartis en distinguant le compte des transactions courantes (échanges de biens, de services, de revenus et de transferts courants), le compte de capital (transferts en capital) et le compte financier (investissements directs ou de portefeuille). C'est un document complexe dont la lecture et l'interprétation requièrent quelques clés<sup>(1)</sup>. Sous une forme un peu simplifiée, la balance se présente comme suit.

## 1. Les différents comptes de la balance des paiements

|                                      | Crédits  | Débits   | Soldes |
|--------------------------------------|--|--|--------|
| 1. Compte des transactions courantes | Exportations de biens ou de services, salaires et revenus du capital et transferts de revenus (des non-résidents vers les résidents)   | Importations de biens ou de services, salaires et revenus du capital et transferts de revenus (des résidents vers les non-résidents)   | S1     |
| 2. Compte de capital                 | Remises de dettes (des non-résidents vers les résidents), ventes de brevets, marques...  | Remises de dettes (des résidents vers les non-résidents), achats de brevets, marques...  | S2     |
| 3. Compte financier                  | Investissements directs et investissements de portefeuille (vente de titres par des résidents à des non-résidents), avoirs <sup>(1)</sup> remboursés par des non-résidents à des résidents ou engagements <sup>(2)</sup> accordés par des non-résidents à des résidents (tels les crédits commerciaux ou les prêts bancaires), vente de produits financiers dérivés par des résidents à des non-résidents, diminution des avoirs de réserve (or monétaire, droits de tirages spéciaux sur le FMI, devises) | Investissements directs et investissements de portefeuille (vente de titres par des non-résidents à des résidents), engagements remboursés par des non-résidents à des résidents ou avoirs accordés par des non-résidents à des résidents (tels les crédits commerciaux ou les prêts bancaires), achat de produits financiers dérivés par des résidents à des non-résidents, augmentation des avoirs de réserve (or monétaire, droits de tirages spéciaux sur le FMI, devises) | S3     |
| 4. Erreurs et omissions nettes       |  |  |        |

(1) Les avoirs financiers sont des titres émis par des non-résidents détenus par des résidents.

(2) Les engagements financiers sont des titres émis par des résidents détenus par des non-résidents.

## Au plan comptable, la balance est toujours équilibrée

Chaque flux d'actifs fait l'objet d'un double enregistrement, au crédit et au débit : le solde de la balance des paiements est, par construction, nul (aux « erreurs et omissions nettes » près, une transaction ne faisant pas toujours l'objet d'un enregistrement simultané au débit et au crédit) ; autrement dit, la balance est nécessairement équilibrée ( $S1 + S2 + S3 +$  « erreurs et omissions nettes » = 0).

Par convention, on note au **crédit** toutes les opérations assimilables à une vente ou entraînant une diminution des avoirs ou une augmentation des engagements (exportation de biens ou de services, revenus et transferts reçus, emprunts auprès des

non-résidents, etc.) ; et au **débit**, toutes les opérations assimilables à un achat ou entraînant une augmentation des avoirs ou une diminution des engagements (importations de biens et services, revenus et transferts versés, prêts accordés à des non-résidents, etc.).

Par exemple, si une entreprise non résidente achète un bien à une entreprise résidente pour une valeur de 1000 euros, payés au comptant et versés sur le compte du résident (dont la banque est résidente), cela donne lieu à l'enregistrement suivant :

- un crédit de 1 000 euros au compte des transactions courantes (opération assimilable à une diminution des avoirs en marchandises) ;
- un débit de 1 000 euros au compte financier (opération assimilable à une diminution des engagements en devises étrangères).

## 2. Balance des paiements de la France (en milliards d'euros)

|   | 2012          | 2013          |
|---|---------------|---------------|
| <b>1. Compte des transactions courantes</b> | <b>- 31,8</b> | <b>- 30,3</b> |
| Biens (balance commerciale)                 | - 54,6        | - 42,5        |
| Services                                    | 24,7          | 18,3          |
| Revenus et transferts                       | - 1,9         | - 6,0         |
| <b>2. Compte de capital</b>                 | <b>0,5</b>    | <b>1,8</b>    |
| <b>3. Compte financier</b>                  | <b>21,0</b>   | <b>14,2</b>   |
| Investissements directs                     | - 14,1        | 5,1           |
| Investissements de portefeuille             | 26,5          | 69,8          |
| Dérivés financiers                          | 14,3          | 16,8          |
| Autres investissements                      | - 1,7         | - 79,0        |
| Avoirs de réserve                           | - 4,0         | 1,5           |
| <b>4. Erreurs et omissions</b>              | <b>10,3</b>   | <b>14,3</b>   |

Source : Banque de France ; Direction générale des statistiques.

Toute inscription au crédit dans les comptes des transactions courantes et de capital doit entraîner une inscription égale au débit du compte financier.

Si la balance des paiements est équilibrée par construction, les différents comptes qui la composent peuvent révéler des déséquilibres donnant des indications sur certaines forces ou faiblesses d'une économie nationale, et orienter les choix de politiques économiques.

La somme des soldes de transactions courantes et de capital, si elle est positive, révèle une **capacité de financement** de la nation : l'épargne intérieure excède l'investissement et permet de couvrir le déficit d'épargne des non-résidents et/ou d'accroître les réserves de changes (opérations qui se traduisent par un déficit du compte financier) ; inversement, une somme négative révèle un **besoin de financement de la nation**.

En France, en 2012 et 2013, les résidents ont globalement un **besoin de financement**, le solde des transactions courantes étant nettement déficitaire : c'est là une faiblesse récurrente de l'économie française. On remarque néanmoins que cette faiblesse est d'abord imputable au **déficit de la balance commerciale** (ligne « biens »), lui-même lié aux **problèmes de compétitivité** rencontrés par l'industrie nationale. Et que, au contraire,

au plan des services, la France dégage un excédent et continue de posséder un point fort.

Concernant le compte financier, on constate en 2012 et 2013 un excédent pour les investissements de portefeuille et les produits financiers dérivés, traduisant des entrées nettes d'épargne de la part des non-résidents en contrepartie du besoin de financement déjà évoqué. Les « autres investissements » voient leur solde devenir nettement négatif en 2013 : les prêts et crédits commerciaux accordés par des résidents à des non-résidents ont nettement dépassé ceux réalisés en sens inverse.

Enfin, on appelle « **solde de la balance globale** » la somme de tous les soldes, moins les avoirs de réserve : lorsqu'il est excédentaire (comme en 2012), cela signifie que la banque centrale a accumulé des réserves de change durant l'année, ce qui se traduit par une augmentation de la masse monétaire ; et inversement, lorsqu'il est déficitaire (comme en 2013). Attention : un solde excédentaire se traduit par un signe négatif sur la ligne « avoirs de réserves » ; un solde déficitaire, par un signe positif.

**Thomas Fabre**

(1) Pour une présentation plus détaillée, voir Chamblay D. (2008), « La balance des paiements », *Cahiers français* n° 345, Paris, La Documentation française, juillet-août.

Les biens collectifs et les externalités sont les deux cas les plus célèbres de défaillance du marché. S'ils appellent l'action publique, celle-ci peut revêtir plusieurs modalités, de la plus interventionniste – la réglementation – à la plus respectueuse des mécanismes de marché – l'attribution de droits de propriété, en passant par l'instauration de taxes et de subventions. Marion Navarro fait le point sur ces configurations sous-optimales en termes de bien-être collectif et sur les moyens d'y faire face en s'appuyant sur l'exemple des problèmes d'environnement.

Problèmes économiques

# Biens collectifs et externalités

L'accroissement des émissions de gaz à effet de serre (GES) depuis la révolution industrielle conduit à une augmentation de la température moyenne à la surface du globe dont les conséquences pourraient être catastrophiques si la tendance se poursuivait. Les scientifiques réunis au sein du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) considèrent qu'il est impératif de limiter cette hausse de température à 2 °C par rapport à la température moyenne de la période antérieure à la révolution industrielle. Il est aussi urgent de protéger certaines ressources naturelles dont beaucoup sont menacées. Notre modèle actuel de croissance n'est pas soutenable et il est nécessaire que les agents économiques modifient leur comportement et tiennent compte des effets environnementaux de leurs actions. La préservation de l'environnement est nécessaire à double titre : d'un point de vue purement économique, le capital naturel est indispensable à la croissance, et du point de vue du bien-être de la population, la qualité de l'environnement est un paramètre essentiel de la qualité de vie. Comment comprendre alors le fait que le comportement spontané des acteurs conduise à une telle dégradation environnementale ? Que faire pour réorien-

© MARION NAVARRO

Professeur de sciences économiques et sociales en CPGE B/L  
au lycée du Parc à Lyon

ter le comportement des acteurs ? Les économistes analysent cette situation en termes de défaillance du marché. Compte tenu des caractéristiques des ressources naturelles et de la façon dont les acteurs déterminent leur comportement, le fonctionnement non régulé des marchés conduit à une allocation des ressources non optimale, c'est-à-dire que l'affectation des ressources disponibles à certains usages n'est pas la plus efficace possible et ne maximise pas le bien-être de la population. Nous nous intéresserons ici à deux cas classiques de défaillance du marché : le cas des biens collectifs et celui des externalités.

## Des biens collectifs en quantité insuffisante

Un bien collectif est un bien **non rival** et **non excluable**. Un bien est dit rival si sa consommation prive les autres de son usage.

Le fait qu'un individu profite d'un service de nettoyage d'une plage publique n'empêche pas un autre agent d'en bénéficier simultanément. Ce service est non rival, ce qui n'est pas le cas d'un service de coiffure : deux personnes ne peuvent simultanément bénéficier du travail d'un même coiffeur. Un bien est excluable si le producteur peut en réserver l'accès à ceux qui en ont payé le prix. Le service de nettoyage d'une plage publique est aussi non excluable. Une fois que la plage a été nettoyée, on ne peut empêcher certains individus de profiter d'une plage propre puisqu'elle est publique et accessible à tous. Le service de nettoyage d'une plage privée est à l'inverse excluable car il est possible de refuser l'entrée de la plage aux personnes qui refuseraient de payer le prix imposé par le propriétaire.

Le marché est défaillant dans la gestion des biens collectifs : trop peu de ressources sont mobilisées pour les produire alors même qu'ils ont de l'intérêt aux yeux de la population. Comment expliquer cela ?

Peu de producteurs vont se lancer dans la production d'un bien non excluable pour ensuite chercher à le vendre. Tout le monde pouvant en bénéficier librement, le producteur n'a aucune garantie de tirer ensuite un revenu lui permettant de couvrir ses coûts. Prenons le cas des feux d'artifice qui sont un exemple de bien collectif. Aucun artificier ne va lancer de lui-même un feu d'artifice et passer ensuite dans la foule des spectateurs pour essayer de tirer une rémunération du service dont ils ont bénéficié. Les consommateurs n'ont pas non plus intérêt à payer à l'avance pour que le bien soit produit car ces biens sont généralement très coûteux à produire et il est très peu probable qu'un consommateur ait une disposition individuelle à payer supérieure ou égale au prix de vente du bien. Dans certains cas, par exemple lors d'un grand mariage, un individu va accepter de payer pour qu'un feu d'artifice ait lieu car le but est justement d'en faire profiter tout le monde gratuitement, mais la plupart des feux d'artifice sont financés par les collectivités locales.

## ZOOM

### LA TRAGÉDIE DES BIENS COMMUNS

Les biens communs sont des biens non excluables mais rivaux. De nombreuses ressources naturelles sont des biens communs comme c'est le cas par exemple des ressources halieutiques. Le fonctionnement du marché sans intervention de l'État conduit à une surexploitation de ses ressources car elles sont accessibles à faible coût (un pêcheur n'a qu'à acheter son équipement, il peut ensuite pêcher autant de poissons qu'il veut gratuitement) alors même qu'elles ont une valeur marchande (particulièrement une fois transformées). Les individus ont donc tendance à surconsommer ce type de

bien et ces derniers étant rivaux, à chaque fois qu'un individu en consomme, le stock de biens disponibles diminue. On risque alors d'observer la disparition de certaines ressources, c'est ce que Garrett Hardin<sup>(1)</sup> appelle la tragédie des biens communs. L'intervention de l'État est souvent nécessaire pour limiter les prélèvements (exemple des quotas de pêche) mais les travaux d'Elinor Ostrom<sup>(2)</sup> ont montré que les acteurs privés pouvaient éviter cette tragédie en mettant en œuvre des stratégies coopératives, ce qui nécessite un capital social et institutionnel important.

**Marion Navarro**

(1) Hardin G. (1968), « The Tragedy of the Commons », *Science*, vol. 162, n° 3859.

(2) Ostrom E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Actions*, Cambridge University Press.

Sans intervention publique, les biens collectifs sont ainsi produits en quantité insuffisante et le bien-être de la population n'est pas maximal. Considérons le cas d'individus qui seraient prêts à payer individuellement 20 euros pour que la plage publique où ils se baignent habituellement soit nettoyée. Le plaisir qu'ils gagneront à être sur une plage plus propre vaut ainsi 20 euros à leurs yeux : ils sont prêts à renoncer à 20 euros d'autres biens (comme l'achat d'un T-shirt) pour que ce service soit produit. Supposons que 1 000 personnes soient concernées, ce qui fait que collectivement, la disposition à payer des utilisateurs de la plage serait de 20 000 euros. Supposons que le coût de nettoyage de la plage soit de 15 000 euros. Il serait logique que ce service soit produit puisque cela accroîtrait la satisfaction des utilisateurs de la plage. Aucune entreprise ne va toutefois nettoyer la plage, ces dernières sachant que les riverains ne paieront pas ensuite pour ce service, l'accès à la plage étant gratuit. Aucun usager n'étant prêt à payer 15 000 euros de sa poche pour que ce service soit produit, le plus probable est donc que la plage reste sale sans intervention publique.

De nombreux services qui contribuent à la préservation de l'environnement sont des biens collectifs. Ainsi en est-il de la politique climatique. En effet, il s'agit d'un bien non rival (le fait qu'un individu bénéficie de la stabilité du climat n'empêche pas un autre d'en bénéficier) et d'un bien non excluable (il n'est pas possible d'exclure un individu ou un pays de l'avantage social que constitue la réduction des émissions de GES par un autre individu ou un autre pays). Par conséquent, personne ne va produire spontanément cette politique climatique. Dans la plupart des cas, la production des biens collectifs est prise en charge par une autorité publique (soit directement, soit à travers une délégation de service public) et financée par les prélèvements obligatoires. Pour la politique climatique, le cas est plus complexe car comme il n'existe pas de gouvernement mondial, il n'est pas

possible d'imposer par la contrainte réglementaire la mise en œuvre de cette politique et seule la coopération entre États peut conduire à une action concertée de réduction des émissions de GES.

## La défaillance du marché en présence d'externalités

Une externalité représente la conséquence du comportement d'un agent économique sur d'autres agents économiques qui ne fait pas l'objet d'une transaction marchande. Les externalités peuvent être positives, si l'effet sur les autres agents conduit à améliorer leur situation, ou négatives si au contraire il la dégrade.

Pour déterminer leur comportement, les agents économiques comparent les coûts qu'ils subissent et les avantages qu'ils retirent de chaque possibilité envisagée. Ainsi, lorsqu'un producteur de voitures se demande s'il doit en produire une de plus, il compare le coût de production de cette unité supplémentaire (coût marginal privé) à la recette qu'il pourra retirer de sa vente (rendement marginal privé) et décidera de la produire si le coût est inférieur à la recette. Il ne tient pas compte du fait qu'en produisant cette voiture, il va émettre des produits polluants qui auront un coût sur d'autres agents économiques (coût marginal externe).

Le coût total d'une activité, qu'on appelle coût social, représente la somme du coût privé de réalisation de cette activité pour l'agent économique qui en est à l'origine et du coût externe. Le rendement total d'une activité, qu'on appelle rendement social, représente la somme du rendement privé et du rendement externe. En présence d'externalités négatives, il existe un coût externe positif et les agents économiques vont sous-estimer le coût de leur action dans leur prise de décision. À l'inverse, en présence d'externalités positives, il existe un rendement externe positif et les agents économiques sous-estiment le rendement de leur action. Trop de res-

sources vont ainsi être affectées à la réalisation d'activités qui génèrent des externalités négatives et trop peu pour des activités qui génèrent des externalités positives. L'allocation des ressources par le marché n'est donc pas optimale. Les ressources disponibles ne sont pas utilisées d'une façon qui maximise le bien-être de la population car les effets externes ne font pas l'objet de transactions marchandes. Les signaux envoyés par les prix sont faussés car ils ne reflètent plus le rendement ou le coût véritable des différentes activités. De nombreux problèmes environnementaux peuvent être compris en mobilisant cette analyse. Les émissions de gaz à effet de serre (GES) sont ainsi une externalité négative générée par la réalisation de nombreuses activités. Les acteurs n'intègrent pas dans leur calcul les véritables coûts de leurs actions, ce qui conduit à une émission de GES supérieure à celle qui serait observée s'ils en tenaient compte.

## Quelle régulation publique pour internaliser les externalités ?

Deux catégories d'instruments sont mobilisables pour modifier les choix individuels en matière de production et de consommation :

– les instruments non économiques, qui renvoient aux règles qui encadrent le comportement des acteurs. La réglementation repose sur l'autorité hiérarchique des administrations publiques, qui peuvent contraindre les agents économiques à adopter ou pas tel ou tel comportement ;

– des instruments économiques qui ont recours à un signal-prix pour inciter les agents à modifier leurs comportements. L'idée est de faire en sorte que des « signaux-prix » permettent aux individus de prendre en compte dans leur calcul économique le coût et le rendement sociaux – et non uniquement privés – de leurs décisions. Il s'agit des systèmes de taxes et subventions, d'une part, et des marchés de quotas, d'autre part.

## La réglementation

Le recours à la réglementation consiste à ce qu'un pouvoir hiérarchique définisse des normes qui s'imposent à tous les agents pour ensuite en contrôler l'application. On emploie souvent l'expression anglo-saxonne de *command and control* pour désigner ce type de politique. La réglementation peut consister à interdire l'utilisation de certains biens au cours du processus de production, limiter la quantité d'émissions polluantes autorisées, etc. Cet instrument semble efficace pour modifier le comportement des acteurs puisqu'en interdisant ou limitant par la réglementation certains comportements qui génèrent des externalités, l'intervention de la puissance publique conduit mécaniquement à une modification de leurs choix. Pour la gestion des externalités positives, l'État peut à l'inverse forcer certains individus à réaliser certaines activités. Le fait de se former est une activité qui génère des externalités positives et l'État contraint ainsi les individus à aller à l'école jusqu'à 16 ans (les motivations de cette décision sont plus larges mais cette décision peut être en partie interprétée en ces termes).

Cet instrument présente néanmoins de nombreux défauts. Prenons le cas des émissions polluantes. La puissance publique peut définir le niveau d'émissions qui serait socialement optimal et imposer des quotas d'émission ensuite aux acteurs privés de telle sorte que la somme des quotas individuels corresponde au volume global d'émissions désiré. Les entreprises devront donc réduire leurs émissions en transformant leurs méthodes de production de façon à moins polluer et cela induira des coûts. Ces coûts ne sont pas les mêmes pour toutes les entreprises et en leur demandant un effort uniforme, l'objectif de réduction des émissions n'est pas atteint de la façon la plus efficace possible. Il aurait été économiquement rationnel de demander aux acteurs pour lesquels les coûts de dépollution sont plus faibles de réduire relativement plus leurs émissions. La puis-

sance publique ne connaissant pas les coûts de réduction des émissions des différents acteurs, elle ne peut définir au niveau individuel les quotas d'une façon appropriée. Il est par ailleurs très difficile d'évaluer le niveau auquel l'État doit fixer le niveau global d'émission car il ne connaît pas la valeur précise du coût externe des émissions. Ce type d'instrument est toutefois adapté pour les émissions polluantes dangereuses aux effets locaux importants où une réglementation stricte est nécessaire car il importe que tous les acteurs, quels qu'ils soient et où qu'ils se situent, respectent ces normes.

## Les systèmes de taxes et subventions

Une alternative à la régulation par des normes consiste à agir sur les prix. En taxant les acteurs responsables d'externalités négatives ou en subventionnant ceux à l'origine d'externalités positives, les agents sont conduits à tenir compte des effets externes de leur action. Une taxe égale au coût marginal externe (taxe dite pigouvienne) et une subvention égale au rendement marginal externe conduiraient à faire converger coût social et coût privé, ainsi que rendement social et rendement privé. Prenons l'exemple simple suivant : supposons qu'en produisant une voiture supplémentaire, un constructeur automobile génère à chaque fois une externalité négative d'une valeur de 20 euros. En imposant une taxe de 20 euros par voiture produite, le producteur automobile va intégrer dans son calcul la valeur des externalités que génère son activité de production. S'il produit 10 000 voitures, il aura à supporter le coût de production de ces 10 000 voitures plus 200 000 euros de taxes ( $20 \times 10\,000$ ).

L'avantage de cet outil par rapport à un système de quotas uniformes est qu'il permet d'atteindre le niveau de pollution souhaité à moindre coût. En effet, si nous reprenons l'exemple précédent, nous voyons facilement que les producteurs d'automobiles pour lesquels le coût de réduction des émissions est

inférieur à la taxe préféreront dépolluer plutôt que de payer la taxe alors qu'à l'inverse les producteurs pour lesquels le coût de dépollution est supérieur ne réduiront pas leurs émissions et préféreront payer. L'objectif de réduction des émissions est atteint à moindre coût, l'effort de réduction n'étant pas réparti de façon uniforme. En imposant un tel système de taxe et subvention, les décisions privées des acteurs conduiraient théoriquement à atteindre un optimum social. En pratique, cela est plus complexe car l'État ne connaît qu'imparfaitement la valeur du coût marginal externe et du rendement marginal externe des différentes activités qui génèrent des effets externes et ne peut parfaitement calibrer les taxes et subventions qu'il met en place. Par ailleurs, si les incitations qu'ont les acteurs à adopter tel ou tel comportement sont modifiées avec la mise en place d'un tel système, leurs choix vont évoluer s'il leur est possible d'adopter un comportement alternatif. Un individu qui habiterait dans une zone géographique reculée où aucun système de transport en commun n'existe serait moins incité à utiliser sa voiture si une taxe sur les émissions polluantes était mise en œuvre mais il continuerait à l'utiliser de la même façon s'il ne peut faire autrement.

## Les marchés de quotas

### *Rétablir les droits de propriété pour rétablir l'efficacité du marché*

Le problème des externalités peut être vu comme découlant d'une mauvaise définition des droits de propriété. Si l'effet externe ne fait pas l'objet d'une transaction marchande, c'est qu'il n'existe pas un droit sur lequel l'acteur qui subit l'effet externe pourrait s'appuyer pour exiger un dédommagement financier pour le préjudice subi. Si un agent économique possède un terrain à proximité duquel il habite et qu'une entreprise veut y déposer des déchets malodorants, cette dernière devra l'indemniser et l'acteur n'acceptera que s'il estime que la somme proposée compense le préjudice subi. À l'inverse, si le terrain n'ap-



partient à personne et qu'il n'y a pas de réglementation publique sur le rejet des déchets, l'entreprise pourra y déposer librement ses déchets et la personne qui habite à proximité ne pourra demander aucune compensation. Ronald Coase proposait ainsi de résoudre le problème des externalités en redéfinissant les droits de propriété et en laissant ensuite les acteurs négocier directement entre eux des arrangements. Le théorème de Coase stipule ainsi que si les coûts de transaction sont nuls (c'est-à-dire s'il n'existe pas de coûts à la négociation, à la conclusion et à la surveillance de l'exécution du contrat, etc.), les agents sont capables de trouver une solution efficace pour internaliser les externalités, sans que l'intervention de l'État soit nécessaire. Ce résultat est valable quelle que soit la façon dont sont répartis initialement les droits de propriété. Reprenons l'exemple précédent. Si la propriété du terrain est attribuée à l'entreprise, le riverain a intérêt à engager une négociation avec la firme afin de trouver un accord dans lequel l'entreprise réduit ses déchets en échange d'une compensation financière. Le riverain sera prêt à payer pour la réduction d'une unité de déchet tant que le bénéfice qu'il retire de la moindre présence de déchet est supérieur à l'indemnité qu'il doit verser à l'entreprise pour qu'elle émette une unité de déchets en moins. Cette dernière va accepter de réduire ses déchets d'une unité tant que la somme versée pour les réduire d'une unité est supérieure au coût marginal de réduction des déchets. Ils vont ainsi négocier un accord tel que le niveau optimal de déchets sur le terrain est atteint. Le marché n'est donc pas défaillant dans ce cas de figure car les droits de propriété ont été bien définis au départ. En fonction de la façon dont les droits de propriété sont répartis initialement, la répartition des revenus entre le riverain et l'entreprise ne sera pas la même (puisque dans le scénario où le terrain appartient au riverain, c'est l'entreprise qui verse une indemnité alors que c'est l'inverse si le terrain appartient à l'entreprise) mais la quantité de déchets présents sur le terrain sera, elle, identique.

### *L'institutionnalisation des marchés de quotas par la puissance publique*

Dans la réalité, les coûts de transaction sont souvent très élevés, notamment quand les effets externes touchent de nombreux individus. Prenons le cas d'une entreprise qui possède un terrain sur lequel elle rejette des déchets toxiques, ce qui a pour conséquence de polluer l'eau que consomment des milliers d'individus situés dans des zones géographiques variées. Il va être très complexe pour ces derniers de négocier avec l'entreprise. C'est pour ces raisons que, en s'inspirant des travaux de R. Coase, certains économistes ont pensé à institutionnaliser la négociation entre les acteurs en créant des marchés auxquels les individus sont obligés d'avoir recours. Les marchés sont des institutions qui n'émergent pas spontanément et sur lesquels s'échangent des titres de propriété. L'État peut ainsi imposer à tous les émetteurs de GES de détenir des titres de propriété correspondant à des quotas d'émission qui limitent la quantité que chaque agent est autorisé à émettre. Ces titres s'échangent sur des marchés spécifiques mis en place par la puissance publique. Supposons que celle-ci donne initialement à certains acteurs des droits (peu importe à qui elle les attribue au départ), chaque droit permettant d'émettre une tonne de GES. Les acteurs détenant des droits et qui émettent des émissions polluantes vont chercher à vendre leurs droits si le coût de réduction des émissions d'une tonne de GES est inférieur au prix du titre sur le marché et à l'inverse les acteurs qui génèrent des émissions mais qui n'ont pas les droits suffisants vont préférer acheter des droits plutôt que de réduire leurs émissions si le coût de la réduction excède le coût d'achat des droits. Du point de vue des personnes qui subissent les préjudices liés à l'émission de GES, ces derniers vont acheter des droits sans intention d'émettre de GES mais pour faire diminuer le niveau global d'émission, si la satisfaction qui découle de la baisse de la pollution excède le coût d'achat des

titres. À l'inverse, les acteurs qui détiennent des titres mais qui ne polluent pas vont les vendre si la satisfaction liée au fait d'avoir un revenu supplémentaire en vendant les titres excède la perte de plaisir liée à l'existence d'une pollution plus forte. Le prix du quota sur le marché va dépendre de l'offre et de la demande de titres et va ainsi à l'équilibre se fixer à un niveau tel qu'il reflète la valeur du préjudice subi par les acteurs du fait de l'émission de GES. Comme dans le cas de la taxe, un tel système va permettre d'internaliser les externalités et ce à moindre coût, les efforts de réduction des émissions n'étant pas répartis de façon uniforme. L'avantage par rapport à la taxe est que l'État n'a pas besoin de connaître la valeur des effets externes pour les acteurs, ces derniers révélant cette information à travers leurs décisions de vente et d'achat de titres. Le volume global d'émissions polluantes ne correspond pas forcément à l'équilibre au volume autorisé qui est défini par le volume de titres émis initialement par l'État car des titres ont pu être sortis du marché par des acteurs motivés par la préservation de l'environnement et qui sont prêts à payer pour que moins de GES soient émis. L'État n'a donc pas besoin de connaître le volume optimal d'émission, il est révélé par le fonctionnement du marché.

### **Des difficultés de mise en pratique**

De tels systèmes n'ont pourtant jamais été mis en pratique d'une façon parfaitement conforme à la théorie, car cela supposerait de pouvoir identifier tous les acteurs à l'origine d'émissions polluantes et de contrôler qu'ils détiennent bien les titres correspondant à leurs émissions. Cela est très complexe à effectuer, notamment dans le cas des émissions des ménages. Par ailleurs, en autorisant tous les acteurs à acheter des quotas sur le marché et pas uniquement les pollueurs, la puissance publique ne contrôle pas le niveau global d'émission (puisque

des titres peuvent être achetés dans le but d'empêcher d'autres agents d'émettre des GES) et les autorités craignent de ce fait que le prix du titre soit trop élevé. Il pourrait en résulter une forte hausse des coûts des entreprises, créant des problèmes de compétitivité si tous les pays n'ont pas adopté un tel système. Les entreprises ont ainsi fait pression pour qu'un tel système ne soit pas adopté. Des marchés de quotas d'émission existent néanmoins et c'est ainsi que, à la suite des accords de Kyoto, un marché européen du carbone a été créé en 2005. Il est réservé à certaines entreprises européennes qui émettent des GES. Le marché n'est donc ouvert qu'aux pollueurs et uniquement à certains d'entre eux. La puissance publique définit un quota annuel global d'émission (qui n'est pas forcément optimal) et répartit au début de chaque année les droits entre les entreprises concernées, droits qui s'échangent ensuite sur le marché européen du carbone. On emploie souvent l'expression *cap and trade* pour désigner une politique qui consiste à fixer un plafond (par exemple pour les émissions de GES), puis à laisser ensuite répartir l'effort de réduction entre les agents par des mécanismes de marché. Le bilan de ce marché est peu favorable. Les quotas ont été attribués de façon trop généreuse et le prix des quotas a été trop bas en moyenne pour constituer une incitation suffisante à réduire les émissions de GES. Par ailleurs, le prix a été trop volatil, ce qui brouille le signal envoyé par les prix et rend difficile pour les entreprises la programmation de leurs investissements destinés à réduire les émissions de GES. Ces problèmes pourraient être résolus en attribuant moins de quotas et en régulant le fonctionnement du marché, en mettant par exemple en place un prix plafond et un prix plancher afin de réduire la volatilité du prix du quota.

## POUR EN SAVOIR PLUS

---

❖ **BIEN F. et MÉRITET S.** (2014), *Microéconomie – Les défaillances de marché*, Pearson, coll. « Bachelor Licence ».

❖ **LAURENT É. et LE CACHEUX J.** (2012), *Économie de l'environnement et économie écologique*, Paris, Armand Colin, coll. « Cursus ».

❖ **ROTILLON G.** (2010), *Économie des ressources naturelles*,

Paris, La Découverte, coll. « Repères ».

❖ **BONTEMS P. et ROTILLON G.** (2013), *Économie de l'environnement*, Paris, La Découverte, coll. « Repères ».

❖ **LÉVÊQUE F.** (2004), *Économie de la réglementation*, Paris, La Découverte, coll. « Repères ».

❖ **OSTROM E.** (2010)[1990], *Gouvernance des biens*

*communs : pour une nouvelle approche des ressources naturelles*, Bruxelles, De Boeck.

❖ **REGARDS CROISÉS SUR L'ÉCONOMIE** (2009), « Les économistes peuvent-ils sauver la planète ? », n° 6, novembre 2009.

❖ **VALLÉE A.** (2011), *Économie de l'environnement*, Paris, Points, coll. « Points Essais ».

---

Depuis les années 1960, le concept d'asymétrie d'information a acquis une importance fondamentale pour expliquer les dysfonctionnements de l'économie de marché. Le fait que des catégories opposées d'individus – salariés et employeurs, prêteurs et emprunteurs, entreprises et législateur – n'aient pas accès à la même information permet ainsi de comprendre le chômage, les défauts du marché du crédit, ou encore les inefficacités de la politique de la concurrence. Anne Corcos et François Pannequin présentent les différents types d'asymétrie d'information – antisélection et aléa moral – ainsi que leurs conséquences, selon que la partie la mieux informée est celle qui conçoit le contrat ou l'autre partie. Ils font également le point sur les réponses à apporter à ces défauts de marché.

Problèmes économiques

# Les asymétries d'information

Dès les années 1960, la littérature économique mettait en avant le fait que l'imperfection de l'information pouvait altérer la bonne marche de l'économie concurrentielle. L'hypothèse d'imperfection de l'information stipule que tous les agents économiques n'ont pas accès à toute l'information disponible ou nécessaire lors du déroulement d'une transaction. Certains agents économiques peuvent être mieux informés que d'autres. Ce type d'imperfection faisait partie de ce que les économistes qualifiaient d'échecs de marché (« *market failures* ») ; une catégorie qui regroupait toutes les situations dans lesquelles le modèle canonique de concurrence pure et parfaite ne pouvait déployer ses propriétés d'efficience bien connues<sup>1</sup>.

Nous allons dans un premier temps présenter la typologie des situations d'asymétrie d'information avant d'aborder leurs conséquences sur les marchés, puis les outils que les économistes ont développés pour y remédier.

© ANNE CORCOS

Professeur des universités, université de Picardie, CURAPP-ESS (UMR 7319)

© FRANÇOIS PANNEQUIN

Maître de conférences à l'ENS-Cachan, CES-Cachan (UMR 8174)

## Une typologie

Le problème d'asymétrie d'information ne se pose que quand les intérêts de la partie informée et ceux de la partie non informée divergent, la partie informée étant alors tentée d'exploiter à son avantage l'information dont elle dispose. On distingue deux grandes familles d'asymétrie d'information : les situations d'**antisélection** (AS) se rencontrent *avant* la signature d'un contrat entre les deux parties, alors que le problème d'**aléa moral** (AM) naît *après* la signature du contrat.

En théorie des contrats, la modélisation de base – la relation principal-agent – confronte les deux parties. Elle rend compte de la rela-

<sup>[1]</sup> Le modèle standard de concurrence pure et parfaite, bien qu'il n'illustre qu'une minorité de marchés réels, possède des propriétés de correspondance entre équilibre et optimum dont l'aspect normatif est susceptible d'éclairer la politique économique.

tion économique qui lie la partie qui conçoit le contrat, qualifiée généralement de principal (P), à l'autre partie, qualifiée d'agent (A). À cette première dichotomie (pré/postsignature du contrat) s'ajoute une seconde dichotomie, selon que la partie la mieux informée est le principal ou l'agent. On est ainsi confronté à quatre grandes familles de situations décrites puis résumées dans le tableau ci-après.

[2] Akerlof G. (1970), « The Market for Lemons », *Quarterly Journal of Economics*, vol. 84.

### Antisélection ou sélection adverse

Les deux premiers cas (AS\_P et AS\_A) relèvent de situations dans lesquelles les agents n'ont pas passé de contrat mais sont en passe de le faire. On parle d'antisélection ou sélection adverse car l'une des parties ne connaît pas précisément les caractéristiques de l'autre partie.

Dans le contexte AS\_A, l'agent est la partie la mieux informée ; la modélisation confronte le principal à un agent qui peut revêtir autant de « types » que de caractéristiques plausibles. Le principal est doté de croyances relatives au type de l'agent. Cette situation se rencontre très fréquemment en pratique. Les marchés d'assurance ont été les premiers à avoir identifié ces situations d'antisélection. Un assureur est en effet en situation d'antisélection s'il est face à un assuré dont il ignore le réel degré d'exposition au risque. Or, cette information constitue un élément majeur de sa tarification puisqu'elle lui permet d'établir des primes « actuarielles », qui couvriront, en moyenne, le montant des indemnités qu'il aura à verser en cas d'accident. En présence d'antisélection, il ne sera alors pas à même de lui proposer un contrat adapté.

On retrouve également des situations d'antisélection sur le marché du travail quand un employeur n'a pas accès à certaines caractéristiques du candidat (efficace/peu efficace ; motivé/peu motivé...). L'employeur encourt donc le risque d'embaucher un salarié qu'il payera un montant supérieur à sa productivité.

Le marché du crédit est également le lieu de situations d'antisélection puisque les

prêteurs (les banques par exemple) ne connaissent pas parfaitement la nature (risquée, solvable) du projet de l'emprunteur. Ils ne peuvent donc pas proposer un taux d'intérêt en relation avec le risque que l'emprunteur leur fait courir.

Dans le second contexte (AS\_P), le principal est la partie qui détient l'information. Ces situations d'antisélection caractérisent plus généralement les biens d'expérience ou les biens de confiance (Nelson, 1970)<sup>2</sup>, pour lesquels la qualité du bien n'est révélée, dans le meilleur des cas, qu'au moment de sa consommation. L'exemple le plus emblématique est celui de G. Akerlof (1970), relatif au marché des voitures d'occasion (marché des « lemons ») sur lequel les vendeurs ont une bien meilleure connaissance de la qualité réelle de leur véhicule que les acheteurs potentiels.

### Aléa moral

La seconde grande famille d'imperfection de l'information caractérise les comportements opportunistes qui peuvent surgir après la signature du contrat entre le principal et l'agent. On retrouve principalement des problématiques d'aléa moral. Une fois encore, on peut distinguer selon que la partie informée est le principal (AM\_P) ou l'agent (AM\_A). Les exemples de situations présentant un risque d'aléa moral sont pléthoriques. Elles se retrouvent une fois encore sur le marché du travail et vont caractériser par exemple l'employé qui « tire au flanc » dans le contexte AM\_A ou l'employeur qui a recours à des stagiaires en leur promettant une embauche qu'il n'a pas l'intention de réaliser dans le contexte AM\_P.

Les exemples relatifs au marché de l'assurance sont également très nombreux : ils vont des agents (AM\_A) qui, une fois assurés ne prennent pas suffisamment de précaution (oublie de fermer les portes, etc.) ou fraudent au moyen de fausses déclarations. Mais on retrouve également cet opportunisme du côté des assureurs (AM\_P) qui peuvent, à l'occasion d'un sinistre, montrer un certain nombre de

réticences, ou allonger les délais avant d'indemniser l'agent qui a subi un sinistre.

Un dernier domaine, également emblématique de ce type de problème – mais la liste ne prétend pas à l'exhaustivité – est celui des experts (principal) face à un agent, non informé (AM\_P) et dont les intérêts peuvent être en conflit. On retrouve ici, par exemple, certains couples garagistes/propriétaire de la voiture, avocat/client, médecin/patient... La situation renvoie au risque que l'expert pose un diagnostic qui serve ses intérêts davantage que ceux de son client.

## Les conséquences économiques de l'asymétrie d'information

Les désordres économiques engendrés par ces deux principaux types d'asymétrie d'information peuvent être majeurs : les allocations optimales de 1<sup>er</sup> rang – atteintes lorsque l'information est parfaite – ne sont plus réalisables dès lors que les agents informés bénéficient d'une rente informationnelle leur permettant d'obtenir davantage de surplus (relativement à la situation d'information symétrique). Ainsi en est-il d'un monopole confronté à deux types de consommateurs : s'il ne connaît pas le consentement à payer de chaque consommateur, une discrimination par les prix consistant à imposer à chaque consommateur un prix égal à son consentement à payer n'est plus réalisable car, pour payer moins cher, les consommateurs dont le consentement à payer est élevé auront intérêt à se faire passer pour des consommateurs à faible consentement.

Comme le montre l'article d'Akerlof (1970), focalisé notamment sur l'exemple du marché des voitures d'occasion, la sélection adverse peut désorganiser un marché, voire aboutir à son inexistence. Cela s'applique aussi aux marchés d'assurance santé. En effet, en l'absence d'obligation d'assurance, des primes d'assurance identiques pour tous pourraient

avoir pour effet d'évincer les « bien portants » trouvant le tarif trop cher. Il ne resterait plus alors sur le marché que les individus « à risque », incitant les assureurs à ajuster leurs tarifs à la hausse. À leur tour, les clients risqués, s'estimant insolubles ou trouvant les prix prohibitifs, pourraient quitter le marché.

En résumé, « la mauvaise qualité chasse la bonne ». De tels processus d'éviction peuvent être plus ou moins marqués et induire potentiellement la disparition du marché.

En ce qui concerne l'aléa moral, le principal n'observe pas les caractéristiques de l'agent et cela peut altérer, voire compromettre leur relation économique. La relation employeur-salarié en offre une illustration, lorsque l'employeur embauche le salarié et le rémunère en vue de l'accomplissement d'une tâche dont le résultat, partiellement aléatoire, dépend de l'effort qu'il mobilise. Dans une économie « parfaite », l'employeur embauche le salarié, celui-ci fournit un effort optimal et les deux participants se partagent une production optimale. Si, pour des raisons réglementaires, le salaire est forfaitaire, donc indépendant du niveau d'effort fourni par le salarié, ce dernier n'aura alors aucune incitation à fournir d'effort s'il est coûteux. À l'extrême, si l'absence d'effort conduit inévitablement à un résultat nul, l'employeur, anticipant qu'un salaire fixe aboutit à l'absence d'effort et donc à l'échec, préférera ne pas embaucher. Cette situation d'aléa moral est évidemment inefficace. À l'inverse, si l'effort fourni par le salarié n'a aucun impact sur le résultat obtenu, l'employeur choisira d'embaucher. Dès lors, pour toutes les situations intermédiaires, pour lesquelles l'effort influence la probabilité de succès à des degrés divers, on conçoit intuitivement que la relation entre l'employeur et le salarié puisse être affectée et déboucher sur une inefficace plus ou moins marquée.

## Quels outils analytiques ?

En réponse à ces dysfonctionnements fondamentaux, la théorie économique a identi-

fié différents outils analytiques inhérents à la théorie des contrats, parfois spécifiques à une situation, parfois communs à plusieurs situations, qui ont profondément modifié l'appréhension des phénomènes économiques. Ils sont résumés dans le tableau ci-après. Un trait commun à tous ces mécanismes contractuels tient au fait que la partie informée bénéficie d'une rente informationnelle dont il convient de comprendre, voire de contrôler les effets.

En exploitant le croisement des distinctions entre asymétries pré- et postsignature du contrat, et asymétries en faveur du principal ou de l'agent, on peut faire ressortir deux familles de mécanismes contractuels : ceux à l'initiative de la partie non informée et ceux à l'initiative de la partie informée.

## Les mécanismes à l'initiative de la partie non informée

En présence de sélection adverse, les agents économiques se caractérisent par leurs types (par exemple productif/peu productif pour un travailleur), et si ces types ne sont pas observables, il est possible de recourir à des mécanismes d'autosélection suscitant, à l'équilibre, la révélation des types. De même, en présence d'aléa moral, l'effort de l'agent n'est pas observable et le principal peut mettre en place des contrats incitatifs afin de soutenir l'effort fourni par l'agent. La théorie des enchères propose également des mécanismes permettant de pallier l'absence d'information du vendeur. Nous examinerons successivement ces différents mécanismes.

