



HAL
open science

A propos du statut épistémologique des expériences en économie

Yves de Curraize, Sylvie Thoron

► **To cite this version:**

Yves de Curraize, Sylvie Thoron. A propos du statut épistémologique des expériences en économie. 2018. hal-01966636

HAL Id: hal-01966636

<https://hal.science/hal-01966636>

Preprint submitted on 28 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A propos du statut épistémologique des expériences en économie

Yves de Courraize* et Sylvie Thoron**

*, **Université Paris-Est, LIPHA (EA 7373), UPEC, F-94000 Créteil, France

10 octobre 2018

Résumé :

Apparue dans l'immédiat après-guerre, voire dans les années 1930, la pratique des expériences en économie a connu un essor important depuis les années 1970, et une consécration plus récente dans les années 2000 à travers l'attribution de prix Nobel en économie expérimentale, ainsi que la reconnaissance des travaux d'Esther Duflo. Dans cet article nous nous intéressons aux deux méthodologies les plus couramment utilisées : celles qui entrent dans la catégorie de l'économie expérimentale et les expériences aléatoires contrôlées (EAC). Partageant le même principe de base qui est de « provoquer une observation dans le but d'étudier certains phénomènes », elles n'en diffèrent pas moins par leur origine et leur statut épistémologique. D'une part l'économie expérimentale, tournée vers la théorie, trouve ses origines dans une tentative réfutationniste dans les années 1970. D'autre part les EAC, tournées vers l'évaluation des politiques publiques, viennent d'une recherche de confirmation. Nous montrons par contre que l'influence conjointe des deux méthodologies a ensuite ébranlé la conception de l'économie comme science inexacte et séparée (Hausman 1992). Economie expérimentale et EAC ont œuvré pour que l'économie devienne une discipline moins généraliste et plus contextualisée. Elles ont aussi obligé les chercheurs, expérimentalistes, randomistes et théoriciens à repenser la définition du domaine de l'économie. Quand certains s'efforcent de le préciser, d'autres tentent de l'élargir, selon différentes conceptions de l'interdisciplinarité. Par contre, il semble que les différentes méthodologies expérimentales ne remettent que peu en question la théorie de la rationalité comme référence normative.

Mots clés : épistémologie, économie expérimentale, expériences aléatoires contrôlées, expériences de laboratoire, expériences de terrain, réfutationnisme, confirmationnisme.

Abstract:

Appeared immediately after the war, or even in the 1930s, the practice of experiments in economics has grown significantly since the 1970s, and has known a more recent consecration in the 2000s through the awarding of Nobel prizes in experimental economics, as well as the recognition of Esther Duflo's work. In this article, we are interested in the two most commonly used methodologies: Experimental Economics and Random Controlled Trials (RCTs). They share the same basic principle of "provoking an observation for the purpose of studying certain phenomena", but they differ in their origin and their epistemological status. On the one hand experimental economics, oriented towards theory, finds its origin in a refutationist attempt in the 1970s. On the other hand RCTs, oriented towards the evaluation of public policies, is coming from a confirmation approach. We show, however, that the joint influence of the two methodologies has subsequently shaken the conception of economics as an inexact and separate science (Hausman 1992). Experimental Economics and RCTs have worked to make economics a less general and more contextualized discipline. They also forced researchers, experimentalists, randomists and theorists, to rethink the definition of the domain of economics. When some of them strive to clarify it, others want to expand it, according to different conceptions of interdisciplinarity. However, it seems that the different experimental methodologies hardly question the theory of rationality as a normative reference.

Key words: epistemology, experimental economics, randomized controlled trials, laboratory experiments, field experiments, refutationism, confirmationism.

Introduction

Expériences en laboratoire, expériences de terrain, expériences aléatoires contrôlées, et même, expériences naturelles, les sciences économiques font aujourd'hui grand usage du mot expérience pour désigner certaines pratiques de recherche. Depuis une vingtaine d'années, ces pratiques ont par ailleurs acquis une reconnaissance qu'elles n'ont jamais eue auparavant et qu'attestent, entre autres, l'attribution du prix Nobel d'économie à Daniel Kahneman et Vernon Smith en 2002, à Alvin Roth en 2012 et à Richard Thaler en 2017, ainsi que la célébrité aussi bien scientifique que politique d'Esther Duflo¹. Dans le même temps, diverses controverses ont bien sûr aussi émergé. Nombre d'entre elles portent sur l'apport des expériences à la façon dont l'économie se définit en tant que science. Notre discipline est-elle devenue plus scientifique parce que s'y pratiquent désormais, à l'instar de la physique ou de la biologie, des expériences ? Par ailleurs, quelle que soit la réponse à cette question, il y a tant de façon de faire de la science que l'on doit aussi se demander, au-delà de cette recherche d'une légitimité scientifique accrue, quelle est la capacité des expériences à faire évoluer la discipline. Avant de pouvoir apporter des éléments de réponse à ces questions, il nous faut en dire plus sur ce que sont ces expériences en économie. Or, il existe une grande diversité de méthodes, au point que l'on peut se demander ce qu'elles ont en commun en dehors du fait que toutes sont appelées « expériences ». Qu'y a-t-il de commun par exemple entre une expérience de laboratoire, basée sur le jeu du dictateur, au cours de laquelle l'expérimentateur demande à des étudiants assis devant un ordinateur de partager des jetons avec un partenaire inconnu et une expérience aléatoire contrôlée dans laquelle les expérimentateurs tirent au sort les villages d'une zone africaine infestée par la malaria dont les habitants pourront bénéficier d'une moustiquaire gratuite ? A l'inverse, quand il s'agit de tester la théorie de l'utilité et de mettre en évidence un effet de dotation, qu'y-a-t-il de différent entre offrir un tasse ou une barre de céréales à des sujets dans un laboratoire ou à des intervenants sur un marché ? Ce qui les rassemble tout d'abord est qu'elles s'apparentent toutes effectivement à des expériences au sens défini en théorie des sciences comme « le fait de provoquer une observation dans le but d'étudier certains phénomènes » (dictionnaire Le

¹ Esther Duflo a été nommée rédacteur fondateur de la revue *American Economic Journal: Applied Economics*. Fin 2012, elle a est nommée au sein du President's Global Development Council, un organisme américain chargé de conseiller le président Barack Obama sur les questions de développement. En 2015, elle est devenue codirecteur de J-Pal, laboratoire Abdul Latif Jameel d'action contre la pauvreté (Abdul Latif Jameel Poverty Action Lab). Par ailleurs, elle détient la première chaire internationale « Savoirs contre la pauvreté » au Collège de France, soutenue par l'Agence française de développement.

Robert). Alors que l'économie a longtemps été considérée uniquement comme une science d'observation qui consacre sa partie empirique à la collecte et au traitement des données, elle se reconnaît de plus en plus comme science expérimentale capable de *construire* les « faits » économiques qu'elle va traiter (Annie Cot et Samuel Ferey, 2016). Construire n'est pas collecter. Contrairement, par exemple, à la collecte de données par le biais d'une enquête, qu'elle soit qualitative ou quantitative, construire des données suppose de manipuler un environnement avant d'observer les conséquences de cette manipulation. A partir de ce point commun on pourrait décliner toutes les façons de construire des données et nous reviendrons dans le détail sur ces différences dans la suite. Dans cette introduction on se limitera à souligner une différence originelle entre l'économie expérimentale et les expériences aléatoires contrôlées (EAC) qui concerne les objectifs poursuivis. L'économie expérimentale s'est constituée comme sous-discipline des sciences économiques en cherchant à tester la théorie microéconomique. Les EAC ont été développées pour évaluer les politiques publiques. Les unes étaient tournées vers la théorie quand les autres s'adressaient au politique. Les méthodes de l'économie expérimentale ont cultivé une proximité avec la théorie microéconomique quand les EAC trouvaient leur légitimité dans la statistique. Ces dernières allant parfois, peut-être pour se démarquer des premières, jusqu'à se qualifier de a-théorique. Pourtant, comme nous tenterons de le montrer, il est arrivé un moment où ces méthodes ne se sont pas seulement développées en parallèle. Elles ont aussi largement échangé, par l'intermédiaire de controverses, ont parfois convergé sur certains points. Ce faisant, elles ont contribué, ensemble, à la recherche d'une scientificité accrue de la discipline et elles l'ont fait évoluer.