### Les mécanismes d'autosélection

Ces mécanismes se fondent essentiellement sur le « principe de révélation » (Myerson, 1979<sup>3</sup>), permettant la révélation des caractéristiques des agents économiques. Il s'agit de mettre en place et de proposer aux individus un même menu de contrat spécifié de telle

sorte que chacun soit incité à ne choisir que le contrat qui lui est destiné. On dit que les individus « s'autosélectionnent ».

Dans l'exemple du monopole confronté à deux types d'acheteurs (ou davantage), comme l'ont montré Maskin et Riley (1984)<sup>4</sup>, la meilleure stratégie du vendeur implique de n'offrir que deux assortiments de consommation différenciés, à la fois par leurs prix et leurs quantités, afin de séparer les acheteurs en fonction de leurs préférences : ceux qui sont prêts à payer cher et ceux dont le consentement à payer est plus faible.

La bonne stratégie revient à proposer un menu de contrats<sup>5</sup> construit à l'aide de deux paramètres (le prix et la qualité ou le prix et la quantité) – tout en respectant des impératifs d'incitation : chaque type d'acheteur doit préférer l'assortiment qui lui est destiné à celui de l'autre type d'acheteur. Dans le cas de deux types d'agents, le menu comporte deux contrats, l'un offrant davantage de quantité (ou de qualité) pour un prix plus élevé, l'autre offrant une quantité (ou qualité) moindre pour un prix inférieur.

Sur un marché d'assurance avec antisélection, comme l'ont montré Rothschild et Stiglitz (1976)<sup>6</sup>, le même type de principe prévaut, en jouant cette fois sur le prix et le montant de couverture (ou la franchise<sup>7</sup>) : à l'équilibre, les assurés qui font face à une fréquence d'accident élevée se tournent vers le contrat proposant une couverture à 100 % à prix élevé, tandis que les assurés à faible fréquence d'accident choisissent un contrat de coût moindre associé à une couverture partielle.

À l'équilibre, l'acheteur qui aime fortement le bien, tout comme l'individu à forte sinistralité bénéficient d'une rente informationnelle qui empêche le vendeur ou l'assureur d'accaparer tout le bénéfice de l'échange tandis que l'acheteur aimant modérément le bien ou l'individu de risque faible n'obtiennent aucun bénéfice net de l'échange.

Les mécanismes d'autosélection revêtent une pluralité de formes car ils peuvent s'inscrire

<sup>[3]</sup> Myerson R. (1979), « Incentive Compatibility and the Bargaining Problem », *Econometrica*, vol. 79.

<sup>[4]</sup> Maskin E. et Riley J. (1984) « Monopoly with Incomplete Information », *Rand Journal of Economics*, vol. 15.

<sup>[5]</sup> Contenant autant d'assortiments que de types d'acheteurs.

<sup>[6]</sup> Rothschild M. et Stiglitz J. (1976), « Equilibrium in Complete Insurance Markets: An Essay on the Economics of Imperfect Information », *Quarterly Journal of Economics*, vol. 90, n° 4, 1976.

<sup>[7]</sup> La franchise d'assurance correspond à la part du sinistre qui reste à la charge de l'assuré. Les assurés qui présentent une fréquence de sinistre plus élevée redoutent davantage les franchises d'assurance puisqu'ils ont plus de chance d'en subir les conséquences.

## Typologie des situations et remèdes

|  |          | Asymétrie en faveur du principal P   | Asymétrie en faveur de l'agent A  |
|--|----------|--|---|
| Asymétrie présignature<br>Antisélection AS | Problème | AS_P : antisélection sur la qualité du bien (marché des biens d'expérience ou marché des biens de confiance)   | AS_A : antisélection liée aux caractéristiques de l'agent : son type (risqué/pas risqué), ses préférences (aime le bien/ne l'aime pas), son goût pour le risque   |
|  | Remèdes  | <ul style="list-style-type: none"> <li>– Signal à l'initiative du principal (certifications, labels, garanties, réglementations, réputation (marques, avis des acheteurs [sur internet])</li> <li>– Réglementations : contrôle technique (auto)</li> </ul> | Principe de révélation : <ul style="list-style-type: none"> <li>– menus de contrats qui permettent que les individus s'autosélectionnent</li> <li>– enchères</li> <li>– signal : diplômes, avis des vendeurs (sur internet)</li> <li>– réglementations : obligations d'assurance</li> </ul> |
| Asymétrie postsignature<br>Aléa moral AM   | Problème | AM_P : experts, garagistes, dentistes, avocats   | AM_A : salarié tire-au-flanc, fraude à l'assurance  |
|  | Remèdes  | réglementations, syndicats, comités...   | <ul style="list-style-type: none"> <li>– théorie des incitations</li> <li>– réglementations : contrôles, expertises</li> </ul>  |

dans le temps (exemple : les garanties supplémentaires proposées pour certains biens durables permettent d'autosélectionner ceux qui anticipent un usage plus intense du bien, alors que les autres se contenteront de la garantie de base du fabricant) et ils peuvent être améliorés par le recours à des dispositifs complémentaires (exemple : en assurance, le bonus-malus tire parti de l'expérience observée – survenue ou non de sinistres –, ce qui permet d'ajuster les termes contractuels en fonction du risque de l'assuré).

### Aléa moral et incitation

Les situations économiques dans lesquelles les résultats obtenus sont aléatoires, mais dépendent également des actions choisies par les agents économiques relèvent de l'aléa moral. C'est typiquement le cas de l'agent commercial dont les résultats de ventes

dépendent à la fois de la conjoncture économique et des actions commerciales.

Lorsque le principal n'observe pas l'effort – qui est coûteux pour l'agent –, celui-ci peut être influencé par la forme du contrat. Dans le cas de l'agent commercial, un salaire fixe l'assure complètement contre les aléas conjoncturels et n'encourage pas les efforts. Le contrat optimal passe alors par un arbitrage entre partage de risque (lié au caractère aléatoire du résultat) et incitation à l'effort. Le principal va inciter l'agent à l'effort (inobservable), en conditionnant ses paiements aux résultats (observables). De nouveau, l'équilibre entre les deux parties relève d'une optimalité de second rang plutôt que de premier rang. Cela tient au fait que l'agent qui tire au flanc bénéficie d'une rente liée au fait que l'effort est inobservable et que le résultat ne le révèle pas parfaitement.



## Les enchères

Les enchères constituent un mécanisme permettant à un vendeur de révéler le consentement à payer du ou des acheteurs<sup>8</sup>. Le cadre principal-agent permet de les représenter : un principal (le commissaire-priseur) fait face à des agents (les enchérisseurs) qui détiennent chacun une information privative : la valeur qu'ils accordent au bien mis en vente.

On distingue les enchères ascendantes (ou anglaises), descendantes (ou hollandaises) et les enchères sous plis scellés (premier et second prix). On doit à Vickrey (1961)<sup>9</sup> plusieurs résultats de base qui ont été largement relayés depuis lors par une recherche foisonnante. Selon le théorème d'équivalence du revenu : lorsque les enchérisseurs sont neutres au risque, l'espérance de revenu du vendeur est indépendante du type d'enchères qu'il choisit d'utiliser. Toujours selon Vickrey (1961), face à des enchères au second prix<sup>10</sup>, il est optimal pour un acheteur d'annoncer le prix maximal qu'il est prêt à payer. Ces résultats, et bien d'autres, obtenus sous l'hypothèse d'imperfection de l'information, ont motivé le recours massif actuel à ce type de procédures.

## Les mécanismes à l'initiative de la partie informée : la théorie du signal

Nous avons vu que certains agents pouvaient tirer parti de l'imperfection de l'information et qu'ils n'avaient alors aucun intérêt à ce que leurs caractéristiques soient publiquement dévoilées. Nous portons maintenant notre attention sur des situations pour lesquelles la divulgation d'une information privée peut être dans l'intérêt des personnes informées (qu'elles soient en position de principal ou agent).

En présence d'asymétrie d'information sur la qualité d'un bien, le vendeur d'un article de bonne qualité peut être lésé alors même que l'acheteur serait prêt à payer davantage s'il avait la certitude d'être en présence d'un

bien de bonne qualité. On conçoit aisément qu'une amélioration de l'information puisse être mutuellement avantageuse. Dans ce but, le vendeur peut envoyer une information susceptible de convaincre l'acheteur de la qualité du produit. Le vendeur peut faire état de consommateurs précédemment satisfaits, produire une certification émise par une autorité indépendante ou encore offrir des garanties. Ces différents éléments renseignent sur la qualité du bien et ont été identifiés par la théorie économique comme des « signaux ».

Pour analyser le rôle joué par de tels signaux, on distingue l'émetteur (le vendeur dans le précédent exemple) et le récepteur (l'acheteur). Pour qu'un signal soit économiquement pertinent, le récepteur doit pouvoir en vérifier l'exactitude, le signal doit être crédible, il faut qu'il soit coûteux pour l'émetteur et ce coût doit différer en fonction de la qualité de l'émetteur (par exemple, offrir une année supplémentaire de garantie est plus coûteux pour un fabricant de moindre qualité).

C'est à Spence (1973)<sup>11</sup> que l'on doit le premier modèle de signal. Dans un contexte où la relation employeur-employé se caractérise par une forte asymétrie d'information, Spence a développé une analyse novatrice consistant à appréhender l'éducation comme un signal d'aptitude au travail. Si l'employeur ne discerne pas, *a priori*, les capacités de son futur employé, il peut s'avérer pertinent d'investir dans un signal de qualité : le diplôme. En inventant les modèles de signaux, dans lesquels c'est la partie informée qui exhibe sa caractéristique, Spence a produit une justification convaincante de l'éducation, investissement en capital humain qui représente aussi le coût que subissent les travailleurs efficaces pour informer l'employeur sur leur aptitude au travail. Cette théorie montre notamment qu'un travailleur efficace peut être amené à surinvestir en éducation afin de signaler ses compétences.

Ces modèles débouchent principalement sur deux catégories d'équilibres : l'équilibre mélangeant dans lequel chaque individu

<sup>[8]</sup> Les procédures d'enchères peuvent, symétriquement, être le fait d'un acheteur désirant organiser la mise en concurrence de producteurs, comme l'illustre le cas des procédures d'appel d'offres.

<sup>[9]</sup> Vickrey W. (1961), « Counterspeculation and Competitive Sealed Tenders », *Journal of Finance*, vol. 16.

<sup>[10]</sup> Celui qui annonce le premier prix remporte l'enchère et paie le second prix annoncé.

<sup>[11]</sup> Spence A. (1973), « Job Market Signaling », *Quarterly Journal of Economics*, vol. 87.

envoi le même signal (un diplôme identique) et l'équilibre séparateur dans lequel les individus choisissent des signaux différenciés en fonction de leur caractéristique (un individu plus doué investit davantage dans l'acquisition d'un diplôme car cela lui coûte relativement moins).

## Tests et prolongements

Les principes présentés ont donné lieu à des prolongements théoriques et des tests empiriques. Ils ont également inspiré des politiques publiques. Ainsi, les modèles de la théorie des contrats se sont progressivement adaptés à la plupart des contextes économiques. D'abord focalisés sur des relations bilatérales, des contextes statiques, des contrats complets, ils se sont aventurés à l'étude des relations multilatérales, des relations dynamiques (exemple : un contrat à plusieurs périodes amène l'agent à révéler davantage d'information) ou encore à l'étude des contextes où les contrats sont incomplets. D'abord essentiellement théorique, cette littérature économique a donné lieu à différents tests empiriques depuis le début des années 2000, suscitant de nouveaux questionnements et l'exploration de nouvelles hypothèses. L'asymétrie d'information a, par exemple, été

testée sur les marchés d'assurance sur lesquels la présence d'antisélection se traduit en théorie par une relation croissante entre « sinistralité » des individus et niveau de couverture. Les résultats des tests empiriques de cette relation sont mitigés et font apparaître des situations de sélection avantageuse (dans lesquelles les individus faiblement risqués choisissent des degrés élevés de couverture, en raison d'une forte aversion au risque).

L'État peut également intervenir de différentes manières pour améliorer l'allocation des ressources en présence d'asymétries d'information. Il peut avoir un rôle de garant en dernier ressort pour pallier les situations d'incomplétude des contrats. Mais il a surtout la possibilité de légiférer en établissant par exemple des obligations d'assurance, en réglementant le fonctionnement de certaines activités (banques, entreprises du bâtiment...), en imposant des labellisations ou des réglementations permettant de mieux observer la qualité des biens (exemple : le contrôle technique). Cependant, si l'État interfère positivement dans la plupart des mécanismes correcteurs étudiés, il se substitue parfois au marché sans être le mieux informé et se trouve lui-même initiateur d'aléa moral.

## POUR EN SAVOIR PLUS

❖ **AKERLOF G.** (1970), « The Market for Lemons », *Quarterly Journal of Economics*, vol. 84.

❖ **ARROW K. J.** (1963), « Uncertainty and the Welfare Economics of Medical Care », *American Economic Review*, vol. 53. Traduction française parue dans *Risques*, n° 26, 1996.

❖ **CAHUC P.** (1998), *La Nouvelle Microéconomie*, Paris, La Découverte, coll. « Repères ».

❖ **PANNEQUIN F.** (2010), « Risque et asymétrie d'information », *Risques : les cahiers de l'assurance*, n° 81/82 (spécial 20 ans), p. 181 à 193. [http://www.ffsa.fr/webffsa/risques.nsf/html/Risques\\_81-82\\_0023.htm](http://www.ffsa.fr/webffsa/risques.nsf/html/Risques_81-82_0023.htm)

❖ **SALANIÉ B.** (2012), *Théorie des contrats*, Paris, Économica, coll. « Économie et statistiques avancées ».

Utilisée en économie à partir du milieu du  $xx^e$  siècle, la théorie des jeux fait désormais partie de la boîte à outils classique de la discipline. Elle permet de prendre en compte les interactions stratégiques entre les agents et explique des situations aussi diverses que la concurrence imparfaite, la surexploitation des ressources environnementales ou le *dumping* fiscal. Nicolas Eber en présente les éléments de base avant de faire le point sur les différents types de jeux et leurs applications économiques.

Problèmes économiques

# La théorie des jeux

© NICOLAS EBER

LARGE – IEP et École de management Strasbourg  
Université de Strasbourg

La théorie des jeux peut se définir comme la théorie mathématique des comportements stratégiques. Elle a été créée dans les années 1920 par les mathématiciens Émile Borel et John von Neumann<sup>1</sup>. Le point de départ de l'utilisation de la théorie des jeux en économie date de 1944 avec la publication du livre intitulé *Theory of Games and Economic Behavior*, fruit de la collaboration de John von Neumann et de l'économiste Oskar Morgenstern. En 1950, le mathématicien John Nash<sup>2</sup> invente un concept d'équilibre qui deviendra la pierre angulaire de la théorie des jeux moderne. Depuis les années 1970, la théorie des jeux est devenue l'un des principaux outils de modélisation en science économique et l'équilibre de Nash son concept central. Cela a d'ailleurs valu à Nash l'attribution du prix Nobel d'économie en 1994.

## Jeu statique et stratégie

Considérez une situation dans laquelle deux individus appelés joueur 1 et joueur 2

doivent choisir en même temps et sans communiquer entre deux actions  $C$  et  $D$ . Les utilités (ou niveaux de satisfaction) des deux joueurs dans les différentes configurations possibles sont indiquées dans le Zoom p. 66.

Dans le cas du jeu 1, on peut noter que, pour chaque joueur,  $C$  est une **stratégie « strictement dominée »** puisqu'elle donne un résultat strictement inférieur à celui obtenu en jouant  $D$  quelle que soit la stratégie adoptée par l'autre joueur. Un joueur rationnel doit donc renoncer à choisir une telle stratégie et opter au contraire pour  $D$ . Ainsi, l'issue « logique » de ce jeu est la combinaison de stratégies  $(D, D)$ .

Dans de nombreux jeux cependant, il n'existe pas de stratégie dominée. Il convient donc de définir un concept de solution plus général.

## Le concept d'équilibre de Nash

### Définition

Un **équilibre de Nash** est un ensemble de stratégies (une par joueur) tel que la stratégie de chaque joueur est une meilleure réponse aux stratégies des autres.

Comment trouver l'équilibre de Nash du jeu 1 ? Dans cet exemple, il suffit d'éliminer

<sup>[1]</sup> Antoine-Augustin Cournot en est parfois considéré comme un précurseur, car sa théorie de l'oligopole datant de 1838 intégrait une analyse stratégique proche de la théorie des jeux moderne.

<sup>[2]</sup> Dont le destin hors du commun a été rendu célèbre par la biographie écrite à la fin des années 1990 par la journaliste Sylvia Nasar et le film (*Un homme d'exception*) qui en est inspiré.

# ZOOM

## JEU N° 1 : DILEMME DU PRISONNIER

|          |   | Joueur 2 |     |
|----------|---|----------|-----|
|          |   | C        | D   |
| Joueur 1 | C | 3,3      | 0,5 |
|          | D | 5,0      | 1,1 |

Par convention, dans chaque case du tableau, le premier chiffre correspond à l'utilité du joueur 1 (en ligne) et le second à l'utilité du joueur 2 (en colonne).

Cette situation correspond à un « jeu », au sens d'une interaction stratégique entre plusieurs individus (les joueurs) libres de leurs choix (leurs stratégies). Puisque les deux joueurs jouent simultanément, on parle de jeu simultané ou encore de « jeu statique ». Par ailleurs, les deux joueurs ne pouvant pas s'engager préalablement à prendre certaines décisions, il s'agit d'un « jeu non coopératif <sup>(1)</sup> ».

Plus formellement, un jeu non coopératif statique est caractérisé par : (i) le nombre de joueurs, (ii) l'ensemble des stratégies à la disposition de chaque joueur, (iii) les résultats du jeu pour chaque joueur dans les différentes configurations possibles, (iv) les préférences des joueurs concernant les résultats du jeu.

Dans l'exemple du jeu 1, il y a deux joueurs (1 et 2), chacun ayant à sa disposition deux stratégies (C et D). Les deux joueurs jouent simultanément (jeu statique) et connaissent parfaitement la structure du jeu (jeu à information complète). Le tableau qui représente le jeu est appelé « matrice des gains » : il donne le gain de chaque joueur dans les quatre configurations possibles. Enfin, les « gains » correspondant ici à des niveaux d'utilité, les préférences des joueurs sont reflétées directement par la valeur de ces gains, chaque joueur recherchant le niveau d'utilité le plus élevé.

[1] Une branche de la théorie des jeux, initiée par von Neumann et Morgenstern et qualifiée de « coopérative », s'intéresse au contraire aux cas où les joueurs cherchent à former des coalitions pour coordonner leurs décisions.

les stratégies strictement dominées puisque, par définition, une telle stratégie n'est jamais une meilleure réponse à la stratégie de l'autre. Comme D est la meilleure réponse que chaque joueur peut faire à la stratégie de l'autre, (D, D) est l'unique équilibre de Nash de ce jeu.

### Propriétés

On peut s'interroger sur les propriétés d'existence, d'unicité et d'optimalité de l'équilibre de Nash. Nash lui-même a montré que sous certaines conditions (notamment celle d'élargir le concept de stratégie aux stratégies « aléatoires » appelées stratégies mixtes), il existe toujours au moins un équilibre. Par ailleurs, de nombreux exemples de jeux montrent la possibilité d'une multiplicité d'équilibres<sup>3</sup>.

Concernant la question de l'optimalité, reprenons le jeu 1. En jouant tous les deux D, chacun des deux joueurs obtient 1 alors qu'ils auraient obtenu chacun 3 s'ils avaient tous les deux joué C. Ce jeu met donc en évidence un « conflit » entre la rationalité individuelle et la rationalité collective, conflit qualifié de « dilemme du prisonnier », en référence à une petite histoire due à Albert Tucker (en 1950) et mettant en scène deux prisonniers interrogés séparément par la police et amenés à se dénoncer et à être condamnés à une peine de prison alors qu'ils auraient pu sortir libres en gardant le silence. Le dilemme du prisonnier montre qu'un équilibre de Nash peut être sous-optimal et illustre particulièrement bien le conflit qui peut exister entre les incitations sociales à coopérer et les incitations privées à ne pas le faire. En économie,

[3] Voir Eber N. (2013, chapitre 4) pour une présentation des principaux exemples de « jeux de coordination » caractérisés par l'existence de plusieurs équilibres de Nash.

<sup>[4]</sup> Si un joueur joue 0 et l'autre joue  $X > 0$ , le premier l'emporte car 0 est *toujours* plus proche que  $X$  de la moitié de la moyenne, à savoir ici  $\frac{1}{2} \times X/2 = X/4$ . Par ailleurs, 0 est le meilleur choix possible face à un adversaire jouant 0 puisque jouer 0 assure l'égalité alors que jouer autre chose conduit à une défaite certaine. Techniquement, 0 est une *stratégie faiblement dominante* et le seul *équilibre de Nash* correspond à l'adoption de cette stratégie par les deux joueurs.

<sup>[5]</sup> Chaque joueur a toujours intérêt à déclarer exactement un euro de moins que le montant choisi par son camarade tant que ce montant est supérieur à 180 euros. De plus, 180 euros est bien la *meilleure réponse* possible à un partenaire ayant choisi 180 euros.

## ZOOM

### JEU N° 2

**Vous jouez avec un partenaire anonyme. Chacun(e) doit choisir un nombre entre 0 et 100. Le vainqueur du jeu est celui/celle dont le nombre sera le plus proche de la moitié de la moyenne des deux nombres.**

**Le vainqueur recevra 10 euros (en cas d'égalité, chacun gagne 5 euros).**

**Que jouez-vous ?**

ce jeu est souvent invoqué pour décrire des situations aussi diverses que l'instabilité d'un cartel, le problème de financement des biens publics, la surexploitation des ressources naturelles ou encore l'adoption de mesures protectionnistes.

### Portée et limites de l'équilibre de Nash : les apports de l'expérimentation

Imaginez une expérimentation du dilemme du prisonnier où des sujets jouent la « version monétaire » du jeu 1, les chiffres dans la matrice des gains représentant des euros et non plus des utilités. Les observations expérimentales montrent qu'une partie des sujets coopère (joue  $C$ ). Il faut faire attention de ne pas interpréter trop vite ce type de résultats comme une réfutation empirique du concept d'équilibre de Nash. En effet, si un joueur joue  $C$ , c'est soit qu'il n'a pas compris le jeu, soit, plus probablement, qu'il a des préférences non exclusivement basées sur ses gains monétaires. Peut-être est-il guidé par des motivations sociales, par exemple par un sentiment de culpabilité qui lui rend peu attractif de gagner 5 euros au détriment de son partenaire ?

D'autres résultats expérimentaux peuvent cependant être plus problématiques du point de vue de la portée empirique de l'équilibre

de Nash. Considérez attentivement le jeu n° 2 (*cf.* Zoom ci-contre).

Avez-vous répondu 0 ? Avez-vous remarqué que 0 est *toujours* le nombre vainqueur, *quel que soit* le choix de l'autre joueur<sup>4</sup> ? Dans ce jeu, il est clair que la rationalité doit conduire les joueurs à choisir 0. Or, l'expérimentation d'un jeu similaire avec des enjeux monétaires réels a montré que moins de 10 % des sujets étudiants (et, plus étonnant encore, moins de 40 % des économistes professionnels !) jouent 0 ! Dans cet exemple, où il paraît peu probable que les joueurs cherchent autre chose que maximiser leur chance de victoire, le décalage entre les observations expérimentales et l'équilibre de Nash ne peut s'expliquer que par une forme de rationalité limitée.

Dans d'autres jeux expérimentaux, il est difficile de distinguer ce qui relève de la rationalité limitée des joueurs de ce qui relève de motivations sociales rationnellement intégrées dans leurs préférences. Par exemple, considérez le jeu n° 3 (*cf.* Zoom p. 68).

Ce jeu est connu sous le nom de « **dilemme du voyageur** ». Sous l'hypothèse que chaque joueur cherche à obtenir le gain monétaire maximum, le seul équilibre de Nash est une annonce de 180 euros de la part des deux joueurs<sup>5</sup>.

Les expérimentations en laboratoire de ce jeu conduisent à des observations très éloignées de l'équilibre de Nash puisque près de 80 % des sujets choisissent la borne supérieure de 300 euros. Est-ce à dire que ce jeu invalide le concept d'équilibre de Nash ? Pas nécessairement, mais il est difficile de dire dans quelle mesure les comportements observés reflètent des motivations sociales rationnellement incorporées dans les préférences individuelles des joueurs et/ou une capacité de raisonnement limitée.

## Les jeux dynamiques

Considérons le jeu n° 4 (*cf.* Zoom p. 68) impliquant deux joueurs (1 et 2). À la première étape du jeu, le joueur 1 doit choisir entre arrêter le

## ZOOM

### JEU N° 3 : DILEMME DU VOYAGEUR

Vous êtes parti en voyage avec un ami. Vous aviez exactement les mêmes affaires dans vos valises. À votre retour, votre valise ainsi que celle de votre ami sont perdues. La compagnie aérienne décide de vous indemniser et vous demande d'estimer la valeur de votre valise. Mais, anticipant votre propension à surestimer cette valeur et tenant compte du fait que vous n'avez aucun moyen de communiquer avec votre ami au moment de faire votre déclaration, elle vous annonce :

« Nous savons que les deux valises avaient exactement le même contenu et nous accepterons une indemnisation comprise entre 180 et 300 euros, mais chacun sera indemnisé du montant égal au minimum des deux évaluations qui seront soumises. En outre, si les deux estimations diffèrent, nous donnerons une prime de 5 euros à la personne ayant déclaré la valeur la plus faible et nous déduirons une pénalité du même montant de 5 euros à la personne ayant déclaré la valeur la plus élevée. »

Votre estimation doit nécessairement être un nombre entier d'euros compris entre 180 et 300.

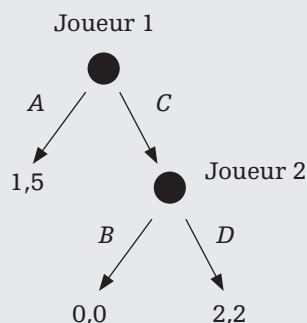
**Quel montant déclarez-vous ?**

jeu (option A) ou le continuer (option C). Si le joueur 1 décide d'arrêter, il obtient une utilité de 1 et le joueur 2 une utilité de 5. S'il décide de continuer, le joueur 2 intervient alors en choisissant entre deux options, B et D : s'il choisit B, les deux joueurs obtiennent 0 ; s'il choisit D, les deux joueurs obtiennent 2.

Dans ce jeu, les joueurs n'agissent plus simultanément, mais l'un après l'autre. On parle

## ZOOM

### JEU N° 4 : EXEMPLE DE JEU SÉQUENTIEL



Chaque point sur la figure est appelé « nœud de décision », chaque flèche représentant un « coup du jeu ». Les gains sont notés pour chaque issue possible du jeu, le premier chiffre correspondant à l'utilité du joueur 1, le second à celle du joueur 2.

de « jeu dynamique » ou encore de « jeu séquentiel ». Ce type de jeu est représenté sous forme dite « extensive », c'est-à-dire sous la forme d'un arbre de décision.

On peut « résoudre » ce type de jeu en appliquant un raisonnement appelé « induction à rebours » (« *backward induction* »). Cette méthode consiste à remonter le jeu *de la fin vers le début*. Que va faire le joueur 2 s'il est amené à agir, c'est-à-dire si le joueur 1 a choisi C ? Il choisira D puisque  $2 > 0$ . Sachant cela, quel est le meilleur choix du joueur 1 ? C'est de choisir C puisque  $2 > 1$ . Ainsi, l'issue « rationnelle » du jeu sera (C, D). Techniquement, cette combinaison de stratégies constitue un **équilibre de Nash « parfait »**<sup>6</sup>.

De nombreux jeux expérimentaux ont cherché à tester la robustesse de l'analyse théorique par l'induction à rebours. Considérez par exemple le problème suivant.

<sup>[6]</sup> Le critère de perfection (en sous-jeux) a été proposé par Reinhard Selten et permet de ne retenir que les équilibres de Nash crédibles séquentiellement, c'est-à-dire reposant sur des menaces crédibles.

## ZOOM

### JEU N° 5 : JEU DE L'ULTIMATUM

Le jeu concerne deux joueurs. Le joueur 1 reçoit 100 euros et doit répartir cette somme entre lui-même et le joueur 2. Le joueur 2 connaît le montant que le joueur 1 doit répartir, mais les deux joueurs ne se connaissent pas et ne peuvent pas communiquer.

Le joueur 1 doit faire une offre au joueur 2 qui peut :

– soit accepter l'offre, auquel cas il reçoit le montant offert et le joueur 1 garde la différence ;

– soit refuser l'offre, auquel cas les deux joueurs ne reçoivent rien (0 euro).

1. Vous êtes le joueur 1. Quel montant (entre 0 et 100 euros) offrez-vous au joueur 2 ?

2. Vous êtes le joueur 2 et vous apprenez que le joueur 1 vous fait une offre de 1 euro. Acceptez-vous son offre ?

<sup>[7]</sup> Ce type de jeu a de nombreuses applications économiques ; il est souvent invoqué dans l'analyse des négociations (par exemple salariales).

<sup>[8]</sup> Pour être précis, dans un tel jeu à information incomplète, on est amené à intégrer les probabilités sur les valeurs des signaux dans le raisonnement des joueurs. C'est pourquoi le concept d'équilibre de Nash doit être « élargi ». L'équilibre de Nash obtenu dans ce type de jeu est qualifié d'équilibre de Nash « bayésien », un concept dû à John Harsanyi.

Le jeu n° 5 (cf. Zoom ci-dessus) correspond au « jeu de l'ultimatum<sup>7</sup> ». La méthode de l'induction à rebours permet d'en identifier les équilibres de Nash parfaits, en supposant, pour faciliter l'analyse, que les offres correspondent nécessairement à un nombre entier d'euros (pas de centimes), et sous l'hypothèse que chaque joueur cherche à obtenir le gain monétaire maximum. On peut noter que le joueur 2 ne devrait jamais refuser une offre positive puisque, s'il la rejette, il obtient 0. À l'équilibre, le joueur 1 n'a donc jamais intérêt à offrir plus de 1 euro puisqu'il anticipe que même ce montant sera accepté. Finalement, il y a deux équilibres de Nash parfaits. Dans le premier, l'offre est de 1 euro et le joueur 2

l'accepte, mais aurait rejeté une offre égale à 0. Dans le second, l'offre est égale à 0 et le second joueur, indifférent entre l'accepter et la refuser, choisit néanmoins de l'accepter.

Le jeu de l'ultimatum a servi de base à un très grand nombre d'études expérimentales. Les résultats montrent la tendance des sujets à s'éloigner de la solution théorique puisque les offreurs (joueurs 1) proposent en moyenne environ 40 euros et les joueurs 2 rejettent en moyenne une fois sur deux les offres inférieures ou égales à 20 euros !

Il est important de noter que **ces résultats expérimentaux n'invalident pas l'analyse théorique**. Ils montrent les limites de l'hypothèse selon laquelle, dans ce type de contexte, les préférences des joueurs les conduiraient simplement à vouloir obtenir le gain monétaire le plus élevé ; manifestement, la plupart des individus intègrent ici dans leurs préférences un souci d'équité.

## Un exemple d'application à l'économie : les enchères à valeur commune

Considérez le jeu n° 6 (cf. Zoom p. 70).

Ce jeu décrit une **enchère à valeur commune**, c'est-à-dire une vente aux enchères (sous pli cacheté et au premier prix) d'un bien dont la valeur exacte est la même pour tous les participants (valeur commune), mais incertaine au moment où ceux-ci doivent faire leur offre.

Techniquement, ce jeu est différent de ceux que nous avons vus jusqu'à présent puisqu'il s'agit d'un « jeu à information incomplète », au sens où les joueurs n'ont pas toutes les informations sur la structure du jeu. Plus précisément, les deux joueurs connaissent leur propre signal, mais ignorent le signal observé par l'autre.

On peut chercher l'équilibre de Nash de ce jeu statique<sup>8</sup>. Appelons  $s_i$  le signal reçu et  $b_i$  l'offre émise par le joueur  $i$  ( $i = 1, 2$ ). On peut montrer qu'à l'équilibre de Nash les joueurs choi-

# ZOOM

## JEU N° 6 : ENCHÈRE À VALEUR COMMUNE

Le jeu concerne deux joueurs, le joueur 1 (« vous ») et le joueur 2. Les deux joueurs reçoivent un signal privé concernant la valeur d'un objet mis aux enchères. Ce signal correspond à une valeur comprise entre 0 et 10 euros. Même s'il ignore la valeur du signal de son partenaire, chaque joueur sait que la valeur de ce signal est distribuée de manière uniforme (ce qui signifie que les signaux des joueurs peuvent prendre n'importe quelle

valeur entre 0 et 10 euros avec la même probabilité). La valeur réelle de l'objet correspond à la **moyenne des deux signaux**.

Chaque joueur doit faire une offre sous pli cacheté (sans savoir quelle est l'offre de l'autre joueur). Celui qui fait l'offre la plus élevée emporte l'enchère et reçoit donc l'objet. Par conséquent, le gain du vainqueur de l'enchère sera égal à la valeur réelle de l'objet diminuée du montant de son offre, alors que le perdant gagne 0 euro. (En cas d'égalité des offres, il y a tirage au sort pour déterminer le vainqueur.)

Votre signal indique une valeur de 8 euros.  
**Quelle offre faites-vous ?**

sissent une offre correspondant exactement à la moitié de la valeur de leur signal, c'est-à-dire  $b_i = s_i/2$  (cf. Zoom). On notera que, puisque la valeur réelle de l'objet est  $(s_1 + s_2)/2$ , le vainqueur de l'enchère fait toujours, à l'équilibre de Nash, un gain positif<sup>9</sup>.

Les résultats expérimentaux montrent que les joueurs font des offres trop élevées par rapport à la solution de Nash. Ce constat est important car il explique pourquoi le vainqueur de ce type d'enchère fait bien souvent des pertes, un phénomène connu sous le nom de « **malédiction du vainqueur** » (*winner's curse*). Prenons un exemple simple pour l'illustrer. Supposons que les signaux des deux joueurs soient  $s_1 = 8$  euros et  $s_2 = 1$  euro. La valeur réelle de l'objet est donc 4,50 euros. Si le joueur 1 pense jouer la prudence en faisant une offre de 5 euros bien en dessous de la valeur de son signal (mais au-dessus de l'offre de Nash), il gagne l'enchère, avec un gain final de 4,50 euros – 5 euros = – 0,50 euro < 0. Il a gagné l'enchère, mais a perdu de l'argent : il est victime de la malédiction du vainqueur ! Ce phénomène a été mis en évidence pour la première fois dans le cas de la vente aux enchères des droits d'exploitation de gisements de pétrole et il est fréquemment invo-

qué au sujet des enchères visant à attribuer les licences de téléphonie mobile ou encore lors de l'attribution des droits de retransmission télévisuelle sur les grands événements sportifs.

Un point important est que ce problème est lié à la rationalité limitée des acteurs. En effet, si tout le monde joue l'équilibre de Nash, personne ne peut être victime de la malédiction ! Il s'agit donc d'un cas où la solution théorique n'est pas évidente et pourrait servir de repère normatif puisque les joueurs perdent de l'argent lorsqu'ils s'en éloignent.

\* \* \*

Une des évolutions majeures de la théorie des jeux moderne réside dans le développement de la **théorie des jeux comportementale** qui, sur la base d'observations expérimentales, cherche à intégrer des facteurs psychologiques dans l'analyse du comportement des joueurs. Comme nous l'avons vu, les observations expérimentales n'invalident pas nécessairement les concepts d'équilibre, mais invitent plutôt à enrichir le modèle comportemental des joueurs, et le cadre conceptuel de la théorie des jeux est suffisamment flexible pour intégrer ces nouveaux paramètres.

<sup>[9]</sup> Dans le cas où  $s_1 > s_2$ , on a :  $b_1 = s_1/2 > b_2 = s_2/2$  ; c'est donc le joueur 1 qui remporte l'enchère et il réalise un gain égal à  $(s_1 + s_2)/2 - s_1/2 = s_2/2$ .



## ZOOM

### RÉSOLUTION DU JEU D'ENCHÈRE À VALEUR COMMUNE

Voici une démonstration partielle (et facultative pour la compréhension du reste de l'article) de l'équilibre de Nash dans le jeu d'enchère à valeur commune, sous l'hypothèse que les joueurs cherchent à obtenir le gain monétaire le plus élevé. Supposons que le joueur 2 choisisse la stratégie  $b_2 = s_2/2$ . La probabilité pour le joueur 1 de gagner avec une offre  $b_1$  correspond à la probabilité que  $b_1$  soit supérieure à  $b_2 = s_2/2$ , c'est-à-dire à la probabilité que  $s_2 < 2b_1$ , égale à  $2b_1/10 = b_1/5$ . La valeur espérée de l'objet conditionnelle à une victoire de l'enchère avec une offre  $b_1$

est :  $(s_1 + E[s_2 | \text{gagner avec l'offre } b_1])/2 = (s_1 + E[s_2 | s_2 < 2b_1])/2 = (s_1 + 2b_1/2)/2 = (s_1 + b_1)/2$ . Par conséquent, la fonction de gain espéré du joueur 1 s'écrit :  $G_1 = (2b_1/10)[(s_1 + b_1)/2 - b_1] = b_1 s_1/10 - b_1^2/10$ . Le gain espéré  $G_1$  est maximum lorsque  $dG_1/db_1 = s_1/10 - 2b_1/10 = 0$ , d'où :  $b_1^* = s_1/2$ . Ce raisonnement suffit à montrer que, si le joueur 2 opte pour la stratégie  $b_2 = s_2/2$ , la meilleure réponse pour le joueur 1 est de choisir  $b_1 = s_1/2$ . Le jeu étant symétrique, on peut faire le même raisonnement pour montrer que  $b_2 = s_2/2$  est la meilleure réponse que le joueur 2 peut faire à un joueur 1 jouant  $b_1 = s_1/2$ . Ainsi, les deux joueurs pratiquant la stratégie  $b_i = s_i/2$  correspond à un équilibre de Nash (bayésien). En toute rigueur, il faudrait montrer, ce qui est un peu plus laborieux, qu'il s'agit bien du seul équilibre.

Nicolas Eber

## POUR EN SAVOIR PLUS

❖ **EBER N.** (2013), *Théorie des jeux*, Paris, Dunod, coll. « Les Topos », 3<sup>e</sup> éd.

❖ **GUERRIEN B.** (2010), *La théorie des jeux*, Paris, Économica, 4<sup>e</sup> éd.

❖ **YILDIZOGLU M.** (2011), *Introduction à la théorie des jeux*, Paris, Dunod, 2<sup>e</sup> éd.

## ↓ COMPLÉMENT

### Quelques applications de la théorie des jeux

#### Théorie des jeux et environnement : la tragédie des biens communs

Considérons un jeu à 4 joueurs. Il y a 40 jetons dans un pot commun. Chacun des quatre joueurs peut retirer entre 0 et 10 jetons de ce pot commun. Le gain de chaque joueur est égal à : (2 euros × nombre de jetons qu'il a retirés) + (1 euro × nombre total de jetons restant dans le pot commun).

Par exemple, si un joueur retire 10 jetons alors que les trois autres n'en retirent que 5, il reste 15 jetons dans le pot commun et le premier joueur gagne  $2 \times 10 + 1 \times 15 = 35$  euros alors que chacun des trois autres obtient  $2 \times 5 + 1 \times 15 = 25$  euros.

Les joueurs devant choisir simultanément le nombre de jetons qu'ils retirent, il s'agit d'un jeu statique qui s'analyse à l'aide du concept d'équilibre de Nash. Sous l'hypothèse que les joueurs cherchent à maximiser leur gain monétaire, ils ont toujours intérêt à retirer le plus de jetons possible. Le seul équilibre de Nash correspond donc à la situation où chacun

retire le maximum (10 jetons), avec un gain individuel de 20 euros. On notera que ce jeu relève d'une logique proche du dilemme du prisonnier puisque les incitations individuelles se font contre l'intérêt général ; en effet, chacun gagnerait le double, soit 40 euros, si personne ne retirait de jeton du pot commun !

Ce type de jeu est utilisé en économie de l'environnement pour conceptualiser la tendance des individus à surexploiter les ressources communes au détriment de la collectivité, phénomène connu sous le nom de « tragédie des biens communs ». En effet, retirer un jeton dans le jeu peut s'assimiler à prélever dans une ressource naturelle, au détriment du groupe.

C'est en étudiant des versions élaborées de ce jeu qu'Elinor Ostrom (Prix Nobel d'économie en 2009) a proposé une vision nouvelle de la tragédie des biens communs. Constatant expérimentalement que face à ce type de jeu, les individus sont capables de faire émerger des normes sociales de coopération, elle développe l'idée qu'une « autogestion » des biens communs peut être, dans certains cas, plus efficace que les solutions basées sur la propriété privée et la compétition marchande ou sur l'action coercitive de l'État.

### **Théorie des jeux et concurrence imparfaite**

La théorie de l'organisation industrielle, développée notamment par Jean Tirole (Prix Nobel 2014), applique la théorie des jeux à l'étude de la concurrence imparfaite, c'est-à-dire des structures de marché dans lesquelles les entreprises cherchent à obtenir un pouvoir de marché par des stratégies de prix, de différenciation de produits, etc. Un exemple classique de ce type d'analyse est le modèle de duopole avec concurrence en prix, parfois qualifié de « duopole de Bertrand » en référence au mathématicien français Joseph Bertrand qui en avait donné l'intuition dès 1883.

On considère deux entreprises  $A$  et  $B$  produisant un même bien au même coût

unitaire  $c$ . On suppose que ces deux entreprises n'ont pas de contrainte de capacité, au sens où elles peuvent produire n'importe quelle quantité (et donc répondre à la totalité de la demande du marché). Les deux entreprises proposent les prix  $p_A$  et  $p_B$  respectivement. Si une entreprise est moins chère que l'autre, elle récupère la totalité du marché ; si les deux prix sont identiques, les deux entreprises se partagent le marché. Quel est l'équilibre de Nash de ce jeu en prix ? Notons tout d'abord que à l'équilibre, les deux prix sont forcément égaux ( $p_A = p_B$ ) ; en effet, si ce n'est pas le cas, l'entreprise la plus chère a toujours intérêt à changer de stratégie pour s'aligner sur son concurrent. De plus, tant que les prix sont supérieurs au coût de production, et donc que les entreprises font des profits positifs, il est toujours profitable pour chaque firme de choisir un prix un centime plus bas que son concurrent (afin de récupérer la totalité de la demande). Ainsi, toute situation dans laquelle les deux entreprises pratiquent un même prix supérieur au coût  $c$  ne peut pas être un équilibre. Au final, l'unique équilibre de Nash est caractérisé par  $p_A = p_B = c$ . Cet équilibre conduit à une tarification au coût unitaire et donc à des profits nuls pour les deux entreprises.