Nous montrerons dans une première section comment chaque méthode, à sa façon, a pu être utilisée dans le but d'accroître la légitimité de la discipline selon les critères de scientificité de la théorie des sciences. Puis, dans une deuxième section, nous montreront comment, au-delà de ces critères, elles contribuent à faire évoluer la discipline. La dernière section conclue.

1. Les expériences comme gage de scientificité

1.1 La science économique, science inexacte et séparée

Depuis que l'économie se constitue comme discipline, dans la deuxième moitié du XVIII^{ème} siècle, elle est à la recherche d'une légitimité scientifique. Selon Daniel Hausman (1992),

l'économie théorique dominante, telle qu'elle s'est développée depuis la deuxième guerre mondiale, trouve cependant son origine plus récemment, dans le *Système de logique* de John Stuart Mill (1843). La méthode déductive issue des travaux de Mill serait ainsi la pierre angulaire de sa légitimité scientifique et de là découleraient ses caractéristiques fondamentales. Hausman (1992) décrit en effet l'économie comme une science *séparée* et *inexacte*. La dénomination de science séparée signifie chez Hausman que l'économie s'est défini un domaine d'application spécifique et autonome des autres domaines du social. Ce domaine serait caractérisé par un ensemble de contextes et de phénomènes qui intéressent les économistes, comme si un sous-ensemble de relations pouvait être extrait de la complexité de la vie sociale. Ainsi John Stuart Mill (1843) définit le domaine de l'économie politique de la façon suivante :

L'économie politique [...] concerne les phénomènes de l'état social tels qu'ils se produisent à la suite de la poursuite de la richesse. Elle fait abstraction totale de toute autre passion ou motivation humaine, sauf celles qui peuvent être considérées comme des principes perpétuellement antagonistes au désir de richesse, à savoir l'aversion pour le travail et le désir de jouissances coûteuses. (1843, Livre VI, Chapitre 9, Section 3)

Cette définition trouve un écho dans une définition plus récente de Marshall (1920) :

L'économie est une étude de l'humanité dans les affaires ordinaires de la vie; elle examine le rôle de l'action individuelle et sociale le plus étroitement liée à la réalisation et à l'utilisation des conditions matérielles du bien-être. (1890, Livre 1, Chapitre 1)

Dans chacune de ces deux définitions, nous retrouvons l'idée que les économistes s'intéressent aux situations dans lesquelles le mobile et les actions des individus sont orientés vers la recherche de la richesse et des conditions matérielles du bien-être, Mill ajoutant que les économistes peuvent faire «abstraction de toute autre passion ou motivation humaine». A partir de cette définition d'un domaine séparé, et sur la base de l'introspection, les économistes ont construit des hypothèses de base, en particulier sur la rationalité des agents telle qu'elle est censée opérer dans ces contextes. L'élaboration de ces hypothèses est cependant aussi le résultat, en partie au moins, des nécessités de l'analyse mathématique. Selon les principes de la méthode hypothétique-déductive, les économistes combinent alors ces hypothèses de base pour établir des relations de causalité. Il en résulte que celles-ci sont considérées comme des lois générales car elles sont censées s'appliquer partout dans le domaine d'application de l'économie. La cohérence interne de la théorie qui est censée découler de cette méthode est un fondement important de la légitimité scientifique de la

discipline. Cependant, toujours selon Hausman, ces lois sont inexactes, et c'est là la deuxième caractéristique fondamentale. Les économistes ne prenant en compte qu'un ensemble restreint de facteurs de causalité, ils admettent que les lois générales ne peuvent pas être parfaitement exactes. Dans un article publié plus tard, Hausman (2008) développe la même idée quand il dit:

Parce que les économistes tentent d'étudier les phénomènes économiques comme constituant un domaine distinct, influencé uniquement par un petit nombre de facteurs causaux, les affirmations de l'économie ne sont vraies que *ceteris paribus* : elles ne sont vraies que s'il n'y a pas d'interférences ou de causes inquiétantes. (Hausman 2008, section 2.4, page 8)

En effet, lorsque la science séparée est confrontée à la réalité, les économistes ne s'attendent pas à ce que les observations coïncident exactement avec la théorie et que les prédictions soient parfaitement exactes. Nous dirons que la différence est due à des «effets parasites» qui sont exogènes à la théorie (et au domaine) car ils obéissent à d'autres logiques, les autres *passions* dont parlait Mill. Ces effets parasites sont donc imprévisibles mais ils sont cependant considérés comme seulement marginalement perturbateurs et les économistes pensent pouvoir les ignorer à moindre coût.

Ce tableau de la science économique tel que brossé par Daniel Hausman au début des années 90 nous semble bien saisir les spécificités de la discipline telle qu'elle s'est développée depuis la deuxième guerre mondiale. Il permet notamment de comprendre ce qui a pu en constituer les fondements scientifiques. Ces fondements ne sont cependant pas établis une bonne fois pour toutes et les économistes n'ont cessé de se poser la question de la légitimité de leurs recherches, en particulier dans les périodes critiques. Nous allons maintenant tenter d'expliquer pourquoi et comment les expériences sont ainsi devenues à un moment donné de nouveaux outils de légitimation scientifiques, alternatifs à la seule cohérence interne de la théorie. Nous verrons que les raisons sont cependant très différentes qui peuvent expliquer le recours aux expériences de laboratoire d'une part et aux EAC de l'autre.

1.2 Expériences de laboratoire et réfutationnisme

Dans un cadre Poppérien, la scientificité d'une discipline, fut-elle formelle, doit être fondée sur la possibilité d'une expérience radicale. Au risque de caricaturer la conception du réfutationnisme à la Popper (1934, 1963), on dira en effet d'une discipline qu'elle est une

science s'il est possible de concevoir une expérience qui permettrait de réfuter ses théories, c'est-à-dire de conclure qu'elles sont fausses. Or, selon Mikaël Cozic (2014), les thèses Poppériennes ont eu une influence incontestable sur les micro-économistes théoriciens, qui auraient bien voulu pouvoir imaginer une expérience radicale. Il explique en particulier comment la théorie des préférences révélées de Samuelson (1938) a été proposée comme tentative de contournement des difficultés liées au caractère inobservable des préférences. Cette proposition a cependant été largement critiquée par Sen (1973) et Hausman (1992, 2000 et 2008) et, à l'issue du débat, il est apparu essentiel de conserver dans la science économique une notion de préférences différente des choix.

C'est dans ce contexte que, à partir de la fin des années 70, les expériences de laboratoire sont apparues comme un instrument privilégié pour tester la théorie microéconomique. Deux articles, sans doute parmi les plus connus en économie expérimentale encore aujourd'hui, sont publiés à ce moment-là. Dans le premier article, Vernon Smith (1976) présente les résultats d'une expérience basée sur un jeu de marché et affirme que la convergence vers un équilibre ne dépend pas nécessairement de la rationalité des acteurs. Ce résultat offre ainsi une alternative à l'entreprise théorique de l'équilibre général. Une approche empirique des institutions et règles des marchés pourrait faire ce que la théorie n'a pas réussi à faire : montrer comment les marchés fonctionnent et sont apurés. Quelques années plus tard, Kahneman et Tversky (1979) présentent les résultats d'une expérience destinée à tester la théorie de l'utilité espérée. Ils concluent que la théorie de l'utilité espérée n'est pas un modèle positif satisfaisant pour décrire le comportement des agents réels et proposent une nouvelle théorie, la théorie des perspectives, comme meilleure alternative. Leur article propose ainsi une autre direction de recherche par rapport à l'approche formelle, axée sur la réévaluation de la rationalité définie en théorie. Ainsi, ces deux articles présentent les résultats d'expériences réalisées en laboratoire et, au cours des deux décennies qui ont suivi, cette méthodologie s'est développée pour devenir la base d'une nouvelle sous-discipline reconnue : l'économie expérimentale.