Ce résultat est souvent qualifié de « paradoxe de Bertrand » car il montre que la présence de deux entreprises sur le marché peut suffire à restaurer l'équilibre concurrentiel (profits nuls). Ce modèle très simple a été développé et enrichi dans de nombreuses directions : introduction de contraintes sur les capacités de production, prise en compte de stratégies de différenciation du produit, etc. Dans sa version la plus simple, il a le mérite de mettre en exergue la logique stratégique de « guerre des prix » et de montrer à la fois l'intérêt (pour les entreprises) et l'instabilité potentielle d'une entente sur les prix.

Nicolas Eber

L'économétrie s'est fortement développée après la Seconde Guerre mondiale. Partie empirique des sciences économiques, elle vise à identifier et quantifier des causalités entre des phénomènes économiques. Luc Behaghel en présente ici les instruments de base. Il commence par rappeler en quels termes se posent les problèmes de causalité, puis expose les outils permettant leur appréhension : la régression simple, la régression multiple et les variables instrumentales.

Problèmes économiques

# Les outils économétriques

© LUC BEHAGHEL

Membre associé, directeur de recherche INRA

Les outils de l'économétrie sont nombreux et souvent techniques. Nous nous concentrons ici sur deux outils de base, la régression et les variables instrumentales, pour les relier à l'objectif fondamental de l'économétrie : mettre en évidence et quantifier des relations causales entre phénomènes économiques<sup>1</sup>.

## Le problème de la causalité : un exemple

Identifier les effets d'une cause, c'est comparer deux états du monde : l'un où la cause est à l'œuvre, l'autre où elle ne l'est pas (le « contrefactuel »). Le problème fondamental de l'identification d'un effet causal est qu'il n'est jamais possible d'observer simultanément deux états du monde : l'un des deux doit être reconstitué. Cette reconstitution du contrefactuel est l'enjeu fondamental de l'identification causale.

Prenons l'exemple d'un traitement médical. L'effet du traitement sur la santé d'un patient est l'écart entre sa santé avec traitement et sa santé sans traitement. Mais jamais le *même*

patient n'est observé *simultanément* avec et sans traitement. Les outils de l'économètre mettent alors en œuvre des comparaisons qui permettent, sous certaines hypothèses, de reconstituer la situation contrefactuelle. L'outil approprié dépend de la situation considérée. Examinons trois cas de figure.

Premièrement, le traitement a été prescrit de façon aléatoire.

C'est l'exemple du test clinique : dans une population de patients suffisamment grande, deux groupes sont tirés au sort : le groupe test reçoit le traitement, le groupe témoin ne le reçoit pas. Par construction, les deux groupes sont au départ comparables en tout, puisqu'ils sont représentatifs de la même population. Il existe bien sûr dans les deux groupes des individus en plus ou moins bonne santé, des individus plus ou moins riches, soutenus par leur famille, etc., et tous ces facteurs sont susceptibles d'affecter leur santé future ; mais ces individus sont en proportions équivalentes dans les deux groupes, si bien qu'en moyenne, leur présence affecte de la même façon le groupe test et le groupe témoin. On peut donc en faire abstraction et déduire directement l'effet du médicament de la relation entre le fait d'avoir été traité ou non (variable explicative) et la santé au bout d'un temps déter-

<sup>[1]</sup> L'analyse de données macroéconomiques (macroéconométrie, ou économétrie des séries temporelles) a développé des outils propres qui ne sont pas présentés ici. Pour une vue d'ensemble des grandes tendances de l'économétrie depuis les années 1970, et pour une discussion des débats qui parcourent la microéconométrie, on pourra se référer à Behaghel (2011). Par ailleurs, Behaghel (2012) propose une présentation détaillée des outils, partiellement reprise ici.

miné (variable de résultat, ou variable expliquée), en utilisant l'outil économétrique le plus simple : la **régression simple**.

Deuxième cas de figure, tous les paramètres entrant dans la décision de prescription du médicament sont connus de l'économètre.

C'est le cas, par exemple, si la décision de prescription est prise sur la base de l'information contenue dans le dossier médical du patient, et si l'économètre a accès au dossier du patient. Dans ce cas, il suffit de comparer un à un des patients dont les dossiers médicaux sont identiques, mais dont l'un est traité, l'autre non. Si, en moyenne, le patient traité est en meilleure santé que son homologue à l'issue du traitement, cette différence de résultat peut être attribuée au traitement. Les outils correspondant à cette approche sont la **régression multiple** et les **estimateurs par appariement** (« *matching* »).

Troisième et dernier cas, un événement fortuit fait que le médicament est plus fréquemment prescrit dans certains groupes de patients que dans d'autres pourtant comparables.

Imaginons par exemple que l'entreprise pharmaceutique qui produit le médicament veuille tester une stratégie de promotion de ce traitement auprès de médecins prescripteurs. Elle invite au hasard la moitié des médecins présents à un congrès à une session d'information sur le médicament. À la fin du congrès, elle fait passer un questionnaire auprès de tous les médecins (invités ou non à la session) les interrogeant sur leurs intentions de prescription, ce qui lui permet de conclure si cette campagne est efficace. Cette campagne de promotion est tournée vers les médecins. Mais elle crée aussi une variation au niveau des patients : les patients de médecins qui ont été informés sur le traitement ont plus de chances de recevoir ce traitement, et ce pour une raison qui n'a rien à voir avec leur état de santé initial (puisque la campagne de promotion a ciblé les médecins de façon aléatoire). S'il s'ensuit une amélioration de leur état de santé par rapport à ceux des patients des médecins qui n'ont pas été invités à la ses-

sion, on peut en déduire que le médicament a un effet bénéfique. Les outils économétriques correspondant à cette approche sont les **estimateurs par variable instrumentale**.

Présentons maintenant ces outils plus en détail.

## La régression simple

La régression simple permet d'analyser la relation entre une variable de résultat (la variable expliquée ou dépendante, notée  $y$ ) et une variable explicative (ou indépendante, notée  $x$ ). La relation est décrite par un modèle linéaire :

$$y = b_0 + b_1 \times x + u,$$

dans lequel  $b_0$  et  $b_1$  sont deux paramètres à estimer, et  $u$  est un terme d'erreur, dû au fait que le lien entre  $y$  et  $x$  est imparfait.  $b_0 + b_1 \times x$  est appelé « **droite de régression** » et correspond aux variations de  $y$  que le modèle peut expliquer.  $u$  correspond à la partie non expliquée, ou résiduelle, qui comporte tous les facteurs explicatifs de  $y$  autres que  $x$ , ainsi que l'erreur de mesure sur  $y$ . L'estimateur de loin le plus utilisé dans ce contexte est l'**estimateur des moindres carrés ordinaires** (MCO) : il consiste à choisir  $b_0$  et  $b_1$  tels que la somme des carrés des résidus ( $u^2$ ) soit minimale, dans un échantillon donné. Visuellement, la droite de régression passe alors au milieu du nuage de points  $(x, y)$  de l'échantillon<sup>2</sup>.

Le rôle important joué par l'estimateur des MCO s'explique par le fait qu'il est facile à calculer et se généralise aisément au cas avec plusieurs variables explicatives (voir ci-dessous). Mais il tient aussi au fait que l'estimateur de  $b_1$  s'interprète directement comme l'effet causal moyen de  $x$  sur  $y$ , sous une hypothèse clé : l'hypothèse de moyenne conditionnelle nulle du terme d'erreur<sup>3</sup>, c'est-à-dire que l'effet moyen des variables omises est le même pour les différentes valeurs de  $x$ <sup>4</sup>. Cette hypothèse est vraie par construction dans le cas d'une expérimentation contrôlée : en vertu du tirage au sort, les variables omises du modèle prennent

<sup>[2]</sup> D'autres critères seraient possibles pour ajuster le modèle aux données, par exemple, minimiser la somme des valeurs absolues des écarts plutôt que la somme des carrés des écarts.

<sup>[3]</sup> Mathématiquement, l'espérance de  $u$  sachant  $x$  est nulle :  
 $E(u|x) = 0$

<sup>[4]</sup> On vérifie alors que l'effet moyen sur  $y$  d'une augmentation d'une unité de  $x$  vaut :  
 $E(y|x = x_0 + 1) - E(y|x = x_0) = b_1 + E(u|x = x_0 + 1) - E(u|x = x_0) = b_1$

en moyenne les mêmes valeurs dans les groupes test et témoin. Mais elle risque fort d'être violée dans d'autres applications sur données non expérimentales. Considérons l'estimation des rendements de l'éducation par le modèle de la régression simple :  $y$  est le salaire,  $x$  le nombre d'années d'études. Le terme d'erreur  $u$  comporte alors toutes les variables explicatives omises du modèle : réseaux sociaux et familiaux, capacités individuelles, etc. L'hypothèse de moyenne conditionnelle nulle du terme d'erreur stipule que ces caractéristiques sont distribuées de la même façon quel que soit le niveau d'éducation considéré. On peut et doit en douter : ceux qui font des études longues sont en moyenne issus de milieux plus favorisés, ce qui facilite l'accès aux emplois plus rémunérateurs ; par ailleurs, on peut penser que les capacités individuelles favorisent l'obtention d'un salaire plus élevé et la poursuite des études. Le modèle de la régression simple conduirait à attribuer aux études non seulement leur effet propre, mais aussi l'effet lié aux différences de milieu ou de capacité, si bien qu'on surestimerait vraisemblablement l'effet causal des études. Un tel biais est appelé biais de la variable omise. C'est le problème que les outils suivants essaient de résoudre.

<sup>[5]</sup> L'appariement exact n'est pas toujours possible. En effet, les variables explicatives sont souvent nombreuses et peuvent prendre un grand nombre de valeurs, si bien que les cellules à considérer sont trop nombreuses et parfois vides. Différents estimateurs ont été développés pour prendre en compte ce problème de dimensionnalité. Leur principe commun est de ramener le problème à un appariement sur une seule variable, appelée le score de propension. On peut en effet montrer que, pour estimer l'effet d'un traitement, comparer des individus qui ont les mêmes caractéristiques observables revient au même que de comparer des individus qui ont la même probabilité de recevoir ce traitement au vu de leurs caractéristiques observables.

<sup>[6]</sup> Ly S. et Riegert A. (2013), « Persistent Classmates: How Familiarity with Peers Protects from Disruptive School Transitions », *PSE Working Papers*, n° 2013-21.

## La régression multiple et les estimateurs par appariement

La régression multiple généralise la régression simple en prenant en compte le fait qu'une variable expliquée dépend de plusieurs variables explicatives potentiellement corrélées, dont il faut donc démêler les effets par une analyse jointe. Dans le cas avec deux variables explicatives, on écrit le modèle statistique :

$$y = b_0 + b_1 \times x_1 + b_2 \times x_2 + u.$$

Par exemple, le salaire  $y$  d'un individu dépend de ses études  $x_1$  et de son milieu d'origine  $x_2$ . L'intérêt de ce modèle est de ne plus avoir

à supposer, pour interpréter  $b_1$  causalement, que la durée des études est indépendante du milieu d'origine, ou que le milieu d'origine n'a pas d'effet sur le salaire, puisque l'effet du milieu d'origine est explicitement pris en compte dans le modèle. En revanche, il faut maintenir que les autres facteurs explicatifs omis (par exemple les capacités individuelles) ne viennent pas perturber la comparaison entre les différents groupes définis par une durée d'étude  $x_1$  et un milieu d'origine  $x_2$ . Autrement dit, l'interprétation causale de  $b_1$  (comme effet des études) et  $b_2$  (comme effet du milieu d'origine) repose sur l'hypothèse d'indépendance conditionnelle  $E(u|x_1, x_2) = 0$ .

Pour estimer le modèle de la régression multiple, on procède de la même façon que pour la régression simple, en minimisant la somme des carrés des résidus (estimateur des MCO). Mais pour bien comprendre la logique sous-jacente, il est utile d'introduire un estimateur proche de l'estimateur des moindres carrés, l'estimateur par appariement (« *matching* »). Sous sa version la plus simple (appariement « exact »), cet estimateur consiste à constituer autant de cellules qu'il existe de valeurs prises par les variables explicatives. Par exemple, si on considère deux catégories d'études (longues et courtes) et six catégories sociales, on constitue douze cellules. L'effet des études longues sur le salaire est alors estimé à partir de comparaisons deux à deux, à catégorie sociale donnée, entre salaire moyen après études longues et salaire moyen après études courtes. On obtient ainsi six estimations, une par catégorie sociale. L'effet moyen toutes catégories confondues s'obtient en faisant la moyenne de ces six effets, pondérés par la part de chaque catégorie sociale dans la population<sup>5</sup>.

La récente étude de Ly et Riegert (2013)<sup>6</sup> constitue une bonne application de la régression multiple. Les auteurs analysent l'effet de conserver les mêmes camarades de classe entre collège et lycée. La difficulté de l'analyse est que les classes ne sont pas constituées au hasard par les proviseurs : on peut

donc craindre que les élèves qu'on maintient ensemble d'une année sur l'autre ne soient spécifiques (par exemple, moins dissipés ou plus motivés). Mais Ly et Riebert ont accès précisément à l'ensemble des données du dossier scolaire dont disposent les proviseurs au moment de la constitution des classes (y compris, par exemple, les notes de comportement). Leur approche consiste alors à constituer des paires d'élèves issus d'une même classe de troisième et qui ont exactement le même dossier scolaire en fin de troisième. Lorsque deux élèves d'une même paire sont alloués à deux classes différentes, ce choix est nécessairement fait au hasard, dans la mesure où ces deux élèves sont indiscernables du point de vue du proviseur. Par conséquent, les élèves envoyés vers une classe de seconde où ils retrouvent des élèves de leur classe de troisième sont en moyenne comparables à leurs homologues, issus de la même paire mais envoyés vers une classe de seconde voisine où ils ne retrouvent pas de camarades de l'année précédente. Les résultats obtenus sur la base de cette comparaison sont frappants : le fait de garder un camarade de classe entre collège et lycée réduit de 13 % le risque de redoubler la seconde (pour des élèves qui étaient en troisième dans la deuxième moitié de la distribution) et augmente d'autant la probabilité d'obtention du baccalauréat. La familiarité avec une partie de ses camarades de classe facilite la transition vers le lycée.

## Les variables instrumentales

L'exemple précédent illustre bien comment l'hypothèse de moyenne conditionnelle nulle du terme d'erreur peut se justifier quand l'économètre a accès à l'ensemble des variables qui déterminent l'entrée dans un « traitement » donné. Mais pour beaucoup d'autres « traitements », le processus de sélection est beaucoup plus difficile à observer et dépend potentiellement de variables inobservables qui font craindre le biais de variable omise. L'exemple type est

celui du choix de faire des études. On peut certes contrôler le milieu d'origine, mais il est plus difficile de mesurer la motivation ou les capacités individuelles. La régression multiple et les modèles d'appariement sont alors inopérants. Les variables instrumentales constituent une alternative. On n'essaie plus de contrôler tous les facteurs de sélection dans un traitement. Plutôt, on fait reposer l'estimation sur une sous-partie des entrées en traitement, celles dont on pense qu'elles sont dues à un événement fortuit (ou du moins sans lien direct avec l'effet à expliquer). C'est ce caractère fortuit ou « exogène » d'une partie du traitement qui permet d'identifier un effet causal sans craindre le problème de la variable omise : il joue un rôle analogue au tirage au sort dans l'approche expérimentale.

Pour isoler cette variation exogène du traitement, on utilise un **instrument**, c'est-à-dire une variable qui remplit deux conditions :

- elle est **pertinente**, au sens où elle a un impact sur la variable explicative dont on souhaite mesurer l'effet ;

- elle n'a pas d'impact direct sur la variable expliquée (cette deuxième condition est la **condition d'exclusion**).

Par exemple, dans le cas des rendements de l'éducation, un instrument est une variable qui explique la durée des études sans avoir d'effet direct sur les salaires. Angrist et Krueger (1991)<sup>7</sup> proposent pour cela d'utiliser, aux États-Unis, le trimestre de naissance. L'idée est que naître à un moment ou à un autre de l'année civile est un événement fortuit, indépendant de tous les déterminants du salaire, mais qui a un effet sur la durée de scolarisation, en raison des règles de scolarisation obligatoire<sup>8</sup>. L'instrument impacte la durée des études : la condition de pertinence est bien remplie. La condition d'exclusion semble également remplie : en quoi le fait d'être né tel ou tel jour de l'année serait-il lié au salaire, sinon *via* la durée des études ? Supposons alors qu'on observe que les personnes nées au premier trimestre

[7] Angrist J. et Krueger A. (1991), « Does Compulsory School Attendance Affect Schooling and Earnings? », *Quarterly Journal of Economics*, vol. 106.

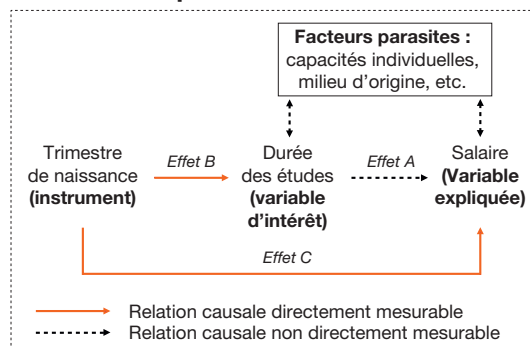
[8] En effet, dans la plupart des États américains, la loi stipule qu'on entre à l'école en septembre de l'année calendaire de ses 6 ans. Ainsi, une personne née un 31 décembre rentrera à l'école en septembre de la même année, à l'âge de 5 ans et 8 mois, mais une personne née le lendemain rentrera en septembre de l'année suivante, à l'âge de 6 ans et 8 mois. Comme la scolarité est obligatoire jusque 16 ans, la première personne devra faire au moins 10 ans et 4 mois d'études, et la seconde seulement 9 ans et 4 mois d'études. Comme il existe une certaine partie de la population qui souhaite faire des études courtes et pour laquelle l'âge minimum de fin d'études est une contrainte, la durée moyenne d'études observée dépend bien, dans la population, de la date de naissance.

de l'année civile fassent en moyenne des études plus longues et aient des salaires plus élevés. Alors, comme on n'a pas de raison de penser que le trimestre de naissance en lui-même explique le salaire, on attribue ces salaires plus élevés aux études. Il n'y a plus lieu de s'inquiéter du biais de variables omises : certes, le salaire dépend aussi du milieu social d'origine, des capacités individuelles, etc. Mais si ces caractéristiques sont en moyenne les mêmes quel que soit le trimestre de naissance, les omettre ne crée pas plus de biais que dans le cas d'une expérience contrôlée.

Le schéma ci-contre illustre ce raisonnement. Les variables instrumentales constituent un chemin détourné pour mesurer l'effet causal A de l'éducation sur le salaire. Une mesure directe n'est en effet pas possible, car toutes sortes de variables omises, non observées, viennent « parasiter » la relation entre ces deux variables. En revanche, il est possible de mesurer l'effet causal B du trimestre de naissance sur la durée des études, ainsi que l'effet causal C du trimestre de naissance sur le salaire. Reste alors à faire une hypothèse cruciale, qui correspond à la condition d'exclusion : le trimestre de naissance doit n'avoir aucun effet direct sur le salaire, c'est-à-dire que tout l'effet doit transiter par la durée des études. Alors, l'effet causal du trimestre de naissance sur le salaire est le produit d'un effet causal du trimestre de naissance sur la durée des études et d'un effet causal de la durée des études sur le salaire ( $C = B \times A$ ). Comme on mesure B et C, on peut déduire A. Angrist et Krueger (1991) observent ainsi que les personnes nées au premier trimestre d'une année calendaire font des études en moyenne d'un dixième d'année plus courtes que les personnes nées plus tard dans l'année (effet B). Ces mêmes personnes gagnent en moyenne des salaires 1 % plus faibles (effet C). Donc un dixième d'année d'études en moins réduit le salaire de 1 % ; par une simple règle de trois, cela implique que le rendement d'une année d'études supplémentaires (effet A) est de

10 % de salaire en plus. On vérifie bien que  $C = B \times A : 1 \% = 0,1 \times 10 \%$ .

### Schématisation du raisonnement de l'estimation par variables instrumentales



Trouver des instruments est difficile. Ainsi, les résultats d'Angrist et Krueger ont été critiqués dans la mesure où il existe, en réalité, une différenciation sociologique de la saisonnalité des naissances : les cadres ont des pics de conceptions pendant l'été et pendant les vacances d'hiver, pics qu'on ne retrouve pas chez les indépendants, ouvriers, employés, chômeurs et inactifs<sup>9</sup>. Pour utiliser le trimestre de naissance comme instrument, il est alors nécessaire d'étudier séparément les différents groupes sociaux ou encore de combiner variables instrumentales et régression multiple. L'outil le plus fréquemment utilisé pour cela est celui des doubles moindres carrés (« *two-stage least squares* »).

## Extensions : doubles différences et discontinuité de la fonction de régression

On le voit, moindres carrés ordinaires et variables instrumentales sont des outils dont l'interprétation doit être pesée avec soin. Mais ils constituent aussi la matrice de multiples approches économétriques. On peut, pour ouvrir la discussion, reprendre l'exemple initial de l'effet d'un médicament, et considérer deux autres situations :

<sup>[9]</sup> Grenet J. (2010), « La date de naissance influence-t-elle les trajectoires scolaires et professionnelles ? Une évaluation sur données françaises », *Revue économique*, vol. 61, n° 3.

Première situation, la prescription du traitement obéit, au moins partiellement, à un critère discontinu.

Par exemple, le médicament est recommandé pour les nouveau-nés qui font moins de 2000 grammes à la naissance. De ce fait, la fréquence du traitement est plus élevée parmi les enfants situés juste au-dessous du seuil. Si leur état de santé est par la suite meilleur que celui des enfants situés juste au-dessus du seuil, on en déduit que le traitement a été bénéfique. Il s'agit là en fait d'une application particulière des variables instrumentales, connue sous le nom de *regression discontinuity design*. La relation d'exclusion nécessaire est que le fait que l'enfant se situe juste au-dessous du seuil de 2000 grammes plutôt que juste au-dessus n'a d'effet sur sa santé que *via* le traitement.

Seconde situation, le médicament a été introduit à des dates différentes à différents endroits.

Par exemple, les autorisations de mise sur le marché conduisent le médicament à être disponible plus tôt dans certains pays que dans d'autres. Si ces pays voient la pathologie visée par le médicament diminuer relativement aux autres pays, alors que les tendances étaient auparavant parallèles, on conclura que le médicament a des effets bénéfiques. Cette stratégie est connue sous le nom de  **doubles différences** , puisqu'elle effectue une double différenciation dans le temps (évolution de la pathologie) et dans l'espace (entre pays). Mais il ne s'agit pas d'autre chose que de moindres carrés ordinaires dans lesquels la variable de résultat, plutôt que d'être mesurée en niveau, est mesurée en tendance, et où l'hypothèse de moyenne conditionnelle nulle du terme d'erreur correspond au fait que les écarts de mise sur marché du traitement sont en moyenne indépendants des tendances qui auraient prévalu en l'absence de médicament.

---

## POUR EN SAVOIR PLUS

---

❖ **BEHAGHEL L.** (2011), « Le développement de l'économétrie », *Cahiers français*, n° 363.

❖ **BEHAGHEL L.** (2012), *Lire l'économétrie*, Paris, La Découverte, coll. « Repères ».

---



Thème central de la science économique, la monnaie fait l'objet de controverses qui se sont révélées durables. Les fonctions de la monnaie suffisent-elles à en donner une définition satisfaisante ? La monnaie est-elle neutre ou active ? Exogène ou endogène à l'activité économique ? Revenant sur ces grandes questions, Alain Beitone montre que les postures quant à la monnaie sont moins simples qu'il n'y paraît et qu'elles se superposent difficilement aux clivages traditionnels de la pensée économique.

Problèmes économiques

# La monnaie dans les théories économiques

## © ALAIN BEITONE

Professeur de sciences économiques et sociales,  
lycée Thiers Marseille

Les questions monétaires sont de longue date au cœur des préoccupations des économistes (et avant eux des philosophes, voire des théologiens). Ce thème suscite des débats assez vifs, mais, comme nous allons le montrer, les clivages ne sont pas toujours ceux que l'on croit.

## Approche fonctionnaliste et approche essentialiste

Pour certains économistes, la monnaie est ce que la monnaie *fait*. La définition de la monnaie est donc construite à partir de la liste de ses fonctions. On cite traditionnellement la fonction d'intermédiaire des échanges, de mesure des valeurs et de réserve des valeurs. Au sein de cette approche, certains privilé-

gient l'une des fonctions par rapport aux autres. Par exemple, pour Carl Menger et les économistes autrichiens, c'est la fonction d'intermédiaire des échanges qui est essentielle. D'autres économistes mettent l'accent sur la fonction de réserve des valeurs car, selon la formule de Keynes, la monnaie est un pont entre le présent et l'avenir.

L'approche **fonctionnaliste** apparaît insuffisante aux économistes qui souhaitent mettre l'accent sur la nature de la monnaie. Cette interrogation, parfois rejetée comme trop « philosophique », fait l'objet d'un regain d'intérêt chez les économistes contemporains, qui se réfèrent alors à des approches historiques et à l'anthropologie économique. Pour un économiste allemand du début du xx<sup>e</sup> siècle, Georg Friedrich Knapp, la monnaie est liée au pouvoir de l'État et du droit. C'est parce que l'État impose le paiement de l'impôt en monnaie (et qu'il accepte donc la monnaie en paiement) que celle-ci est acceptée au sein de l'ensemble de la communauté de paiement. On parle de conception **nominaliste** de la monnaie et

d'approche « **chartaliste** », pour signifier que la valeur de la monnaie ne repose pas sur une « substance » (les métaux précieux) mais sur une validité proclamée par l'État. Cette approche conduit à insister sur le lien entre monnaie et souveraineté politique. De nombreux travaux, tout en relativisant la survalorisation de l'État par Knapp, mettent l'accent sur le fait que la monnaie doit être comprise en relation avec la société, avec la cohésion du groupe social au sein duquel elle circule. À la suite de Simmel, on peut souligner que la monnaie est une composante du lien social et, avec Michel Aglietta, que « la monnaie est le mode de socialisation des sujets sous la forme de l'auto-organisation » (Aglietta, 1988, p. 93).

À l'approche de la monnaie comme objet remplissant des fonctions s'oppose donc la conception de la monnaie comme **rapport social**. Le débat doit être abordé avec prudence, car, par exemple, un texte célèbre du très orthodoxe Milton Friedman traite de la monnaie de pierre dans l'île de Yap et développe une conception de la monnaie comme rapport social<sup>1</sup>. Soulignons aussi que pour les économistes autrichiens, comme pour nombre d'hétérodoxes, la monnaie est une institution et non un simple objet.

## Monnaie exogène et monnaie endogène

La formule « *helicopter Ben* » (et le dessin humoristique associé) pour désigner Ben Bernanke, alors président du conseil des gouverneurs de la Réserve fédérale des États-Unis<sup>2</sup>, illustrent bien la conception de la **monnaie exogène**. La monnaie serait, dans cette perspective, soit le résultat des découvertes d'or (au temps de l'étalon-or), soit le fait de la décision de la banque centrale : elle aurait son origine à l'extérieur de l'économie. Dans cette perspective, défendue par de nombreux économistes, de David Ricardo à Milton Friedman en passant par Jacques Rueff, ce qui importe, c'est la quantité de monnaie. Cette approche cor-

respond aux thèses de la Currency School qui aboutissent à la réforme monétaire de 1844 en Grande-Bretagne (*Bank Charter Act* ou *Peel's Act*). Celle-ci limite la création de monnaie en subordonnant toute émission supplémentaire de billets à une couverture en or à 100 %. Dans cette approche quantitative, il faut soit imposer une contrainte métallique stricte grâce à l'étalon-or (J. Rueff), soit imposer une règle intangible à la banque centrale en matière de création monétaire (c'est la fameuse « règle des  $k\%$  » de M. Friedman). En se fondant sur le mécanisme du multiplicateur de crédit, les tenants de la théorie de la monnaie exogène considèrent qu'il suffit de contrôler la quantité de monnaie centrale pour contrôler l'ensemble de la création de monnaie par le système bancaire.

À l'inverse, les conceptions de la **monnaie endogène** considèrent que ce qui est premier est le crédit accordé par les banques aux entreprises qui souhaitent financer leur production. Dans ses travaux postérieurs à la *Théorie générale*, Keynes intègre cet aspect à travers le « motif de finance » dans l'analyse de la demande de monnaie. C'est grâce à l'impulsion que constitue la création de monnaie que les entreprises vont pouvoir produire et distribuer des revenus, qui vont alimenter une demande adressée à l'économie (débouchés pour la production). Au total, les entreprises, à partir des produits de la vente, vont pouvoir rembourser le crédit initial. Il s'agit d'une analyse en termes de flux : la monnaie est considérée comme une dette et la dynamique de l'économie est assurée par la succession des opérations de création et de destruction de monnaie. Cette approche est déjà formulée par le **Banking School** au XIX<sup>e</sup> siècle. Les auteurs de ce courant, notamment Thomas Tooke, considèrent que l'inflation a des causes réelles plutôt que monétaires et mettent en avant la « loi du reflux » en vertu de laquelle toute monnaie créée (et injectée dans le circuit économique) finit par revenir dans les banques au moment du remboursement. Ce qui importe

[1] Friedman M. (1993), *La Monnaie et ses pièges*, Paris, Dunod.

[2] La formule et le dessin font référence à la politique de *quantitative easing* conduite par la banque centrale des États-Unis en réaction à la crise financière mondiale de la fin des années 2000.

## ZOOM

### MULTIPLICATEUR DE CRÉDIT ET DIVISEUR DE CRÉDIT

L'approche en termes de **multiplicateur de crédit** est celle qui, dit-on, a convaincu les banquiers qu'ils créaient bien de la monnaie. Dans cette approche, on suppose que les banquiers disposent de réserves excédentaires, en monnaie banque centrale, c'est-à-dire d'une quantité de monnaie centrale excessive par rapport à ce qui est nécessaire pour faire face aux retraits des clients et aux règlements interbancaires. On montre alors que l'octroi d'un prêt correspondant à cette réserve excédentaire va déclencher un effet en chaîne au sein du système bancaire, de telle sorte que, en fin de compte, la quantité totale de monnaie créée est un multiple du crédit initial. Le multiplicateur est égal à l'inverse du taux de réserve que les banques respectent pour disposer en permanence d'une quantité suffisante de liquidités. Dans cette conception, l'existence d'avoirs en monnaie banque centrale (base monétaire) est une condition nécessaire pour que les banques de second rang créent de la monnaie. La banque centrale a donc dans ce cas la maîtrise de la création monétaire, puisqu'elle décide de la création de la base monétaire.

Dans l'approche en termes de diviseur de crédit, on considère que la banque commence par accorder des crédits en sélectionnant des clients solvables. Ce faisant, elle crée de la monnaie de banque sans se préoccuper de disponibilités préalables en monnaie centrale. Ce n'est que dans un second temps qu'elle cherche, s'il y a lieu, à se procurer les liquidités qui lui sont nécessaires. Dans cette seconde représentation de la création monétaire par les banques, la banque centrale n'a pas la maîtrise de la création monétaire. En effet, lorsque les banques manquent de liquidité et qu'elles ne parviennent pas à se la procurer sur le marché interbancaire, la banque centrale est placée devant le fait accompli. La monnaie de banque a déjà été créée. Refuser de fournir des liquidités conduirait à un risque d'illiquidité pour une banque particulière, voire pour le système bancaire dans son ensemble. Dans cette approche, ce qui est premier est l'octroi de crédit par les banques de second rang, donc la contrainte de liquidité qui pèse sur ces banques est une contrainte souple. Réciproquement, comme le montre l'expérience récente de la politique monétaire, si la banque centrale fournit en abondance des liquidités aux banques de second rang, cela ne garantit pas que ces dernières vont accorder des crédits aux agents non bancaires.

Alain Beitone

donc, c'est la qualité de la monnaie (plutôt que sa quantité), c'est-à-dire la solvabilité des agents auxquels les banques accordent des crédits. De plus, on considère que les banques commencent par accorder des crédits et se préoccupent ensuite d'assurer leur liquidité en se procurant la monnaie banque centrale nécessaire aux règlements interbancaires. La création de monnaie doit être pensée en termes de diviseur de crédit<sup>3</sup>. La conception endogène de la monnaie est défendue par les post-keynésiens (N. Kaldor, J. Robinson, M. Lavoie, H. Minsky, etc.), les

« circuitistes » français (A. Parguez, F. Poulon, etc.), les régulationnistes (M. Aglietta, R. Boyer, etc.) et les conventionnalistes (A. Orléan, etc.).

## La neutralité de la monnaie et la théorie quantitative

### Les versions successives

Une longue tradition de l'analyse économique considère que la monnaie est neutre. Les for-

<sup>[3]</sup> Le Bourva J. (1962), « Création de monnaie et diviseur de crédit », *Revue économique*, vol. 13, n° 1.

## ZOOM

### MONNAIE ENDOGÈNE ET HYPOTHÈSE D'INSTABILITÉ FINANCIÈRE

C'est à partir de la conception endogène de la monnaie que Hyman Minsky a formulé son « hypothèse d'instabilité financière ».

En effet, si la création de monnaie résulte d'une demande de crédit des agents économiques et d'une offre de crédit des banques, il importe de s'interroger sur le comportement des uns et des autres.

Minsky montre que trois types de financement sont concevables : le financement prudent est le financement où les flux de revenu de l'emprunteur peuvent rembourser le principal et l'intérêt de la dette ; le financement spéculatif est le financement où le flux de revenu de l'emprunteur peut rembourser l'intérêt mais pas le principal (l'emprunteur doit alors s'endetter à nouveau à l'échéance pour rembourser sa dette) ; enfin le financement Ponzi est un financement où les flux de revenus ne permettent de

rembourser ni le principal ni la dette. L'analyse de Minsky conduit à démontrer, à partir du concept de « paradoxe de la tranquillité », que dans une situation d'endettement prudent avec de faibles taux d'intérêt, certains agents au profil plus risqué vont emprunter pour bénéficier de ces taux faibles. Les banques vont accepter de leur prêter pour bénéficier de rendements un peu plus élevés. Au total, le taux moyen d'endettement augmente et avec lui le niveau de risque. Au bout d'un certain temps, les niveaux d'endettement et de risque sont devenus trop forts, les banques contractent leurs crédits. Certains agents font défaut parce qu'ils ne peuvent pas se réendetter pour rembourser les crédits antérieurs. La période de boom marquée par la hausse du prix des actifs s'interrompt, certains agents vendent leurs actifs pour se désendetter (c'est le « moment Minsky »), ils précipitent la baisse des prix, ce qui aggrave la situation, etc.

L'analyse de Minsky met donc en évidence l'instabilité endogène des économies de marché et le lien entre marché, financement et spéculation.

Alain Beitone

mules ne manquent pas depuis la métaphore de la « monnaie-voile » jusqu'à l'affirmation selon laquelle « les produits s'échangent contre des produits ». La monnaie, dans cette perspective, ne change rien de fondamental à l'économie ou, pour le dire autrement, il n'y a pas de différence de nature entre une économie monétaire et une économie de troc. Les variations de la quantité de monnaie affectent les prix nominaux, mais pas les prix relatifs et par conséquent ne changent rien aux choix des agents (consommateurs ou producteurs). La monnaie facilite les échanges, mais elle n'affecte pas les grandeurs réelles de l'économie. On dit qu'il y a dichotomie entre la sphère monétaire et la sphère réelle. Cette conception dichotomiste conduit à la formulation de la théorie quantitative de la monnaie que l'on peut faire remonter à Jean

Bodin (*Réponse au paradoxe de Monsieur de Malestroit*, 1568) et qui a été reformulée par Irving Fisher au début du xx<sup>e</sup> siècle. Dans la version la plus simple, on peut écrire :

$$M.V = P.T$$

où  $M$  est la quantité de monnaie en circulation ;  $V$ , la vitesse de circulation de la monnaie ;  $P$ , le niveau général des prix, et  $T$ , le volume des transactions. Cette égalité est nécessairement toujours vraie. Elle n'a donc aucune portée explicative. Elle ne devient explicative que sous plusieurs hypothèses :

- la vitesse de circulation est stable ;
- il y a dichotomie (c'est-à-dire que les variables monétaires  $M$ ,  $V$  et  $P$  n'ont pas d'effet sur  $T$ ) ;
- la causalité va de  $M$  vers  $P$  (et non l'inverse).

Dans ces conditions (et dans ces conditions seulement), tout accroissement de  $M$  plus rapide que l'accroissement de  $T$  conduit à une hausse des prix.

Comme l'a fait remarquer le grand historien de la pensée économique Mark Blaug, il existe un lien nécessaire entre la loi des débouchés de Jean-Baptiste Say et la théorie quantitative de la monnaie. Si les agents sont rationnels et si la monnaie ne peut pas perturber leurs décisions, alors le système des prix relatifs déterminé par les marchés conduit à ce que tous les marchés soient simultanément en équilibre : il ne peut pas y avoir de surproduction générale et durable. On peut donc aussi faire un rapprochement avec l'analyse walrasienne relative à l'équilibre général des marchés qui repose sur une représentation des échanges marchands dans laquelle la monnaie n'a aucune place : il suffit qu'un bien quelconque soit choisi comme unité de compte afin de faciliter les échanges.

Au début des années 1960, J. Le Bourva pensait que cette approche quantitativiste avait définitivement perdu la partie dans les débats des économistes à propos de la monnaie. Sur ce point, il se trompait. On a assisté à un premier retour en force de la théorie quantitative avec les analyses de M. Friedman. Pour l'économiste de Chicago, « l'inflation est toujours et partout un phénomène monétaire ». La hausse de la quantité de monnaie peut avoir un effet réel à court terme, mais pas à long terme. En effet, les agents économiques ont des anticipations adaptatives : à court terme, les variations nominales de prix et de revenu peuvent les conduire à modifier leurs comportements (effets réels), mais à long terme, ils vont constater qu'ils se sont trompés et ils vont corriger leurs choix en fonction des grandeurs réelles de l'économie (et non plus des grandeurs nominales). La monnaie est donc active à court terme et neutre à long terme. Cette position est encore aujourd'hui celle de nombreux macroéconomistes.

La défense de la neutralité de la monnaie est radicalisée par les tenants de la théorie

des anticipations rationnelles. Pour eux, les agents sont capables en permanence de traiter au mieux toute l'information disponible sur la base du modèle économique pertinent. Ils ne sont donc jamais victimes d'illusion monétaire et prennent leurs décisions en fonction des grandeurs réelles de l'économie. La monnaie est neutre à court et à long termes (on dit parfois qu'elle est « superneutre »).

### **Neutralité monétaire et monnaie endogène : deux postures incompatibles**

L'idée de neutralité de la monnaie est évidemment incompatible avec les théories qui mettent l'accent sur son caractère endogène. Si la monnaie est une dette qui permet de s'acquitter de toutes les dettes, c'est parce qu'elle est créée en réponse à une demande de crédit des agents économiques. On ne saurait séparer une sphère réelle et une sphère monétaire. Schumpeter le soulignait avec force, l'analyse monétaire s'oppose radicalement à une analyse « réelle », car la prise en compte de la monnaie interdit de rendre compte de l'économie avec un modèle de troc. Les auteurs qui, comme Marx et Keynes, rejettent la loi de Say, rejettent donc aussi toute idée de neutralité de la monnaie. Les économistes autrichiens adoptent sur ce point une position similaire. Von Mises qualifie d'hérésie la théorie quantitative de la monnaie. Hayek, quant à lui, parle d'effet Cantillon pour indiquer que toute augmentation de la quantité de monnaie se traduit par une modification de la structure des prix relatifs, ce qui modifie les décisions des agents. Hayek, comme Gunnar Myrdal, s'appuie sur les analyses de Knut Wicksell pour qui l'écart entre le taux monétaire de l'intérêt et le taux d'intérêt naturel déclenche des effets cumulatifs qui peuvent conduire à une modification de la longueur du détour de production et à de l'inflation (lorsque le taux monétaire est inférieur au taux naturel) ou à du chômage (lorsque le taux monétaire est supérieur au taux naturel). Les variables monétaires ont donc, pour les économistes suédois comme

## ZOOM

### LA THÉORIE QUANTITATIVE À L'ÉPREUVE DE L'HISTOIRE

La théorie quantitative de la monnaie dispose de solides fondements académiques, mais elle est devenue aussi une sorte de poncif, de discours du sens commun, qui conduit à penser que toute augmentation de la quantité de monnaie provoque de l'inflation.

Pourtant, quelques épisodes historiques marquants conduisent à relativiser fortement cette idée reçue. Le premier est l'hyperinflation allemande des années 1920. Le travail fondateur d'Albert Aftalion (*Monnaie, prix et change*, 1927) montrait déjà sur la base d'un solide appareil statistique que l'augmentation de la quantité de monnaie était consécutive à la hausse des prix et non l'inverse. C'est la dépréciation du mark qui produit de l'inflation importée. Le mouvement est amplifié par les anticipations des agents économiques, conduisant à l'augmentation de la masse monétaire. L'analyse proposée par A. Orléan et M. Aglietta met l'accent sur le fait que l'inflation doit être comprise en liaison avec la conflictualité sociale globale. C'est le contexte international (la question du traité de Versailles et des réparations) et le contexte intérieur (écrasement de la révolution spartakiste, agitation d'extrême droite, grèves insurrectionnelles, attentats politiques) qui sape la légitimité de la république de Weimar. Si la monnaie est une institution qui exerce la violence sociale, la perte de la valeur de la

monnaie (l'inflation) est la manifestation de la crise de l'institution monétaire. La faible légitimité du pouvoir politique conduit à des difficultés croissantes de collecte de l'impôt. La fuite en avant dans l'inflation est la manifestation d'une profonde crise du lien social.

Le second exemple, c'est la situation de l'économie mondiale depuis 2008. En effet, les banques centrales, surtout dans les pays industrialisés, ont très fortement accru la taille de leur bilan. Afin d'éviter une crise de liquidité et de soutenir l'activité économique, elles ont monétisé des créances (qui se retrouvent à l'actif de leur bilan) et émis en contrepartie de la monnaie centrale (qui se trouve au passif de leur bilan). Elles ont, ce faisant, refinancé pratiquement sans limitation les banques de second rang et à des taux d'intérêt pratiquement nuls. Pour autant, cette création abondante de monnaie n'a pas conduit à l'inflation. Bien au contraire, dans la zone euro, c'est la déflation qui menace. Il n'y a donc pas de relation mécanique entre quantité de monnaie et inflation. L'explication est assez simple : l'existence de réserves excédentaires en monnaie banque centrale ne suffit pas à convaincre les banques d'accorder des crédits lorsque leurs anticipations de croissance sont pessimistes et lorsqu'elles pensent que l'octroi de crédit serait trop risqué. Réciproquement, les ménages et les entreprises ne sollicitent pas de crédit lorsqu'ils pensent que la demande va continuer à stagner et que le chômage va rester élevé.

**Alain Beitone**

pour les économistes autrichiens, des effets réels : la monnaie n'est pas neutre. On le voit, le clivage entre les économistes libéraux et les autres n'est pas pertinent à propos du quantitativisme et de la neutralité de la monnaie.

## Monnaie, marché et institutions

Pour comprendre la nature de la monnaie et son rôle dans le système économique, il faut partir de la compréhension de ce qu'est une

économie marchande. C'est ce que souligne M. Aglietta : « Concevoir la monnaie comme le médiateur de la socialisation des sujets économiques, c'est affirmer que l'analyse de la monnaie et celle de l'économie marchande sont un seul et même problème. On est aux antipodes des théories naturalistes de l'économie qui ne voient dans la monnaie qu'un intermédiaire technique commode des échanges, ou même une marchandise particulière. » (Aglietta, 1988, p. 98).

Une économie marchande est en effet une économie dans laquelle des producteurs décident indépendamment les uns des autres de l'usage qu'ils vont faire des ressources productives. Dans le langage de Marx, on peut dire que des producteurs privés décident de l'allocation du travail social. Rien ne garantit *a priori* que cet usage du travail social soit conforme aux attentes de la société. Bien sûr, si le producteur produit pour son propre usage, cela n'a aucune importance, mais s'il produit pour vendre sur le marché, la question de la validation sociale des travaux privés se pose. On est en présence d'un problème de coordination : il s'agit de rendre cohérentes entre elles des décisions qui ne le sont pas nécessairement *ex ante*. C'est ici qu'intervient la monnaie : elle est une institution qui permet de surmonter l'opposition entre le privé et le social : un instrument de validation sociale des travaux privés. Pour M. Aglietta, « la monnaie est un rapport global entre des centres de décision économiques et la collectivité qu'ils forment grâce auquel les échanges entre ces agents acquièrent une cohérence<sup>4</sup> ». Lorsqu'un producteur individuel (qui peut être une firme de grande taille) présente sa production sur le marché (c'est le stade de la réalisation dans le vocabulaire de Marx), il se confronte, à travers la monnaie, à une procédure de validation. Réciproquement, la monnaie qu'il reçoit en paiement, si sa production est validée, constitue un droit sur l'ensemble de la production sociale. Selon la judicieuse formule de G. Simmel, la monnaie est « une lettre de change sur laquelle le nom de l'intéressé n'est pas porté ». Autrement dit, la monnaie est une créance d'un montant déterminé sur l'ensemble des biens et services qui résultent de l'utilisation du travail social. La monnaie permet notamment de rendre commensurables des biens et des services singuliers : elle permet d'en mesurer la valeur à travers les prix. Un débat subsiste sur ce point. Pour Marx et pour les classiques, la monnaie ne joue ce rôle que parce que la monnaie a, directement ou indirectement, une valeur travail et qu'elle permet ainsi d'éva-

<sup>[4]</sup> Aglietta M. (1986), *La Fin des devises clés*, Paris, La Découverte, Coll. « Agalma », p. 17.

luer le temps de travail dépensé pour produire les marchandises. Pour les Autrichiens, la valeur est subjective et la monnaie permet de surmonter et de mettre en relation les préférences subjectives des agents. Pour les auteurs conventionnalistes, comme A. Orléan, il n'y a pas de valeur intrinsèque des marchandises ; la valeur, comme la monnaie, sont des conventions sociales. Mais le point commun de toutes ces analyses est bien que la monnaie est une institution indispensable à la coordination des décisions des agents dans une économie de marché décentralisée.

## Monnaie, crédit et production

Cette analyse de la monnaie comme institution, lien social et procédure de coordination permet de comprendre le lien entre monnaie, crédit et production. Mais il faut pour cela produire une analyse en termes de circuit monétaire de production, qui est, selon B. Schmitt, commune à Marx, à Keynes et aux Autrichiens : la monnaie et la production sont nécessairement articulées. Ce sont les banques et les crédits qu'elles accordent qui constituent le premier temps (et l'impulsion) du circuit. Un entrepreneur s'endette nécessairement, comme le soulignait Schumpeter. La monnaie est donc une dette (de l'emprunteur) et une créance (de la banque). En décidant d'octroyer le crédit, la banque valide par anticipation la production que l'entrepreneur se propose de réaliser. Elle fait l'hypothèse, après évaluation du projet de l'entrepreneur et des risques qu'il comporte, que les moyens de production ainsi financés conduiront à une production qui sera validée par le marché (c'est le « pari bancaire »). Si les choses se passent comme prévu, à l'issue du cycle de production, l'entrepreneur est en mesure de rembourser sa dette, qui disparaît en même temps que la créance correspondante et la monnaie est détruite. Dans cette perspective, les banques ne sont pas de simples intermédiaires ; elles jouent un rôle décisif dans la sélection des projets productifs qui seront mis en œuvre. Elles exercent ainsi la fonc-

tion essentielle de création de monnaie sans laquelle la croissance économique est impossible. Comme le soulignait Evsey Domar, la croissance ne dépend pas du niveau de l'investissement mais de l'augmentation de l'investissement : la simple mobilisation de l'épargne préalable ne permet, au mieux, que de maintenir le niveau de la production. Il n'y a croissance que si, *ex ante*, du fait de la création monétaire des banques, l'investissement est supérieur à l'épargne. On ne peut séparer

la création monétaire et la dynamique de la production.

La crise de 2008 et ses conséquences durables en termes de croissance rappellent le rôle décisif de la monnaie et du crédit dans la dynamique économique : la compréhension des phénomènes monétaires est indispensable à la régulation de l'économie.

---

## POUR EN SAVOIR PLUS

---

❖ **AGLIETTA M.** (1988), « L'ambivalence de l'argent », *Revue française d'économie*, n° 3.

❖ **AGLIETTA M. et ORLÉAN A.** (2002), *La Monnaie entre violence et confiance*, Paris, Odile Jacob.

❖ **COUPPEY-SOUBEYRAN J.** (2012), *Monnaie, banques, finance*, Paris, PUF, coll. « Licence », 2<sup>e</sup> éd.

❖ **DELAPLACE M.** (2013), *Monnaie et financement de l'économie*, Paris, Dunod, 4<sup>e</sup> éd.

❖ **LAVIGNE M. et POLLIN J.-P.** (1997), *Les Théories de la monnaie*, Paris, La Découverte, coll. « Repères ».

❖ **MINSKY H.** (2013), *L'Hypothèse d'instabilité financière*, Montreuil-sur-Breche, Diaphane.

❖ **MISHKIN F.** (2014), *Monnaie, banque et marchés financiers*, Paris, Pearson, 10<sup>e</sup> éd.

❖ **PLIHON D.** (2013), *La Monnaie et ses mécanismes*, Paris, La Découverte, coll. « Repères », 6<sup>e</sup> éd.

❖ **TUTIN CH.** (2014), *Les Grands Textes de la pensée monétaire*, Paris, Flammarion, coll. « Champs ».

---



Agent économique central, l'entreprise a paradoxalement longtemps été réduite dans la théorie économique à un agent individuel maximisant mécaniquement son profit. Avant une époque récente, seules des analyses en marge de l'approche dominante se sont intéressées à son organisation, sa « gouvernance » ou ses fonctions.

Après un rappel de ces travaux fondateurs, Olivier Weinstein fait le point sur les théories récentes de l'entreprise. Si une nouvelle analyse dominante s'est construite autour de la conception de la firme comme « nœud de contrats », des approches alternatives telles que la firme comme « système de compétences » se sont développées en parallèle.

Problèmes économiques

# L'entreprise dans la théorie économique<sup>1</sup>

© OLIVIER WEINSTEIN

Professeur émérite à l'université Paris-XIII

L'entreprise n'a occupé, jusqu'aux années 1970, qu'une place très marginale dans la tradition dominante des enseignements économiques. Qu'il s'agisse de la théorie de l'équilibre général, des théories des marchés ou de la plus grande partie de l'économie industrielle, la firme est réduite à peu de choses : une « firme point », c'est-à-dire assimilée à un agent individuel, sans prise en considération de son organisation interne, et une « firme automate » qui, supposée parfaitement rationnelle comme tout agent économique, ne fait que transformer, de manière efficiente, des facteurs de production en produits et s'adapter mécaniquement à des contraintes techniques et des environnements donnés. Cette vision de la firme se comprend relativement à ce qu'a été pendant longtemps l'objet central de la micro-économie : l'étude des marchés et des méca-

nismes de prix. Pour trouver des analyses s'intéressant à la firme en tant que telle, il fallait considérer des travaux en marge de la théorie économique standard.

## En marge de la théorie économique standard : trois approches de la firme

### Berle et Means : la « révolution managériale »

Une première analyse majeure de l'entreprise se trouve chez Adolf Berle et Gardiner Means. Leur ouvrage publié en 1932, *L'Entreprise moderne et la propriété privée*<sup>2</sup>, point de départ de ce que l'on a appelé la « révolution managériale », a eu une influence considérable. La thèse centrale qui a été retenue du livre est que le développement de la grande société par actions et la dispersion de la propriété entre un grand nombre d'actionnaires tend à entraîner la séparation de

<sup>[1]</sup> Ce texte développe et actualise l'article : « Comment la théorie économique tente d'apprivoiser l'entreprise », *Alternatives économiques*, hors-série n° 43, 1<sup>er</sup> trimestre 2000. Il a été publié dans *Cahiers français* n° 345, Découverte de l'économie. Concepts, mécanismes et théories économiques, Paris, La Documentation française, juillet-août 2008.

<sup>[2]</sup> Berle A. A. et Means G. C. (1932), *The Modern Corporation and Private Property*, New York, Macmillan.

la propriété et du contrôle de l'entreprise ; le pouvoir de décision passe alors des actionnaires aux « managers ». C'est à partir de là que va être posée la question centrale de la « gouvernance d'entreprise », qui est revenue sur le devant de la scène depuis une vingtaine d'années. Indépendamment de cette thèse, l'ouvrage de Berle et Means est important parce qu'il offre un mode de théorisation de l'entreprise marqué par trois traits :

- les caractéristiques, le fonctionnement et le comportement de la firme se comprennent en considérant les rapports entre différents groupes, aux intérêts propres : les actionnaires et les managers, mais également les salariés et les fournisseurs de crédits ;
- une question centrale est celle de savoir qui contrôle l'entreprise ;
- le cadre institutionnel, en l'occurrence le système de la société par actions et la constitution de marchés financiers, joue un rôle essentiel dans la structuration de la firme.

### Cyert et March : l'approche « béhavioriste » de la firme

L'ouvrage de Richard Cyert et James March de 1963, *A Behavioural Theory of the Firm*<sup>3</sup>, qui ouvre la voie à la théorie dite « béhavioriste », constitue une deuxième étape essentielle dans l'histoire des théories de la firme. C'est là que s'affirme la vision de l'entreprise comme une organisation complexe, constituée de groupes aux intérêts divers, dans des rapports simultanés de conflits et de coopération. L'analyse de Cyert et March ajoute à cela un autre aspect essentiel qui sera repris et développé par les évolutionnistes : la firme est le lieu d'apprentissages collectifs. Ainsi apparaissent les deux dimensions clés autour desquelles vont se construire les théories actuelles de la firme : d'un côté l'analyse des modes de gestion des conflits d'intérêts, de l'autre les conditions de constitution d'une capacité collective à produire.

## L'approche de Chandler

Il faut enfin évoquer les travaux d'Alfred D. Chandler, le grand historien américain de l'entreprise. Dans trois ouvrages majeurs<sup>4</sup>, il offre des matériaux et des réflexions essentielles qui ont eu une influence notable sur la théorisation de la firme. Tout d'abord, une caractérisation de l'entreprise moderne comme institution complexe, fondée sur un système de coordination administrative de nature hiérarchique, ce qui rejoint l'analyse de Coase que nous allons voir plus loin. Ensuite, une analyse des liens entre les changements dans la technologie et la concurrence et les transformations de l'entreprise. Enfin, il met en évidence l'importance des formes de propriété et de contrôle et des conditions de coopération entre agents pour expliquer la diversité nationale des formes d'entreprises et des formes de capitalisme.

## La nouvelle orthodoxie : la firme comme « nœud de contrats »

### Firme et coûts de transaction

Les analyses précédentes se situent en dehors du noyau central de l'enseignement économique. Elles sont sur bien des points en opposition avec l'orthodoxie néoclassique. À partir des années 1970, la théorie économique de la firme va être profondément renouvelée, sur une autre base : la redécouverte d'un article de Ronald Coase de 1937, qui n'avait connu jusque-là que peu d'audience<sup>5</sup>. Coase soulève la question de « la nature de la firme » : qu'est-ce qu'une firme, et pourquoi les firmes existent-elles ? Coase propose un type de réponse qui reste, dans sa forme générale, couramment admise : la firme constitue un mode de coordination économique alternatif au marché. Alors que sur le marché la coordination des comportements des individus se fait par le système de prix, la firme se caractérise par une coor-

[3] Cyert R. M. et March J. G. (1962) [1963], *A Behavioural Theory of the Firm*, Oxford, Wiley-Blackwell.

[4] Chandler A. D. (1972), *Stratégies et structures de l'entreprise*, Les Éditions d'organisation. Chandler A. D. (1988), *La Main visible*, Economica. Chandler A. D. (1990), *Scale and Scope*, Cambridge (Mass.), Harvard University Press.

[5] Coase R. (1937), « The Nature of the Firm », *Economica*, novembre. Traduction française : « La nature de la firme », *Revue française d'économie* n° 2.

dination administrative, par la hiérarchie. Il reste alors à se demander pourquoi le recours à une coordination administrative peut être nécessaire. La réponse de Coase est que la coordination par les prix entraîne des coûts, ignorés dans les analyses standards du marché, ce que l'on appellera par la suite des coûts de transaction. Quand ces coûts sont supérieurs aux coûts d'organisation interne, la coordination dans la firme s'impose. On trouve ainsi chez Coase deux thèses, qui seront fortement discutées par la suite : (1) firme et marché constituent deux modes de coordination profondément différents ; (2) ce qui caractérise fondamentalement la firme, c'est l'existence d'un pouvoir d'autorité ; la firme est en effet une organisation hiérarchique.

L'analyse de Coase constitue le point de départ de la conception de la firme aujourd'hui dominante chez les économistes : la vision contractuelle. Dans cette perspective, la firme s'analyse comme un système de relations contractuelles spécifiques entre agents. Au centre de ces analyses se trouvent les problèmes résultant de la divergence des intérêts des parties, supposées guidées par des motivations égoïstes, et des asymétries d'information : un agent peut disposer d'informations que les autres n'ont pas, et en tirer un profit personnel, au détriment des autres<sup>6</sup>. Le problème est alors de trouver le système contractuel le plus efficient, en fonction de différents paramètres, et en particulier de contraintes techniques et de la nature des informations détenues par les parties.

On peut identifier trois grandes variantes de cette vision contractuelle : (1) la vision défendue par ceux qui se situent dans un cadre néoclassique renouvelé, incluant principalement la **théorie de l'agence** ; (2) la **théorie des coûts de transaction**, développée principalement par O. Williamson ; (3) la théorie la plus récente, la **théorie des contrats incomplets**, qui tente de reformuler la théorie des coûts de transaction.

## Droits de propriété et théorie de l'agence

Les néoclassiques tentent de rendre compte des problèmes de la firme sans remettre en question leur vision du monde et leur approche de l'économie, c'est-à-dire une représentation fondée sur des comportements individuels supposés parfaitement rationnels et, au moins pour les néoclassiques les plus strictement libéraux, la démonstration, ou l'affirmation, qu'un système de relations « libres » entre agents conduit à l'optimum social. Cela a été fait en développant une théorisation qui s'appuie sur deux corps d'analyse complémentaires, les droits de propriété et la théorie de l'agence. La théorie des droits de propriété est au cœur de l'approche néoclassique des institutions. Son objet est de montrer comment les systèmes de droits de propriété agissent sur les comportements individuels et sur l'efficacité des systèmes économiques, en insistant sur les vertus des droits de propriété privés. Dans ce cadre, la firme est caractérisée par une structure particulière de droits de propriété, définie par un ensemble de contrats. Un « bon » système de droits de propriété est celui qui permet de profiter des avantages de la spécialisation, et du fait que les différents agents ne détiennent pas les mêmes informations et connaissances, et qui assure un système efficace d'incitation. Dans un article célèbre, Armen Alchian et Harold Demsetz (1972)<sup>7</sup> tentent de démontrer sur ces bases que la firme capitaliste « classique », l'entreprise individuelle, est la forme d'organisation la plus efficace quand la technologie impose le « travail en équipe », c'est-à-dire quand le produit résulte de la coopération de différents agents, sans qu'il soit possible de mesurer la contribution individuelle de chacun.

La théorie de l'agence est aujourd'hui le cadre d'analyse standard des questions d'organisation dans les approches néoclassiques ; elle complète la théorie des droits de propriété. Elle repose sur la notion de relation d'agence, qui sert à formaliser les relations entre des

<sup>[6]</sup> Sur ce point, voir dans ce même numéro l'article d'Anne Corcos et François Pannequin, p. 58.

<sup>[7]</sup> Alchian A. et Demsetz H. (1972), « Production, Information Costs, and Economic Organization », *American Economic Review* vol. 62, décembre.

individus aux intérêts différents, et à déterminer des contrats incitatifs optimaux adaptés aux situations les plus diverses. L'application de la théorie de l'agence à l'analyse de la firme est marquée par l'article fondateur de Michael Jensen et William Meckling (1976)<sup>8</sup>. En approfondissant l'analyse des propriétés des structures contractuelles de la firme, ce courant de pensée se propose de démontrer l'efficacité des formes organisationnelles caractéristiques du capitalisme contemporain, et en particulier de la société par actions. C'est dans le cadre de la théorie de l'agence que se sont développées les analyses récentes sur la « *corporate governance* » et les thèses qui soutiennent la supériorité du modèle anglo-saxon pour lequel les managers doivent défendre les intérêts des seuls actionnaires (le modèle « *shareholder* »). Parallèlement, les théoriciens des droits de propriété se sont attachés à démontrer l'inefficacité de l'entreprise publique et de la firme autogérée. Notons ici un aspect majeur de cette vision qui inspire des conceptions ultralibérales. La firme étant caractérisée comme un « nœud de contrats » entre les détenteurs des différents facteurs de production, les auteurs que nous venons de citer s'opposent à Coase sur deux points : (1) il n'y a dans la firme aucune relation d'autorité, mais simplement des rapports contractuels libres ; (2) il n'y a pas d'opposition entre firme et marché : la firme n'est pas fondamentalement différente d'un marché, elle est un « marché privé ». Cette conception conduit à nier toute spécificité à la relation d'emploi : le contrat de travail est censé être similaire à un contrat commercial<sup>9</sup>. Pour apprécier ce point de vue, il est utile de rappeler un point clé des analyses de Marx, que l'on retrouve en économie du travail : le salarié de l'entreprise capitaliste ne vend pas son travail, c'est-à-dire des prestations de service comme le fait un travailleur indépendant, il loue sa force de travail. La firme acquiert par contrat le droit d'utiliser à son gré les compétences du salarié, et de le diriger. Il y a bien ainsi entre l'employeur et l'employé une *relation d'autorité*, comme le

dit Coase en 1937, et comme l'a montré Simon dans un article fondamental de 1951<sup>10</sup>.

## La théorie des coûts de transaction

La théorie de Williamson est compatible avec cette vision des choses et se situe directement dans le prolongement de Coase<sup>11</sup>. Elle développe l'analyse des coûts de transaction pour expliquer, en particulier, dans quels cas la firme s'impose comme mode de coordination, c'est-à-dire dans quel cas l'intégration d'une activité dans la firme sera préférée au recours au marché, l'extériorisation.

Williamson se distingue des approches néo-classiques que nous venons de voir, par ses hypothèses sur le comportement des agents économiques, et sur les caractères des contrats. Il reprend la théorie de la rationalité limitée d'Herbert Simon : les agents ont des capacités cognitives limitées, ils ne peuvent pas, dans des environnements complexes, envisager tous les événements possibles et calculer parfaitement les conséquences de leurs actes. En conséquence, les contrats seront, le plus souvent, des **contrats incomplets**, qui n'envisagent pas tous les événements possibles. Le problème est alors de savoir ce qui va se passer, après signature d'un contrat, en cas d'événement imprévu. L'incomplétude des contrats laisse une marge de manœuvre aux parties, ce qui va permettre les comportements opportunistes, la manipulation de l'information par les agents. C'est là que se situe, pour Williamson, le problème essentiel : l'opportunisme, et la manière de s'en protéger, sont au centre des choix organisationnels. Ce problème se pose tout particulièrement quand, pour une transaction, les agents doivent réaliser des investissements spécifiques, non réutilisables en dehors de la transaction, qui les rendent dépendants l'un de l'autre. Chaque partie peut alors craindre que l'autre s'approprie le bénéfice de la transaction, qu'il y ait « *hold-up* ». C'est essentiellement dans ce cas où une transaction implique des investissements fortement spécifiques que, selon la théorie des coûts de transaction, la coordination dans

[8] Jensen M. C. et Meckling W. H. (1976), « Theories of the Firm: Managerial Behaviour, Agency Costs, and Ownership Structure », *Journal of Financial Economics*, vol. 3, n° 4, octobre. Voir aussi : Jensen M. C. (1998), *Foundations of Organizational Strategy*, Cambridge (Mass.), Harvard University Press.

[9] C'est le point de vue défendu par Alchian et Demsetz, et repris par Jensen et Meckling.

[10] Simon H. A. (1951), « A Formal Theory of the Employment Relationship », *Econometrica*, vol. 19.

[11] Voir, par exemple : Williamson O. E. (1994), *Les Institutions de l'économie*, InterÉditions.

la firme sera préférée à la coordination par le marché. À la suite de R. Gibbons, on peut estimer que la théorie des coûts de transaction propose en fait deux théories de la firme : la théorie du « *hold-up* », et une « théorie de l'adaptation ». Pour cette dernière, le propre de la firme, fondée sur un rapport d'autorité, est de donner à l'une des parties le pouvoir de prendre de manière discrétionnaire les décisions adaptées aux événements, et donc de permettre une adaptation de l'organisation, sans renégociation. Son avantage est d'accroître la capacité d'adaptation de l'organisation à son environnement.

La firme, pour Williamson, est ainsi un système contractuel particulier, un « arrangement institutionnel » caractérisé par un principe hiérarchique selon lequel c'est la direction de l'entreprise qui a le pouvoir de prendre les décisions en cas d'événement non prévu par les contrats, et qui permet de limiter les risques liés à l'opportunisme. La grande importance donnée dans cette vision à la spécificité des actifs est très discutée, d'un point de vue empirique aussi bien que théorique (Coase notamment s'est opposé sur ce point à Williamson). Il est également permis de se demander en quoi l'internalisation limiterait les comportements opportunistes, et si la firme (comme le marché) peut fonctionner sans un minimum de confiance, incompatible avec des comportements purement opportunistes.

### **Théorie des contrats incomplets et droits de propriété**

La théorie plus récente des contrats incomplets, développée en particulier par Oliver Hart<sup>12</sup>, se propose essentiellement de reformuler la théorie des coûts de transaction et de l'intégrer dans le cadre analytique de la nouvelle microéconomie néoclassique, en conservant une hypothèse de rationalité parfaite. Elle propose une formalisation différente, en particulier en ce qui concerne la cause et les implications de l'incomplétude des contrats<sup>13</sup>. Elle y ajoute la nécessité de prendre en compte les droits de propriété.

Ce qui conduit Oliver Hart à insister sur un point essentiel : l'incomplétude des contrats donne une grande importance à la définition des rapports de pouvoir dans les relations contractuelles. La firme est définie comme un ensemble d'actifs (non-humains), soumis à une propriété unifiée, et à un contrôle unifié. La définition des arrangements institutionnels et des systèmes de droits de propriété vise à « allouer le pouvoir entre les agents ». Cette allocation du pouvoir est liée à la définition des droits de propriété sur les facteurs de production, et en particulier sur le capital : on retrouve ainsi, comme le note O. Hart lui-même, un aspect des thèses marxistes !

Il reste que, comme le dit B. Holmstrom<sup>14</sup>, la formalisation proposée implique que c'est un individu, et non pas la firme, qui est censé détenir les actifs, et passer un contrat avec d'autres parties. Cette vision qui reste dans une logique de contrats entre individus rend difficile la compréhension de la firme comme une entité cohérente<sup>15</sup>, et ainsi la prise en compte de la production proprement dite. C'est là une des limites majeures des approches contractuelles.

## **La recherche d'une vision alternative**

### **La firme comme système de compétences**

Les théories contractuelles abordent, à leur manière, une dimension de l'entreprise : les modes de gestion des conflits d'intérêts entre les parties. L'autre dimension, tout aussi importante, reste absente : la firme a pour fonction de produire des marchandises, elle repose sur la constitution d'une capacité collective à produire, à gérer, à innover. Depuis une trentaine d'années se sont multipliées des analyses dites « fondées sur les ressources » ou « fondées sur les compétences » qui se focalisent sur cette deuxième question. Elles se présentent comme alternatives ou complémentaires des approches contractuelles. La théorie évolutionniste de la firme peut être

[12] Voir, par exemple : Hart O. (1995), *Firms, Contracts and Financial Structure*, Oxford, Clarendon Press.

[13] Sur ce qui différencie théorie des coûts de transaction et théorie des droits de propriété, on peut se reporter à Gibbons R. (2005), *op. cit.*, et Williamson O.E. (2000), « The New Institutional Economics: Taking Stock, Looking Ahead », *Journal of Economic Literature*, vol. 38, n° 2.

[14] Holmstrom B. (1999), « The Firm as a Subeconomy », *Journal of Law, Economics and Organization*, vol. 15, n° 1.

[15] Voir sur cette question : Biondi Y., Canziani A. et Kirat T. (eds.) (2007), *The Firm as an Entity*, New York, Routledge.

rattachée à ce courant. Les travaux qui se situent dans cette perspective sont très divers et ne constituent pas une théorie unifiée. Ils se distinguent cependant tous des approches contractuelles par les questions qu'ils traitent et par leurs fondements théoriques.

## Les compétences au cœur des performances de la firme

L'objet premier de ces analyses est d'expliquer pourquoi certaines firmes ont durablement des performances supérieures, ou plus généralement, chez les évolutionnistes, « pourquoi les firmes diffèrent durablement dans leurs caractéristiques, comportements et performances<sup>16</sup> ». Une réalité essentielle, que les approches contractuelles peuvent difficilement expliquer. La réponse à cette question va être recherchée dans l'analyse des dynamiques d'accumulation de connaissances et de compétences spécifiques par les firmes. Chaque firme détient des compétences qui lui sont propres, que les autres firmes ne peuvent pas acquérir rapidement, parce qu'elles sont difficiles à imiter et qu'elles ne peuvent être acquises sur le marché. Et cela en particulier parce que les compétences reposent en partie sur des connaissances tacites, non formalisées, qui sont difficilement transférables entre individus ou entre organisations. Ainsi, l'activité et la compétitivité de chaque firme repose sur un ensemble de compétences « foncières » (*core capabilities*). On pourrait dire que l'on a ainsi une vision de la firme comme « nœud de compétences » plutôt que comme nœud de contrats. Ces analyses se situent pour la plupart dans une perspective dynamique, donnant une grande importance aux changements technologiques et aux processus d'évolution. D'où l'accent mis par certains sur les compétences dynamiques, qui donnent à la firme la capacité de se transformer, de suivre et d'impulser les changements technologiques et les transformations de l'environnement. Chandler, qui s'est fait le défenseur d'une théorisation de la firme à partir des compétences, contre la théorie des coûts de transaction, se situe dans cette perspective.

Ce type d'approche de la firme peut donner une réponse à la question de Coase sur le choix entre firme et marché, totalement différente des réponses contractuelles. Une firme serait conduite à choisir entre l'internalisation d'une activité et le recours au marché, essentiellement en fonction des compétences qu'elle détient. Ce qui signifie que deux firmes pourront, de manière rationnelle, faire des choix différents. Cette manière d'aborder la question des frontières de la firme est celle que l'on trouve dans un article ancien de George B. Richardson (1972)<sup>17</sup>. Plus généralement, il faut admettre que les firmes pourront, y compris dans un même secteur, avoir des formes d'organisation et de gouvernance différentes. Ce qui va à l'encontre de l'idée dominante selon laquelle il y aurait toujours, dans un contexte donné, un mode d'organisation efficient unique qui devrait s'imposer à tous.

## Formation et évolution des compétences

L'accent placé sur les compétences de la firme conduit à s'interroger sur les conditions dans lesquelles ces compétences se forment et évoluent. Cette question est abordée depuis longtemps par la théorie des organisations ; elle est particulièrement développée par les évolutionnistes dans un cadre théorique qui se veut une alternative au paradigme néoclassique. L'analyse repose d'abord sur une théorisation des comportements individuels construite dans la lignée de H. Simon et J. March, et mobilisant des avancées des sciences cognitives. Elle donne une place centrale à l'étude des processus d'apprentissage. Elle aborde tout particulièrement les apprentissages organisationnels, afin de comprendre comment se construisent dans la firme des apprentissages collectifs et des compétences collectives qui vont se matérialiser dans un ensemble de « routines organisationnelles ». L'analyse de la firme renvoie ainsi, comme dans les approches contractuelles, aux modes de coordination entre les individus et les groupes, mais il s'agit d'une coordination « cognitive » visant à combiner les connaissances et les compé-

[16] Dosi G. et Marengo L. (1994), « Some Elements of an Evolutionary Theory of Organizational Competences » in England R. W. (ed.): *Evolutionary Concepts in Contemporary Economics*, Ann Arbor, University of Michigan Press.

[17] Richardson G. B. (1972), « The Organisation of Industry », *Economic Journal*, n° 82.

tences individuelles et à favoriser les apprentissages, par opposition à une coordination « politique » qui vise à rendre compatibles les intérêts des individus.

## Forces et faiblesses de l'approche par les compétences

L'accent placé par ces approches sur les problèmes de connaissances et d'apprentissage, sur la dimension cognitive des organisations est à la fois leur force et leur faiblesse. Leur force, car elles touchent ainsi à une question essentielle pour comprendre ce que sont les firmes et comment elles fonctionnent, et cela tout particulièrement quand la capacité d'innovation est la condition de leur survie. Leur faiblesse, dans la mesure où cela conduit à ignorer, le plus souvent, les dimensions conflictuelles des rapports économiques, et le fait que les firmes capitalistes sont des organisations particulières, dont la finalité n'est pas tant la production pour elle-même, que la recherche du profit.

Le problème central que doit traiter la théorie de la firme est celui de la combinaison de la dimension cognitive et de la dimension « politique ». On peut trouver des éléments qui vont dans cette direction dans les analyses de la firme japonaise, chez M. Aoki notamment, qui montrent comment les systèmes d'incitation et les modes d'organisation et de gouvernance peuvent favoriser les apprentissages et l'innovation.

## Autres alternatives

Le dépassement des approches contractuelles peut se faire dans des directions très différentes de celle des approches fondées sur les compétences. On évoquera pour terminer deux axes de réflexion qui touchent à des questions essentielles. Il y a tout d'abord la question du pouvoir, qui ne fait qu'affleurer dans certains travaux

contractualistes. Il est possible, tout en restant sur le terrain des approches contractuelles, de construire une théorisation de la firme centrée sur les questions de pouvoir. C'est ce qu'ont fait les radicaux américains en analysant la firme à partir d'une double fonction : une fonction de coordination et d'allocation, seule prise en compte dans la vision néoclassique, et une fonction disciplinaire, assurée par un système de pouvoir. En faisant de celle-ci la fonction principale, on est conduit à une tout autre vision de l'entreprise et de sa logique d'organisation, et l'on peut soutenir, que, à l'encontre des thèses des théoriciens de l'agence, la firme capitaliste, et le système de droits de propriété qui la fonde sont profondément inefficients<sup>18</sup>.

On peut enfin chercher à resituer l'analyse de la firme et de ses formes dans le cadre des systèmes institutionnels où elle s'insère. C'est ce que fait, d'une certaine manière, la théorie des conventions en analysant la diversité des formes de rationalité et des modes de coordination. Ce qui lui permet d'avancer une lecture originale des conditions de la coordination dans différents modèles d'entreprises, en développant les implications de l'incomplétude des contrats, et notamment du contrat de travail<sup>19</sup>. C'est ce qu'a fait, d'une autre manière, la théorie de la régulation en analysant la firme fordiste comme composante d'un certain système de formes structurelles, ou encore les travaux d'analyse institutionnelle comparée<sup>20</sup>. Cela conduit à mettre en évidence la dimension historique des institutions et des formes d'organisation et à considérer la diversité des formes de firme et des « modèles productifs »<sup>21</sup>, notamment selon les contextes nationaux. Il y a dans ces différentes voies des éléments essentiels pour comprendre les traits fondamentaux de la firme capitaliste et les formes qu'elle prend dans la période actuelle.

## POUR EN SAVOIR PLUS

❖ **CORIAT B. et WEINSTEIN O.** (1995), *Les Nouvelles théories de l'entreprise*, Paris, Librairie générale française.

❖ **EYMARD-DUVERNAY F.** (2004), *Économie politique de l'entreprise*, Paris, La Découverte.

[18] Voir par exemple, Bowles S., Gintis H. et Gustafsson B. (eds.) (1993), *Market and Democracy : Participation, Accountability and Efficiency*, Cambridge, Cambridge University Press.

[19] Voir notamment, Orléan A. (ed.) (2004), *Analyse économique des conventions*, Paris, PUF, coll. « Quadrige » et Eymard-Duvernay F. (2004), *Économie politique de l'entreprise*, Paris, La Découverte.

[20] Voir notamment, Aoki M. (2006), *Fondements d'une analyse institutionnelle comparée*, Paris, Albin Michel.

[21] Boyer R. et Freyssinet M. (2000), *Les Modèles productifs*, Paris, La Découverte.

Les banques sont des acteurs essentiels de l'économie puisqu'elles jouent un rôle-clé dans son financement. Sous l'effet de la libéralisation financière amorcée dans les années 1980, elles ont profondément diversifié leurs activités. Celles en lien avec les marchés de capitaux ont connu un essor considérable relativement aux activités traditionnelles, ce qui apparaît nettement dans l'évolution de la structure du bilan des établissements bancaires. La taille des bilans s'est aussi fortement accrue, de telle sorte que les grands groupes du secteur ont aujourd'hui en Europe un poids comparable au PIB de leur pays d'origine. Après avoir rappelé le rôle et les activités des banques, Jézabel Couppey-Soubeyran fait le point sur leurs évolutions récentes. Celles-ci posent un certain nombre de problèmes, qui se sont notamment illustrés lors de la crise financière de 2007-2008.

Problèmes économiques

# Comment (dys)fonctionnent les banques ?<sup>1</sup>

À quoi servent les banques ? Comment ont-elles fait évoluer leurs activités au cours des dernières décennies ? Sont-elles responsables de la crise ? Les réformes parviendront-elles à renforcer la sécurité du secteur et à le remettre au service de l'économie réelle ? Autant de questions déterminantes pour l'avenir des banques européennes et pour la croissance de nos économies.

## À quoi sert une banque ?

En théorie, la banque remplit trois fonctions essentielles. Elle collecte, gère les dépôts des clients et leur fournit des moyens de paiement (chéquier, carte bancaire, ordre de virement...). Elle gère aussi leur épargne et réduit les risques qui s'y rapportent en les mutualisant. Enfin, elle finance leurs projets d'investissement ainsi que ceux des entreprises.

© JÉZABEL COUPPEY-SOUBEYRAN

Économiste, maître de conférences,  
université Paris-I Panthéon-Sorbonne

Ces trois fonctions sont absolument indispensables pour le bon fonctionnement de l'économie. Si les moyens de paiement venaient à manquer ou si les dépôts ne pouvaient plus circuler, les échanges seraient rapidement empêchés. Et si les épargnants n'avaient pas d'intermédiaires à qui confier leur épargne, leur épargne financière s'en trouverait sans doute réduite et moins bien orientée vers le financement des entreprises. Sans les banques, il serait aussi beaucoup plus difficile pour les ménages et de nombreuses entreprises de financer leurs projets autrement qu'en s'autofinançant.

Les crédits bancaires sont également vitaux pour les petites et moyennes entreprises

[1] Ce texte est en grande partie adapté du livre : Couppey-Soubeyran J. et Nijdam Ch. (2014), *Parlons banque en 30 questions*, Paris, La Documentation française, collection « Doc en Poche ».



(PME), en raison d'un accès aux marchés financiers limité ou inadapté. Et même pour les grandes entreprises qui se financent sur les marchés de titres, ce sont souvent les banques qui, en bout de chaîne, détiennent une partie des titres émis. Au total, les financements bancaires au sens large (qu'il s'agisse des crédits octroyés ou des titres que les banques détiennent sur les entreprises émettrices) représentent encore 70 % du financement externe des entreprises en Europe.

En pratique, les banques exercent plusieurs types de métiers plus ou moins orientés vers les particuliers et les entreprises.

Les services de **banque de détail** (ou banque commerciale) s'adressent à une clientèle de particuliers et de PME et consistent à gérer des dépôts, fournir des moyens de paiement – chéquier, carte bancaire, virement... – et octroyer des crédits.

Les services de **banque de financement et d'investissement** (BFI) s'adressent à une clientèle de très grandes entreprises et leur permettent de réaliser des transactions sur les marchés financiers.

Au-delà, les banques ont élargi leurs services à la gestion d'actifs consistant à gérer des portefeuilles de titres et d'OPCVM pour le compte de clients épargnants ou d'entreprises. La plupart des groupes bancaires français offrent aussi des services d'assurance : c'est la bancassurance.

Depuis la fin des années 1990, en France tout particulièrement, mais plus largement aussi dans la plupart des autres pays d'Europe, ces métiers sont le plus souvent logés sous un même toit, celui de la banque dite « universelle ». Ses mérites sont débattus. Certains y voient une structure productive plus efficace, permettant de réaliser des économies d'échelle (un coût moyen plus faible grâce à une échelle d'activité plus grande) et d'envergure (un moindre coût global en produisant plusieurs services ensemble plutôt que séparément). Les études empiriques sont toutefois loin d'être unanimes sur le sujet. Et pour d'autres, au contraire, la banque uni-

verselle additionne les risques de la banque de détail et de la banque d'investissement, met en danger les dépôts en les exposant aux risques de turbulence des marchés financiers, et crée des conflits d'intérêt entre les deux métiers.

## Comment les banques ont-elles fait évoluer leur activité depuis les années 1980-1990 ?

Les banques ont considérablement fait évoluer leur activité depuis la fin des années 1980. Les années 1990-2000 ont vu ces changements s'accélérer. La principale transformation réside dans un adossement de plus en plus important aux marchés de capitaux. Cela a profondément modifié tant la taille que la structure des bilans des banques (cf. Zoom p. 96).

Jetons un coup d'œil pour commencer sur la structure, c'est-à-dire la composition des bilans bancaires. Au passif, la part des ressources de marché s'est accrue au cours des dernières décennies, tandis que celle des dépôts s'est réduite (sur ce terrain, les banques françaises ont subi la concurrence des assurances-vie et des SICAV monétaires, ainsi que celle de leurs propres filiales). À l'actif, les activités de marché ont gagné du terrain sur les activités plus traditionnelles de prêts. Dans les grandes banques universelles cotées en bourse, les prêts à la clientèle de proximité ne représentent plus qu'un gros tiers de l'activité, amplement déployée à l'étranger sur les marchés de titres et de produits dérivés. Les banques mutualistes ont, en revanche, maintenu leurs activités de détail, ce qui les complexait avant la crise, et au contraire aujourd'hui rassure leur clientèle.

Ce faisant, c'est aussi la structure des revenus des banques qui a changé. Les métiers exercés par les banques engendrent deux principaux types de revenus : des intérêts et des commissions. Quand la banque octroie un crédit, elle fait payer à l'emprunteur un taux débiteur. La banque paie elle-même des intérêts quand

# ZOOM

## QU'EST-CE QU'UN BILAN BANCAIRE ?

Un bilan est un document effectuant la synthèse des emplois (à l'actif) et des ressources (au passif) d'une entreprise, actif et passif étant par construction égaux. Le bilan d'une banque commerciale est principalement constitué à l'actif de titres et de crédits octroyés aux clients et au passif de dépôts de la clientèle et de fonds propres, les dépôts étant largement plus importants que les fonds propres (qui représentent moins de 10 % de la taille du bilan).

Cette structure de bilan particulière fait apparaître la fonction des banques et les risques inhérents à cette activité : alors qu'une large part des dépôts

collectés est à vue ou à court terme, beaucoup de crédits octroyés sont à long terme. La banque pratique donc la « transformation d'échéance » (c'est son métier !) et se trouve ainsi exposée à un risque d'illiquidité, si elle ne peut faire face aux demandes de retraits de fonds de ses clients\*.

[Ce risque d'illiquidité s'est significativement accru au tournant des années 2000, en lien avec l'augmentation de la part des dettes de marché au bilan des banques, contractées à des échéances très courtes. Très instables, ces ressources soumettent les banques à la nécessité de les renouveler sans cesse et à un risque de panique non plus à ses guichets mais sur le marché interbancaire.]