Pourquoi ces expériences de laboratoire ont-elles un tel écho dans la profession à ce moment-là ? Vernon Smith travaillait depuis longtemps sur ces questions² quand il a pu publier ses résultats au début des années 70 dans la revue *Econometrica*, si reconnue. Notre hypothèse est que les années 70 ont constitué une période de doute particulièrement favorable à l'émergence d'une nouvelle méthodologie. Jusque-là, la théorie de l'équilibre général avait semblé pouvoir constituer une base solide pour la science économique, en parfaite adéquation avec l'idée que les économistes théoriciens pouvaient se faire de leur discipline, et que nous reconnaissons dans la description de Daniel Hausman (1992). Elle a prospéré dans les années 50 et 60, jusqu'à ce que, avant le milieu des années 70, un théorème prouvé par Sonnenschein et étendu par Mantel puis Debreu, marque une rupture. Les trois principaux contributeurs à la théorie de l'équilibre général ont alors montré que, malgré la définition axiomatique prudente de la rationalité, il n'est pas possible de garantir que les agents d'une économie atteindront l'équilibre. Ce résultat a signé l'échec d'une entreprise intellectuelle lancée cent ans plus tôt par Léon Walras et dont l'objectif était de construire une théorie parfaitement cohérente du fonctionnement général d'une économie. Cet échec a ouvert la voie au développement de la théorie des jeux et de l'approche de l'équilibre partiel en économie. La construction de la théorie des jeux est bien plus ancienne et commence juste après la deuxième guerre mondiale, mais elle n'a occupé un rôle central en économie que dans les années 80. C'était une condition nécessaire au développement des expériences. Ce point de rupture a selon nous participé à la réorientation de la profession vers une approche plus empirique de la microéconomie car la cohérence interne du modèle ne pouvant être prouvée, la méthode déductive purement théorique se trouvait être mise en doute. Il reste à expliquer pourquoi ce sont les expériences de laboratoire qui sont d'abord apparues comme une alternative intéressante.

Pour que, dans une approche réfutationniste, l'expérience puisse faire la preuve qu'elle permet de tester une théorie, il faut que l'on puisse indéniablement admettre, en observant les résultats de l'expérience, que la théorie est fautive. Il faut donc que l'expérience soit en quelque sorte la traduction empirique de la théorie. Le laboratoire expérimental est apparu comme le cadre adéquat. Tout d'abord, c'est parce que les économistes concevaient leur

² Vernon Smith a commencé ses travaux en économie expérimentale quand il était à la Krannert School de Purdue University, où il est resté de 1955 à 1967.

discipline comme une science inexacte et séparée au sens de Hausman qu'ils ont pu imaginer utiliser le laboratoire pour tester la théorie microéconomique. En particulier, comme Charles Plott (1991) le dit explicitement, les lois économiques générales étaient sensées s'appliquer partout, y compris dans le laboratoire :

Les théories générales doivent s'appliquer à des cas particuliers simples. La technologie du laboratoire peut être utilisée pour créer des économies simples (mais réelles). Ces économies simples peuvent alors être utilisées pour tester et évaluer la capacité prédictive des théories générales lorsqu'elles sont appliquées aux cas particuliers. De cette manière, une fusion des théories générales et des données est réalisée. (Plott 1991, page 902)

Par ailleurs pour que l'expérience puisse prétendre tester la théorie, il faut un contrôle parfait des éléments du protocole. Or le premier avantage du laboratoire, que Harrison et List (2004) décrivent comme un « environnement stérile », est qu'il permet de se débarrasser des effets parasites mentionnés précédemment.

Nous ne voyons pas la notion d'«environnement stérile» comme négative, à condition que l'on reconnaisse son rôle dans le processus de recherche. En un sens, cette stérilité nous permet de mettre en relief les effets des traitements exogènes sur le comportement. (Harrison et List 2004, page 1009)

Deuxièmement, le protocole est sensé recréer entièrement l'environnement dans lequel les sujets vont devoir prendre leurs décisions et accroître ainsi le contrôle. L'économie expérimentale a ainsi pu devenir une sous-discipline reconnue de l'économie en bâtissant sa légitimité autour d'un certain nombre de règles devant garantir le plus grand contrôle du protocole expérimental. Sur le site de l'Association Française d'Economie Expérimentale (ASFEE)³, on peut lire par exemple que « l'observation des faits dans leur environnement naturel ne permet pas d'isoler avec précision les multiples facteurs susceptibles de les provoquer, ni de quantifier leur influence respective » ; ou encore « [qu']une expérience typique consiste à créer un environnement contrôlé afin de créer artificiellement une situation reflétant les conditions de la théorie économique ou une situation hypothétique. » Les incitations monétaires en particulier, sont devenues une condition indispensable alors qu'elles ne l'étaient pas au début. Que l'on pense aux expériences de Kahneman et Tversky des années 70 qui reposaient sur des scénarios hypothétiques. Cette condition est notamment devenue une marque distinctive des expériences de laboratoire en économie par rapport aux expériences en psychologie. Ces dernières sont par ailleurs aujourd'hui souvent tout à fait

³ <http://www.asfee.fr/>

comparable aux premières. La justification des incitations monétaires est double. Elles sont nécessaires pour recréer les incitations matérielles spécifiques au domaine de l'économie, mais surtout, elles sont destinées à s'assurer que les sujets jouent le jeu qui leur est présenté. Tout est fait pour que l'expérience se déroule en vase clos.

La théorie est alors testée grâce à un protocole qui permet de comparer les comportements observés et les prédictions théoriques. Voyons ici deux exemples. Selon le modèle théorique de la rationalité les préférences sont fixes. Kahneman, Knetsch et Thaler (1990) ont remis en question cette hypothèse en mettant en évidence un effet de dotation qui était déjà bien connu par ailleurs des psychologues. Leur protocole était constitué de trois traitements. Dans un premier traitement des étudiants recevaient des tasses à l'effigie de leur université (des « mugs ») qu'ils pouvaient échanger contre des barres de chocolat (groupe 1). Dans un deuxième traitement d'autres étudiants recevaient des barres de chocolat qu'ils pouvaient échanger contre des mugs (groupe 2). Dans un troisième traitement enfin des étudiants ne recevaient rien mais devaient donner leur préférence entre les barres de chocolat et les mugs (groupe 3). Le troisième traitement permet de déduire les prédictions théoriques. Or, 56% des étudiants du groupe 3 ont opté pour un mug. Si la composition des trois groupes est équivalente, ce que l'on peut espérer quand les étudiants sont assignés aux différents groupes de manière aléatoire, et si l'on suppose que les préférences ne dépendent pas de la situation de départ (l'étudiant a un mug ou une barre de chocolat ou rien) on doit s'attendre à ce que, dans les deux autres traitements, il y ait environ 50% d'échanges. Or les résultats sont les suivants : 89% des sujets du groupe 1 choisissent de conserver leur mug, 90% des sujets du groupe 2 préfèrent garder leur barre de chocolat. Les auteurs en déduisent un effet dotation : quand les individus entrent en possession d'un bien, ils modifient la façon dont ils l'évaluent. Ils accordent plus d'importance à un bien quand ils le possèdent déjà que quand ils doivent l'acquérir. On voit dans cet exemple comment le test est construit à partir de la comparaison des traitements et de la référence à la prédiction théorique.

Voyons maintenant un autre exemple d'expérience, basée cette fois-ci sur un jeu stratégique ou jeu d'interaction. Le jeu de contribution au bien public a constitué la base de nombreux protocoles expérimentaux. Il s'agit cette fois-ci, dans le laboratoire, de mettre en relation un groupe de plus de deux sujets, toujours physiquement isolés dans leurs boxes, mais

communiquant par l'intermédiaire de leurs écrans selon des règles précisément décrites dans les instructions par l'expérimentateur. Chaque membre du groupe reçoit une somme au début de l'expérience et doit alors décider d'utiliser, ou pas, tout ou partie de cette somme pour contribuer à un pot commun. Ce qui n'est pas mis dans le pot commun est conservé comme paiement privé. Les instructions expliquent alors que le pot commun permet de générer un surplus, ou bien public, qui sera ensuite réparti également entre les membres. Formellement, le total des contributions du groupe est multiplié par un coefficient qui est toujours inférieur au nombre de membres dans le groupe. On observe alors la somme des contributions des sujets au pot commun, que l'on compare ici non pas à un mais à deux points de référence. Le premier est toujours la prédiction théorique donnée par l'équilibre de Nash. Dans les jeux le plus couramment utilisés, à l'équilibre de Nash la somme des contributions est nulle. Ainsi, comme c'est le cas dans les jeux d'interaction les plus célèbres, l'équilibre de Nash n'est pas efficace. Aussi, on utilise un deuxième point de référence appelé optimum social. Il correspond au surplus le plus important qui peut être produit, lorsque chaque sujet verse dans le pot commun l'intégralité de la somme qui lui a été allouée. Pour permettre aux sujets de se familiariser avec les règles, l'expérience est répétée plusieurs fois. A chaque étape, chaque sujet a la même somme à sa disposition et doit décider d'une contribution. Théoriquement, cette répétition n'a pas de conséquence sur les contributions à l'équilibre qui restent nulles. On observe cependant que les contributions sont au départ de l'ordre de 40 à 60% de la dotation, puis diminuent pour se rapprocher du résultat théorique. Dans cet exemple le test consiste à comparer la somme des contributions observée à ces deux points de référence et fait apparaître qu'elle est inférieure à l'optimum social mais supérieure à l'équilibre de Nash. Par rapport à ce traitement de base, la littérature tente alors de répondre à deux questions : pourquoi les sujets contribuent-ils plus qu'à l'équilibre de Nash et peut-on imaginer des conditions dans lesquelles leur contributions se rapprocheraient encore de l'optimum social. Pour se faire d'autres traitements sont proposés qui sont comparés au traitement de base.