\* Courderc N. et Montel-Dumont O. (2009), *Des Subprimes à la récession. Comprendre la crise*, Paris, La Documentation française/France info, p. 86.

elle se procure des ressources nécessaires à son refinancement et rémunère l'épargne de ses clients. Ces services de banque de détail sont très rémunérateurs en commissions : la banque facture ainsi notamment ses services de gestion de compte et de mise à disposition des moyens de paiement. Dans ses activités d'investissement et de financement, la banque perçoit aussi des commissions (conseils aux entreprises pour s'introduire en bourse, réaliser des fusions-acquisitions...) et des intérêts (crédits, achats de titres...). Le produit net bancaire (PNB) mesure l'ensemble de ces revenus, nets des charges de refinancement associées aux activités de la banque : c'est le chiffre d'affaires d'une banque.

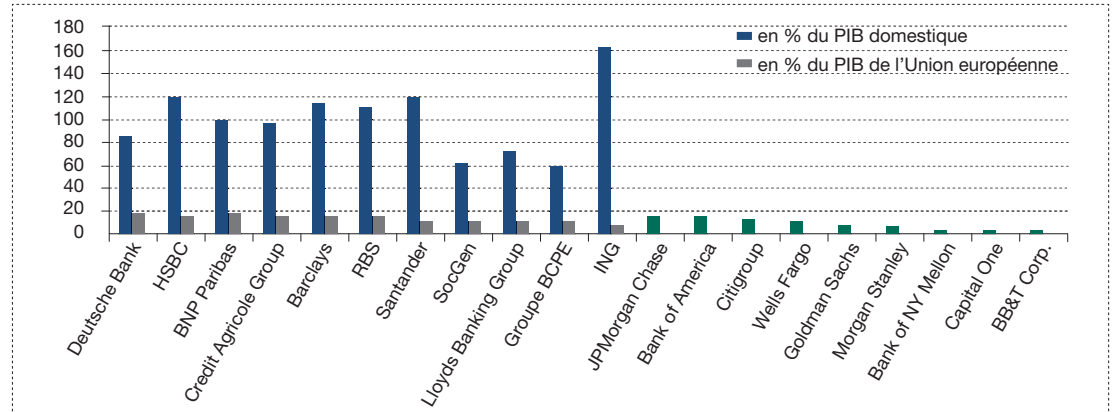
Le PNB se répartit ainsi entre une marge d'intérêts et des commissions. Que sait-on de cette répartition ? Comparés aux autres grandes banques européennes, les revenus des grandes banques françaises proviennent relativement moins de la marge d'intérêt (environ 50 % du PNB des banques françaises contre 60 à 65 % pour près de la moitié des grandes banques européennes), et relativement plus des commissions (environ 30 % du PNB des banques françaises contre 20 à 25 %

pour la moitié des grandes banques européennes). Au sein des activités de détail, qui représentent elles-mêmes environ 70 % du PNB total de l'ensemble des banques françaises, les commissions clients représentent entre 40 % et 50 % du PNB.

Si l'essor des activités de marché a largement contribué à l'essor des commissions, il faut souligner que les activités de détail ne sont pas en reste en la matière. Les services facturés à la clientèle de détail concernent principalement la gestion de compte (mise à disposition de moyens de paiement : chéquier, carte bancaire, virement...) et l'octroi de crédits. La part des commissions dans les revenus d'activité des banques a fortement augmenté depuis les années 1990, en particulier celles qui se rapportent à la gestion de comptes et aux moyens de paiement. Alors même que l'activité de détail a vu sa part diminuer dans les bilans bancaires jusqu'à la crise, les revenus qui en sont issus ont continué de croître, du fait de l'augmentation des frais bancaires. Cette hausse ne s'est sensiblement ralentie que depuis la crise.

Toutes ces activités et les revenus qui s'y rapportent sont, pour une part qui a significati-

## 1. Combien pèsent les plus grosses banques ?



Note : ce graphique mesure l'actif total des groupes bancaires en pourcentage du PIB national. En France par exemple, le total de bilan de BNP-Paribas pèse l'équivalent de 100 % du PIB. Source : rapport Liikanen (2012).

vement augmenté, réalisés à l'étranger. En ce qui concerne les banques françaises, par exemple, BNP réalise 68 % de son chiffre d'affaires en dehors de la France, Société générale 57 %, Crédit agricole 34 % et Natixis 47 %. Enfin, outre la structure, c'est la taille des bilans bancaires qui a radicalement changé. Les banques universelles européennes sont devenues extraordinairement grosses. Cela a été parfaitement bien illustré par le rapport

Liikanen en 2012, et plus récemment par le rapport du comité européen du risque systémique (ESRB), en juin 2014 : les grands groupes bancaires européens pèsent chacun, en total de bilan, à peu près l'équivalent du PIB de leur pays d'origine. Plus généralement, la taille des bilans bancaires a triplé, voire quadruplé dans les principaux pays de l'OCDE entre le début des années 1990 et la veille de la crise.

### ZOOM

#### QU'EST-CE-QUE LA TITRISATION ?

La titrisation consiste à transformer des créances en titres financiers pour les vendre à des investisseurs sur le marché. Ce faisant, les banques transfèrent à ces derniers le risque de crédit associé aux créances sous-jacentes. Les banques américaines ont massivement recouru à cette technique pour externaliser les risques liés aux crédits « subprimes » qui ont déclenché la crise financière de 2007-2008. Les banques européennes ont également participé au processus en achetant en bout de chaîne des titres issus de la titrisation.

Problèmes économiques

## Les raisons d'un tel essor

L'expansion de l'activité des banques doit beaucoup à la capacité décuplée qu'elles ont eu de prendre des risques au tournant des années 1990-2000. D'abord, elles ont profité des marchés et des intermédiaires non bancaires pour leur transférer les risques qu'auparavant elles conservaient, soit par la titrisation (cf. Zoom), soit en achetant des protections (les dérivés de crédit permettent de transférer le risque de défaut des emprunteurs). Or, quand on n'a plus besoin d'assumer les risques pris, on en prend davantage ! Ce n'est pas seulement dans les activités de marché que les banques ont pris plus de risque, c'est aussi dans leur traditionnelle activité de crédit. Et ce n'est pas qu'aux États-Unis, la patrie du crédit *subprime*, que le crédit s'est emballé.

# ZOOM

## LA TAILLE N'EST PAS LE SEUL CRITÈRE DE « SYSTÉMICITÉ »

Le Comité de Bâle retient cinq critères de classification d'un établissement financier en institution financière d'importance systémique mondiale ou G-SIFIs (Global Systematically Important Financial Institutions) :

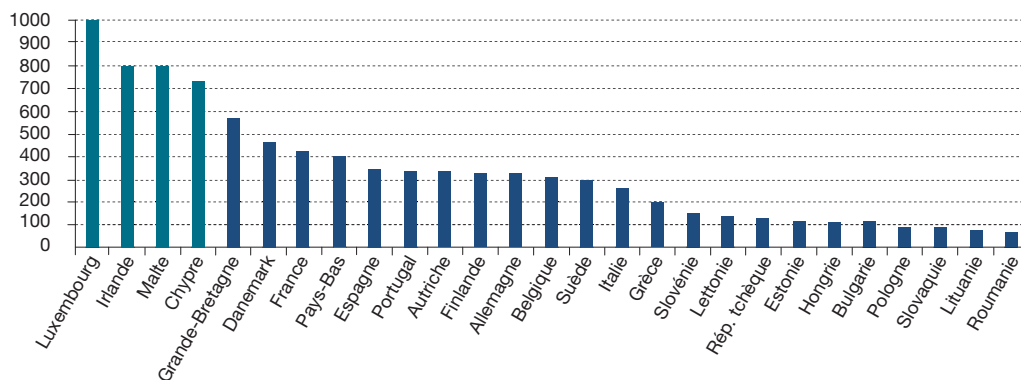
- 1) L'importance des activités transfrontières.
- 2) La taille du bilan et des engagements hors bilan.
- 3) Les interconnexions avec les autres institutions financières.
- 4) Le degré de substituabilité des activités exercées.
- 5) La complexité des activités (activités de marchés, produits dérivés, taille des actifs [difficiles à évaluer]).

Jézabel Couppey-Soubeyran

Tout cela s'est opéré dans un contexte très permissif. D'abord au niveau de la politique monétaire, très accommodante à partir du début des années 2000 (suite au krach Internet de 2000, puis aux attentats du 11 septembre 2001) : les banques ont bénéficié du niveau bas des taux d'intérêt. Ensuite, au niveau de la régulation financière. Les activités de marché des banques ont été jusqu'à la crise trop peu exigeantes en fonds propres (et sans aucun garde-fou en matière de liquidité). Les activités de crédit l'étaient davantage mais la titrisation d'un côté, les modèles internes de l'autre ont permis aux banques de desserrer la contrainte, tout en nourrissant un formidable excès de confiance. Cette prise de risque accrue est bien entendu l'un des facteurs de la crise actuelle.

En prenant du poids, de nombreux groupes bancaires sont devenus systémiques. L'Europe compte à peu près la moitié des 29 banques systémiques listées par le Conseil de stabilité financière (Financial Stability Board). Concrètement, même si la taille n'est pas l'unique critère de systémicité (*cf.* Zoom), plus le bilan d'une banque pèse lourd, plus les dommages collatéraux en cas de faillite sont importants : beaucoup d'entreprises se retrouveront à court de financement, beaucoup de déposants à court de moyens de paiement, d'autres banques seront touchées à leur tour sur le marché interbancaire, etc. C'est pourquoi les pouvoirs publics se refusent à laisser tomber une banque si elle est trop importante. Le principe « *too big to fail* » (TBTF) part d'une bonne

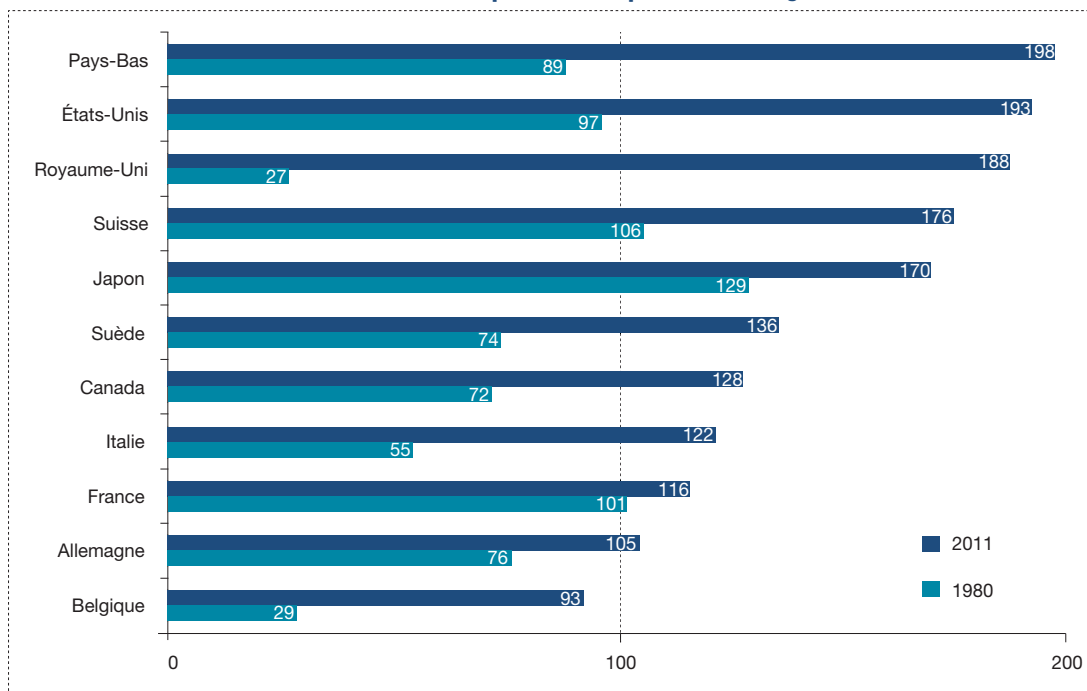
## 2. Le poids du secteur bancaire (en % du PIB)



Source : rapport Liikanen (2012).

Nota bene : axe vertical coupé – Luxembourg : 2400 % !

### 3. L'emballage du crédit (crédit au secteur privé – entreprises et ménages – en % du PIB national)



Source : *L'Économie mondiale 2013*, La Découverte, collection « Repères ».

intention : éviter le risque systémique que la faillite d'un grand établissement ferait courir. Mais l'enfer est pavé de bonnes intentions ! Comme les investisseurs et les agences de notation accordent de la valeur à cette garantie de sauvetage, les banques TBTF sont mieux notées et s'endettent à moindre taux. C'est comme si elles recevaient une subvention des pouvoirs publics. Le FMI a estimé le montant pour celles des pays de la zone euro : entre 90 et 300 milliards de dollars en 2011-2012.

Et grâce à cette subvention implicite, une banque TBTF peut encore se développer. Plus elle croît, plus les pouvoirs publics veulent éviter qu'elle ne tombe, plus elle est donc subventionnée. Un formidable cercle vicieux dont nos gouvernants peinent à sortir, d'autant qu'ils ont eux-mêmes encouragé la formation de grands établissements qu'ils voient (et défendent) comme leurs champions nationaux.

## Un secteur bancaire plus gros contribue-t-il davantage à l'économie réelle ?

Le secteur bancaire a grossi, au point que le total des actifs qu'il gère représente près de quatre fois la création de richesse annuelle d'un pays comme la France, six fois au Royaume-Uni, trois fois en Allemagne. Sans même parler de ce qu'il est dans les paradis fiscaux européens : 24 fois le PIB au Luxembourg, huit fois en Irlande et à Malte, près de sept fois à Chypre. Dans les années 1990, de nombreuses études empiriques montraient un lien positif fort entre finance et croissance. Depuis la crise, elles parviennent à un résultat bien plus mitigé : la finance profite à la croissance jusqu'à un certain seuil seulement, au-delà duquel la relation devient négative. Ce seuil, souvent mesuré en rapportant le crédit au PIB (qui mesure le flux de richesse créée

en un an dans un pays), serait de l'ordre de 110 % dans les pays avancés. Le fait est qu'il est dépassé à peu près partout dans ces pays.

Le crédit s'est emballé, c'est un fait, et il n'a pas contribué comme il aurait dû à l'économie réelle. Cela peut sembler paradoxal. Cela l'est moins dès que l'on essaie d'identifier la destination des crédits avant la crise. Quand le crédit coulait à flot avant la crise, que finançait-il ? L'immobilier pour une grande part, la consommation également, mais l'investissement des entreprises... fort peu. C'est là un point à souligner, d'autant que la crise n'a guère changé les choses en la matière. Les crédits aux entreprises représentent ainsi à peine 10 % du bilan agrégé du secteur bancaire français, ceux aux PME à peine 5 %.

En se développant, les banques n'ont donc pas forcément contribué davantage à l'économie réelle. Elles sont aussi devenues plus fragiles car elles ont financé leur croissance par la dette, contractée à des échéances de plus en plus courtes. De quoi se retrouver subitement à cours de liquidité et dans l'incapacité d'éponger par elles-mêmes des pertes éventuelles. Les exigences de fonds propres que les régulateurs ont imposées aux banques à partir de la fin des années 1980 (accords de Bâle 1 puis Bâle 2 avant ceux plus conséquents de Bâle 3), sous la forme d'un ratio pondéré par les risques, ne les a pas empêchées d'accroître considérablement leur levier d'actif. Tout en satisfaisant l'exigence réglementaire de fonds propres, les banques européennes ont fonctionné avec en moyenne 97 % de dette à leur bilan (soit une part de fonds propres en pourcentage de leur actif total – sans les pondérations – de l'ordre de 3 %). Cet écart constaté au niveau des grands groupes entre ratio de solvabilité standard et ratio de levier (fonds propres sur actif total) ou, ce qui revient au même, entre actif total et

actifs pondérés par les risques (*risk weighted assets* – RWA) est vérifiable dans les résultats de l'évaluation globale des bilans à laquelle a procédé la Banque centrale européenne. Ce n'est toutefois pas sur cet aspect que la BCE a choisi d'axer sa communication le 26 octobre 2014... Elle a de loin préféré rassurer les marchés en insistant sur le grand nombre de banques ayant réussi le *stress test*. On constatera néanmoins que la solidité du secteur bancaire européen s'apprécie en des termes très différents selon la mesure que l'on retient de la solvabilité.

\* \* \*

Aujourd'hui, en Europe, pratiquement plus aucun État n'a les moyens de sauver individuellement ses banques en cas de problème. Les superviseurs nationaux sont comme des nains face aux géants que sont devenues les banques transfrontières. L'Union bancaire arrive à point nommé ! D'une part, elle transfère la supervision des grands groupes bancaires à la bonne échelle, celle de l'Europe, du moins celle de la zone euro, en confiant la tâche à la BCE. D'autre part, elle va permettre de réintroduire progressivement l'idée même qu'une faillite, du moins un démantèlement, est possible en cas de problème.

La crise bancaire et financière a coûté très cher à l'Europe : entre octobre 2008 et octobre 2012, la Commission européenne a approuvé environ 5050 milliards d'euros d'aides d'État (aides directes et garanties de toutes sortes) en faveur des banques, ce qui représente environ 40 % du PIB de l'Union européenne. La réglementation des banques a été « un peu » renforcée, les marchés « un peu » réorganisés, la supervision « un peu » réagencée : « un peu » signifie que nous ne sommes pas à l'abri d'une prochaine crise.

## POUR EN SAVOIR PLUS

❖ **COUPPEY-SOUBEYRAN J.**  
et **NIJDAM CH.** (2014), *Parlons  
banque en 30 questions*, Paris,

La Documentation française,  
collection « Doc en poche ».

❖ **SCIALOM L.** (2013),  
*Économie bancaire*, Paris, La  
Découverte, coll. « Repères ».

# ↓ COMPLÉMENT

## L'esprit des accords de Bâle

### Pourquoi réglementer les banques ?

Une banque n'est pas une entreprise comme les autres. D'une part, elle investit l'argent de ses clients qui ne sont guère en capacité de lui imposer qu'elle le gère dans leur intérêt plus que dans le sien. D'autre part, la faillite d'une banque entraîne pour la collectivité des coûts bien supérieurs à ceux de n'importe quelle autre faillite d'entreprise. La difficulté d'une banque à vite fait de s'étendre à d'autres puisque toutes sont liées financièrement sur le marché interbancaire. Par contagion, une faillite bancaire peut ainsi s'étendre à l'ensemble du secteur et devenir une crise systémique. De plus, le dérèglement ou pire l'interruption des services bancaires sont très paralysants pour l'économie : sans moyens de paiement et sans financement, les échanges ne peuvent plus se dérouler normalement. Le régulateur doit autant veiller à la protection des déposants qu'aux problèmes d'externalités qui viennent d'être mentionnés. La réglementation des banques répond à ces deux objectifs et les exigences de fonds propres en sont devenues l'instrument principal.

### Risque de crédit, risques de marché et risque opérationnel

Le comité de Bâle s'est en tout premier lieu préoccupé du **risque de crédit**. C'est le risque auquel la banque s'expose quand elle octroie un crédit. Si l'emprunteur fait défaut, la banque subit une perte. Pour absorber cette perte sans risquer elle-même de faire défaut vis-à-vis de ses créanciers, la banque doit disposer d'un capital suffisant. Exiger des banques qu'elles détiennent un montant de fonds propres proportionnel aux risques qu'elles prennent est une mesure qui permet donc de réduire le risque d'insolvabilité de la banque. La réglementation des fonds propres oblige les banques à détenir plus de fonds propres qu'elles ne s'y contraindraient par elles-mêmes.

Les premiers accords de Bâle en 1988 ont recommandé l'adoption du **ratio Cooke** : un

rapport d'au moins 8 % entre les fonds propres et les actifs exposés au risque de crédit (en pondérant ces actifs par le risque).

L'activité risquée des banques ne se réduit cependant pas aux crédits accordés. De plus en plus actives sur les marchés de titres et de produits dérivés depuis le début des années 1990, les banques sont aussi exposées aux **risques de marché** (risque de variations du cours des titres détenus, risque de taux, risque de change...). Quant à l'incorporation des avancées technologiques dans les services bancaires, elle a renforcé le **risque opérationnel** (risque de panne informatique, de défaut de procédure, d'erreur humaine, de fraude, etc.). L'exigence de fonds propres a donc dû être étendue, d'abord aux risques de marché, puis au risque opérationnel avec les accords de Bâle 2.

### Une organisation en trois « piliers »

Trois piliers ont dès lors été définis : un pilier « **exigences minimales de fonds propres** » (maintenue à 8 % pour couvrir les trois catégories de risques – risque de crédit, risque de marché et risque opérationnel), un pilier « **surveillance prudentielle** » permettant aux superviseurs, s'ils le jugent nécessaire, d'exiger plus que la norme, un pilier « **discipline de marché** » exigeant de la transparence et de la communication d'informations. Les accords de Bâle 2 ont aussi donné aux banques le choix entre une approche réglementaire standard et une approche dite avancée les autorisant, sous certaines conditions, à utiliser leurs modèles internes d'évaluation des risques, ce qui a permis à certaines d'entre elles de réduire sensiblement la charge en fonds propres. Les accords de Bâle 3 ont toutefois rapidement emboîté le pas du fait de la crise. Ils rehaussent en quantité et en qualité l'exigence de fonds propres, introduisent également un simple ratio de fonds propres non pondéré par les risques (ratio de levier) et deux ratios de liquidité pour éviter aux banques de succomber à une rupture subite de liquidité\*.

\* Couppey-Soubeyran J. (2013), « De Bâle 2 à Bâle 3 : la nouvelle réglementation bancaire internationale », *Cahiers français* n° 375, *Quelle finance après la crise ?*, Paris, La Documentation française.

Les ménages partagent leurs revenus entre consommation et épargne. Qu'est-ce qui détermine leurs décisions ? En microéconomie, elles sont expliquées par un modèle calqué sur la théorie standard du choix de consommation à deux biens. Dès lors qu'ils ont accès à un marché des fonds prêtables, les ménages arbitrent entre consommation présente et consommation future et déterminent donc un profil intertemporel de consommation et d'épargne. Nicolas Drouhin présente ici les hypothèses et les principaux résultats de ce modèle.

Problèmes économiques

# Consommation/épargne : les choix des ménages

C'est au début du xx<sup>e</sup> siècle que l'économiste américain Irving Fisher a donné le premier une modélisation rigoureuse de la décision d'épargne. La version la plus aboutie de sa théorie a été publiée en 1930, sous le titre *The Theory of Interest*<sup>[1]</sup>. Cet ouvrage a eu un retentissement important à l'époque, mais, sans doute trop novateur, il a ensuite été un peu oublié. Ce n'est que dans les années 1950, avec les travaux de Modigliani et Brumberg (1954) et de Milton Friedman (1957)<sup>[2]</sup> que le modèle de choix intertemporel s'est largement diffusé dans la communauté des économistes au point de devenir l'une des clés de voûte de l'analyse économique moderne.

## Un modèle à deux périodes

Nous étudions le cas d'un ménage qui vit deux périodes, la période 1 et la période 2, au cours desquelles il perçoit des revenus monétaires notés respectivement  $w_1$  et  $w_2$ . Le couple constitue ce que les économistes appellent le profil intertemporel de revenu, mais aussi la dotation initiale. De la même manière, nous noterons  $c_1$  et  $c_2$ , les dépenses de consomma-

tion respectives de chaque période. Le couple  $(c_1, c_2)$  sera appelé profil **intertemporel de consommation**. Il est choisi par les ménages et c'est ce choix qui va déterminer la quantité de revenu épargnée à chaque période. L'épargne de la période 1 va alors être égale, par définition, à la différence  $(w_1 - c_1)$ .

On suppose que le ménage a accès au marché des fonds prêtables. S'il souhaite consommer plus que son revenu en première période (épargne négative), il doit emprunter, emprunt qu'il remboursera, majoré des intérêts, en période 2. Au contraire, un ménage qui souhaite consommer moins que son revenu en période 1 aura une épargne positive qu'il prêtera en période 1 à d'autres agents qui lui rembourseront en période 2 le montant emprunté majoré des intérêts. Le contrat de prêt définit le taux d'intérêt qui va jouer un rôle important dans la décision de consommation et d'épargne.

© NICOLAS DROUHIN

École normale supérieure de Cachan,  
Centre d'économie de la Sorbonne (UMR CNRS 8174)

[1] Fisher I. (1930), *The Theory of Interest, as Determined by Impatience to Spend Income and Opportunity to Invest it*, New York, Macmillan.

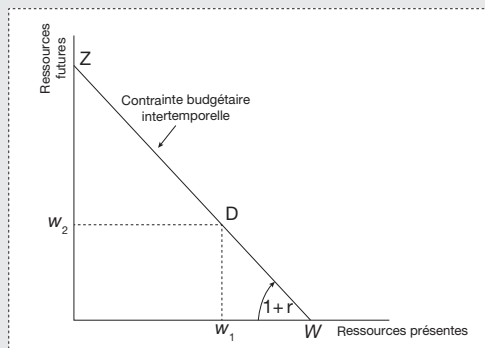
[2] Modigliani F. et Brumberg R. (1954), « Utility Analysis and the Consumption Function: An Interpretation of Cross Section Data », in *Post Keynesian Economics*, Rutgers University Press. Friedman M. (1957), *A Theory of the Consumption Function*, Princeton (New Jersey), Princeton University Press.



# ZOOM

## LA CONTRAINTE BUDGÉTAIRE INTERTEMPORELLE

### 1. La contrainte budgétaire intertemporelle



L'axe des abscisses représente les quantités de ressources en période 1 et l'axe des ordonnées les quantités de ressources en période 2. Le profil intertemporel de revenu de l'agent est donc représenté par le point de coordonnées  $(w_1, w_2)$ , point que nous dénommons  $D$ , pour « dotation initiale ».

Le point  $W$  correspond à la richesse de l'agent, qui, comme nous l'avons souligné précédemment, correspond au cas extrême dans lequel l'agent déciderait de consommer l'intégralité de ses ressources durant la première période. À l'opposé, le point  $Z$  correspondrait au choix inverse d'un agent choisissant de consommer toutes ces ressources en deuxième période ( $c_1 = (1+r)w_0 + w_1$ ). Bien évidemment, tous les choix intermédiaires sont possibles. Par exemple, si, partant de sa dotation initiale, un agent souhaite renoncer à une unité de consommation présente, il peut obtenir  $(1+r)$  unités de consommation future à la place. Ainsi, grâce à l'accès au marché des fonds prêtables, tous les points du segment de droite  $[ZW]$  constituent-ils des profils de consommations possibles, ceux entre lesquels l'agent a la possibilité de choisir. Ils ont pour caractéristique d'avoir tous la même valeur actualisée, égale à la richesse de l'agent. C'est la contrainte budgétaire intertemporelle. Formellement :

$$c_0 + \frac{c_1}{1+r} = w_0 + \frac{w_1}{1+r}$$

Nicolas Drouhin

Pour un ménage ayant accès au marché des fonds prêtables, quelle est la quantité maximale de dépense de consommation qui peut être réalisée en période 1 ? En fait, le ménage pourrait consommer la totalité de son revenu de première période, majoré du montant maximum qu'il peut emprunter, c'est à dire  $w_1 + w_2 / (1+r)$  puisque, dans ce cas, le ménage devra rembourser  $w_1$ , soit la totalité de son revenu de seconde période. Le montant  $w_1 / (1+r)$  correspond au pouvoir d'achat présent d'un revenu futur, ce que les économistes et les actuaires dénomment **valeur actualisée**. Le montant  $W = w_1 + w_2 / (1+r)$  définit ce que les économistes appellent la richesse. C'est la somme des valeurs actualisées des revenus présent et futurs, à laquelle il faut éventuellement ajouter le patrimoine

initial des ménages. On remarque que la définition économique de la richesse est plus générale que celle du sens commun. L'agent est riche non seulement de son revenu présent et de son patrimoine, mais également de tous ses revenus futurs. Ainsi, toutes les aptitudes, capacités et connaissances qui permettront dans l'avenir de recevoir un revenu font partie intégrante de la richesse. On remarquera également que la richesse dépend du taux d'intérêt.

En présence d'un marché des fonds prêtables, les ménages peuvent avoir, à chaque période, une dépense de consommation différente de leur revenu. Ils peuvent emprunter ou prêter de l'argent. Mais, au bout du compte, tous les montants empruntés ou prêtés devront être remboursés et, *in fine*, seront consommés. Les

ménages restent donc contraints par la valeur actualisée de l'ensemble des ressources perçues au cours de leur existence, c'est-à-dire leur richesse. Le graphique 1 représente l'ensemble des paniers de consommation possibles pour les ménages (cf. Zoom p. 103).

## Les préférences des ménages

Comme dans le modèle microéconomique de base des choix de consommation entre deux biens, on suppose que les ménages sont capables de classer l'ensemble des profils intertemporels de consommation par ordre de préférence. Pour connaître cet ordre de préférence, on peut demander à un ménage d'exprimer son choix entre deux profils  $C$  et  $C'$ . En fonction de leurs goûts, il pourrait répondre qu'ils préfère  $C$  à  $C'$  ou au contraire  $C'$  à  $C$ , ou enfin, troisième possibilité, qu'ils sont indifférents entre  $C$  et  $C'$ .

Le graphique 2 est un exemple de carte des préférences. Chaque point du graphique est un profil de consommation intertemporel. Chaque ligne courbe du graphique est une courbe d'indifférence. Par exemple,  $C^b$  et  $C^c$  appartiennent à la même courbe d'indifférence : le ménage est indifférent entre les profils de consommation  $C^b$  et  $C^c$ . Évidemment, cette « carte » est spécifique à un ménage particulier.

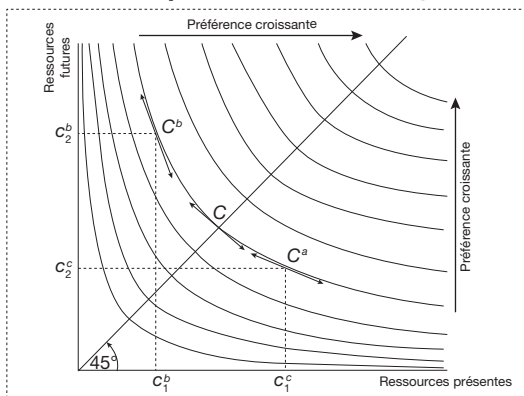
La carte d'indifférence représentée au graphique 2 possède un certain nombre de propriétés :

- les courbes d'indifférence ne se coupent pas ;
- elles sont ordonnées entre elles ;
- elles sont convexes.

Les préférences d'un ménage peuvent être caractérisées par trois paramètres : le taux marginal de substitution, l'élasticité de substitution intertemporelle et le taux de préférence pour le présent.

Le **taux marginal de substitution** (TMS) répond à la question suivante : « à la marge, si je renonce à une petite quantité

## 2. La carte des préférences d'un ménage

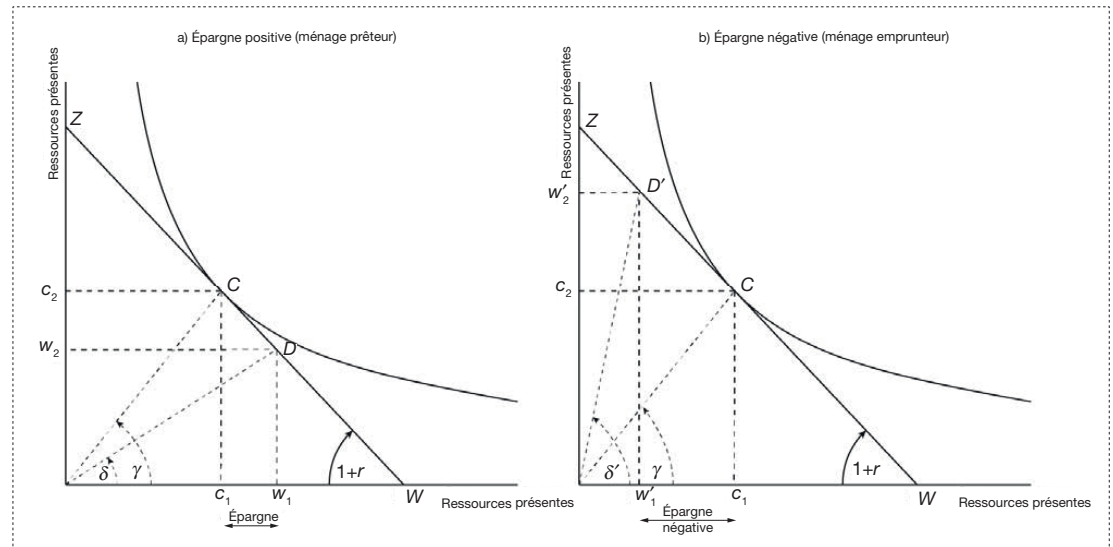


de dépense de consommation présente, de combien dois-je augmenter ma dépense de consommation future pour que ma satisfaction reste la même ? » Cette notion a une interprétation graphique simple, il s'agit de la valeur absolue de pente de la tangente à la courbe d'indifférence en  $C$ . Sur le graphique 2, on voit que le taux marginal de substitution tend à décroître lorsque l'on se déplace de la gauche vers la droite, le long d'une courbe d'indifférence.

L'inverse de la vitesse à laquelle le TMS décroît définit la deuxième notion importante, celle d'**élasticité de substitution intertemporelle**, qui caractérise l'intensité de la réponse des agents à une variation du taux d'intérêt : si le taux d'intérêt augmente de 1 %, quelle est la variation de la consommation présente du ménage ? Si l'élasticité de substitution intertemporelle est très forte, cela signifie que les courbes d'indifférences sont presque plates (le TMS décroît très lentement). Au contraire, si l'élasticité de substitution intertemporelle est faible (proche de zéro) cela signifiera que les courbes d'indifférence sont très coudées. En pratique, cette élasticité va caractériser l'intensité de la réponse des agents à une variation du taux d'intérêt.

On peut calculer le taux marginal de substitution, le long de la première bissectrice, c'est-à-dire pour tous les profils correspon-

### 3. Le choix de consommation et d'épargne



danst à une même dépense de consommation aux deux périodes. En général, ce taux est supérieur à 1. Cela signifie que, pour ces profils de consommation, si les ménages renoncent à une unité de dépense de consommation courante, il faut leur donner plus d'une unité de consommation future en plus, pour qu'ils restent indifférents. Le pourcentage de consommation qu'il faut leur donner en plus définit le **taux de préférence pour le présent**. Un ménage souhaitera d'autant plus consommer en première période que sa préférence pour le présent est forte.

### Le choix des ménages

En superposant les graphiques 1 et 2, on peut représenter le profil de consommation préféré du ménage. C'est ce que nous faisons dans les graphiques 3a et 3b.

Le graphique 3a reprend exactement la contrainte budgétaire du graphique 1 et les préférences du graphique 2. Le raisonnement est le suivant : le ménage va choisir parmi les profils de consommation de la contrainte budgétaire (le segment  $WZ$ ) celui qu'il préfère, c'est-à-dire celui qui se trouve

sur la courbe d'indifférence la plus « en haut à droite » parmi celles qui ont au moins un point commun avec la contrainte. On comprend que c'est le profil de consommation  $C$  qui va être choisi par le ménage. En effet, les autres points de la contrainte sont situés sur des courbes d'indifférence situées en dessous de celle qui passe par  $C$  et sont donc moins satisfaisants pour le ménage. Les courbes d'indifférence située au-dessus de celle qui passe par  $C$  vont quant à elles au-delà de la contrainte budgétaire intertemporelle : elles sont donc inaccessibles compte tenu des ressources.

On voit immédiatement qu'au point  $C$  est vérifiée une propriété géométrique importante : la contrainte budgétaire est tangente à la courbe d'indifférence, ce qui implique que la courbe d'indifférence et la contrainte budgétaire ont la même pente. La pente de la contrainte budgétaire est égale à  $-(1+r)$ , la pente de la courbe d'indifférence est égale à  $-TMS$ , on retrouve là l'application directe au modèle de choix intertemporel de l'égalité fondamentale de la théorie microéconomique de la demande : l'égalité du TMS et du rapport des prix.

Un point remarquable qui découle de ce qui précède est que, lorsque les agents ont accès au marché des fonds prêtables, la pente du profil de consommation, déterminé par le rapport entre la consommation présente et future ne dépend pas du profil de revenu. Cette propriété n'est pas particulièrement intuitive, et c'est justement une des contributions importantes de la théorie du choix intertemporel que de permettre de la comprendre (*cf.* Zoom p. 103).

Un résultat essentiel de la théorie de la consommation et de l'épargne est que le signe de la pente du profil intertemporel de consommation est uniquement déterminé par la différence entre le taux d'intérêt ( $r$ ) et le taux de préférence pour le présent ( $\theta$ ). Ce résultat est extrêmement intuitif. Le taux de préférence pour le présent incite les agents à consommer plus dans le présent que dans le futur. Mais, de l'autre côté, le taux d'intérêt incite les agents à être patients. Plus le taux d'intérêt est élevé, plus les ménages sont incités à différer leur consommation dans le futur. On comprend bien que c'est la différence de ces deux taux qui va déterminer quel effet l'emporte. Si  $\theta > r$ , les ménages consommeront plus en début de vie qu'en fin de vie. Au contraire, si  $\theta < r$  les ménages consommeront moins en début de vie qu'en fin de vie.

Cet effet va être amplifié ou amorti selon la valeur de l'élasticité de substitution intertemporelle du ménage considéré. Dans le cas où l'élasticité intertemporelle de substitution est forte ( $\sigma > 1$ ), l'écart entre le taux d'intérêt et le taux de préférence pour le présent va avoir un effet amplifié sur la pente du profil de consommation. À l'inverse, lorsque l'élasticité est faible ( $\sigma < 1$ ), l'effet va être considérablement amorti, le cas limite étant que lorsqu'elle est proche de zéro : le profil de consommation sera alors quasiment plat, le ménage choisissant de garder une consommation constante tout au long de la vie.

En revanche, les ménages restent soumis à la contrainte budgétaire intertemporelle.

La richesse affecte donc le niveau de la consommation : un ménage plus riche pourra consommer plus aux deux périodes, mais le rapport  $c_2/c_1$  restera inchangé.

Pour résumer, la décision de consommation et d'épargne est expliquée par la volonté des agents de répartir leur consommation de manière régulière au cours du temps. Le profil de consommation intertemporel possède deux caractéristiques essentielles :

- la pente qui dépend du taux d'intérêt et des préférences des agents, caractérisées par le taux de préférence pour le présent et l'élasticité intertemporelle de substitution ;
- le niveau qui dépend de la richesse.

Si le profil de consommation ne dépend en rien de la pente du profil de revenu, il en va différemment de l'épargne.

Revenons au graphique 3. Dans le cas  $a$ , on voit que la consommation choisie en période 1,  $c_1$ , est inférieure au revenu de période 1,  $w_1$ . L'épargne est donc positive.

On pourrait faire le même raisonnement à partir des pentes des profils de consommation et de revenu. Dans le graphique 3a, la pente du profil de revenu (mesurée par l'angle  $\delta$ ) est plus faible que la pente du profil de consommation (mesurée par l'angle  $\gamma$ ). Puisque le profil de consommation et le profil de revenu ont la même valeur actualisée, si le revenu croît moins vite que la consommation, le ménage va être épargnant en période 1 et consommer plus que son revenu en période 2.

À l'inverse, dans le graphique 3b,  $c_1 > w_1$ , l'épargne est négative, le ménage est emprunteur en première période. Le même raisonnement que précédemment peut être fait sur les pentes des profils de consommation. Quand le revenu croît moins vite que la consommation, le ménage va vouloir emprunter en première période. En seconde période, il remboursera son emprunt et sa consommation sera alors plus faible que son revenu.

## Effet d'une variation du taux d'intérêt ou du revenu sur la consommation et sur l'épargne

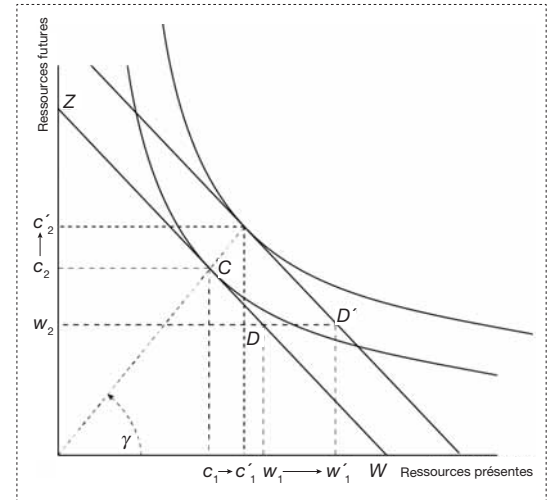
La politique fiscale et les transferts sociaux sont susceptibles de modifier les revenus des ménages. De la même manière, la politique monétaire va affecter les taux d'intérêt. Ou encore, la variation du taux de taxation des revenus de l'épargne va influencer sur les intérêts après impôts réellement perçus par les ménages. Quels effets peut-on en attendre ?

Le graphique 4 illustre l'effet d'un changement du revenu courant. La dotation initiale est changée en  $D'$ . La contrainte budgétaire est déplacée vers la droite. Comme on raisonne à taux d'intérêt constant, la pente de la contrainte budgétaire intertemporelle est inchangée. Ce déplacement traduit une augmentation de la richesse. Le niveau de la consommation aux deux périodes va augmenter. La pente du profil de consommation n'étant pas affectée par la variation du revenu, le rapport consommation future/consommation présente reste inchangé. Cela veut dire que l'augmentation de la richesse va être répartie entre des augmentations de consommation aux deux périodes.

Ainsi, si le revenu courant augmente d'un euro, la richesse augmente également d'un euro. Une fraction de cet euro va se transformer en une consommation supplémentaire en période 1, alors que le reste, majoré des intérêts, va être transformé en de la consommation supplémentaire en période 2. Cela implique que « la propension marginale à consommer » va être positive et inférieure à l'unité, puisque seule une fraction de l'euro supplémentaire sera consommée en période 1. C'est la loi psychologique fondamentale énoncée pour la première fois par Keynes (1936)<sup>3</sup>.