Ainsi la tentative réfutationniste a produit un corpus d'expériences incluant de nombreuses observations contraires aux prédictions de la théorie. Ces « anomalies » par rapport à la théorie économique standard se sont révélées robustes, ont été répliquées et sont devenues de véritables faits stylisés. Dans les expériences basées sur des jeux d'interaction, les résultats pouvaient être présentés, selon un schéma devenu standard, en partant du constat que les

observations ne coïncidaient pas avec l'équilibre de Nash. Ainsi les sujets sont plus généreux dans leurs contributions aux jeux de bien public, aux jeux d'ultimatum et aux jeux de la confiance. De même, mettre en évidence le fait que les décisions des sujets réels ne satisfont pas les axiomes des préférences, que les préférences ne sont pas transitives, ou que les sujets présentent un effet de dotation, est apparu comme une tâche relativement facile. C'est en tout cas devenu l'objet de l'économie comportementale dont le programme de recherche, inspiré de la psychologie cognitive, s'est appliqué à inventorier l'ensemble des biais par rapport à la décision rationnelle dont sont victimes les sujets réels. Notons que dans cette littérature, contrairement à la littérature sur les expériences dans les jeux d'interaction, la rationalité définie en théorie de la décision constitue une référence normative. Ce qui a amené Gerardo Infante, Guilhem Lecouteux et Robert Sugden (2016) à dire que cette approche relève parfois de ce qu'ils appellent une *purification préférentielle* :

Une réponse commune à ces problèmes consiste à traiter les déviations par rapport à la théorie conventionnelle du choix rationnel comme des erreurs et à tenter de reconstruire les préférences sur lesquelles les individus auraient agi s'ils avaient raisonné correctement. Nous soutenons que cette approche de purification préférentielle utilise implicitement un modèle dualiste de l'être humain, dans lequel un agent rationnel interne est piégé dans une coque psychologique externe. (Infante, Lecouteux et Sugden 2016, page 1)

L'économie comportementale se construit donc en gardant comme référence la théorie du choix rationnel.

Clairement, ces résultats expérimentaux contraires aux prédictions théoriques n'ont donc pas abouti à une réfutation au sens de Popper en économie. Par ailleurs, comme nous le verrons dans la deuxième section, les résultats de l'économie expérimentale ont provoqué une réponse défensive des théoriciens ouvrant un débat au sein de la profession. Ce débat a cependant ébranlé la conception de l'économie comme science inexacte et séparée et enclenché un certain nombre d'évolutions autant dans la théorie que dans les méthodes expérimentales. Voyons auparavant une histoire qui s'est développée dans un premier temps en parallèle, celle des EAC.

1.3 Expériences aléatoires contrôlées et confirmationnisme

Il existe des approches de la scientificité en économie alternatives au réfutationnisme. Pour Milton Friedman (1953), la scientificité d'une théorie se mesure uniquement à l'aune de sa capacité

prédictive. On parle aussi de confirmationnisme pour désigner cette approche. Une différence fondamentale avec l'approche précédente est que Friedman a une conception instrumentale de la théorie. Les hypothèses qui sont à la base des modèles économiques ne décrivent pas dans le détail la réalité économique. La représentation de l'équilibre d'un marché à l'intersection des courbes d'offre et de demande ne prétend pas représenter le comportement de personnes réelles. De même, la représentation du comportement d'un consommateur à travers un programme de maximisation d'une fonction d'utilité n'est pas supposée décrire le raisonnement d'un individu réel. Dès lors, puisque les hypothèses qui sont à la base d'une théorie n'ont pas prétention à être réalistes, elles ne méritent pas d'être testées. Tester par exemple la transitivité des préférences sur des sujets réels ne présente aucun intérêt. Pour reprendre une terminologie utilisée par l'épistémologue Alan Chalmers (1987), les sciences économiques ainsi conçues entrent dans la catégorie des sciences instrumentales. Alors que « pour le réaliste, la science cherche à formuler des descriptions *vraies* de ce qu'est réellement le monde. [...] La composante théorique d'une science [instrumentale] ne décrit pas la réalité. Les théories y sont vues comme des instruments conçus pour relier entre eux deux séries d'états observables.» (Chalmers 1987, page 234). Dans sa *Méthodologie de l'économie positive*, Milton Friedman (1953) est peut-être celui qui est allé le plus loin dans la défense du caractère instrumental de la théorie économique lorsqu'il y affirme qu'« en général plus une théorie est significative plus ses postulats [...] sont irréalistes » (1953, page 11). Il ne s'agit pas là d'une simple provocation car Friedman explique plus loin le sens de cette assertion. L'hypothèse doit effectivement être irréaliste dans cette conception car, si elle doit capter un élément déterminant de la réalité, elle ne doit pas capter toute la réalité. Elle doit seulement capter de la réalité ce qui est pertinent pour le phénomène que le scientifique veut étudier. En conséquence, la réalité que l'on observe est méconnaissable dans les hypothèses de la théorie. Notons ici que la théorie macroéconomique à laquelle Friedman fait essentiellement référence correspond tout autant que la théorie microéconomique à la description de Hausman. Elle se fixe comme objectif de mettre en évidence des relations de causalité conçues comme des lois générales valides, aux inexactitudes près, sur un domaine séparé.

Selon cette conception de la scientificité de la science économique à la Friedman (1953), ce n'est donc pas le réalisme des postulats qui compte. « [...] Le seul test de la validité d'une

hypothèse est la confrontation de ses prévisions avec l'expérience. » L'expérience à laquelle Friedman fait référence n'est bien sûr pas l'expérience de laboratoire, mais l'expérience tirée du traitement des données d'observation du passé. La macroéconomie est ainsi traditionnellement accompagnée du traitement économétrique, en particulier de séries chronologiques. Ces analyses servent tout aussi bien à tester les relations qu'à quantifier les paramètres et à calibrer les modèles macroéconomiques. L'estimation est le préalable à la prévision. L'approche instrumentaliste de Friedman a certainement connu ses détracteurs, l'économétrie des données d'observations n'en constitue pas moins le moyen de validation empirique le plus utilisé en économie. La confiance dans les modèles théoriques augmente bien avec leurs capacités prévisionnelles.

La critique de Lucas a grandement fragilisé ce système de validation. Au milieu des années 70 la recherche en macroéconomie est dans la lignée de l'épistémologie Friedmanienne. Il s'agit de construire de grandes fonctions macroéconomiques comme la consommation, l'investissement, la demande de travail, la demande de monnaie, qui constituent les éléments de base des modèles. Or, selon Lucas, comme l'expliquent Michel De Vroey et Pierre Malgrange (2007), « on peut donc utiliser les modèles macroéconomiques pour prévoir le fonctionnement normal d'une économie mais pas pour prédire les conséquences d'un changement institutionnel de l'environnement. L'apport de Lucas [1976] a été de montrer que si l'on pouvait considérer que ces modèles étaient satisfaisants dans leur utilisation pour la prévision, il n'en allait pas de même pour l'évaluation de politiques économiques alternatives par des simulations. » Parce que les politiques économiques des gouvernants changent les comportements des gouvernés, le fonctionnement de l'économie à un moment donné peut changer qualitativement. En conséquence, Lucas préconise la construction de ce qu'il appelle des « modèles structurels » qui décrivent explicitement la rationalité des agents et permettent donc d'endogénéiser les changements de comportement et leur utilisation pour l'évaluation des politiques publiques. Ainsi, l'économie appliquée produit depuis longtemps des études estimant l'impact de politiques publiques en utilisant pour cela des données d'enquêtes, déclaratives ou encore des données administratives. Par exemple l'économie du travail a fortement investi ce champ, notamment depuis les années 1970 à la suite des travaux de Heckman (1974).