L'effet d'une augmentation du taux d'intérêt sur la consommation est plus compliqué, car le taux d'intérêt agit simultanément sur la pente du profil intertemporel de consom-

### 4. Effet de hausse du revenu courant



mation et sur son niveau, puisque il apparaît dans le calcul de la richesse. L'effet sur la pente est simple à prévoir : plus le taux d'intérêt est élevé, plus la pente du profil de consommation est importante. C'est l'**effet de substitution**, qui correspond au fait que la consommation future étant relativement moins chère, les ménages vont souhaiter y consacrer une plus grande partie de leur richesse. À l'opposé, l'effet d'une hausse du taux d'intérêt sur le niveau de la consommation est ambigu. Si l'on regarde la contrainte de budget vital, on se rend compte que le taux d'intérêt rend relativement moins chère la consommation future, ce qui accroît le pouvoir d'achat des ménages, mais cela réduit la valeur présente des revenus futurs, diminuant par là même la richesse. Quel effet l'emporte ? En fait, il est facile de comprendre que les ménages épargnants nets en première période vont connaître un enrichissement relatif du fait de la hausse du taux d'intérêt et vont avoir tendance à consommer plus aux deux périodes alors que les ménages emprunteurs en première période vont être relativement appauvris et avoir tendance à consommer moins aux deux périodes. C'est l'**effet de richesse** qui vient compléter l'effet de substitution décrit précédemment. Dans

<sup>[3]</sup> Keynes J. M. (1936) [1988], *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, Paris, Payot. Pour une discussion approfondie du rapport entre la théorie générale de Keynes et la théorie du choix intertemporel, on se rapportera à Drouhin (2007).

# ZOOM

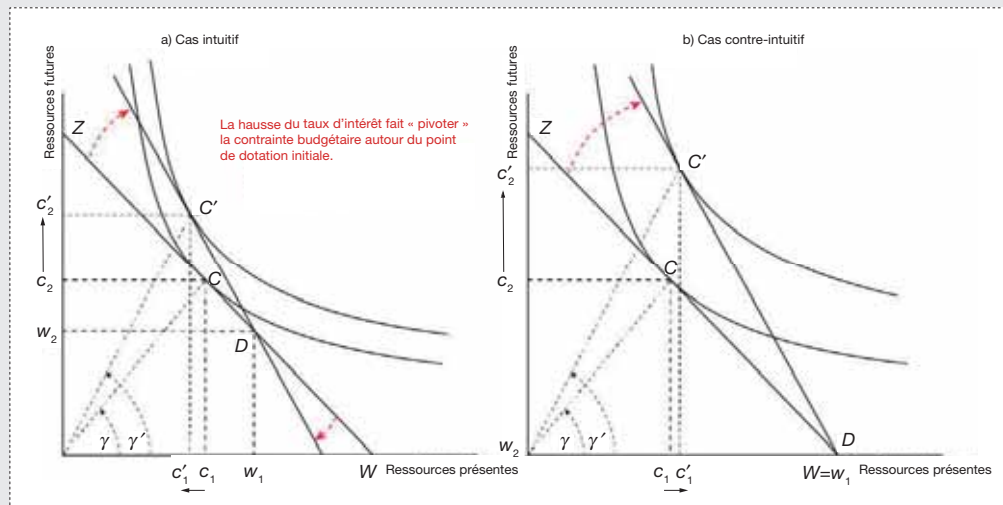
## L'IMPACT D'UNE HAUSSE DU TAUX D'INTÉRÊT SUR LA CONSOMMATION ET L'ÉPARGNE

Commençons par analyser le graphique 5a. Nous raisonnons à profil de revenu donné, c'est-à-dire à dotation initiale donnée ( $D$  est donc invariant). En revanche, le taux d'intérêt définit la pente de la contrainte budgétaire auquel appartient le point  $D$ . Puisque  $D$  est invariant, cela signifie que la hausse du taux d'intérêt se traduit par un pivotement de la contrainte budgétaire autour du point de dotation initiale. Quel va être l'effet sur la consommation ? L'effet de substitution va se traduire par un déplacement de la consommation vers un rayon formant un angle plus élevé par rapport à l'axe des abscisses [ $\gamma \rightarrow \gamma'$ ]. C'est l'intersection de ce rayon avec la nouvelle contrainte budgétaire qui détermine le nouveau point de consommation  $C'$ . Dans le graphique 5a, la nouvelle consommation en période 1 est plus faible après la hausse du taux d'intérêt. Comme le revenu de première période est invariant, l'épargne augmente.

Le ménage du graphique 5b a les mêmes préférences que celui du graphique 5a (les courbes d'indifférences sont les mêmes). Il a la même richesse initiale (les deux contraintes budgétaires de départ sont confondues). La seule différence est la structure de la richesse. Le ménage du graphique 5a perçoit des ressources aux deux périodes alors que celui du graphique 5b dispose de l'intégralité de sa richesse en période 1. Le point de dotation initiale est donc situé sur l'axe des abscisses. Ce serait typiquement le cas d'un ménage « rentier » qui disposerait d'un patrimoine initial important et ne travaillerait pas. Ce ménage épargne plus, puisqu'il doit financer l'intégralité de sa consommation de période 2 par son épargne. Il est donc relativement plus enrichi par la hausse du taux d'intérêt. L'effet de substitution est le même pour les deux ménages puisque les préférences sont les mêmes. C'est donc sur le même rayon [d'angle  $\gamma'$ ] que va se situer la nouvelle consommation. Mais la contrainte budgétaire pivotant autour d'un point de dotation initiale « beaucoup plus à droite », la hausse du taux d'intérêt implique que la consommation en période 1 est plus élevée qu'initialement. Donc l'épargne diminue avec la hausse du taux d'intérêt. Dans le graphique 5b, l'effet de richesse l'emporte donc sur l'effet de substitution.

Nicolas Drouhin

### 5. Effet d'une hausse du taux d'intérêt sur la consommation et sur l'épargne de ménages initialement épargnant



le cas des ménages initialement emprunteurs, les deux effets jouent dans le même sens : la hausse du taux d'intérêt implique une baisse de la consommation présente et donc une hausse de l'épargne (c'est-à-dire une baisse du montant emprunté). En revanche, pour un ménage initialement prêteur (c'est-à-dire dont l'épargne est positive en première période), effet de substitution et effet de richesse jouent en sens inverse. L'effet total va être ambigu. Le graphique 5 illustre ces possibilités (cf. Zoom p. 108).

La meilleure compréhension des effets possibles d'une hausse du taux d'intérêt sur l'épargne, incluant la possibilité d'un effet

négatif, est une contribution essentielle de la théorie du choix intertemporel à l'analyse économique. Elle a de nombreuses implications pratiques. Ainsi, il n'est pas du tout évident qu'une hausse de la fiscalité sur les revenus du patrimoine (assimilable à une baisse du taux d'intérêt effectivement perçu par les ménages) implique une baisse de l'épargne, contrairement à ce qui est généralement affirmé. Dans le cas, réaliste, d'une élasticité de substitution intertemporelle faible, c'est au contraire une hausse de l'épargne qui est prévisible pour les plus hauts patrimoines.

---

## POUR EN SAVOIR PLUS

---

❖ **DROUHIN N.** (2007), « Choix intertemporel et loi psychologique fondamentale », *Recherches économiques de Louvain*, vol. 73, n° 3.

---

Le rôle économique de l'État est majeur dans les pays développés puisque la dépense publique y représente entre 35 % et 57 % du PIB. Si certains domaines d'intervention – principalement les fonctions régaliennes – font l'unanimité, d'autres – dépenses sociales, production de biens et services collectifs – suscitent plus de débats. Laurent Simula rappelle les principaux domaines et justifications de l'action publique avant de faire le point sur sa définition et sa mise en œuvre. Il rappelle que si les défaillances du marché appellent l'intervention des pouvoirs publics, il est important de considérer aussi les défaillances de ces derniers.

Problèmes économiques

# L'intervention publique

L'État au sens large regroupe l'État central, les collectivités territoriales et les administrations de sécurité sociale. Il constitue aujourd'hui un acteur économique majeur dans l'ensemble des pays développés. Pourtant, cela n'a pas toujours été le cas, jusqu'à récemment. S'il appartient aux historiens et aux sociologues d'expliquer cette évolution, la science économique permet de comprendre pourquoi la main invisible du marché ne peut se passer de la main visible de l'État.

## Une intervention multidimensionnelle

L'intervention publique est un phénomène pluriel. Il est en effet possible de distinguer six domaines principaux d'action (cf. Zoom p. 111).

En fonction de l'importance relative de chacun de ces domaines, on obtient différents modèles d'intervention publique, allant de l'État minimal centré sur les fonctions régaliennes à l'interventionnisme le plus poussé. Il est intéressant à ce titre de considérer le volume et la structure des dépenses publiques. Celles-ci peuvent être classées en trois catégories :

– les dépenses de fonctionnement servent à la bonne marche des services publics, à l'exemple des dépenses courantes de personnel et d'entretien ;

– les dépenses d'investissement visent à renouveler ou à accroître le capital public (achat de matériels et de mobiliers, constructions de bâtiments et d'infrastructures) ;

– les dépenses de transfert modifient la répartition primaire des revenus par le marché. On peut distinguer les transferts en nature (prise en charge des dépenses de médicaments ou de soins, enseignement gratuit...) et les transferts en numéraire, subventions aux entreprises et revenus sociaux versés aux ménages (retraite, allocations familiales, minima sociaux...).

En 2012, la dépense publique représentait entre 35 % et 57 % du PIB parmi les États membres de l'Union européenne. La carte 1 permet de distinguer divers modèles d'intervention et souligne la position particulière occupée par la France.

De façon plus générale, l'importance de la dépense publique constitue un phénomène récent, comme en témoigne son évolution sur plus d'un siècle (voir graphique). Avant la Première Guerre mondiale, la dépense de l'État représentait, en France comme dans les autres pays occidentaux, entre 10 % et 15 % du PIB. L'intervention publique se limitait alors essentiellement aux fonctions régaliennes

© LAURENT SIMULA

Université d'Uppsala, Suède



## ZOOM

### LES DOMAINES PRINCIPAUX DE L'INTERVENTION PUBLIQUE

#### 1. Définir et appliquer des règles du jeu.

Selon la formule de Lacordaire, entre le fort et le faible, c'est la liberté qui opprime et la loi qui libère. Des lois, une justice, une police et une armée chargée de la sécurité nationale sont nécessaires à l'échange marchand. L'État s'assure également du respect des règles de la concurrence et supervise certaines activités (professions réglementées, marchés financiers...).

**2. Émettre et gérer la monnaie.** Cette fonction est généralement assurée aujourd'hui par une banque centrale qui conduit la politique monétaire selon les règles fixées par la puissance publique (les traités au sein de la zone euro).

**3. Lever l'impôt et dépenser les recettes publiques.** Les politiques fiscale et budgétaire modifient les arbitrages des ménages et des entreprises, influencent la productivité globale de l'économie (dépenses en infrastructures, éducation, recherche...) ainsi que le niveau de la demande agrégée.

**4. Produire des biens et des services.** L'État au sens large produit des biens et services

non marchands pour les ménages et les entreprises. Des services publics et des entreprises jouant un rôle économique essentiel ou stratégique peuvent également être contrôlés par l'État. Par exemple, l'État français contrôle directement ou indirectement près de 87 % du capital d'Areva, dont les activités sont principalement liées à l'énergie nucléaire.

**5. Contribuer à la résolution de problèmes et agir comme médiateur.** Les pouvoirs publics peuvent favoriser l'émergence de solutions, par exemple lorsque des entreprises sont en difficulté, lors de fusions-acquisitions ou de délocalisation. Leur marge d'action est cependant limitée comme le montre la fermeture des hauts fourneaux lorrains d'ArcelorMittal (2012-2014).

**6. Négocier des accords avec d'autres pays.** Ces accords peuvent concerner chacun des domaines que nous venons de mentionner. On peut penser en particulier à l'ensemble des accords qui ont permis la création du « marché intérieur » de l'Union européenne, aux règles régissant le commerce international dans le cadre de l'Organisation mondiale du commerce (OMC) ou encore aux négociations relatives à la réduction des émissions de gaz à effet de serre (protocole de Kyoto).

Laurent Simula

que sont la loi, la justice, la police, la défense et la monnaie. C'est avec la crise des années 1930 puis au cours des Trente Glorieuses que l'intervention publique dans l'économie a connu une progression marquée. Jusqu'à la fin des années 1970, de nombreux États intervenaient directement dans l'économie, dans le secteur de l'énergie, des transports ou des banques. Ce modèle n'était pas très propice à la recherche de l'efficacité et à l'innovation. Il a progressivement cédé la place, dans les années 1980 et 1990, à une intervention plus indirecte, fondée sur deux volets : régulation des monopoles naturels (réseau ferré, d'électricité, de distribution du gaz) et politique

de la concurrence afin d'éviter les pratiques déloyales contraires à l'intérêt public. Au-delà des missions régaliennes, l'intervention publique dans les pays développés est donc loin de se limiter à la dépense. Cette évolution correspond à la fois à une redéfinition des objectifs et des instruments d'action des pouvoirs publics.

## Pourquoi intervenir ? Les trois piliers de l'action publique

La justification de l'intervention publique est avant tout d'ordre normatif et s'appuie sur

trois piliers principaux : allocation des ressources, stabilisation de l'activité et redistribution. Les fonctions d'allocation et de stabilisation ont avant tout pour objectif la meilleure efficacité de l'économie. La fonction de redistribution permet l'obtention d'une répartition des richesses plus *équitable*. Ces trois champs d'action soulignés par Richard Musgrave<sup>1</sup> ne sont pas cloisonnés et, en pratique, de nombreuses mesures ont des effets dans chacun d'entre eux. Il est cependant utile de les examiner un à un afin de mieux cerner la raison d'être de l'action publique.

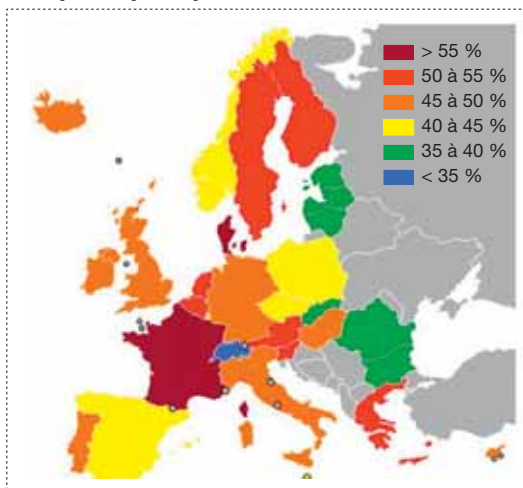
## Allocation des ressources

Le premier théorème de l'économie du bien-être énonce que l'équilibre d'un marché de concurrence parfaite constitue un optimum de Pareto. Cela signifie que, sous cette hypothèse, le libre jeu du marché conduit à un état de société dans lequel on ne peut améliorer le bien-être d'un individu sans détériorer celui d'un autre. À première vue, on pourrait en conclure avec Pangloss que « tout va pour le mieux dans le meilleur des mondes possibles pourvu qu'on le laisse aller tout seul ». Cette conclusion serait pourtant erronée. D'une part, le critère d'efficacité de Pareto ne dit rien sur la *répartition* des ressources. Ainsi, une allocation dans laquelle une caste dirigeante accaparerait toute la richesse constitue un optimum de Pareto. D'autre part, le premier théorème de l'économie du bien-être repose sur des hypothèses qui ne correspondent pas à la réalité des marchés. Leur remise en cause offre un fondement solide à l'intervention publique.

### Monopole naturel

Certains agents économiques, comme les grandes entreprises, peuvent disposer de pouvoirs de marché. Ils ne prennent pas alors les prix comme des données, mais participent directement à leur formation. Il peut s'agir d'un monopole naturel dans un secteur où la production induit des coûts fixes élevés. On peut penser à la distribution de l'électricité ou au réseau ferré. S'il

## 1. Dépense publique en 2012 (en % du PIB)

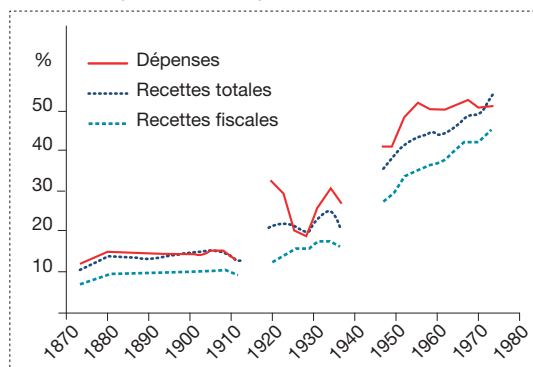


Source : Eurostat.

est plus efficace de disposer d'une seule entreprise gestionnaire que de plusieurs (dupliquer le réseau est très coûteux), rien ne garantit que cette entreprise cherche à minimiser ses coûts ou à développer des innovations sources de progrès. Il est donc important pour l'État de réguler les monopoles naturels. Cette régulation peut être directe (nationalisation) ou indirecte, cette seconde option étant aujourd'hui privilégiée dans la plupart des pays développés. Il s'agit en particulier pour le régulateur de s'assurer que le monopole naturel, détenteur des infrastructures ferroviaires (Réseau ferré de France) ou de distribution électrique (ERDF et RTE) par exemple, pratique un juste prix d'accès à son réseau. Cette tarification constitue en effet un élément essentiel à l'exercice d'une saine concurrence, que ce soit dans les domaines du transport de fret et de voyageurs ou dans la production d'électricité. En France, cette régulation est confiée à des autorités administratives indépendantes, à l'image de l'ARAF pour le rail et de la CRE pour l'énergie. Il existe cependant de nombreuses grandes entreprises dont la taille n'est aucunement justifiée par la technologie qu'elles utilisent. Celles-ci peuvent adopter un comportement prédateur afin

<sup>[1]</sup> Musgrave R. (1959), *Theory of Public Finance*, New York, McGraw Hill.

## 2. Dépenses et recettes publiques en France, 1872-1974 (en % du PIB)



Source : André Ch. et Delorme R. (1991), « Deux siècles de finances publiques : de l'État circonscrit à l'État inséré », *Revue d'économie financière*, vol. 1, hors-série.

de limiter l'entrée de firmes concurrentes, pourtant porteuses d'idées nouvelles et potentiellement sources de progrès. La politique de la concurrence a pour objectif de limiter les pratiques déloyales faisant obstacle au fonctionnement sain des marchés. Il peut s'agir par exemple de l'interdiction des ententes ou des abus de position dominante, ou encore de l'interdiction de la vente à perte et de l'encadrement des ventes liées<sup>2</sup>. La politique de la concurrence est l'un des piliers de la construction européenne. Elle fait partie des compétences communes attribuées à la Communauté économique européenne (CEE) par le traité de Rome de 1957.

### Externalités

Par ailleurs, en présence d'externalités, les arbitrages individuels ne tiennent pas compte des effets de la production ou de la consommation sur les autres agents. Les externalités sont « négatives » lorsqu'elles diminuent l'utilité de ces autres agents. On peut penser à une entreprise pétrochimique qui pollue l'atmosphère. Dans ce cas, les externalités conduisent à un niveau de pollution supérieur à l'optimum de Pareto. En ce sens, il y a « trop » de pollution. Mais les externalités peuvent également augmenter l'utilité des autres agents. Elles sont alors « positives ». La recherche développement génère souvent

des connaissances qui bénéficient à la collectivité dans son ensemble, soit immédiatement soit lorsque les brevets tombent dans le domaine public. Ces externalités sont au fondement des théories de la croissance endogène. Le problème est que le bénéfice marginal social de ces activités est supérieur au bénéfice marginal privé. Par conséquent, le libre jeu du marché conduit à un niveau d'activités inférieur à celui qui prévaudrait à l'optimum de Pareto. Il existe différentes façons de corriger les externalités. L'État peut redéfinir les droits de propriété, en donnant par exemple aux riverains d'une usine le droit à un air propre. Il peut également fixer des quotas ou introduire une fiscalité correctrice afin d'égaliser les coûts marginaux privés et sociaux dans le cas des externalités négatives. Au contraire, il peut subventionner les agents en cas d'externalités positives, afin d'inciter l'agent (investisseur ou consommateur) à augmenter son activité jusqu'au niveau qui maximise le bien-être social.

### Biens publics

Les biens publics soulèvent des difficultés proches des externalités. Il s'agit de biens ou services consommés collectivement, sans rivalité ni exclusion entre les usagers, à l'exemple du savoir et des techniques (à l'issue de la période de protection par les brevets). Dans la mesure où de nombreux agents peuvent profiter de ces biens sans en payer le prix, le marché produit spontanément ces biens en quantité insuffisante par rapport au niveau Pareto-optimal. Les pouvoirs publics peuvent financer la production de ces biens, ou bien créer des institutions chargées de protéger ceux qui les produisent des comportements de « passagers clandestins », sans pour autant porter préjudice au bon fonctionnement de la concurrence et à l'innovation.

### Défauts d'information

Tous les agents ne disposent pas de la même information<sup>3</sup>. Lorsqu'une entreprise souhaite recruter un nouvel employé, elle ne peut

<sup>[2]</sup> Pratique commerciale consistant à subordonner la vente d'un produit à l'achat d'un autre.

<sup>[3]</sup> Sur les asymétries d'information, voir dans ce même numéro l'article d'Anne Corcos et François Pannequin, p. 58.

observer directement sa productivité et est donc contrainte d'inférer celle-ci sur la base de certains signaux, comme ses diplômes ou bien les lettres de recommandation adressées par ses précédents employeurs. L'employeur dispose donc d'une information imparfaite en ce qui concerne le « type » (ici la productivité) des employés potentiels qu'il interviewe. Il fait face à un problème d'antisélection. En outre, une fois le contrat d'embauche conclu, rien ne garantit que l'employé se consacrera sérieusement aux tâches qui lui seront confiées. L'employeur dispose d'une information imparfaite en ce qui concerne le comportement ultérieur à la signature du contrat. On parle alors d'aléa moral. En raison de ces deux problèmes informationnels, l'entreprise peut hésiter à embaucher lorsque les formalités de licenciement sont très contraignantes. L'intervention publique peut alors consister à réformer les contrats de travail. Les pays d'Europe du Nord ont ainsi réformé le marché du travail afin d'autoriser une plus grande facilité de licenciement pour les entreprises tout en accordant des indemnités plus importantes aux salariés licenciés qui bénéficient en outre d'une formation professionnelle personnalisée (flexisécurité).

Les problèmes d'information peuvent même aboutir à l'effondrement et à la disparition de certains marchés, comme l'a souligné George A. Akerlof, prix Nobel d'économie 2001. Le système de marché est alors incomplet. Il est par exemple difficile pour un étudiant d'emprunter pour financer ses études dans une université française. L'État peut alors proposer des bourses à destination de ces étudiants, ou encore créer des dispositifs spécifiques de prêt<sup>4</sup>, afin de corriger l'incomplétude spontanée des marchés.

### **Problèmes de rationalité intertemporelle**

Enfin, les agents peuvent prendre à court terme des décisions qui ne maximisent pas leur utilité sur l'ensemble de leur cycle de vie. Par exemple, les jeunes actifs ne sont pas nécessairement conscients de l'importance de disposer d'une épargne (privée ou collec-

tive) leur garantissant une protection contre les risques santé ou vieillesse. Les pouvoirs publics peuvent donc organiser un dispositif obligatoire (épargne forcée) afin de corriger les problèmes liés à l'horizon temporel trop court de certains ménages.

Toutes ces défaillances du marché, d'ordre essentiellement microéconomique, justifient une intervention publique structurelle. Ces mesures de régulation ont pour objectif d'accroître l'efficacité des marchés de façon durable et d'élever le *trend* de croissance de l'économie à long terme.

### **Stabilisation de l'activité**

Il convient de distinguer le *trend* de croissance, correspondant au niveau de croissance « potentiel » ( $Y^P$ ), du niveau de l'activité économique à court terme ( $Y^T$ ). La différence  $Y^T - Y^P$  correspond à l'écart de production (*output gap*), positif en cas de surchauffe de l'économie et négatif en cas de crise. Les politiques publiques de stabilisation cherchent à minimiser cet écart. Tout comme les politiques corrigent l'allocation des ressources, elles ont pour maître mot l'efficacité. Cependant, alors que les premières ont à dessein d'élever le niveau de croissance potentiel, les secondes ont pour objectif de faire coïncider croissance effective et croissance potentielle à court terme. Elles sont donc essentiellement conjoncturelles.

Ces fluctuations de court terme peuvent résulter d'une part de mauvaises anticipations par les agents. Les sentiments et émotions humaines, les « esprits animaux » selon l'expression de Keynes, peuvent conduire à des prises de décisions peu rationnelles. Les fluctuations peuvent également résulter de rigidités microfondées. Ces « viscosités » limitent les ajustements de prix à court terme, sur le marché du travail (salaires), des biens (prix) et de la monnaie (taux d'intérêt). Par exemple, en raison de l'asymétrie d'information sur le marché du travail, un employeur peut fixer le salaire à un niveau supérieur à la productivité de ses salariés, pour deux raisons tout à fait rationnelles : attirer les

<sup>[4]</sup> On peut citer sur ce point les travaux de Robert Gary-Bobo et Alain Trannoy. Voir par exemple : Gary Bobo R. et Trannoy A. (2008), « Efficient Tuition Fees and Examinations », *Journal of the European Economic Association*, vol. 6, n° 6.

<sup>[5]</sup> Les salariés qui tireraient au flanc ne trouveraient pas d'emploi aussi bien rémunéré par la suite.

<sup>[6]</sup> Mirrlees J. (1986), « The Theory of Optimal Taxation », in Arrow K. J. et Intrilligator M. D. (eds), *Handbook of Mathematical Economics*, vol. 3, Amsterdam, Elsevier.

meilleurs éléments, d'une part, et encourager l'effort, d'autre part<sup>5</sup>. Le salaire ne jouant plus le rôle de variable d'ajustement sur le marché du travail, le marché du travail s'ajuste par les quantités, ce qui est source de déséquilibres macroéconomiques.

La macroéconomie contemporaine utilise le modèle **offre globale/demande globale**, qui intègre ces diverses rigidités, afin d'appréhender quelles politiques de stabilisation sont efficaces. Ce modèle distingue notamment deux types de choc, à l'origine de l'écart de production. Les chocs d'offre affectent la relation entre production et prix, à l'exemple des chocs pétroliers ou technologiques. Les chocs de demande modifient quant à eux la relation entre demande et prix. La réunification allemande après 1989 constitue un exemple de ce type de choc, la reconstruction dans les nouveaux Länder ayant créé une forte demande supplémentaire. L'efficacité relative des politiques budgétaire, monétaire et fiscale dépend de la nature des déséquilibres macroéconomiques ; il est donc essentiel de les identifier clairement afin de rendre l'intervention publique efficace.

### Redistribution des richesses

Une allocation des richesses efficace peut être jugée inéquitable selon les critères de justice sociale retenus par la puissance publique. L'intervention publique peut alors chercher à la rendre plus équitable, par exemple, en maximisant le bien-être des plus mal lotis (« maximin »).

Elle peut tout d'abord mobiliser le système socio-fiscal (impôts et prestations sociales). Le problème est que les transferts monétaires et les impôts modifient dans la plupart des cas les arbitrages des agents. Par exemple, l'impôt sur le revenu modifie les arbitrages travail/loisir. En raison de l'effet de substitution, un agent riche peut avoir tendance à travailler moins ou à réduire son niveau d'effort. L'agrégation de ces effets induit une perte sèche pour la société. Il existe ainsi un arbitrage entre l'équité et l'efficacité, entre une société très

égalitaire et une société pauvre. La théorie de la fiscalité optimale fondée par James A. Mirrlees, Prix Nobel d'économie 1996, a pour objectif de trouver le meilleur compromis entre les deux objectifs, compte tenu d'une conception donnée de la justice sociale.

Dans certains cas cependant, les politiques de redistribution peuvent contribuer à une meilleure efficacité. On peut citer les mesures facilitant l'accès des ménages pauvres aux soins médicaux et à l'éducation, qui corrigent les inégalités, mais contribuent aussi à l'augmentation de leur productivité et permettent donc des gains d'efficacité. En outre, de façon plus générale, il semble y avoir sur le long terme une corrélation positive entre le niveau de la croissance potentielle et la faiblesse des inégalités : les sociétés les plus riches sont moins inégalitaires que les plus pauvres.

## Comment intervenir ? De la définition à la mise en œuvre des politiques publiques

Le cheminement logique qui préside à la détermination de l'intervention publique procède en trois temps. Il convient tout d'abord de s'entendre sur les objectifs poursuivis par le décideur public. Ces objectifs peuvent notamment résulter d'un processus démocratique. Il s'agit ensuite d'examiner les différents moyens d'atteindre les fins poursuivies, compte tenu des contraintes matérielles, institutionnelles ou informationnelles. Enfin, il reste à identifier la meilleure solution réalisable, celle qui satisfait au mieux les objectifs au sein de l'ensemble des possibles. On suit ainsi la démarche préconisée par James Mirrlees (1986)<sup>6</sup> : « Une bonne façon de gouverner est de s'entendre sur les objectifs, découvrir ce qui est possible et optimiser. »

### Le socle normatif : la définition des objectifs

L'approche « bien-être » considère que le bien-être social procède de l'agrégation des

utilités individuelles. Les différents poids sociaux utilisés lors de ce calcul dépendent de l'aversion de la société envers les inégalités. Si cette aversion est infinie, par exemple, tous les agents auront un poids social nul, à l'exception des plus mal lotis. On obtient alors la solution « maximin » qui s'inspire de la philosophie de John Rawls. Si, à l'inverse, le poids social est identique pour tous les individus, l'objectif social correspondra à la simple maximisation de la somme des utilités individuelles. On retrouve alors l'utilitarisme de Jeremy Bentham. Amartya Sen, Prix Nobel d'économie 1998, a souligné le rôle des libertés substantielles en matière de justice sociale. Ces libertés sont par exemple la faculté d'échapper à la famine, à la malnutrition, à la morbidité évitable et à la mortalité prématurée, aussi bien que les libertés qui découlent de l'alphabétisation, de la participation politique ouverte ou de la libre expression<sup>7</sup>.

Si la conception de la justice constitue le point de mire de l'intervention publique, il est en pratique souvent nécessaire de définir des objectifs intermédiaires. Il peut par exemple s'agir du niveau de vie, du plein emploi, de la stabilité des prix, de la répartition des revenus ou de l'accès de tous aux services essentiels (à l'exemple de la CMU en France). Ainsi, le Full Employment and Balanced Growth Act de 1978 définit les objectifs de la politique économique des États-Unis : croissance stable, plein emploi, inflation faible, équilibre budgétaire et équilibre de la balance des paiements. Ces différents objectifs intermédiaires sont parfois contradictoires et il est alors nécessaire de les hiérarchiser.

### Les contraintes positives : des instruments en nombre limité et à l'efficacité circonscrite

Une fois les objectifs définis, il convient de faire l'inventaire des instruments d'action disponibles et de tenir compte de leurs limites respectives. La politique monétaire (action sur les taux d'intérêt) et budgétaire

(dépense publique et fiscalité) constituent les deux volets de l'intervention macroéconomique conventionnelle. On trouve à leurs côtés un ensemble riche d'instruments microéconomiques : réglementation, dans le domaine de la concurrence, par exemple ; structure des prélèvements directs et indirects, sur les ménages et sur les entreprises ; transferts au titre de la sécurité sociale ; investissements publics dans des domaines stratégiques, etc.

L'efficacité de ces instruments est limitée par certaines contraintes. Par exemple, la politique de régulation des monopoles est soumise à deux difficultés, comme l'ont montré Jean-Jacques Laffont et Jean Tirole, lauréat du prix Nobel d'économie 2014<sup>8</sup>. D'une part, le régulateur a une information très incomplète sur les capacités du monopole, en matière de réduction des coûts et d'innovation. D'autre part, le monopoleur peut craindre d'être exploité par le régulateur dans le futur s'il divulgue certaines informations. Ce sont aussi des problèmes de crédibilité qui ont motivé le transfert de la politique monétaire à des banques centrales indépendantes. Au sein de la zone euro, la Banque centrale européenne a pour seule mission le maintien de la stabilité des prix. Une contrainte institutionnelle interdit l'utilisation de cet instrument à des fins concurrentes.

De façon plus large, l'efficacité des instruments de l'action publique dépend de l'environnement institutionnel. Douglas C. North, Prix Nobel d'économie 1993, a souligné l'importance des institutions publiques, mais aussi de droit privé, comme les syndicats, dans le processus de croissance. Ces institutions influencent le fonctionnement des marchés, notamment du marché du travail.

### L'État stratège

L'intervention publique doit disposer d'un nombre d'instruments au moins égal au nombre d'objectifs visés. Il est sinon impossible d'atteindre tous les objectifs simultanément (règle de Tinbergen).

[7] Sen A. (2000), *Un nouveau modèle économique. Développement, justice, liberté*, éditions Odile Jacob.

[8] Laffont J.-J. et Tirole J. (1993), *A Theory of Incentives in Procurement and Regulation*, Cambridge, MIT Press.

En pratique, les objectifs sont souvent bien plus nombreux que les instruments. L'intervention publique apparaît dès lors comme une succession d'arbitrages entre différentes fins, mais également entre différents horizons, en fonction des préférences du décideur public pour le plus ou moins long terme. On peut ainsi penser à l'endettement public, qui peut être bénéfique à court terme mais qui représente une lourde charge pour les générations futures lorsqu'il atteint des niveaux excessifs. Pour chaque politique envisagée, l'État doit opérer une analyse coûts/bénéfices. Aux côtés des coûts directs, il convient d'ajouter les coûts d'opportunité, les ressources consacrées à une politique donnée n'étant pas disponibles pour conduire une politique alternative.

L'État peut opérer soit des arbitrages à la marge soit des arbitrages structurels. Ainsi, la courbe de Phillips, très populaire dans les années 1970, décrit une relation négative entre l'inflation et le taux de chômage. Un arbitrage à la marge consiste à prendre des mesures pour se déplacer le long d'une courbe donnée. Un arbitrage structurel reposera sur un ensemble de politiques susceptibles de déplacer la courbe elle-même et de changer ainsi de régime.

## Des défaillances du marché aux défaillances de l'État

Le processus que nous venons de décrire, allant de la définition des objectifs à la mise en œuvre des politiques publiques, n'est pas aussi lisse qu'il y paraît. Certaines difficultés sont inhérentes à la prise même de

décision en démocratie. Il est par nature difficile de dégager l'intérêt général à partir des intérêts particuliers, comme l'a montré la théorie du choix social à la suite de Kenneth Arrow. En outre, la définition des objectifs peut être influencée par certains *lobbies*. À cela s'ajoutent des délais peu compréhensibles. La prise de décision en démocratie repose sur un processus relativement lourd. En raison de problèmes d'agenda, certaines mesures, appliquées trop tard, peuvent s'avérer moins efficaces que prévu et même parfois avoir l'effet inverse du but recherché. D'autres aspects sont liés au fonctionnement même des administrations et des organismes publics. Comme toute forme d'organisation, ces derniers peuvent avoir tendance à se « bureaucratiser » et adopter des comportements « routiniers » qui en limitent l'efficacité. Frank Kafka décrit ainsi dans *Le Château* une administration dans laquelle « les gens travaillent à leur propre confusion ».

\* \* \*

S'il n'existe pas un modèle unique d'intervention publique, certaines formes d'intervention se révèlent plus efficaces que d'autres. Les nouvelles théories de la régulation, auxquelles Jean-Jacques Laffont et Jean Tirole ont énormément contribué, ont souligné que le « sens commun » conduisait souvent à l'adoption de mesures sous-optimales. L'évaluation, en amont comme en aval, apparaît dès lors un élément indispensable à la légitimité de l'intervention publique.

## POUR EN SAVOIR PLUS

❖ BÉNASSY-QUÉRÉ A., COEURÉ B., JACQUET P. et PISANI-FERRY J. (2012), *Politique économique*, Bruxelles, De Boeck.

❖ BOZIO A. et GRENET J. (dir.) (2010), *Économie des politiques publiques*, Paris, La Découverte, coll. « Repères ».

❖ SIMULA L. et SIMULA L. (2014), *La dissertation économique. Préparation aux concours*, La Découverte, Paris, coll. « Grand Repères ».

Lieu abstrait de rencontre entre offreurs et demandeurs, le marché est théorisé par les économistes comme un mode de coordination particulièrement efficient : une unique variable, le prix, permet en effet d'ajuster les intérêts d'une multitude d'agents. S'intéressant avant tout aux conditions qui permettent aux marchés de bien fonctionner, les économistes ont tendance à délaisser la question de leur élaboration et de leur place au sein de la société, réduisant même parfois les liens sociaux à des liens marchands ou à des liens imparfaits qui gagneraient à être conformes à l'idéal du marché.

François Vatin présente ici l'approche du marché développée par la sociologie économique : mêlant les conceptions des économistes et des sociologues, elle analyse la diversité de l'échange, fonctionnant selon des mécanismes hybrides relevant tout autant de la logique économique que des valeurs et conventions sociales.

Problèmes économiques

# Marché et marchés : une approche institutionnelle des modalités de l'échange économique

## Marché et société

© FRANÇOIS VATIN

Université Paris-Ouest-Nanterre-La Défense

### L'approche des économistes

Pour la science économique, le marché est un concept, sans cesse retravaillé depuis trois siècles. Il vise à concevoir les modalités de l'échange dans une situation abstraite où s'ajustent autour d'une seule variable – le prix – les intérêts opposés des vendeurs et des acheteurs. Le plus souvent, l'économiste contemporain, héritier de la pensée de Léon Walras, ajoutera au terme de « marché » les adjectifs

« pur et parfait » pour définir les normes selon lesquelles doit s'exercer la concurrence (atomicité des agents offreurs et demandeurs, homogénéité des produits, liberté d'entrée et de sortie, transparence de l'information). C'est à ces conditions exclusives, en effet, que l'on aurait affaire à un « vrai » marché, conforme à la théorie. La formulation d'un tel concept n'empêche aucunement les économistes d'être



conscients par ailleurs que les formes sociales d'échange observables sont plus ou moins proches d'un tel schéma et en sont parfois même très éloignées.

Une telle conceptualisation ne s'oppose pas forcément au point de vue du sociologue économiste, si on y voit un idéal-type au sens de Max Weber. Elle devient problématique en revanche dans le cadre de deux régimes argumentatifs que pratiquent parfois les économistes :

– si on attache au concept une valence historique universelle, comme si les liens entre les hommes avaient toujours été des liens marchands, l'étaient tous aujourd'hui et devaient toujours l'être jusqu'à la fin des temps (« naturalisation » du marché) ;

– si l'idéal théorique est mué en idéal social, avec l'idée qu'il faudrait en toutes circonstances rapprocher la réalité du concept en transformant les liens sociaux concrets « imparfaits » en véritables liens marchands conformes à l'idéal théorique (« normalisation » du marché).

## L'approche du sociologue

Pour le sociologue, les « faits économiques » sont des « faits sociaux » comme les autres et relèvent en ce sens d'une investigation « positive » : il faut les décrire pour en expliquer la cohérence et la façon dont ils s'insèrent dans la dynamique globale des sociétés<sup>1</sup>. Le marché est en ce sens une « institution » sociale, selon le sens qu'Émile Durkheim a donné à ce terme : « toutes les croyances et tous les modes de conduite institués par la collectivité<sup>2</sup> ». Les institutions sont le produit de l'histoire des sociétés ; elles sont ce dont les individus sociaux héritent, qu'ils ne peuvent modifier par simple effet de leur volonté. Elles se présentent donc à eux comme des « contraintes », ce qui justifie la formule si controversée de cet auteur selon laquelle il faut « prendre les faits sociaux pour des choses<sup>3</sup> ». Mais le caractère social de cette

nature la plonge dans l'historicité, rétrospective comme prospective (le problème du « changement social »). Or c'est précisément cette historicité que tend à gommer la théorie économique en proposant, sous la figure du marché, une algèbre universelle et intemporelle des relations sociales.

La leçon d'Émile Durkheim rejoint alors celle de Karl Marx quand il critiquait le « fétichisme de la marchandise » des économistes classiques<sup>4</sup>. Ceux-ci auraient, selon ce dernier, pris les « relations entre les hommes » pour des « relations entre les choses » ; ils hypostasieraient la société capitaliste de marché qui serait pour eux la forme universelle et intemporelle des relations sociales ; ils naturaliseraient en ce sens le marché comme une forme indépassable d'organisation économique de la société. Par cette vive critique, Marx ne répudiait pas pour autant la théorie économique. Il considérait au contraire celle-ci, dans sa forme la plus achevée (selon lui, celle de Ricardo), comme une clé de lecture incontournable de la société de son temps. Il entendait seulement historiciser les catégories naturalisées par Ricardo.

## L'approche de la sociologie économique

Dans le sillage de Marx, la sociologie économique, quand elle s'intéresse au marché, ne peut ignorer la conceptualisation des économistes. Mais elle prend pour une question, ce que les économistes posent comme un postulat : l'existence même du « marché ». Il faut ici faire une remarque qui, derrière son apparence terminologique, est en fait de première importance. Si l'on considère, comme il est fait communément, que la théorie économique est la science qui étudie la formation des prix sur un marché, alors, toute économie est nécessairement « marchande<sup>5</sup> ». *A contrario*, pour fonder son programme, la sociologie économique exige une définition plus large de l'économie, comme agencement des relations entre les hommes visant

[1] Steiner Ph. et Vatin F. (2013), « Le fait économique comme fait social », in *Traité de sociologie économique*, Paris, PUF, 2<sup>e</sup> édition, introduction, p. 1 à 11.

[2] Durkheim É. (1901) [1988], *Règles de la méthode sociologique*, Paris, Flammarion, coll. « Champs », préface à la 2<sup>e</sup> édition, p. 90.

[3] Durkheim É. (1895), *Les Règles de la méthode sociologique*, op. cit., chapitre 2.

[4] Marx K. (1867) [1993], *Le Capital*, livre I<sup>er</sup>, chapitre 1<sup>er</sup>, § 4, « Le caractère fétiche de la marchandise et son secret », Paris, PUF, p. 81 à 95.

[5] Il faut ici rappeler la définition canonique de la science économique par Lionel Robbins en 1932 : « Science qui étudie le comportement humain comme une relation entre des fins et des moyens rares qui ont des usages alternatifs. » La notion de marché n'est pas explicitement présente dans cette définition qui pose la théorie économique comme la science de l'allocation rationnelle (optimale) des moyens aux fins. Mais, selon la théorie néoclassique, le marché constitue l'instrument efficace d'une telle allocation optimale. Notons toutefois que l'économiste polonais Oskar Lange avait montré, dès les années 1930, la convergence du marché pur et parfait et de la planification parfaite.

à la reproduction matérielle de la société. Cette définition couvre les champs classiquement reconnus comme « économiques » : la production, la distribution, la consommation, et ouvre la possibilité que les fonctions économiques soient assurées par des dispositifs marchands ou par d'autres dispositifs sociaux<sup>6</sup>.