Notons cependant que l'argumentation de Lucas qui vise à promouvoir la construction de fondements microéconomiques de la macroéconomie repose sur l'hypothèse que la rationalité des agents reste exogène et constante au moment où, pendant la deuxième partie des années soixante-dix, cette rationalité commence à être contestée en théorie et dans les expériences. Nul doute que ce contexte a été particulièrement favorable au développement des EAC comme alternative aux modèles structurels pour l'évaluation des politiques publiques. Celles-ci ont en effet revendiqué une approche dite « pragmatique » ou même « a-théorique » pour ne pas avoir à fonder leur légitimité sur un modèle microéconomique contesté. Dans cette perspective, Hélène Erkel-Rousse (2014), dans son introduction au numéro spécial d'Economie et Prévision consacré à l'évaluation des politiques publiques, indique que « le premier dossier de ce numéro se focalise sur les méthodes micro-économétriques a-théoriques d'évaluation d'impact des politiques. »

Ainsi nous défendons ici l'idée que les EAC participent d'un critère de scientificité de type confirmationniste, dans la lignée de l'approche Friedmanienne décrite au début de cette section. On pourrait voir une contradiction entre l'objectif de cette dernière approche qui est d'évaluer une théorie et le fait que les spécialistes des EAC se revendiquent parfois comme étant a-théorique. Soyons clair, ce pragmatisme ne signifie pas que ces spécialistes ne font pas référence à des modèles. Le choix des mesures évaluées par les EAC se justifie par des croyances sur des relations de cause à effet, que ces croyances soient basées sur une théorie détaillée ou pas. Par contre alors que Friedman pensait qu'il est nécessaire de proposer un modèle permettant de reproduire la relation causale, les EAC les plus a-théorique se contentent du point de départ : la mesure, et de ses effets attendus. Mais pour Friedman, comme pour les spécialistes des EAC, c'est uniquement la relation de causalité qui doit être testée, pas les rouages du modèle, qui n'est qu'instrumental.

La légitimité recherchée dans les EAC est empirique et les sources d'inspiration de cette méthode se trouvent plus dans la statistique que dans la théorie économique. Pour mieux le comprendre comparons les EAC aux modèles structurels à travers un objet d'étude commun à ces deux méthodes : l'estimation de l'effet de réformes visant à accroître l'offre de travail, et en conséquence le taux d'emploi. Ce sujet est depuis longtemps l'objet de débats entre des courants politiques qui, considérant le niveau de l'emploi, dénoncent une insuffisante

incitation à l'activité et d'autres qui proposent des mesures redistributives, mais susceptibles de comporter des effets désincitatifs. Pour éclairer le débat, les méthodes structurelles procèdent en plusieurs étapes. Une fois les paramètres structurels estimés, ceux-ci sont utilisés pour simuler l'effet des réformes affectant le niveau des variables auxquelles ces paramètres sont appliqués. Dans notre exemple, une fois la sensibilité à l'incitation financière estimée de manière générale, on simule l'impact de la réforme sur le niveau de l'incitation financière, puis à partir du paramètre estimé, l'impact final sur le taux d'emploi. La principale difficulté tient dans l'estimation du paramètre structurel qui se heurte au fait qu'une partie des données nécessaires n'est pas observable. Si, moyennant quelques hypothèses, on peut connaître le niveau d'incitation à l'emploi d'un actif occupé en comparant son salaire observé à ce qu'il pourrait percevoir s'il était sans emploi, il est en revanche beaucoup plus difficile de connaître ce niveau pour un individu sans emploi dont le salaire *potentiel* n'est donc pas observable. Pour ce faire, l'économètre doit donc tout d'abord estimer le salaire potentiel à partir de variables (qualification, expérience, etc..) observables cette fois ci pour tous les individus. Mais ces variables ne sont pas suffisamment fines, et une partie de ces déterminants sont eux-mêmes inobservables. Face à ce qu'il ne peut observer, l'économètre doit s'appuyer alors sur des hypothèses qui ne sont pas toujours testables⁴. Ainsi ce type d'étude doit composer avec la nature des données et palier à un manque d'information qui peut conduire à des estimations imprécises ou biaisées.

Qu'en est-il dans le cas d'une estimation de l'impact d'une mesure agissant sur l'incitation à l'activité (type impôt négatif, prime pour l'emploi) par une EAC ? Il s'agit alors d'attribuer la mesure, pour la durée de l'expérimentation, à une partie seulement d'un échantillon d'individus. Le principal critère de validité, qui détermine le degré de confiance, est celui de l'affectation aléatoire : les individus qui bénéficient de la mesure composent le *groupe test* doivent être choisis au hasard, les autres constituant alors *le groupe témoin*. Il faut également que l'échantillon dispose de suffisamment d'individus pour assurer la précision des estimations. Cette méthode vise à construire deux groupes dont les caractéristiques seront en

⁴ A ce sujet, voir l'article de Roux (2015), qui traite des approches structurelles et non structurelles en microéconométrie de l'évaluation des politiques publiques. L'auteur analyse leurs oppositions mais aussi leurs complémentarités. Certains passages concernent le caractère testable ou non testable des hypothèses requises dans les modèles structurels. L'auteur précise par exemple qu'il n'est pas possible de tester une hypothèse formulée sur la distribution des erreurs lorsque celle-ci porte sur une variable latente.

moyenne identiques, à l'exception du bénéfice de la mesure étudiée qui ne concerne que le groupe test. Idéalement, le caractère aléatoire du tirage exclut toute différence entre les deux groupes qui proviendrait soit du choix de l'expérimentateur, soit de celui des individus eux-mêmes. Il assure également, selon la loi des grands nombres, une proximité des caractéristiques moyennes des individus des deux groupes, d'autant plus forte que la taille de l'échantillon est grande. Dans ce cadre, si la question posée est de savoir si le transfert social modifie le taux d'emploi des individus qui en bénéficient, l'EAC donne au décideur une réponse qui s'appuie sur une expérience dotée d'une forte validité interne. En effet, si les deux groupes sont vraiment identiques, la simple différence de taux d'activité entre les deux groupes estime bien l'effet de la seule différence qu'il y a entre eux : le bénéfice du dispositif testé. Cette méthode permet, dans le cadre théorique idéal présenté ici, de disposer d'un contrôle total des effets autres que ceux induits par le dispositif étudié. Ainsi les décideurs pourraient accorder une confiance plus élevée aux résultats des EAC qu'à ceux issus de la microéconométrie structurelle. Heckman et Smith (1995) le relevaient, tout en regrettant la perte de légitimité des modèles structurels :

Selon une perception répandue dans la communauté des politiques publiques, les modèles structurels sont incapables d'expliquer le comportement. La nouvelle priorité est de déterminer si des programmes spécifiques «fonctionnent», en ce sens qu'ils ont un impact moyen positif, plutôt que de tenter d'améliorer la connaissance des paramètres structurels. (Heckman et Smith 1995, p.86)

A la suite de Lalonde (1986), de nombreux auteurs ont en effet critiqué la sensibilité des modèles structurels aux hypothèses qu'ils requièrent. Heckman (1992) a cependant très tôt rappelé que les EAC reposent aussi sur un certain nombre d'hypothèses. On doit par exemple supposer que le contexte local, voir les caractéristiques de l'institution qui met en place l'expérience, ne sont pas trop spécifiques, ce qui nuirait à la généralisation du résultat. La question de la validité externe des EAC a depuis été traitée en détail dans la littérature (voir par exemple Bédécarrats, Guérin et Roubaud, 2017) : ce qui est valable dans le cadre de l'expérience ne le serait pas forcément si celle-ci était généralisée et pérennisée. En effet, le protocole même des EAC, construites sur de petits échantillons pour évaluer une mesure pendant une période réduite est une source d'obstacles à la validité externe. En particulier, la mise en place à grande échelle de la mesure peut générer des effets dits de composition ou d'équilibre. Aider par exemple des individus au chômage dans leur recherche d'emploi peut avoir des effets importants quand la mesure est limitée à une EAC mais ces effets peuvent

disparaître quand la mesure est généralisée à l'ensemble des chômeurs qui se retrouvent en compétition pour un nombre d'emplois inchangé.

La critique la plus importante concerne cependant le tirage aléatoire qui est le principal critère de validité des EAC. Si l'idée d'un tirage au sort est séduisante, sa mise en œuvre est difficile pour de multiples raisons. Le respect du tirage aléatoire n'est pas garanti, car l'intérêt des institutions est plutôt de contrôler l'affectation au dispositif testé. C'est une des principales critiques des auteurs qui observent concrètement la mise en place des EAC (Devaux Spartakis, 2014). Il faut aussi tenir compte du fait que toutes les personnes tirées au sort n'accepteront pas de d'être incluses dans le dispositif (celui-ci n'est pas obligatoire). On voit donc que les EAC n'échappent pas à un problème de sélection, que celle-ci soit due à l'intervention des institutions impliquées ou à la volonté des sujets de l'expérience. Les caractéristiques du groupe test risquent ainsi de s'éloigner de celles de la population visée par l'expérience. Heckman (1992) identifie par ailleurs un « biais de randomisation », pour désigner le fait que les individus peuvent changer leur comportement par le fait même qu'ils sont dans le cadre d'une expérimentation. Rappelant que les EAC ont été initialement utilisées dans les années 1930 pour tester l'effet de nouvelles techniques sur les rendements agricoles, Heckman nous fait remarquer que les hommes et les femmes qui participent aux expériences menées aujourd'hui, ne sont pas inertes comme des parcelles agricoles.

Enfin, les EAC permettent a priori d'éviter une version biaisée du confirmationnisme qui consiste à choisir les cas à étudier dans une recherche systématique de la confirmation. Esther Duflo parle d'un « biais de consensus : plus les résultats renforcent des résultats connus, plus ils ont de chances d'être publiés ». Les EAC présentent en effet un certain nombre de garde fous contre ce biais. Nous l'avons vu, les choix de modélisation propres à d'autres méthodes, notamment structurelles, et l'ajout d'hypothèses spécifiques dans lesquels certains choix arbitraires peuvent alimenter le confirmationnisme sont minimisés dans les EAC. Il y a aussi une simple raison pratique : les EAC sont relativement coûteuses et leur protocole une fois lancé est peu réversible. Il n'est donc guère possible de réitérer de nombreuses fois une expérience en en modifiant le protocole pour rechercher la confirmation. De fait un grand nombre de publications rendent compte d'hypothèses infirmées par les expériences. Bon

nombre d'expérimentations de dispositifs de politique publique par des EAC concluent par exemple à leur absence d'efficacité.

2. Les conséquences du développement des méthodes expérimentales pour l'évolution de la discipline

Dans la section précédente, nous avons défendu l'idée que les expériences de laboratoire et les EAC résultent d'une même recherche de scientificité, mais aux origines et logiques différentes. On perçoit cependant aujourd'hui des éléments certains de convergence. D'un côté l'économie expérimentale a assoupli son exigence de contrôle de l'environnement et légitimé les expériences dites de terrain. De l'autre, Ester Duflo a revendiqué le recours à la théorie économique pour construire les EAC. Aujourd'hui, la frontière entre expériences de terrain et EAC est floue. Si les unes restent davantage tournées vers la théorie quand les autres ont toujours pour objectif l'évaluation des politiques publiques, les méthodologies sont parfois difficiles à distinguer. Cette convergence reflète les évolutions de la discipline qui, d'une part, abandonne en partie sa prétention à décrire des lois générales pour se concentrer sur l'analyse de relations contextualisées et d'autre part, a ré-ouvert le vieux débat sur la définition de son domaine.

2.1. Convergence des méthodes pour une discipline moins généraliste et plus contextualisée

La tentative réfutationniste, portée par les expériences de laboratoire et décrite dans la section précédente, a provoqué une première réponse défensive de la part des théoriciens qui ont accusé les expériences en laboratoire d'être trop irréalistes. En adressant cette critique à l'économie expérimentale, les théoriciens ont cependant exposé leurs propres contradictions. Tout d'abord, l'affirmation selon laquelle les conditions expérimentales ne correspondent pas au domaine d'application de la théorie, n'est pas compatible avec la conception généraliste de Charles Plott. Il fallait donc admettre que le laboratoire ne peut plus être considéré comme un cas particulier d'une théorie générale, parce que les lois économiques ne sont pas générales. En effet, en conséquence, la dernière décennie a vu le développement d'une nouvelle conception d'un comportement qui dépend du contexte dans lequel les décisions sont prises. En 2007, Levitt et List affirmaient déjà :

Les actions que prennent les personnes sont affectées par un ensemble de situations relationnelles, de normes sociales, de cadres, d'expériences passées et de leçons tirées de ces expériences. [...]Par conséquent, le chercheur expérimental perd souvent une partie du contrôle de l'ensemble du contexte dans lequel le sujet prend des décisions. (Levitt et List 2007, page 162)

Ainsi le laboratoire ne pourrait plus être considéré comme l'environnement parfaitement contrôlé qu'espéraient les expérimentateurs réfutationnistes. Cette critique des expériences de laboratoire a eu pour conséquence une plus grande légitimité accordée aux expériences hors laboratoire. Une littérature basée sur des expériences de terrain du type «Lab in the field» s'est ainsi développée dans les revues les plus reconnues de la discipline depuis les années 2000. Selon Harrison et List (2004) il est bien difficile aujourd'hui d'effectuer une catégorisation de ces expériences et lorsqu'ils tentent de le faire c'est sans oublier de souligner à quel point les frontières restent floues. Ils définissent cependant les expériences de terrain en prenant comme référence les expériences de laboratoire et en envisageant à partir de là les différentes façons de se rapprocher d'environnements économiques « naturels ».

En quête d'une plus grande pertinence, les économistes expérimentaux recrutent des sujets sur le terrain plutôt que dans la classe, utilisent des biens réels plutôt que des valeurs induites et utilisent un contexte plutôt qu'une terminologie abstraite décrites dans les instructions. (Harrison et List 2004, pp. 1009-1010)

D'un point de vue méthodologique, comme dans l'approche expérimentale traditionnelle, les expériences de terrain utilisent la comparaison des traitements d'un même protocole pour mettre en évidence, par déduction, une relation de causalité. Par ailleurs, la distinction entre une expérience de terrain qui, à la fois, recruterait des sujets sur le terrain et utiliserait des biens réels dans un contexte réel et une EAC devient ténue. D'autant plus si les sujets sont affectés aux différents traitements de façon aléatoire. La différence reste peut-être au niveau des objectifs de l'expérience, compréhension d'un mécanisme de comportement ou évaluation d'une mesure de politique publique. Nous verrons plus loin que même cette différence semble s'estomper.

De leur côté, les EAC sont souvent critiquées pour leur caractère a-théorique et pour le fait qu'elles ne peuvent évaluer que des micro-projets et restent donc à un niveau trop particulier. Ainsi, selon Angus Deaton et Nancy Cartwright (2018) :

Les ECR peuvent jouer un rôle dans l'acquisition de connaissances scientifiques et de prévisions utiles, mais elles ne peuvent le faire que dans le cadre d'un programme cumulatif, en se combinant à d'autres méthodes, notamment le développement conceptuel et théorique, pour découvrir non pas «ce qui fonctionne» mais «pourquoi les choses fonctionnent». (Deaton et Cartwright 2018, page 2)

On doit cependant reconnaître que certains randomistes, notamment au sein du J-Pal, utilisent la théorie et développent des programmes de recherche dans lesquels toute une série d'expériences sont menées dans des environnements variés et qui allient modélisation, expériences de terrain et évaluation des politiques publiques. On peut citer à titre d'exemple les travaux de Banerjee, Chandrasekhar, Duflo et Jackson (2013) sur la diffusion de la microfinance, dans lesquels l'évaluation de projets de microfinance est accompagnée d'une recherche sur les mécanismes sociaux et comportementaux qui s'appuie sur un modèle de réseau de diffusion de l'information. Ajoutons que l'ambition de ce type de programme n'est pas seulement d'affiner l'évaluation des politiques publiques en utilisant des modèles et des fondements théoriques. Elle est aussi de participer à l'évolution de la théorie. Ce qui, il est vrai, constitue un changement radical par rapport aux objectifs originaux et montre une volonté de rapprocher les méthodologies EAC de l'économie expérimentale. Dans son discours au collège de France, déjà en 2009, Esther Duflo explique que « l'équipe de recherche [...] peut proposer plusieurs « traitements »⁵ comparés dans la population » et parle *d'expérimentation créative* pour décrire cette nouvelle conception des EAC. Elle décrit ainsi l'évolution de la méthode :