La sociologie économique peut alors rendre compte de contextes sociaux où la distribution des biens et des positions n'est pas régie, ou pas exclusivement régie, par des dispositifs marchands. De plus, si elle peut admettre la conceptualisation du marché par les économistes comme un cas-limite, elle doit l'inscrire dans un champ de formes sociales possibles et ne lui donner aucune prééminence *a priori*. Il lui faut à la fois comprendre l'émergence du marché comme forme sociale, son caractère plus ou moins central dans l'organisation des sociétés et la multiplicité de ses formes concrètes, dans l'histoire, comme dans le présent. Cela nécessite aussi de penser l'hybridation entre le « marché » comme catégorie idéaltypique et d'autres catégories sociologiques, relevant de l'ordre du politique, du symbolique, du technique, etc. Enfin, la sociologie ne saurait accorder au marché et, *a fortiori*, au « marché parfait », aucune valeur normative *a priori* : il y a d'autres modalités possibles d'échange entre les hommes que le marché et un marché « parfait » ou, plutôt, un marché qui s'en rapprocherait asymptotiquement, n'est pas forcément plus souhaitable qu'un autre.

Sur cette base, il est possible de distinguer deux questions emboîtées :

- celle de la place historique du marché parmi d'autres institutions économiques dans l'organisation des sociétés ;
- celle des formes différenciées du marché, notamment dans les sociétés contemporaines où cette institution domine la reproduction sociale.

## Institution historique du marché : la sociologie économique polanyienne

### Des sociétés traditionnelles aux sociétés capitalistes : le « désencastrement » du marché

La première question reste attachée au nom de Karl Polanyi, déjà cité<sup>7</sup>. Dans ses travaux d'histoire et d'anthropologie économique, il a insisté sur l'ancienneté historique des formes marchandes, mais, aussi, sur le statut radicalement nouveau du marché dans les sociétés contemporaines. Depuis la plus haute antiquité et dans des sociétés de nature très différente, les hommes ont noué des relations marchandes. Mais, souligne Polanyi, celles-ci ne dominaient pas l'ensemble de la vie sociale. Les marchés constituaient des « poches » au sein de sociétés, dont la reproduction reposait pour l'essentiel sur d'autres institutions sociales. Il distingue à cet égard trois modalités fondamentales des économies « traditionnelles » : l'autarcie (la vie domestique), la réciprocité (le don-contre-don) et la redistribution (par une autorité centrale qui prélève des richesses). Les marchés proprement dits, quand ils existent, sont, selon sa formule qui a connu une grande postérité, « encastrés » (« *embedded* ») dans la société ; leur pouvoir régulateur sur l'ensemble de la société reste faible.

Ce qui caractérise en revanche la société capitaliste moderne pour Karl Polanyi, c'est le principe du « marché autorégulateur ». Dès lors, le marché devient l'institution sociale dominante, chargée d'allouer « mécaniquement » (c'est-à-dire par le jeu impersonnel de l'offre et de la demande) l'ensemble des ressources économiques aux personnes. Polanyi montre que pour que le marché soit ainsi autorégulateur, il faut élargir la catégorie de « marchandise ». Dans les sociétés traditionnelles, les marchandises ne sont que les biens produits. La production elle-même, régie par

<sup>[6]</sup> Polanyi K. (posthume, 1977) [2011], *La Subsistance de l'homme. La place de l'économie dans l'histoire et la société*, Paris, Flammarion, 2011, et notamment le chapitre 2 : « Les deux significations du terme économique. »

<sup>[7]</sup> Outre l'ouvrage cité supra, voir, Polanyi K. (1944) [1983], *La grande transformation*, Paris, Gallimard, ainsi que le riche recueil d'articles parus entre 1922 et 1964, réunis dans ses *Essais*, Paris, Le Seuil, 2008.

les principes d'autarcie et de réciprocité (entraide), échappe au règne du marché. Pour que le marché devienne autorégulateur, il faut donc que ce que les économistes appellent les « facteurs de production » deviennent également des marchandises. Polanyi distingue trois facteurs de production : le travail (la vie productive des hommes), la terre (la nature) et la monnaie (les instruments de paiement libératoires émis par une instance politique). Pour que le marché devienne autorégulateur, il faut que ces trois biens soient traités comme des marchandises. Or, nous dit Polanyi, ces biens ne sont pas « produits » ; ils sont traités comme des marchandises sans en avoir les propriétés constitutives. Ce ne sont que des marchandises « fictives ».

### Un processus socialement destructeur

Le développement du marché autorégulateur correspond pour Polanyi à un processus de « désencastrement » du marché. Ce processus est pour lui socialement destructeur, car le marché n'est pas capable d'assurer harmonieusement la reproduction de ces biens. Le marché autorégulateur a un caractère parasitaire ; il se développe aux dépens de ce dont il se nourrit. Polanyi rejoint en quelque sorte la célèbre formule de Marx, qu'il a critiqué mais auquel il doit beaucoup : « La production capitaliste ne développe la technique et la combinaison du procès de production social qu'en ruinant dans le même temps les sources vives de toute richesse : la terre et le travailleur.<sup>8</sup> » On peut donc, en s'appuyant sur Polanyi, développer une critique des effets pervers de la « société de marché » (c'est-à-dire, non d'une société où existent des marchés, mais d'une société où l'ensemble de la régulation sociale est confiée au marché) :

- critique de la marchandisation du travail (salarariat ou « question sociale ») ;
- critique de la marchandisation de la nature (écologie ou « question naturelle ») ;
- critique de la marchandisation de la mon-

naie (dissolution du politique dans la financiarisation).

Selon Polanyi, le règne du marché autorégulateur aurait provoqué des contre-mouvements sociétaux (« autoprotection de la société »), se traduisant par un retour de la domination de l'instance politique, dont la forme exacerbée est le totalitarisme des années 1930 (fascisme, nazisme et communisme soviétique). Un peu comme son contemporain John Maynard Keynes, il était à la recherche d'une « troisième voie » permettant le maintien d'une société politiquement libérale<sup>9</sup>, ce qui suppose que le marché soit régulé et donc « ré-encastré » dans le social. Ce projet social correspond assez bien aux modes d'organisation qui se sont mis en place aux États-Unis dans le cadre du New Deal, puis en Europe occidentale après la Seconde Guerre mondiale. Pensons, par exemple, à la signification, en France, de l'établissement d'un « salaire minimum », norme sociale définie en fonction d'une mesure des besoins vitaux (salaire minimum interprofessionnel garanti : Smig établi en 1950), puis incorporant l'idée d'une juste redistribution des fruits de la croissance (salaire minimum interprofessionnel de croissance : Smic, établi en 1970). *A contrario*, le « néolibéralisme », né des « contre-révolutions libérales » au Royaume-Uni (période du gouvernement de Margaret Thatcher) et aux États-Unis (présidence de Ronald Reagan), peut être vu comme un nouveau désencastrement du marché, lequel est gros de risques politiques, si on adhère à la thèse polanyienne.

### Une thèse radicale

La limite de l'argumentaire polanyien résulte toutefois de sa radicalité. Même si, comme on l'a vu, des dispositifs d'autoprotection de la société contre le règne exclusif du marché ont été mis en place tout au long des XIX<sup>e</sup> et XX<sup>e</sup> siècles, et, notamment, dans la période qui a succédé à la seconde guerre mondiale, il faut bien admettre que, globalement, les relations marchandes n'ont cessé de s'étendre, tant sur un plan spatial (le capitalisme conquérant de nouveaux espaces), qu'en ce qui concerne la

<sup>[8]</sup> Marx K., *op. cit.*, p. 567.

<sup>[9]</sup> Keynes J. M. (1936) [1988], *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, Paris, Payot, chapitre 24.

nature des biens susceptibles d'être « marchandisés » : extension mondiale du salariat, privatisation généralisée des terres, mais aussi émergence de nouvelles marchandises « virtuelles » comme les assurances, les biens mobiliers, etc. Or, en dépit des inquiétudes de Polanyi et de ses nombreux disciples, la société « tient », ce qui n'est pas dit ici pour invalider les craintes fondées de ces auteurs sur les risques liés à la marchandisation généralisée des sociétés, mais comme un fait que la sociologie économique se doit de prendre en considération.

La sociologie économique contemporaine est alors amenée à regarder la question d'une autre manière. La limite de l'analyse de Polanyi est, d'une certaine manière, d'être restée trop proche de la pensée des économistes, en posant la notion de marché comme un concept homogène au statut incertain. Le marché s'oppose-t-il radicalement aux trois institutions économiques traditionnelles (autarcie, réciprocité et redistribution) ou doit-il être considéré comme une quatrième institution ? Si le marché est une institution, il est par nature lui-même « encastré » et l'idée de « désencastrement » devient conceptuellement problématique. Mais si le marché n'est pas une institution, quel peut-être le statut de ce concept pour la sociologie économique ? Les recherches récentes de sociologie économique adoptent donc plutôt la première posture. Elles se détachent alors de la catégorie abstraite de marché pour étudier *les* marchés dans leurs formes variées, en considérant les capacités socialement intégratrices de ces dispositifs à la fois matériels et politiques, qui ne sont donc plus, selon ce schéma, considérés comme « désocialisateurs » par nature.

## La sociologie des marchés

### La diversité des marchés

L'investigation empirique (l'enquête) qui caractérise la méthode des sociologues par rapport à celle des économistes conduit à

faire éclater la catégorie de marché : peut-on en effet considérer comme relevant d'un même modèle, les marchés de quartier ou de village où l'on va acheter fruits et légumes, ceux des valeurs mobilières qui se déroulent dans les « salles de marché » des grandes banques et autres institutions financières, ou le « marché du carbone » organisé entre États en application du protocole de Kyoto de 1997 ? Oui, diront les économistes, si, dans chaque cas, on peut mettre en évidence un mécanisme de concurrence qui se traduit par l'établissement d'un prix et de quantités échangées à ce prix. Mais le sociologue ne peut en rester là. Il doit observer les acteurs en présence et la nature des liens qui s'établissent entre eux : ceux de la ménagère et du maraîcher sur le marché du village ne ressemblent guère à ceux de deux institutions étatiques sur le marché du carbone. Il doit aussi observer les formes pratiques, techniques, de l'échange : la présence ou non de produits, les modalités de leur évaluation et certification, la temporalité des transactions (les transactions sur le marché hebdomadaire du village ou les échanges à la vitesse de la lumière entre les traders), etc.

C'est à l'examen de toutes ces questions, posées sur les terrains les plus divers, que s'est attelée la sociologie des marchés. Il n'est pas question ici de présenter de manière exhaustive ce courant de recherche, ni même de tenter d'en faire un bilan raisonné, mais seulement de citer quelques cas exemplaires qui permettent d'exposer la nature des questions en jeu et l'heuristique de la démarche<sup>10</sup>.

### Quelques cas exemplaires

#### *Le marché aux fraises solognot*

Il faut d'abord citer l'étude de Marie-France Garcia-Parpet sur le marché aux fraises solognot<sup>11</sup>. Cet exemple est particulièrement éclairant pour notre sujet puisqu'il porte sur le passage d'une forme de marché à une autre. L'auteur montre en effet comment on est passé d'un marché « de gré à gré » entre les producteurs et des intermédiaires qui se

[10] Voir, pour une revue riche et pédagogique de la littérature, Steiner Ph. (2011), *La Sociologie économique*, Paris, La Découverte, 4<sup>e</sup> éd.

[11] Garcia-Parpet M.-F. (1986), « La construction sociale d'un marché parfait : le marché au cadran de Fontaines-en-Sologne », *Actes de la Recherche en sciences sociales*, vol. 65, novembre.

connaissaient et nourrissaient entre eux des liens complexes obligés (puisque la familiarité sociale n'empêchait pas la divergence des intérêts économiques, à court terme en tout cas) à un marché impersonnel, reposant sur un dispositif technique : le « cadran ». Dans ce nouveau cadre, les acheteurs et les vendeurs ne se fréquentent plus ; plus exactement, les relations sociales qu'ils peuvent entretenir ne peuvent plus interférer avec le « mécanisme du marché », c'est-à-dire l'établissement d'un prix et les appariements entre offres et demandes. On voit ici à l'œuvre une logique de « performance économique » (transformation de l'économie réelle par la mise en œuvre de l'économie théorique<sup>[12]</sup>), puisque l'acteur social qui a promu cette nouvelle organisation (jeune diplômé d'une école de commerce) a explicitement cherché à rapprocher les modalités d'échange sur ce marché de l'idéal théorique que constitue le marché walrassien. Cette étude montre donc que le marché « walrassien » ne constitue qu'une forme de marché parmi d'autres, que son instauration (ou, plus exactement, l'instauration de relations marchandes qui s'en inspirent) est le produit de décisions « politiques », qu'elle exige des dispositifs matériels particuliers (ici le « cadran »), et, *in fine*, qu'elle reconfigure les rapports de force entre offreurs et demandeurs, mais aussi l'ensemble de l'organisation sociale autour de ce marché, tout au long de la « filière » qui mène les fraises des serres des producteurs aux assiettes des consommateurs.

### **Le marché de l'assurance-décès**

Un problème différent est posé par l'étude de Viviana A. Zelizer sur l'instauration du marché de l'assurance décès aux États-Unis dans la seconde moitié du XIX<sup>e</sup> siècle. En effet, dans le cas des fraises, le produit existe et son caractère marchand n'est pas mis en doute ; il s'agit seulement du passage d'une forme de transaction à une autre. En revanche, dans le cas de l'assurance-décès, l'enjeu est la construction même de la marchandise, qui est un bien immatériel, et de

son acceptabilité sociale : peut-on « jouer de l'argent » sur la mort ? On voit ici à l'œuvre la dimension morale du marché qui renvoie aux interrogations polanyiennes. Quels sont les biens qui peuvent être soumis aux « lois du marché » et quels sont ceux qu'il faudrait au contraire protéger contre ce mécanisme social aveugle d'allocation des ressources<sup>[13]</sup> ? L'étude de Zelizer montre que la question ne peut se résoudre au seul sein de l'institution marchande. C'est un changement sociétal dans la considération même du statut de la mort qui a pu rendre possible le marché de l'assurance-vie aux États-Unis<sup>[14]</sup>.

### **Donneurs et receveurs d'organes**

Dans le même ordre d'idées, Philippe Steiner a étudié les mécanismes de circulation des organes entre un « donneur » et un « receveur » en montrant précisément que nos sociétés « libérales » avaient, en ce cas d'espèce, mis en place des dispositifs technico-juridiques complexes pour éviter l'allocation marchande de la ressource extrêmement rare et demandée que constituent les organes<sup>[15]</sup>.

### **Un marché scolaire ?**

Si, assurément, le champ des biens susceptibles d'être mis sur le marché est, dans nos sociétés, plus large qu'il ne l'a probablement jamais été dans l'histoire de l'humanité, tous les biens ne sont pas marchandisables et d'autres dispositifs sociaux d'allocation des ressources existent. Ainsi, faut-il se méfier également de l'usage métaphorique de la notion de marché pour désigner tout mécanisme d'allocation de ressources. On a souvent coutume, par exemple, de parler de « marché scolaire » pour rendre compte de l'appariement entre des élèves, définis par leurs attributs sociaux (origines sociales, niveaux scolaires), et les établissements, plus ou moins cotés. Ici, aussi, la société connaît un problème d'allocation de ressources rares (les places dans les « bonnes » écoles), mais les mécanismes sociaux d'allocation de ces ressources ne sont pas, en France, pour l'essentiel, de nature marchande, ne serait-ce que parce qu'ils ne

[12] Muniesa F. et Callon M. (2013), « La performativité des sciences économiques », in Steiner Ph. et Vatin F. (dir.), *op. cit.*

[13] La nature des biens dont la marchandisation est rejetée varie suivant les sociétés. On comprend les tabous qui s'attachent au corps ou à la mort des hommes, mais peut jouer aussi le statut social des animaux, voire même de leur produit (traditionnellement les Maures rejetaient la commercialisation du lait), sans parler du cas de la terre, dont l'appropriation privée et surtout le transfert de propriété par le marché ont été rejetés longtemps dans la plupart des sociétés.

[14] Zelizer V. A., (1979), *Morals and Markets. The Development of Life Insurance in the United States*, New Brunswick, Transaction books. Soulignons l'euphémisation qui consiste, encore aujourd'hui, à dénommer « assurance-vie » ce qui est en fait « assurance-décès ».

[15] Steiner Ph. (2010), *La Transplantation d'organes : un commerce nouveau entre les êtres humains*, Paris, Gallimard.

s'appuient pas sur des prix. Rien n'interdit, pourtant, la mise en place d'un marché du produit éducatif, qui existe effectivement aux marges de notre système d'enseignement (ne serait-ce que pour le soutien scolaire). Mais, pour le moment en tout cas, qu'on le regrette ou que l'on s'en félicite, l'organisation sociale de l'espace éducatif français reste fondée sur des logiques non marchandes.

### Les marchés financiers

Un dernier exemple mérite d'être cité dans cette brève revue, celui des marchés financiers modernes. On est ici au plus proche, en apparence tout au moins, du marché walrasien idéaltypique, puisque l'information circule à la vitesse de la lumière dans les réseaux électroniques et que les transactions s'opèrent en fonction d'une seule variable : le prix. Pourtant, ces marchés ont des particularités sociales significatives. D'une part, comme on l'a vu, ils reposent, tel le marché au cadran des maraîchers solognots, sur des infrastructures matérielles, mais ici beaucoup plus complexes : ces fameuses « salles de marché » qui sont des nœuds de réseaux informatiques particulièrement denses, dotés de multiples écrans et autres outils de communication. Si les biens échangés sur ces marchés apparaissent de plus en plus « virtuels », au sens, non pas d'une déconnexion, mais d'une médiatisation croissante par rapport aux biens économiques « tangibles », en revanche, les transactions ne peuvent s'opérer que grâce à la mise en œuvre d'un appareillage technique qu'il faut analyser comme on le ferait d'un quelconque outil industriel. Or, comme l'ont montré les études de Donald McKenzie, cet appareillage est de plus en plus doté d'automatismes marchands incorporés dans les logiciels qui sont fondés sur

les théories économique-financières mathématisées du comportement rationnel sur un marché de valeurs<sup>16</sup>. Moins que nulle part ailleurs, le marché n'apparaît ici comme la résultante mécanique du libre jeu spontané des acteurs, puisque ce jeu est étroitement encadré par des formules résultant de la théorie, laquelle est censée rendre compte de la façon dont les acteurs agiraient en situation de marché pur.

\* \* \*

Les marchés, même ceux qui se rapprochent le plus des modèles élaborés par la théorie économique, apparaissent donc toujours, au regard de l'enquêteur sociologue, comme des institutions sociales spécifiques appuyées sur des dispositifs matériels. Le marché des économistes (la rencontre mathématique d'une courbe d'offre et d'une courbe de demande) n'est que la formulation résumée, et après coup, de mécanismes complexes d'appariement entre acteurs économiques. L'établissement du prix lui-même dépend de ces formes marchandes : le rapport de force entre vendeurs et acheteurs de fraises est modifié par l'introduction du cadran, comme l'établissement des valeurs financières est modifié par l'introduction des formules de mathématiques financières dans les ordinateurs des salles de marché. Pensées comme des relations sociales comme les autres, les relations marchandes, multiples dans leur forme, ne peuvent être considérées comme désocialisatrices par nature. Pour autant, toutes les sociétés, même la nôtre, mettent des bornes à l'emprise des relations marchandes, qui ne constituent pas la seule modalité possible d'allocation des biens, et, *a fortiori*, la seule forme concevable de relations d'échanges entre les hommes.

[16] McKenzie D. et Millo Y. (2003), « Construction d'un marché et performance théorique. Sociologie historique d'une bourse de produits dérivés financiers », *Réseaux*, n° 122.

## POUR EN SAVOIR PLUS

❖ **STEINER PH.** (2011), *La sociologie économique*, Paris, La Découverte, coll. « Repères ».

❖ **STEINER PH. et VATIN F.** (2013), *Traité de sociologie économique*, Paris, PUF, coll. « Quadrige ».

Condition nécessaire à l'efficacité des marchés, la concurrence est également un processus fragile et autodestructeur. En effet, sans intervention publique, l'activité économique peut déboucher spontanément sur la concentration et l'accaparement du marché par quelques firmes. Or, les situations de monopole et d'oligopole sont dommageables à long terme pour la collectivité car les entreprises en position dominante pratiquent des prix plus élevés tout en étant moins incitées à innover. La politique de la concurrence, qui vise à éviter les abus de position dominante et les ententes anticoncurrentielles, fait néanmoins partie des politiques publiques les plus contestées, comme le montre Frédéric Marty. On lui reproche notamment de réduire la motivation des entreprises les plus innovantes, incitées à prospérer par les perspectives de surprofit liées au dépassement des concurrents. Tout l'enjeu des politiques de concurrence est ainsi de sanctionner les pratiques déloyales sans réduire les incitations à l'efficacité.

Problèmes économiques

# Marchés, concurrence et réglementation

© **FRÉDÉRIC MARTY**

CNRS, groupe de recherche en droit, économie et gestion  
UMR 7321 université Nice Sophia Antipolis

Le marché peut être aussi bien appréhendé comme un ordre spontané que comme un construit social<sup>1</sup>. De la même façon, la concurrence peut être conçue comme un équilibre ou comme un processus. Dans cette seconde hypothèse, on peut considérer soit que le marché conduit inexorablement à la concentration et donc à son propre blocage, soit qu'il est autorégulateur. Dans le premier cas, la politique de la concurrence est une condition essentielle à la pérennité de l'ordre concurrentiel. Dans le second, elle doit principalement veiller à la préservation de la « contestabilité » des marchés.

<sup>[1]</sup> Sur ce point, voir dans ce même numéro l'article de François Vatin, p. 118.

Ces différentes approches s'accordent sur le fait que le fonctionnement même du processus de concurrence puisse être altéré, soit par les comportements des acteurs, soit par la structure même du marché. Le premier facteur de risque relève de la politique de concurrence. Il s'agit de prévenir et de sanctionner les stratégies de marché susceptibles de compromettre la pérennité de l'ordre concurrentiel. Le second facteur relève de la politique de régulation au sens large. Il s'agit de corriger les défaillances de marché, voire de construire des marchés concurrentiels.

## La concurrence comme un processus instable ?

La concurrence est à la fois l'aiguillon essentiel de l'efficacité économique et un proces-

sus fragile, dont la dynamique même peut conduire à son propre blocage.

Tout d'abord, les acteurs du marché ont rationnellement intérêt à échapper à la concurrence. Ils pourront réaliser un surprofit en abusant d'un pouvoir de marché résultant d'une position dominante. Une situation de monopole, ou dans une moindre mesure une configuration de cartel, permettent de dégager un surprofit par rapport à l'équilibre de concurrence pure et parfaite où chaque facteur de production n'est rémunéré qu'au niveau de son coût marginal. Échapper à la concurrence permet également de maîtriser les risques de marché. Comme le notait John Hicks en 1935, le plus grand profit du monopole n'est pas tant un surcroît de rentabilité que la garantie d'une vie paisible, à l'abri des risques concurrentiels<sup>2</sup>.

La concentration du pouvoir économique a plusieurs conséquences négatives. Elle se traduit tout d'abord par l'imposition de prix supérieurs aux prix de concurrence pure et parfaite, entraînant à la fois une perte sèche de bien-être dans l'économie et un transfert de richesses entre les agents au détriment de ceux dépourvus de pouvoir de marché. En termes dynamiques, elle conduit également à un coût économique majeur en amoindrissant les incitations des firmes à innover.

Apparue pour la première fois aux États-Unis en 1890 avec le Sherman Act<sup>3</sup>, la politique de concurrence doit contrecarrer cette tendance naturelle et inexorable. Les deux premières sections de cette loi constituent la matrice des politiques de concurrence qui se généraliseront au xx<sup>e</sup> siècle. La première sanctionne les ententes entre concurrents et la seconde les entreprises qui acquièrent, maintiennent ou étendent une position de monopole sur d'autres bases que celles relevant de leurs propres mérites. La législation européenne a repris cette même structuration, tout en s'en distinguant sur plusieurs aspects. Il ne s'agit tout d'abord pas d'un simple transplant de la législation américaine. Ses racines sont internes et peuvent

être au moins pour partie rattachées à l'école ordolibérale allemande (Gerber, 1998), qui a une vision plus pessimiste du processus de concurrence et admet une intervention étatique pour en assurer la pérennité. Ensuite, ce n'est pas la position dominante qui est sanctionnée en elle-même mais son abus, lequel peut passer, au-delà des pratiques d'éviction des concurrents, par l'imposition de prix excessifs. Or, dans le cadre américain, une firme qui a obtenu une position de monopole par ses propres mérites (notion recouvrant une efficacité supérieure, un avantage lié à l'arrivée plus précoce sur le marché, voire la simple chance) est libre de pratiquer des prix de monopole. Cette liberté est vue comme essentielle à la fois pour récompenser les investissements et les prises de risques passés et pour inciter des concurrents potentiels à entrer sur le marché et donc à revigorer le processus de concurrence.

## Les politiques de concurrence comme outil de dissuasion contre les atteintes au marché

Face aux risques liés à la concentration du pouvoir économique, les finalités des politiques de concurrence se déclinent en trois volets.

### Dissuader toute pratique préjudiciable en termes de bien-être du consommateur

La politique de concurrence a en premier lieu un rôle dissuasif. Il s'agit, par l'imposition de sanctions pécuniaires, de dissuader les entreprises d'acquiescer une position dominante par d'autres moyens que ceux relevant d'une concurrence par les mérites ou d'en abuser au détriment des consommateurs. L'activation des règles de concurrence peut également permettre de remédier à de telles situations au travers de mesures correctives

<sup>[2]</sup> Hicks J.R., (1935), « Annual Survey of Economic Theory: The Theory of Monopoly », *Econometrica*, vol. 3, n° 1.

<sup>[3]</sup> Pour une présentation générale des politiques de concurrence, voir Combe E. (2005) et Marty F. (2010).



de nature comportementale ou structurelle. Il peut s'agir, par exemple, d'injonctions conduisant à cesser la commercialisation d'une offre ayant pour effet d'étendre une position dominante d'un segment de marché à un autre (*via* des offres groupées). Il peut également s'agir de remèdes structurels induisant des cessions d'actifs dès lors que l'abus serait considéré comme consubstantiel à une structure de marché donnée.

### Garantir l'accès au marché

La politique de concurrence peut également viser à garantir l'accès des firmes au marché. Il s'agit de faire pièce aux stratégies de « forclusion », c'est-à-dire de verrouillage anticoncurrentiel, que pourraient mettre en œuvre des opérateurs dominants. Les injonctions en termes d'accès à des réseaux (théorie dite des facilités essentielles) ou des licences obligatoires de brevets sur la base de conditions dites FRAND (*fair, reasonable and not discriminatory*) font écho à cette logique.

### Maintenir la diversité de l'offre

Il peut également s'agir de défendre la liberté de choix du consommateur. Même si une situation de monopole peut s'avérer plus efficace qu'une configuration concurrentielle, la politique de concurrence peut veiller à laisser disponibles des offres alternatives pour le consommateur. La décision de la Commission européenne dans l'affaire Microsoft en 2004 témoignait de ce souci. L'efficacité dynamique (maintien des incitations à innover du fait de la menace concurrentielle) peut être invoquée en contrepartie de l'acceptation d'une moindre efficacité de court terme. Dans certains domaines, à l'instar des médias, la politique de concurrence peut porter des objectifs non exclusivement économiques, tel le maintien d'un degré suffisant de pluralisme.

## Quels risques concurrentiels et quels outils de politiques publiques ?

### Les risques concurrentiels

Les atteintes au marché peuvent prendre des formes diversifiées. Les modalités les plus évidentes sont les pratiques anticoncurrentielles.

Il peut tout d'abord s'agir de pratiques unilatérales. Une firme porte atteinte au processus concurrentiel sans escompter une réaction coopérative de ses concurrents. Il s'agit, en droit de l'Union européenne, des abus de position dominante. Ceux-ci peuvent prendre deux formes ; les abus d'exploitation et les abus d'éviction. Les premiers recouvrent les situations dans lesquelles les firmes tirent profit de leur pouvoir de marché pour pratiquer des prix supérieurs aux niveaux concurrentiels. Les seconds correspondent à des stratégies d'éviction des concurrents. Ils peuvent passer par des pratiques tarifaires, telles des stratégies de prédation ou de compression des marges, des remises de fidélité auxquelles les concurrents ne peuvent répondre, ou encore par des stratégies non tarifaires tenant, par exemple, à des refus d'accès à des actifs essentiels, à des offres groupées, à des stratégies de préemption ou de forclusion basées sur l'obtention de clauses d'exclusivité, etc. Il s'agit, au final, de stratégies permettant d'évincer du marché un concurrent *a priori* aussi efficace que soi.

Les pratiques anticoncurrentielles peuvent également être coordonnées entre concurrents. Il peut s'agir d'accords verticaux (par exemple des accords de distribution exclusive) ou d'accords horizontaux allant de l'échange d'informations entre concurrents jusqu'au cartel, forme la plus aboutie de coordination horizontale. Les ententes anticoncurrentielles sont les pratiques les plus préjudiciables au consommateur et ne peuvent trouver de justification en termes de gains d'efficience. À ce titre, elles sont les plus lourdement sanctionnées.

## Les outils des politiques de concurrence

### *Sanctions pécuniaires*

En matière de pratiques anticoncurrentielles, les sanctions pécuniaires constituent un outil de dissuasion privilégié. Le montant optimal de l'amende doit être égal au surprofit lié à la pratique anticoncurrentielle pondéré par la probabilité de détection et de sanction. En d'autres termes, si la probabilité d'être sanctionné est de un sur dix, la sanction optimale devra atteindre dix fois le surprofit. Dans les faits, le montant de la sanction est basé sur un pourcentage de la valeur des ventes liées à l'activité en cause et plafonnée à 10 % du chiffre d'affaires du groupe auquel appartient l'entreprise. Pour renforcer cet effet dissuasif, l'antitrust américain s'appuie sur le versement de dommages et intérêts et sur la criminalisation de certaines pratiques anticoncurrentielles. Ce volet pénal fait courir un risque de sanctions pécuniaires pour les personnes physiques ainsi que de peines privatives de liberté. Sur le continent européen, une directive de novembre 2014 sur les actions en réparation des dommages liés aux pratiques anticoncurrentielles vise également à renforcer l'effet dissuasif de la sanction en favorisant le dédommagement des victimes des pratiques anticoncurrentielles.

### *Remèdes comportementaux et structurels*

Des remèdes comportementaux et structurels, sur le modèle des mesures correctives dans le domaine du contrôle des concentrations, peuvent également être mobilisés. Dans le cadre de procédures contentieuses, il s'agit d'injonctions ; dans le cadre de procédures négociées, il s'agit d'engagements volontaires pris par les entreprises pour obtenir la clôture de la procédure sans une décision reconnaissant formellement une infraction aux règles de concurrence. Au travers de ces mesures correctives, les firmes dominantes doivent ajuster leurs stratégies pour prévenir l'éviction de concurrents (octroi de licences d'exploitation de brevets, renonciation à des

contrats de long terme...) ou accepter des cessions d'actifs pour réduire leur puissance de marché.

### *Contrôle des concentrations et encadrement des aides publiques*

Il convient également de prendre en considération les deux autres volets de la politique de concurrence, à savoir le contrôle des concentrations et l'encadrement des aides publiques. Le contrôle préalable des concentrations vise à éviter que les firmes n'atteignent une situation de monopole ou ne bénéficient d'une configuration de marché si étroitement oligopolistique qu'un équilibre collusif même tacite puisse être aisément obtenu. Le rôle de la politique de la concurrence est d'interdire de telles opérations de concentrations ou de les conditionner à des pratiques de nature à éviter qu'elles ne puissent porter des dommages irréversibles à la concurrence. L'encadrement des aides d'État est une spécificité européenne. Il vise à juguler deux types de distorsions de concurrence découlant de soutiens publics à des entreprises. Celles-ci peuvent résider en des avantages concurrentiels induits au profit des firmes aidées ou en des risques de processus de concurrence fiscale entre les États membres. Pour autant, les aides publiques ne sont pas interdites. Elles peuvent être autorisées dans la mesure où elles soutiennent des objectifs relevant de la stratégie de croissance de l'Union, visent à compenser des handicaps économiques régionaux ou des coûts liés à la prise en charge de missions d'intérêt économique général. Il s'agit alors pour la Commission européenne d'apprécier la nécessité de l'aide et de jauger sa proportionnalité.

## Les politiques de concurrence en débat

### *Une politique récente et contestée*

Malgré leur place dans la politique européenne, les politiques de concurrence furent longtemps contestées et ne jouissent que d'une faible adhésion.

Un recul historique montre que les cartels, qui constituent pourtant la pratique anti-concurrentielle par excellence, ont fait l'objet d'une longue tolérance. Un arrêt de la Cour suprême allemande du 4 février 1897 reconnut leur licéité sur la base de la liberté contractuelle. De nombreux économistes les défendirent en raison des gains d'efficacité liés à la concentration et à la rationalisation des capacités de production ou encore comme une condition nécessaire à la compétitivité internationale. Après la grande guerre, la coordination entre les firmes fut défendue sur la base d'une concurrence régulée appelée à se substituer à une concurrence « coupe-gorge ». Cette approche, qui entretient quelques liens avec la notion de cartel de crise, trouva une consécration aux États-Unis dans les premières années du New Deal avec le NIRA (National Industry Recovery Act) avant d'être abandonnée dès 1937 au profit d'une relance de la politique antitrust.

Des tensions analogues existent pour les pratiques unilatérales. En 1891, la décision de la Chambre des Lords dans l'affaire *Mogul Steamship* mit en doute la possibilité même de sanctionner une entreprise pour avoir évincé du marché une firme concurrente. Les lois antitrust américaines n'interdisent pas à une firme dominante de pratiquer les prix qu'elle désire, fussent-ils des prix de monopole. La difficulté de caractériser une pratique comme anticoncurrentielle peut faire craindre le risque de multiplications de décisions sanctionnant à tort des entreprises qui n'ont pratiqué qu'une concurrence par les mérites.

Ainsi, la politique de concurrence peut-elle s'avérer porteuse de risques juridiques pour les entreprises dominantes. Cet argument fut celui de l'École de Chicago aux États-Unis dans les années 1960 et 1970 pour contester l'activisme des juridictions chargées de l'application du Sherman Act et pour plaider en faveur d'un cadre plus sécurisant pour les opérateurs dominants.

La politique de concurrence a ainsi pu, et peut toujours apparaître, comme une anti-

politique industrielle, susceptible d'entraver le développement de champions nationaux (ou européens) ou de pénaliser les firmes par rapport à des concurrents étrangers. La politique de la concurrence pourrait également être vue comme une entrave au développement d'entreprises innovantes, notamment dans le secteur des technologies de l'information et de la communication qui se caractérise par un très fort rythme d'innovations, des rendements croissants, de forts effets de réseaux et donc par des situations de quasi-monopole. Une intervention trop précoce ou imposant des remèdes disproportionnés ou dont les effets seraient irréversibles pourrait donc porter préjudice au processus de concurrence et donc au consommateur.

Cependant, malgré les risques induits, la politique de concurrence peut s'avérer essentielle pour la défense des intérêts du consommateur. À court terme, elle le protège de prix abusifs ou de conditions contractuelles déséquilibrées ; à long terme, elle prévient l'apparition de positions dominantes irréversibles. Dans le domaine des hautes technologies logicielles, le fait de disposer d'une large base de clientèle peut rendre irrémédiablement plus efficace une firme que ses concurrents, même si ces derniers sont devenus plus innovants. De la même façon, l'opérateur dominant peut être naturellement choisi par des partenaires commerciaux (par exemple en matière de droits audiovisuels) pour négocier des accords d'exclusivité qui risquent de structurer l'industrie en autant de silos étanches et donc de verrouiller le consommateur en lui imposant de forts coûts potentiels de changements d'opérateur.

### **De la défense de l'ordre concurrentiel à la construction du marché ?**

Cependant, il n'y a pas loin entre la défense de l'ordre concurrentiel par la prévention de situation de dominance irréversible et une possible tentation de constructivisme concurrentiel. En d'autres termes, alors que

la politique de concurrence vise sur le principe à sanctionner *ex post* des pratiques qualifiées d'anticoncurrentielles, il est possible de s'interroger sur une possible action *ex ante* passant par un jeu sur les structures de marché et sur les degrés de liberté des firmes dominantes. La pratique décisionnelle de la Commission européenne, et surtout de la Cour de justice de l'Union européenne participe d'une telle logique de construction du marché intérieur (Gerber, 1998). L'application des règles de la concurrence dans le domaine de l'énergie montre comment la Commission s'est saisie de cette politique pour peser sur la structuration du marché, en obtenant au travers de décisions, souvent issues de procédures négociées, des remèdes comportementaux ou structurels, tels des cessions d'actifs de production ou encore des réseaux de transport de l'énergie, la conduisant ainsi à construire une structure de marché qu'elle ne pouvait obtenir des États membres par voie de directives.

### De l'insuffisance du recours aux règles de concurrence pour garantir l'ordre concurrentiel

Cependant, si l'action au travers des règles de concurrence peut dissuader les firmes d'engager des pratiques anticoncurrentielles et à la marge corriger certaines imperfections de marché, elle peut s'avérer insuffisante dès lors que les structures mêmes du marché font obstacle au développement du processus de concurrence ou que l'on doit faire face à des défaillances de marché.

Il peut en aller ainsi lorsque le fonctionnement spontané du marché ne permet pas de parvenir à une solution souhaitable en termes collectifs. Par exemple, dans le domaine de l'électricité, les centrales sont appelées par ordre de coût marginal de production croissant (ordre de préséance économique). En d'autres termes, les centrales les moins efficaces ne sont mobilisées qu'un nombre extrêmement réduit d'heures pendant l'année, voire jamais si la demande n'atteint pas son

pic théorique. Ce faisant, ces capacités ne couvrent pas obligatoirement leurs coûts et devraient en toute logique être démantelées. Pour autant, elles protègent potentiellement le système électrique du risque de *black-out*. Le fait qu'elles demeurent en réserve fournit donc un service essentiel à la collectivité qui n'est pas rémunéré par le marché. Une intervention publique est nécessaire pour les maintenir en service.

Une régulation publique intervenant *ex ante* peut donc être indispensable pour prévenir les dommages à la concurrence, pour permettre au processus de concurrence de se développer et pour porter des objectifs relevant de l'intérêt général et dépassant donc le strict bien-être du consommateur. Il en est par exemple ainsi dans le domaine de l'internet. La position de marché de Google a suscité des appels à une régulation publique, non seulement sur la base d'une possible disparition progressive de la contestabilité même de sa position de marché, mais aussi sur la base de sa capacité à réguler lui-même le marché. Ainsi, l'intervention publique pourrait-elle être défendue non seulement sur la base de la prévention d'un dommage concurrentiel (en termes de bien-être du consommateur), mais également en termes de prévention d'un possible risque de régulation du marché (c'est-à-dire des conditions tarifaires ou techniques d'accès au marché) par des pouvoirs économiques privés. Le débat sur la neutralité de l'internet actuellement en cours aux États-Unis et la prise de position du Président Obama pour une régulation publique témoignent de la nécessité dans certaines configurations d'aller au-delà de la seule politique de concurrence.

La régulation sectorielle participe d'une logique de politique industrielle allant au-delà de la défense de l'ordre concurrentiel mais portant également des objectifs de développement des infrastructures ou de soutien aux investissements. À ce titre, la politique de concurrence seule est insuffisante pour garantir le bon fonctionnement à long terme de certaines industries, notamment quand

subsistent des segments en monopole naturel ou existent des dimensions d'intérêt général que les signaux du marché ne peuvent relayer. L'adage « concurrence autant que possible, État (ou régulation) autant que nécessaire » s'applique alors pleinement. Cette complémentarité est d'autant plus importante qu'elle peut prendre des formes diverses dans le temps et selon les industries. Dans certains cas, comme celui des télécommunications, il est possible que la régulation mise en place au moment de l'ouverture à la concurrence s'efface peu à peu au profit de la seule application *ex post* des règles de concurrence. Dans d'autres secteurs, à l'inverse, le besoin de régulation sectorielle pourra être pérenne. C'est comme nous l'avons vu le cas dans le secteur de l'énergie. Plus généralement, plus les dimensions de biens publics, d'externalités et de monopoles naturels sont fortes, moins la libéralisation se traduit par une dérégulation. En d'autres termes, la libéralisation crée un besoin en régulation que la seule application des règles de concurrence ne peut satisfaire.

<sup>[4]</sup> Easterbrook F. et Fischer D. (1996), *The Economic Structure of Corporate Law*, Harvard University Press.

Il apparaît donc que malgré les défauts de la régulation publique et les possibilités d'instrumentalisation des règles de concurrence, les logiques d'autorégulation des marchés ne puissent réellement être envisagées. En 1996, Easterbrook et Fisher s'interrogeaient sur l'efficacité de la régulation en matière de sanction aux manquements des opérateurs en matière de diffusion d'informations financières au marché, considérant que l'anticipation d'un dommage réputationnel irréversible suffit à discipliner les firmes<sup>4</sup>. En d'autres termes, la menace de sanction du marché suffirait à les dissuader. Au vu de l'expérience de la crise financière de 2007-2008 et des ententes anticoncurrentielles sanctionnées sur les marchés du LIBOR et de l'EURIBOR par la Commission européenne, la nécessité de protéger l'ordre de marché par la politique de concurrence et, si nécessaire, par une régulation publique spécifique apparaît indispensable.

## POUR EN SAVOIR PLUS

❖ **COMBE E.** (2005), *Économie et politique de la concurrence*, Paris, Dalloz/Sirey, coll. « Précis Dalloz ».

❖ **GERBER D.** (1998), *Law and Competition in Twentieth Century Europe – Protecting Prometheus*, Oxford, Clarendon Press.

❖ **LAFFONT J.-J. et TIROLE J.** (1993), *A Theory of Incentives in Regulation and Procurement*, Cambridge, MIT Press.

❖ **LÉVÊQUE F.** (2004), *Économie de la réglementation*, Paris, La Découverte, coll. « Repères ».

❖ **MARTY F.** (2010), « La politique de la concurrence », in Montel-Dumont O., *La Politique économique et ses instruments*, Paris, La Documentation française, coll. « Notices ».

Le volume et la rapidité des transactions ainsi que la réglementation, perçue comme minimale, font que les marchés financiers sont souvent considérés comme relativement proches de l'idéal concurrentiel. S'ils possèdent en effet, par rapport aux autres marchés, des propriétés qui les en rapprochent – un fort degré d'atomicité et de transparence –, ils ne correspondent pas non plus parfaitement au modèle de la concurrence pure et parfaite, qui reste, comme nous le rappelle Yves Jégourel, une abstraction.