La première génération d'études expérimentales concernait des recherches évaluant des programmes simples, ayant pour objectif de répondre à la question « qu'est-ce qui marche ? » [...] Au-delà des résultats eux-mêmes, ces premières expériences ont progressivement fait émerger la valeur principale de l'approche expérimentale : elle permet non seulement d'évaluer un programme donné mais en outre, en travaillant de manière proche avec l'organisation en charge du programme, de construire des expériences pour tester une ou plusieurs théories. (Duflo 2009, pages 55-56)

2.2.Retour sur le domaine

2.2.1 Préciser le domaine

La réouverture du vieux débat sur le domaine de l'économie est une autre conséquence importante du développement des méthodes expérimentales. Nous avons déjà expliqué pourquoi la réaction défensive des théoriciens à l'économie expérimentale devait être

⁵ Les guillemets sont dans le texte cité.

accompagnée d'une remise en cause de la possibilité d'établir des lois générales. Nous allons discuter ici d'une autre contradiction que cette réaction soulève par rapport à la conception de l'économie comme science inexacte et séparée. Comme nous l'avons vu, selon cette conception, les économistes ne prétendaient pas tout représenter. Ils admettaient que des facteurs exogènes pouvaient interférer avec les causalités économiques et considéraient ces interférences comme des effets parasites qui ne font pas partie du domaine de l'économie. Cependant, lorsque Francesco Guala (2005) rapporte et discute la critique du manque de réalisme du laboratoire, il considère que ce qui n'est pas (encore) dans la définition du domaine - et dans la théorie - n'est plus de l'ordre d'un parasitage mais est nécessaire pour que la théorie fonctionne :

[...] Les conditions expérimentales sont trop irréalistes, ou en quelque sorte pas « correctes » pour pouvoir effectuer des tests théoriques, car certaines conditions et facteurs clés qui font fonctionner la théorie dans son domaine d'application prévu [...] ne sont pas représentées en laboratoire. (Guala 2005, page 185)

Et s'il conclue qu'il faut admettre que le domaine d'application de l'économie n'est pas précisément défini, il ne veut pas pour autant en blâmer la théorie :

[On ne peut pas imposer] l'exigence que le domaine d'une théorie soit écrit explicitement dans ses hypothèses. [...] Le domaine de la théorie économique est constamment redéfini et mis au point après bien des efforts empiriques. (Guala 2005, page 194)

Ce faisant, Guala propose un programme de recherche dans lequel l'interaction entre la théorie et les expériences devrait permettre de *préciser* le domaine de l'économie. Il illustre sa proposition en prenant pour exemple l'hypothèse des préférences découvertes proposée par Charles Plott (1996). Pour interpréter une série de faits stylisés observés en laboratoire qui ne sont pas cohérents avec les prédictions théoriques, Plott affirme que, pour être rationnel, un individu doit bénéficier d'une phase d'apprentissage.

L'hypothèse suggère que l'attitude, comme les attentes, les croyances, l'aversion au risque, etc., sont découvertes, tout comme d'autres éléments de l'environnement. Les gens acquièrent une compréhension de ce qu'ils veulent à travers un processus de réflexion et de pratique. (Plott 1996, page 227)

Ainsi, par exemple, l'hypothèse des préférences découvertes fournit une interprétation permettant d'expliquer le comportement des sujets dans les expériences de laboratoire basées sur des jeux de contribution au bien public décrites plus haut. La diminution des contributions au cours de l'expérience serait due au fait que les sujets utilisent les premières étapes pour découvrir leurs préférences pour le bien public et devenir rationnels.

Considérons comme autre exemple, faisant intervenir cette fois plusieurs méthodologies, le travail de List sur l'effet de dotation. Si l'article de Daniel Kahneman, Jack Knetsch et Richard Thaler (1990) qui met en évidence l'effet de dotation est important dans la littérature en économie, c'est parce que ses auteurs en donnent également une interprétation théorique. Selon eux, l'effet de dotation s'explique par l'aversion aux pertes, qui renvoie à la tendance des individus à préférer éviter les pertes à l'acquisition de gains équivalents. L'aversion aux pertes est une caractéristique de la théorie des perspectives, qui, comme nous l'avons vu, a été proposée par Kahneman et Tversky (1979) comme alternative à la théorie standard de l'utilité espérée. Lorsque John List (2004) propose une expérience de terrain sur l'effet de dotation, son objectif affiché dans le titre de l'article est de tester la théorie néoclassique par rapport à la théorie des perspectives. Son protocole est aussi proche que possible de l'expérience de laboratoire originale décrite plus haut, mais il se déroule sur un marché réel, un marché où l'on trouve des cartes de sports, avec de vrais participants. Ceux-ci reçoivent un mug (ou une barre de céréales) pour les remercier d'avoir bien voulu répondre à un questionnaire et se voient ensuite offerte la possibilité de changer pour une barre de céréale (inversement un mug). Deux types de personnes sont par ailleurs recrutées, d'une part de simples consommateurs, qui viennent sur le marché pour acheter des cartes de sport pour leur propre collection privée et d'autre part, des commerçants qui viennent sur le marché pour vendre des cartes. List observe que ces derniers sont moins susceptibles de présenter un effet de dotation que les premiers et conclue que «les consommateurs apprennent à surmonter l'effet de dotation».

Le programme qui consisterait à préciser le domaine de l'économie répond aussi à une critique émanant d'une littérature qui s'est attachée à montrer les risques liés à une mauvaise connaissance des limites de ce domaine. Ainsi, la vaste littérature sur les motivations intrinsèques et extrinsèques⁶ a été mobilisée par des économistes critiques pour montrer que toutes les relations sociales ne fonctionnent pas comme dans le domaine de l'économie et que confondre les domaines peut amener à mettre en place des politiques publiques qui se

⁶ Selon cette théorie issue de la psychosociologie, les individus auraient deux types de motivations qui les pousseraient notamment à travailler ou plus généralement à participer à différentes activités. Le salaire est alors rangé dans la catégorie des motivations extrinsèques alors que d'autres motivations plus complexes, façonnées en particulier par les normes sociales intégrées par l'individu, sont appelées motivations intrinsèques.

veulent incitatives mais peuvent avoir des résultats contre-productifs. A l'origine de cette littérature on trouve cependant des exemples qui ne sont pas des expériences à proprement parler. Il s'agit de politiques publiques ou de mesures qui ne sont pas mises en place par des expérimentalistes mais par des acteurs de la vie économique et dont les effets sont simplement observés et relevés par les économistes pour la singularité de leur caractère « non attendu ». Ainsi, c'est en comparant les systèmes de don de sang en Grande Bretagne et aux Etats-Unis, que l'économiste britannique Richard Titmuss (1970) a pu observer que les dons de sang sont plus nombreux en Grande Bretagne, et interpréter la différence comme une conséquence de la rémunération des donneurs aux Etats Unis. Les incitations monétaires décourageraient le don de sang, qui ne ferait donc pas partie du domaine de l'économie.

2.2.1. Elargir le domaine

L'idée générale du programme de recherche proposé par Guala (2005) consiste à garder la théorie inchangée et à utiliser des observations provenant d'expériences, en laboratoire ou sur le terrain, pour définir les conditions dans lesquelles le comportement des sujets est compatible avec la théorie. Bien souvent, cependant, comme nous allons le voir dans la suite, l'objectif est au contraire de modifier la théorie afin de la rendre applicable à davantage de circonstances. Il s'agit d'une réponse au programme original de l'économie comportementale décrit plus haut, qui, puisant dans la psychologie cognitive, vise à inventorier les biais de comportement des sujets réels par rapport à la théorie du choix rationnel. Une littérature s'est alors développée qui a proposé de modifier le modèle théorique de la rationalité pour inclure des dimensions jusque-là absentes, comme les émotions et les normes sociales. L'ambition est alors d'étendre le domaine de l'économie en intégrant dans la théorie les autres « passions » que J. S. Mill voulait ignorer. Les économistes se sont ainsi tournés vers les autres disciplines des sciences sociales, la psychologie et la sociologie, mais aussi les neurosciences, pour comprendre les ressorts de ce qui était considéré auparavant comme relevant des effets parasites.