Problèmes économiques

# Marchés financiers : des marchés proches de l'idéal concurrentiel ?

Les marchés financiers sont souvent perçus dans l'imaginaire collectif comme un monde déconnecté de la sphère dite « réelle » où la réglementation est minimale et dans lequel les profits sont considérables. Symboles, pour certains, d'un libéralisme triomphant, ils incarneraient en cela un idéal concurrentiel. La faiblesse des coûts de transaction, la rapidité d'exécution des ordres inférieure parfois à la milliseconde, la libre fluctuation des prix et la forte volatilité qui en découle traditionnellement laissent en effet à penser que la loi de l'offre et de la demande joue pleinement et en permanence, plus que sur n'importe quel autre marché. Bien que cette vision ne soit pas totalement infondée, la réalité est cependant plus complexe pour trois raisons intimement liées aux notions mêmes de marchés financiers et de concurrence.

Il faut, en premier lieu, prendre acte du fait que les marchés financiers ne forment pas un « tout ». Ils fonctionnent selon des mécanismes différents et poursuivent des ambitions variées. Les marchés de capitaux ne

regroupent ainsi pas uniquement les marchés obligataires ou d'actions, mais concernent également celui des « produits dérivés », permettant de gérer les risques financiers<sup>1</sup>, ou le marché des changes, appelé « Forex<sup>2</sup> » sur lequel s'échangent des devises contre d'autres devises. Les marchés financiers ne proposent de plus pas tous le même mode de transaction : ils ne reposent en effet pas nécessairement sur le principe d'une Bourse et n'ont pas toujours une réalité physique. On oppose de ce point de vue les marchés organisés et les marchés de gré à gré. Reconnaissons d'ailleurs que la notion de marché, qu'il soit financier ou non, peut être largement théorique : il n'est que le lieu, réel ou non, de rencontre d'une offre et d'une demande. La notion de « marché de l'immobilier » n'a par exemple de sens que pour parler d'une offre

© YVES JÉGOUREL

Maître de conférences, université de Bordeaux

<sup>[1]</sup> Voir Jégourel (2012) sur ce sujet.

<sup>[2]</sup> Par contraction de « FOREIGN EXchange market ».

et d'une demande de logements et faire émerger un prix d'équilibre, dont la valeur informationnelle n'est en définitive que très relative. Certains marchés financiers, enfin, ont un rayonnement international et ont valeur de référence, alors que d'autres n'ont, au mieux, que pour seule vocation d'orienter les flux d'épargne nationale, limitant leur importance dans le système économique.

Rappelons, en second lieu, que le concept de « concurrence » est, à l'inverse de celui de « marché financier », particulièrement précis, pour le moins dans sa forme absolue que les économistes néoclassiques appellent la concurrence pure et parfaite (CPP). Pour qu'un marché fonctionne sous un tel régime, il faut en effet que plusieurs conditions soient réunies :

- atomisticité : un nombre important d'acheteurs et de vendeurs présents, sans comportement collusif, de telle sorte qu'aucun d'entre eux n'ait le pouvoir d'influencer durablement les prix ;

- homogénéité : le produit échangé sur un tel marché est de qualité homogène, poussant les acteurs présents sur le marché à ne fonder leur choix d'achat ou de vente que sur le prix ;

- transparence : l'information est parfaitement disponible, sans coût, et aucun agent économique ne dispose d'une information que les autres n'ont pas ;

- absence de barrière à l'entrée ou à la sortie : tout agent peut entrer ou sortir sans délai de ce marché ;

- libre circulation des facteurs de production : les facteurs capital et travail, nécessaires à l'activité productive, sont libres d'entrer ou de sortir du marché en fonction des conditions de rémunération (rendements et salaires) offertes.

À la lecture de cette énumération, on comprendra aisément que peu de marchés répondent à des critères aussi précis. Dans cette optique, il s'agit en réalité de s'interroger sur la capacité des marchés financiers à approcher un tel idéal. Cela revient notamment à déterminer

si leur accès est aisé, si les transactions sont rapides, si la diffusion de l'information est efficace et si aucun agent économique présent n'a de réelle influence sur les prix.

Il importe enfin de préciser à quels niveaux les mécanismes concurrentiels sont observés : entre les places financières qui cherchent à attirer émetteurs et investisseurs ou entre acteurs des marchés financiers eux-mêmes, qu'il s'agisse des émetteurs, des investisseurs ou des spéculateurs ? Dans le cas des places financières qui sont animées par des entreprises à part entière telles que Euronext, force est de reconnaître que les hypothèses de concurrence parfaite ne sont pas vérifiées. C'est au contraire une concurrence imparfaite de type oligopolistique ou monopolistique qui s'observe, où chaque bourse tente de parvenir à une situation de monopole sur un produit grâce au processus de différenciation. Une action ou obligation cotée à New York ne le sera *a priori* pas à Paris et ne sera de toute façon pas libellée dans la même monnaie. Les produits financiers permettant de gérer les risques de prix sur matières premières agricoles seront, quant à eux, massivement présents à Chicago, tandis que la place financière de référence pour gérer le risque de prix sur les métaux se situe à Londres. Lorsque ce questionnement porte sur les relations entre les acteurs de ces marchés et non sur celles caractérisant les marchés eux-mêmes, l'analyse devient en revanche plus complexe.

## Des marchés financiers : une accessibilité relative

Bien que la notion de financiarisation de l'économie, traduisant le fait que les marchés financiers occupent une place croissante dans le système économique des pays industrialisés et émergents, soit bien connue des médias et des étudiants en économie, il importe de rappeler que l'accès aux marchés financiers n'est pas libre et que le recours à un intermédiaire est bien souvent nécessaire.

C'est ici que la distinction entre marchés organisés fonctionnant sur le principe d'une Bourse, et marchés dit de gré à gré ou OTC (*over the counter*) prend tout son sens.

## Marchés OTC

À la différence des marchés organisés, les marchés OTC n'imposent pas le recours à une chambre de compensation comme intermédiaire obligataire entre un acheteur et un vendeur : les transactions sont ainsi réalisées directement entre intervenants. Le marché des changes, qui compte parmi les marchés OTC les plus importants, est d'une taille considérable avec un volume de transactions journalier de près de 5 300 milliards de dollars en 2013, selon les données de la Banque des règlements internationaux (BIS, 2014). L'accès à ce marché est en théorie libre puisque tout opérateur qui achète ou vend des devises est considéré comme appartenant à ce marché. En pratique néanmoins, toute transaction impose le recours à ce qu'il est convenu d'appeler un teneur de marché, dont le rôle est précisément de se porter systématiquement contrepartie des offres d'achat ou de vente de devises. Ce rôle est assumé par la plupart des banques commerciales, expliquant ainsi pourquoi le marché des changes offre un degré de concertation des intervenants élevé : seulement 9 % des transactions étaient en effet, en 2013, réalisées par des contreparties non financières.

## Marchés organisés

Sur les marchés organisés, qu'il s'agisse des marchés d'actions, d'obligations ou de certains produits dérivés, toute opération d'achat ou de vente impose de recourir à un membre de la Bourse. Pour un particulier comme une entreprise, acheter un titre financier est un processus *a priori* simple, mais qui nécessite de trouver au préalable un courtier, appelé *broker*. Si les marchés financiers nationaux sont, de ce point de vue, faciles d'accès car ce rôle d'intermédiation est assumé par les banques, tel n'est pas le cas pour des marchés

étrangers. À titre d'exemple, il serait ainsi compliqué, sinon impossible, pour un épargnant français d'investir directement sur des bourses chinoises. La théorie financière met en outre en évidence que quiconque souhaite placer ses capitaux en Bourse doit diversifier ses supports d'investissements. Ainsi, si un individu peut dans l'absolu accéder à des actions peu coûteuses, il sera contraint, s'il souhaite optimiser son portefeuille financier, de multiplier les investissements, conduisant *de facto* à un accroissement du capital nécessaire pour accéder efficacement au marché boursier. C'est la raison pour laquelle la plupart des épargnants privilégient traditionnellement l'achat de parts d'organismes des placements collectifs en valeurs mobilières (OPCVM), permettant de bénéficier avec un capital réduit de cet effet de diversification, plutôt que d'investir directement en Bourse. Le recours au secteur bancaire est, une fois encore, souvent nécessaire.

Pour une entreprise ou un État, se financer sur les marchés financiers – c'est-à-dire émettre des actions ou des obligations – est également un processus qui peut s'avérer long et complexe. Les conditions légales sont en effet multiples et d'autant plus contraignantes que le marché sur lequel ces titres sont émis est important. Dans le cas des marchés d'actions français, trois segments peuvent être distingués : Eurolist, Alternext et le Marché libre. Si ce dernier, comme son nom l'indique, est relativement accessible, les deux autres imposent un certain nombre de critères pour une admission à la cote : au moins 25 % du capital social (ou 5 % si le montant est équivalent à au moins 5 millions d'euros) de l'entreprise doit être émis pour une admission sur les marchés réglementés européens d'Euronext, contre 2,5 millions d'euros pour le marché Alternext dédié aux entreprises de plus petite taille ; la production de deux et trois ans d'états financiers certifiés, en norme IFRS dans le cas d'Eurolist, ainsi que l'édition d'un prospectus d'information visé par l'Autorité des marchés financiers (AMF). Autant d'éléments qui démontrent que l'accès aux



marchés financiers n'est pas simple, contredisant ainsi l'hypothèse de concurrence pure et parfaite.

### Marchés dérivés

La difficulté à accéder directement aux marchés financiers est encore plus grande lorsque les marchés dérivés sont considérés. Les produits négociés sur ces marchés sont en effet relativement techniques et peuvent, lorsqu'ils sont utilisés à mauvais escient, générer des niveaux de pertes bien supérieurs aux sommes investies. Gérer, par des instruments tels que les contrats à terme ou les options, le risque de prix qui découle de l'achat et/ou de la vente d'une matière première, impose en particulier des compétences techniques importantes et peut pousser une banque, membre de ce marché, à ne pas accorder l'accès à l'un de ses clients. Celui-ci doit par ailleurs disposer d'une surface financière suffisamment grande pour que l'utilisation de ces produits lui soit possible. Un opérateur souhaitant, à titre d'exemple, se protéger contre le risque de blé en utilisant un contrat à terme négocié sur la place financière de Paris Euronext, devra traiter un minimum de cinquante tonnes, tandis qu'un importateur souhaitant se protéger contre l'appréciation du dollar américain ne pourra le faire pour un montant inférieur à 20000 euros.

<sup>[3]</sup> London International Bank Offered Rate.

### Sortie des marchés financiers

La question de la sortie des marchés financiers se doit également d'être posée. Si tout opérateur peut en théorie librement sortir d'un marché sur lequel il est précédemment entré, le coût d'une telle opération doit également être pris en compte. Celui-ci dépend intrinsèquement de la liquidité du marché considéré, définie comme la capacité qu'a un intervenant à revendre un actif, sans que cette opération ne se traduise par une modification sensible de son prix. Une fois encore, des différences sensibles existent entre marchés financiers. Ce coût est *a priori* faible sur des marchés organisés, mais peut être important

sur les marchés financiers de gré à gré peu liquides, tels que celui des actions non cotées.

## Un prix d'équilibre librement déterminé, des marchés transparents mais des produits hétérogènes

### Atomicité

Si l'accessibilité des marchés financiers demeure relative, conduisant à rejeter l'hypothèse de concurrence pure et parfaite, tel n'est pas le cas pour la dynamique des prix qui ne semble pas pouvoir être influencée par un ou plusieurs acheteurs. L'atomicité du marché semble ici vérifiée. Sur le marché des changes, le volume de transactions est notamment tel qu'aucun intervenant ne peut prétendre influencer les cours de change, à l'exception des banques centrales. Il faut en outre rappeler que toute tentative de manipulation des cours, si elle est avérée, tombe sous le coup de la loi. Il est néanmoins nécessaire de garder en mémoire que l'histoire financière récente est émaillée d'exemples de comportements collusifs qui limitent la portée de cette affirmation. Ce qui a été appelé le « scandale du Libor<sup>3</sup> » a en effet mis en exergue que certaines banques s'étaient engagées dans une stratégie d'entente afin de fausser ce taux d'intérêt interbancaire de référence. Défini sur une base déclarative comme une moyenne des transactions réelles réalisées par dix-huit banques, le Libor reflète les conditions financières auxquelles les banques se prêtent leurs liquidités à court-terme : toute élévation de ce taux étant révélateur d'un accroissement du risque de l'emprunteur, les banques l'ont artificiellement maintenu à un niveau faible en 2008 afin de minimiser, dans la perception des observateurs, les fragilités financières qu'elles connaissaient alors. Des manipulations de marché ont également pu être observées sur les marchés financiers de matières premières, notamment ceux du cacao et des métaux de base (cuivre, aluminium, nickel).

Certaines banques, membres du London Metal Exchange (LME), marché londonien de référence pour ces métaux, ont ainsi été suspectées de jouer sur le niveau des stocks, élément déterminant du niveau des prix de ces matières premières.

## Homogénéité

Il n'est de plus pas certain que l'hypothèse d'homogénéité des produits soit vérifiée sur les marchés financiers, en dépit de l'existence d'une norme internationale de rentabilité. Il faut, pour le comprendre, rappeler les similitudes existant entre les marchés de biens et services et les marchés de capitaux. Sur le marché automobile par exemple, l'hypothèse de concurrence pure et parfaite n'est pas vérifiée car le bien échangé peut être différencié par la qualité (on parle alors de différenciation verticale) ou par la variété (différenciation horizontale). Dans le cas des marchés financiers, et notamment des marchés d'actions et obligataires, il en est de même puisque le « produit » est unique, intrinsèquement lié à la nature de l'émetteur. Cela ne signifie pas pour autant que la comparaison entre actifs financiers n'a pas de sens, bien au contraire. Les investisseurs fondent en effet leur choix sur le couple rendement-risque que l'actif financier leur procure. Si une obligation d'État français est, par nature, différente de celle émise par le Trésor britannique ou italien, celles-ci seront néanmoins systématiquement comparées au regard du rendement qu'elles offrent et du niveau de risque qu'elles impliquent, au même titre que le sont les automobiles de marques différentes, mais appartenant à une gamme similaire. Un investisseur compose ainsi son portefeuille en choisissant les titres qui satisferont à ses objectifs de rendement, compte tenu de son aversion au risque.

## Transparence

Le sujet demeure assez largement débattu, mais la transparence des marchés semble en

revanche garantie sur les marchés organisés. Grâce notamment à la présence d'agences dédiées à la production et à la diffusion d'informations financières, à la multiplication des indices boursiers visant à offrir aux investisseurs une information agrégée sur la dynamique globale des prix et en raison de la centralisation des ordres d'achat et de vente au sein de la chambre de compensation, tout agent peut, en temps réel, comparer l'ensemble du prix des actifs cotés en Bourse. Il peut ainsi réaliser des opérations d'arbitrage entre deux places financières et de ce fait favoriser la cohérence des prix. En effet, si une même action cotée sur deux Bourses différentes, Paris et New York, venait à avoir des prix différents, les opérateurs l'achèteraient immédiatement, là où elle est la moins chère, pour simultanément la revendre au prix le plus élevé, contribuant de ce fait à un rééquilibrage entre places. Une fois encore, ce constat évolue lorsque les marchés OTC sont considérés. Les actifs étant par définition non standardisés et les opérations qui y sont réalisées étant, à la différence des marchés organisés, de nature strictement privée, leur prix ne peut être obtenu, empêchant *de facto* toute tentative de comparaison entre produits. La crise des *subprimes* de 2007 s'est ainsi transformée en crise financière mondiale une année plus tard car un des produits dérivés les plus en vogue alors, le *Credit Default Swap* (CDS) était échangé sur un marché OTC particulièrement opaque. Au moment de la faillite de Lehman Brothers, en septembre 2008, personne ne savait en effet combien de CDS avaient été émis et par qui. Le paradoxe veut pourtant qu'un tel produit ait pour fonction de venir indemniser un investisseur lorsqu'un ou plusieurs émetteurs des obligations qu'il détient fait défaut. Les opérateurs du marché ne connaissant pas, par nature, le montant des indemnités requises à la suite des faillites à répétition, il n'en fallait pas plus pour que la panique se diffuse. C'est une des raisons pour lesquelles les autorités financières nationales et internationales se sont attachées depuis à imposer aux marchés

OTC une compensation centrale des transactions, les rapprochant *de facto* des marchés organisés.

\* \* \*

Les éléments précédents démontrent qu'il est difficile de conclure de façon radicale sur le type de concurrence prévalant sur les marchés financiers. La grande diversité des mécanismes et des objectifs qui les caractérisent explique assez largement cette indétermination. Si les marchés organisés peuvent présenter certaines caractéristiques de la concurrence pure et parfaite, et notamment la pleine disponibilité des prix, tel ne semble

pas être le cas pour les marchés OTC qui demeurent, par nature, moins transparents. La concurrence y est forte, mais elle n'est ni pure ni parfaite. Si l'on porte ce raisonnement au niveau des places boursières elles-mêmes, l'hypothèse de concurrence imparfaite – existence d'oligopoles ou de monopoles locaux – apparaît en revanche raisonnable. Ce constat en demi-teinte a néanmoins pour mérite de rappeler que l'hypothèse de concurrence pure et parfaite, l'idéal néoclassique d'un marché sans dysfonctionnement, relève bien plus de la théorie que de la réalité.

## POUR EN SAVOIR PLUS

---

❖ **BANK OF INTERNATIONAL SETTLEMENT** (2014), *Triennial Central Bank Survey, Global Foreign Exchange Market Turnover in 2013*, disponible sur le site <http://www.bis.org>.

❖ **JÉGOURELY**. (2012), « Les produits dérivés : outils d'assurance ou instruments dangereux de spéculation ? », *Cahiers français* n° 361, Comprendre les marchés financiers, Paris, La Documentation française.

---

Loin de l'idéal-type de la concurrence parfaite, le marché du travail connaît des déséquilibres durables que l'insuffisance de la demande ou le coût excessif du travail ne peuvent expliquer à eux seuls. Des caractéristiques propres à ce marché permettent de mieux comprendre la persistance du chômage. François Fontaine met ainsi en évidence son caractère frictionnel : tandis que les créations et destructions d'emplois sont massives, l'information imparfaite, les coûts liés à la mobilité professionnelle et l'inadéquation entre l'offre et la demande de travail font que des personnes ne trouvent pas de travail alors que des emplois restent vacants. Dans ce contexte, les institutions du marché du travail telles que l'assurance chômage ne sont pas un frein à son fonctionnement mais s'avèrent au contraire nécessaires.

Problèmes **économiques**

# Le marché du travail, un marché en mouvement

Comme dans beaucoup de champs de l'économie, l'usage des données microéconomiques a révolutionné notre compréhension du marché du travail. Avant les années 1980, il était impossible d'observer à l'échelle d'un pays la croissance ou la disparition des entreprises, ou les mobilités des individus entre entreprises. Le chômage était analysé soit comme la résultante d'un coût du travail trop élevé, soit comme la conséquence d'une demande de biens et de services insuffisante. Il s'agissait en réalité de deux visions complémentaires : les entreprises ne produisent que s'il y a une demande à satisfaire et que si la satisfaction de cette demande leur est profitable. Du point de vue du cycle des affaires, on considérait qu'il y avait des périodes où les licenciements constituaient la majeure partie des mouvements sur le marché du travail et d'autres, plus favorables, où les embauches étaient dominantes.

© **FRANÇOIS FONTAINE**

Professeur à l'université Paris-I  
membre de l'Institut universitaire de France

Or, nous savons désormais que nos économies sont en permanence le théâtre de réallocations massives, c'est-à-dire de mobilités d'une entreprise à l'autre avec parfois un passage au chômage, et de destructions et de créations d'emplois simultanées. Avant même la crise économique, on estimait que, chaque jour, environ 10 000 emplois étaient détruits tandis que 10 000 autres étaient créés<sup>1</sup>. Plus étonnant, les firmes qui licencient ou ne renouvellent pas certains contrats sont aussi des firmes qui créent des emplois. En moyenne sur une année pour une embauche nette, une firme se sera séparée de quatre travailleurs et en aura embauché cinq. Par ailleurs, les salariés sont mobiles et certains changent d'em-

<sup>[1]</sup> On pourra consulter avec profit l'ouvrage de Cahuc P. et Zylberberg A. (2004).

ploi de manière répétée durant leur carrière, voire au cours d'une même année. Enfin, les entrées au chômage sont fréquentes même en période de boom économique et il faut souvent y rester plusieurs mois avant de retrouver un emploi.

Si l'on pense à ces larges mouvements de main-d'œuvre, les manières dont on expliquait par le passé le fonctionnement du marché du travail ne semblent plus satisfaisantes. Elles ne disent rien sur le temps que passent les travailleurs au chômage alors que ces durées ont un impact important sur le bien-être des travailleurs. Elles n'expliquent pas pourquoi une même firme embauche et licencie, et pourquoi certains postes restent vacants. En outre, elles restent silencieuses sur le rôle que les institutions du marché du travail, et notamment de l'assurance chômage, peuvent jouer dans ces mouvements incessants.

## Le marché du travail est un marché frictionnel

Sentant les limites des approches qui, à leur époque, étaient considérées comme standards, plusieurs économistes ont cherché dans les années 1970 à proposer une description plus réaliste du marché du travail<sup>2</sup>. Celle-ci, sans repousser nécessairement les explications que nous avons déjà données, insiste particulièrement sur deux écarts à la concurrence parfaite : l'asymétrie de l'information<sup>3</sup> et l'existence de coûts à échanger, c'est-à-dire en l'occurrence à chercher un emploi ou, pour les employeurs, à embaucher. On parle de *frictions* du marché du travail.

Cette nouvelle manière de voir fut popularisée quelques années plus tard par P. Diamond, D.T. Mortensen et C. Pissarides. Elle valut à ces trois économistes un prix Nobel en 2010. Elle part d'un raisonnement très simple. Lorsqu'un individu est à la recherche d'un emploi, il ne dispose pas d'une vision parfaite des postes qui sont disponibles et

qui pourraient lui convenir. En outre, la collecte de telles informations prend du temps. Enfin, accepter un poste est en soi une prise de risque et cela implique parfois de supporter des coûts importants, par exemple parce qu'il faut déménager. De même, un employeur n'observe pas sans effort les personnes qui pourraient convenir pour le poste qu'il propose. C'est à la fois un problème de localisation des candidats et d'évaluation de la qualité des candidats.

Embaucher prend du temps, trouver un emploi aussi. Ainsi, coexistent sur le marché du travail des emplois vacants et des chômeurs. Les données dont nous disposons aujourd'hui montrent que cela est vrai au niveau du marché du travail pris dans son ensemble, mais aussi au sein d'un même secteur ou d'une même profession. Il ne s'agit à proprement parler ni d'un problème de demande, ni d'un problème de rigidité des salaires. Si l'on souhaite améliorer le fonctionnement du marché du travail et rapprocher demandeurs d'emplois et emplois vacants, il faut donc approfondir notre compréhension de ces imperfections que les économistes nomment frictions.

## Les destructions créatrices

Comme nous l'avons évoqué, même lorsque l'environnement économique est stable, des entreprises disparaissent tandis que d'autres sont créées. Parallèlement, au sein des entreprises qui poursuivent leurs activités, certains emplois sont détruits, de nouvelles personnes sont embauchées et des employés changent de tâches. Cela vient du fait que nos économies fonctionnent en partie par destructions créatrices. L'idée n'est pas nouvelle, elle a été popularisée par l'économiste autrichien Joseph Schumpeter dans les années 1940 même s'il n'existait pas de données à l'époque pour la tester. Elle est somme toute très simple : des activités ou des produits en perte de vitesse sont poussés vers la sortie par de nouvelles activités et de nouveaux produits. Tandis que certaines professions

<sup>[2]</sup> George Stigler, prix Nobel 1982, a été l'un des premiers à avoir appliqué la théorie des frictions de recherche au marché du travail dans un article de 1962. Une brève présentation de ces développements peut être trouvée dans le discours prononcé lors de la remise de son prix Nobel en 2010 par Dale T. Mortensen.

<sup>[3]</sup> Sur les différents types d'asymétrie d'information, voir le texte d'Anne Corcos et François Pannequin, p. 58.

deviennent obsolètes, d'autres sont créées pour s'adapter aux nouvelles activités ou parce que certaines innovations ont changé la manière de produire et de travailler. C'est en principe une source de croissance si les activités à forte valeur ajoutée prennent le pas sur les autres, mais en pratique, ce processus pose un certain nombre de difficultés. Celles-ci, comme nous allons le voir, trouvent leurs origines dans les imperfections du marché du travail que nous avons déjà évoquées.

Pour le comprendre, supposons au contraire que le marché du travail soit parfaitement concurrentiel. Cela suppose que les travailleurs observent tous les postes disponibles et que leurs mobilités ne sont pas coûteuses. Cela suppose en outre, hypothèse discutable sur laquelle nous reviendrons, que les salaires reflètent la productivité des emplois, c'est-à-dire de la valeur ajoutée qu'ils génèrent. Dans ce cas, un salarié qui perd son emploi, car celui-ci est obsolète, ira vers l'un des emplois nouvellement créés pour lequel il est le mieux rémunéré et donc (du fait de l'hypothèse sur les salaires) où il est le plus productif<sup>4</sup>. Le processus de destruction créatrice est efficace car les réallocations vont des postes les moins productifs vers ceux qui le sont le plus.

Que change l'existence d'imperfections sur le marché du travail ? Tout d'abord, comme trouver un emploi n'est pas instantané, une grande partie des réallocations de main-d'œuvre se fait par un passage des salariés par le chômage. Si cette transition est trop lente, les coûts sur le bien-être et la santé peuvent être importants. En France, la durée moyenne des épisodes de chômage était de 14 mois en 2012, à égalité avec la moyenne des pays de l'Union européenne, mais deux fois plus élevée qu'en Norvège (sources : INSEE et OCDE). Cette lenteur peut provenir du fait que les travailleurs au chômage peinent à identifier les emplois où ils peuvent postuler, par exemple parce que leurs compétences ne sont plus en adéquation avec les emplois disponibles. Elle peut aussi venir du fait que ce ne sont pas les mêmes régions qui perdent des emplois et qui en créent. Les changements

d'emploi impliquent alors des mobilités qui sont coûteuses (déménagement, contraintes familiales). *In fine*, ces difficultés limitent de manière importante les réallocations sur le marché du travail, même en période où l'activité économique est favorable, et vont venir aggraver l'impact des baisses d'activité.

Si l'on s'intéresse maintenant à l'efficacité des réallocations, il est probable qu'elles ne se font pas toujours dans le bon sens, c'est à dire qu'elles ne correspondent pas nécessairement à des départs d'entreprises à faible valeur ajoutée vers des entreprises à forte valeur ajoutée. Le marché du travail étant frictionnel, les travailleurs ne vont pas nécessairement vers les « meilleures » entreprises. Ils cherchent avant tout à travailler ou, lorsqu'ils ont déjà un emploi, à améliorer leurs salaires et conditions de travail. Dit autrement, ils acceptent toute offre qui améliore la situation dans laquelle ils se trouvent. Ainsi, de nombreux travaux montrent que des firmes pourtant peu productives survivent et continuent à embaucher<sup>5</sup>. Les imperfections du marché du travail font que leurs salariés ont du mal à trouver d'autres postes et que les chômeurs n'ont pas nécessairement la possibilité de refuser leurs offres d'emploi. Elles bénéficient donc d'une main-d'œuvre à moindre coût qui leur permet de rester profitables et donc de survivre. À l'inverse, certaines entreprises, pourtant à plus forte valeur ajoutée, auront du mal à trouver les employés dont elles ont besoin et n'auront pas la croissance qu'elles auraient pu avoir si le marché du travail était moins frictionnel.

## Les salaires, des prix pas comme les autres

Les salaires affectent la demande de travail des entreprises, mais ils affectent aussi les choix des travailleurs en termes de secteur, de profession ou de nombre d'heures travaillées. C'est pourquoi l'efficacité des réallocations suppose que, lorsqu'un travailleur observe les salaires des emplois auxquels

<sup>[4]</sup> Ici, le terme « productif » se rapporte à la valeur ajoutée supplémentaire que l'embauche de ce travailleur génère pour l'entreprise. Elle dépend de son efficacité au travail mais aussi des technologies de production de l'entreprise et de la profitabilité du secteur sur lequel elle opère.

<sup>[5]</sup> Voir par exemple Bartelsman E. J. et Doms M. (2000), « Understanding Productivity: Lessons from Microdata », *Journal of Economic Literature*, vol. 38, n° 3.

il peut prétendre, ceux-ci reflètent les différences de productivité entre les emplois.

Ce serait le cas si le marché du travail était un marché en concurrence parfaite c'est-à-dire sans frictions. Dans cette situation, un demandeur d'emploi observe l'ensemble des postes disponibles et les salaires qui y sont associés. Il peut donc choisir l'employeur qui offre les meilleures conditions. Cette possibilité de choisir le meilleur emploi disponible exacerbe la compétition entre les employeurs. Elle devient telle que la rémunération à laquelle ce travailleur va pouvoir prétendre correspond au salaire maximal qu'il peut obtenir sans que l'entreprise ne perde de l'argent. Le salaire est alors égal à ce que le travail du salarié rapporte à l'entreprise (sa productivité).

Comment tester ce qu'il en est en réalité ? Il suffit de remarquer que si le salaire est égal à la productivité, il est alors fonction des technologies utilisées pour produire et d'éléments observables, comme l'éducation, l'expérience, l'âge, ou inobservables, comme la motivation, qui affectent la productivité. Or, les études empiriques sur la question ont montré que les caractéristiques observables n'expliquent qu'une part limitée des différences de salaires entre individus (entre un tiers et la moitié en fonction des études). Même si l'on prend en compte le fait que l'on n'observe pas tous les déterminants de la productivité et que l'on contrôle ce que les statisticiens nomment l'hétérogénéité inobservée, il reste entre un tiers et la moitié des différences de salaires à expliquer. Dit autrement, des individus, identiques sont payés différemment<sup>6</sup>.

Éclairant cette énigme empirique, de nombreux travaux théoriques ont montré que, si la productivité affecte bien les rémunérations, la manière dont elle se reflète dans le salaire dépend fondamentalement des imperfections du marché du travail. L'idée est simple. Pour l'illustrer, imaginons le cas inverse de la concurrence parfaite, c'est-à-dire un univers dans lequel il n'existe qu'un seul employeur et un grand nombre

de demandeurs d'emploi. Pour embaucher, il suffira à l'employeur de fixer un salaire supérieur à ce que l'individu touche au chômage. Il n'y a donc pas de raison pour que le salaire reflète la productivité du salarié. Si l'on généralise un peu l'idée, il est aisé de comprendre que plus la compétition entre employeurs pour conserver leur main-d'œuvre est importante et plus le salaire est proche de la productivité. De manière générale, les employeurs prennent en compte les opportunités extérieures de leurs salariés et, en fonction de celles-ci, le salaire oscille entre le montant des allocations chômage et la productivité de l'emploi.

Lorsque l'information sur les postes est imparfaite et que les mobilités sont coûteuses, le marché du travail peut donc être vu comme une loterie. Certains travailleurs vont avoir la chance de trouver un employeur proposant des salaires élevés ou auront une opportunité de carrière qui leur permettra de changer d'entreprise et d'améliorer leurs conditions salariales. D'autres ne bénéficieront pas de cette chance et auront des salaires plus faibles et des carrières salariales moins dynamiques. Les imperfections du marché du travail sont donc une source importante d'inégalité salariale entre travailleurs à la productivité pourtant identiques.

## Une vision mesurée du fonctionnement du marché du travail

On le voit bien, le salaire n'est pas un simple prix de marché permettant d'égaliser offre et demande de travail ; il est l'expression des imperfections de marché qui se trouvent au cœur des analyses modernes. Celles-ci ne sont pas antithétiques avec les approches keynésiennes, liant étroitement la demande de travail et la demande de biens et services, ni avec la vision néoclassique où la demande de travail dépend du salaire et où le chômage découle des rigidités sur les salaires.

<sup>[6]</sup> Il ne s'agit pas de discrimination car celle-ci repose sur des éléments observables (comme le sexe) qui sont donc contrôlés en première étape. Les études existantes sont assez techniques, on pourra trouver une présentation abordable des principaux résultats dans Kramarz F. (2004).

D'un côté, la demande de biens et services détermine la profitabilité d'une embauche. Elle pousse donc l'entreprise à ouvrir des postes qui seront d'abord vacants avant d'être éventuellement pourvus ou, au contraire, à réduire le nombre d'employés. De l'autre, les salaires sont bien un facteur pris en compte par l'entreprise lorsqu'elle crée ou détruit des emplois. Ceux-ci sont rigides dans le sens où ils ne s'ajustent pas parfaitement aux chocs de productivité. L'existence d'éléments comme les allocations chômage les poussent à la hausse mais, à l'inverse, les imperfections du marché du travail permettent à l'employeur de verser des rémunérations plus faibles que ce qu'ils verseraient en concurrence parfaite.

Un accroissement de la demande de biens et services peut créer des emplois, de même qu'une baisse du coût du travail. Cependant, leurs effets sont limités par la présence de frictions de recherche d'emploi qui ralentissent les sorties du chômage et allongent le temps nécessaire pour pourvoir les postes ouverts. Ce constat appelle à compléter les politiques d'offre et de demande habituelles par des mesures pouvant affecter les frictions en elles-mêmes en modifiant la *structure* du marché du travail.

## La place fondamentale des institutions du marché du travail

Plusieurs institutions du marché du travail ont cette capacité. Nous prendrons ici l'exemple du service public pour l'emploi. Sous ce terme sont regroupés à la fois l'assurance chômage, qui indemnise les chômeurs, et l'ensemble des services qui sont proposés aux demandeurs d'emploi et aux entreprises désireuses d'embaucher.

Dans un marché du travail où chaque jour 10 000 emplois sont détruits, l'existence d'une assurance chômage est essentielle. Comme nous l'avons évoqué, les réallocations, pourtant cruciales pour le bon fonctionnement de l'économie, conduisent souvent à un passage par le chômage, dont les

durées sont particulièrement longues dans beaucoup de pays. L'assurance chômage est donc nécessaire pour limiter les pertes de revenus qui en résultent. Elle est obligatoire dans la majeure partie des pays et est souvent financée par une taxe sur les salaires (6,4 % du salaire brut en France). Ce mode de financement permet une mutualisation des risques sur le marché du travail. Cette mutualisation est essentielle car de nombreux salariés ne disposent pas d'un niveau d'épargne suffisant pour financer leurs épisodes de chômage.

Premier effet direct sur les mobilités, des allocations suffisamment généreuses permettent aux demandeurs d'emploi de ne pas accepter la première offre venue. Convenablement assurés, ils peuvent chercher des emplois qui correspondent à leurs compétences. C'est *a priori* une bonne chose car l'adéquation entre le salarié et le poste qu'il occupe favorise la stabilité de la relation d'emploi et limite le risque de retourner rapidement au chômage. C'est aussi positif pour la productivité globale de l'économie. Dans ce sens, l'assurance chômage permet de financer une recherche d'emploi efficace.

Réduire le service public pour l'emploi aux allocations chômage qu'il attribue serait une erreur. En effet, il joue aussi le rôle d'intermédiaire sur le marché du travail. Il collecte les offres d'emplois des entreprises et les met en relation avec les demandeurs d'emploi. Il aide à réduire les imperfections d'informations sur les postes et les candidats et donc les frictions du marché du travail. C'est une dimension importante qui a longtemps été négligée. Ce n'est plus le cas aujourd'hui et *Pôle emploi*, qui est le principal acteur du service public pour l'emploi en France, cherche à améliorer cette activité, notamment en renforçant son offre de services auprès des entreprises<sup>7</sup>. On notera que des pays comme le Danemark ou le Royaume-Uni ont dès les années 1990 pris conscience de l'importance de cette mission d'intermédiation.

<sup>[7]</sup> Ces objectifs sont définis dans la convention tripartite État-Unédic-Pôle emploi 2015-2018, qui a été signée le 18 décembre 2014.



Au-delà de l'intermédiation, l'une des missions de ce service public est d'accompagner les demandeurs d'emploi, c'est-à-dire de les conseiller sur la meilleure manière de chercher un emploi, sur les secteurs à prospecter, de les aider à rédiger leur *curriculum vitae* et de les orienter si nécessaire vers les formations les plus pertinentes. Cet accompagnement est très important pour le bon fonctionnement du marché du travail. Il accentue l'efficacité des recherches d'emploi en réduisant les imperfections d'information et en améliorant l'adéquation entre les postes à pourvoir et les candidats. C'est une autre dimension par laquelle le service public pour l'emploi limite les frictions de recherche d'emploi.

Cela explique sans doute que le renforcement de l'accompagnement est l'une des rares politiques pour l'emploi dont l'efficacité est avérée par un grand nombre d'études portant sur de nombreux pays<sup>8</sup>. Nous ne donnerons ici qu'un exemple. La France a expérimenté en 2006 un dispositif appelé « suivi mensuel personnalisé ». Destiné aux demandeurs d'emploi présentant un risque faible à modéré de chômage, il mettait en place, dès le quatrième mois passé sur les listes de l'agence pour l'emploi, un entretien mensuel avec un conseiller. L'évaluation existante a

montré que la durée du chômage de ceux qui bénéficiaient d'un renforcement de l'accompagnement diminuait très significativement. Les chômeurs bénéficiant de ce suivi renforcé ont ainsi vu leur taux de sortie du chômage augmenter d'environ 30 % durant la première année où ils étaient inscrits au chômage<sup>9</sup>.

\* \* \*

Le marché du travail est un marché en mouvement continu, traversé par de gigantesques flux de main-d'œuvre. Ceux-ci sont essentiels au bon fonctionnement de l'économie mais se heurtent aux imperfections du marché, en particulier à ce que les économistes ont appelé les frictions de recherche d'emploi. Ces frictions ralentissent ces mouvements, les rendent coûteux en termes de bien-être et réduisent leur contribution à la croissance économique. Cette compréhension relativement récente du fonctionnement du marché du travail permet de saisir les limites des politiques d'offre et de demande telles qu'elles étaient conçues par le passé. *A contrario*, elle redonne toute leur place à des politiques qui, comme le renforcement de l'accompagnement des chômeurs, modifient directement le fonctionnement du marché.

<sup>[8]</sup> Un panorama des études sur la question est proposé par Fontaine F. et Malherbet F. (2013).

<sup>[9]</sup> L'évaluation a été menée par Thomas le Barbanchon et Maël Fontaine (Document d'étude n°175 de la DARES).

## POUR EN SAVOIR PLUS

❖ CAHUC P. et ZYLBERBERG A. (2004), *Le chômage, fatalité ou nécessité*, Flammarion.

❖ FONTAINE F. et MALHERBET F. (2013), *Accompagner les demandeurs d'emploi*, Paris, Presses de Sciences Po.

❖ KRAMARZ F. (2004), « Mobilité et salaires : une longue tradition de recherche », *Économie et Statistique*, n° 369-370, Paris, INSEE.

A retourner à la **Direction de l'information légale et administrative (DILA)** – 26, rue Desaix 75727 Paris Cedex 15

# BULLETIN D'ABONNEMENT ET BON DE COMMANDE

## Comment s'abonner ? Où acheter un numéro ?

Sur [www.ladocumentationfrancaise.fr](http://www.ladocumentationfrancaise.fr)  
(paiement sécurisé).

Sur papier libre ou en remplissant  
ce bon de commande  
(voir adresse d'expédition ci-dessus).

En librairie, à la librairie  
de la Documentation française,  
29/31 quai Voltaire – 75007 Paris  
et en kiosque pour l'achat  
d'un numéro.

## Je m'abonne à *Problèmes économiques* (24 numéros par an dont 2 numéros spéciaux)

France métropolitaine 76,50 €  Tarifs étudiants et enseignants (France métropolitaine) 51,00 €

Europe 103,10 €  DOM-TOM-CTOM HT 96,40 €  Autres pays HT 122,70 €

Supplément avion 19,30 €

## Je commande un numéro de *Problèmes économiques* 4,90 €

Je commande le(s) numéro(s) suivant(s) :

Pour un montant de  € Participation aux frais de port<sup>(1)</sup> + 4,95 €

Soit un total de  €

- Par chèque bancaire ou postal à l'ordre  
du comptable du B.A.P.O.I.A – DF  
 Par mandat administratif (réservé aux  
administrations)

Par carte bancaire N°

Date d'expiration

N° de contrôle

(indiquer les trois derniers chiffres situés au dos  
de votre carte bancaire, près de votre signature)

Raison sociale

Nom  Prénom

Adresse

(bât, étage..)

Code postal  Ville  Pays

Téléphone  Courriel

Ci-joint mon règlement de ..... €

Date

Signature

### <sup>(1)</sup>Pour les commandes de numéro seulement

Informatique et libertés – Conformément à la loi du 6.1.1978, vous pouvez accéder aux informations vous concernant et les rectifier en écrivant  
au département marketing de la DILA. Ces informations sont nécessaires au traitement de votre commande et peuvent être transmises à des tiers,  
sauf si vous cochez ici



## En vente en kiosque,

chez votre libraire,  
sur [www.ladocumentationfrancaise.fr](http://www.ladocumentationfrancaise.fr)  
et par correspondance :  
DILA, 26 rue Desaix – 75727 Paris cedex 15

Abonnement à 6 numéros : 50 €

Tarif spécial étudiants et enseignants : 42 €

## HORS-SÉRIE

FÉVRIER 2015 NUMÉRO 7

# comprendre L'ÉCONOMIE

## 1. Concepts et mécanismes

Rédigé par des enseignants et des universitaires, ce premier tome de la série « Comprendre l'économie » présente de façon simple et non formalisée les savoirs fondamentaux des sciences économiques.

Le numéro commence par une présentation de la discipline et de ses grands courants de pensée pour se concentrer ensuite sur ses outils et ses acteurs. Une dernière partie s'intéresse aux mécanismes des marchés et à leurs dysfonctionnements éventuels, en présentant les marchés les plus emblématiques, tels que les marchés de capitaux et du travail.

Prochain numéro à paraître :

Comprendre l'économie

2. Questions économiques contemporaines

Directeur de la publication

Bertrand Munch

Direction de l'information

légale et administrative

Tél. : 01 40 15 70 00

[www.ladocumentationfrancaise.fr](http://www.ladocumentationfrancaise.fr)

Imprimé en France par la DILA

Dépôt légal 75059, février 2015

DF 2PE39430

ISSN 0032-9304

CPPAP n° 0518B05932

9 €