Une démarche, très fréquente dans cette littérature, consiste à introduire de nouvelles variables dans la fonction d'utilité pour représenter ces différents effets. Ainsi, dans les théories des préférences sociales, la fonction d'utilité d'un individu dépend du paiement reçu par d'autres individus. Elle dépend d'une différence entre le paiement de l'individu et celui

des autres dans les modèles d'aversion aux inégalités de Fehr et Schmidt (1999) ou Bolton et Ockenfels (2000), ou d'autres fonctions comme la somme des paiements de tout un groupe d'individus ou le paiement le plus faible dans le modèle concurrent de Engelmann et Strobel (2004) ou encore d'autres fonctions représentant des principes de justice comme chez Cappelen, Drange Hole, Sørensen et Tungodden (2007). Levitt et List (2007) ont proposé un modèle, qu'ils ont voulu plus complet, dans lequel la fonction d'utilité comporte une première composante qui représente le coût ou l'avantage moral non pécuniaire associé à une action et une deuxième composante représentant la richesse associée à cette action. Toute une panoplie d'émotions dites sociales (colère, gratitude, empathie, honte...) ont aussi fait l'objet de modélisations, simples ou sophistiquées (voir par exemple pour une revue de cette littérature Eyal Winter 2014 et Emmanuel Petit 2015). L'idée de base que l'on retrouve dans les différentes contributions à cette littérature est que, lorsqu'un individu perçoit un paiement matériel, la valorisation qu'il en fait dépend du contexte social (voir par exemple les travaux de James Andreoni 1990 et Atsu Amegashi 2006 expliqués dans Petit 2015). L'émotion modélisée est alors caractérisée par des propriétés particulières de la fonction d'utilité. Ainsi, les normes sociales et des émotions ont été intégrées au prix de modifications ad hoc, en préservant l'hypothèse selon laquelle l'objectif de l'individu est de maximiser l'utilité qu'il tire d'un paiement.

Notons que ces tentatives d'extension du domaine de l'économie en enrichissant le modèle théorique ne doivent pas être confondues avec ce que proposait un économiste comme Garry Becker. L'approche Békérienne vise en effet à élargir le domaine sans changer la théorie et n'est en aucun cas liée à une approche expérimentale. Il s'agit alors de modéliser des phénomènes sociaux tels que le mariage ou la criminalité, considérés auparavant comme relevant des autres disciplines, comme des phénomènes obéissant à une rationalité économique.

2.2.3 Une interdisciplinarité pragmatique dans les EAC

La relation au domaine est différente dans le cas des EAC. Leur pragmatisme longtemps revendiqué a fait qu'elles relèvent parfois davantage d'une technique ou d'un art plutôt que d'une science théoriquement fondée. Un cadre théorique ex ante n'est pas alors considéré

comme nécessaire à l'établissement d'une relation de causalité. Il n'intervient le plus souvent que secondairement, au moment de l'interprétation. Cependant, même dans les évolutions plus récentes qui ont orienté certains chercheurs vers une recherche plus fine, associant théorie, expériences et évaluation, le problème se pose différemment. Parce que l'objectif est avant tout d'apporter une réponse en termes de transformations sociales, les choix théoriques peuvent être éclectiques. Pour Duflo (2009) par exemple, les théories utilisées « peuvent être des théories développées par des économistes (c'est souvent le cas quand les économistes font partie de l'équipe), des sociologues, des spécialistes des sciences politiques, des psychologues. »

Dès lors, les limites du domaine pris en compte semblent artificielles et les apports venant de différentes disciplines des sciences sociales apparaissent assez naturellement. Les emprunts interdisciplinaires sont parfois mobilisés dès la conception de l'expérience pour définir les variables qui seront utilisées pour interpréter le résultat final. Par exemple la notion d'estime de soi, issue de la psychologie, est une des variables d'intérêt de l'expérimentation d'un dispositif d'aide à la création d'entreprises (Crépon, Duflo, Huillery, Parienté, Seban, 2014), le dispositif « créajeunes ». L'étude relève un effet ambigu de ce dispositif sur cette variable : il induit moins de très fortes estimations de soi, au profit des niveaux forts et moyens, différences qui disparaissent ensuite à long terme.

C'est aussi parfois la méthodologie même qui se trouve influencée par les autres disciplines. Par exemple, le *Guide méthodologique pour l'évaluation des expérimentations sociales*, disponible sur le site du Fonds d'expérimentation pour la jeunesse indique que « la phase quantitative peut toutefois être précédée d'une phase qualitative, permettant par exemple de repérer dans les propos des acteurs des effets non attendus, qui seront ensuite mesurés ». Or, cette phase qualitative renvoie principalement aux techniques issues de la sociologie. L'éthique et le droit sont également mobilisés dans les expériences contrôlées dont la mise en œuvre s'inscrit dans un cadre juridique précis et suppose le respect de règles éthiques. Nous sommes loin alors d'une perspective millénaire axée sur le développement d'une science séparée, mais dans l'intention d'estimer au mieux l'effet total d'une mesure incluant ceux qui résultent aussi bien de déterminants économiques et non économiques.

Enfin, les emprunts à d'autres disciplines sont aussi mobilisés au moment de l'interprétation des résultats. Par exemple les nombreuses études montrant par « testing »⁷ la réalité des pratiques discriminatoires s'appuient sur des hypothèses typiquement économiques comme la notion de discrimination statistique. La discrimination statistique consiste à utiliser un critère comme un signal donnant une information globale sur un groupe dont on ne connaît qu'imparfaitement les caractéristiques⁸. Des hypothèses extérieures au domaine de l'économie sont toutefois aussi utilisées. Ainsi, dans une étude sur l'effet du type de moyen de transport utilisé sur l'accès à un poste de contrôle de gestion (L. du Parquet, E. Duguet, Y. L'Horty, P. Petit et F. Sari, RFE, 2011), les auteurs montrent que les femmes titulaires d'un permis moto perdent leur avantage dans l'accès à ce type de poste, alors qu'elles sont dans d'autres cas plutôt favorisées :

Un homme avec un permis moto est un homme comme les autres. Mais une femme avec un permis moto n'est plus une femme comme les autres. Elle devient l'égale d'un homme aux yeux de l'employeur et semble perdre tout attrait professionnel, au sens où elle perd son avantage d'accéder à un entretien d'embauche. (Parquet, Duguet, L'Horty, Petit et Sari 2011 page 46)

Ainsi les auteurs semblent ici appuyer leur interprétation sur la notion de genre, qui désigne la part socialement construite des différences entre femmes et hommes, et est issue de la sociologie.

2.3 Une approche normative partagée

Pour une partie de la profession, qui ne considère pourtant pas la théorie comme une bonne description du comportement des acteurs dans le domaine, il n'est nul besoin de tenter de changer ni la définition du domaine, ni la théorie. La théorie du choix rationnel n'est alors pas sensée décrire le comportement tel qu'il est mais tel qu'il devrait être. Comme nous l'avons déjà évoqué plus haut, l'économie comportementale interprète en effet les écarts entre

⁷ Dans ce domaine, les études par « testing » consistent par exemple à répondre à des offres d'emploi par des CV identiques en tout point, sauf sur le critère que l'on suppose discriminant : sexe, consonance du nom, lieu d'habitation etc...Ce type d'étude relève des EAC dans la mesure où un tirage aléatoire guide l'attribution des CV aux différents employeurs.

⁸ Par exemple, appliquée à la discrimination selon le lieu de résidence, Petit, Bunel, Ene et L'Horty (2016) donnent la définition suivante de la discrimination statistique: « Les employeurs observent imparfaitement la productivité des candidats à l'emploi. Or, évaluer la productivité d'un candidat a un coût non nul pour l'employeur. Le lieu de résidence joue alors le rôle de signal. Les employeurs peuvent par exemple s'appuyer sur des croyances ou connaissances relatives à la tenue vestimentaire, la façon de s'exprimer, la qualité des réseaux, la fiabilité et le sérieux professionnels, la motivation dans le travail, etc. »

observations au sein du laboratoire et prédictions théoriques en termes de biais cognitifs, ce qui suppose que la rationalité définie en théorie de la décision constitue une référence normative. En conséquence, le courant du Nudge, qui trouve ses origines dans l'économie comportementale et un ouvrage célèbre de Thaler et Sunstein (2008), propose de penser les politiques publiques en tenant compte de ces biais et en tentant de les éliminer, pour amener les individus à prendre des décisions dans *l'intérêt général et/ou pour leur propre bien-être*. On serait tenté de discuter d'une définition de l'intérêt général et de celui de l'individu dans ce cadre. Nous nous contenterons de dire ici que la théorie du choix rationnel est considérée comme une référence normative qui permet de décrire les comportements qui seraient les plus souhaitables *du point de vue des objectifs d'une politique publique*. De plus, ce paternalisme libéral veut orienter les individus sans que soit altérée leur liberté de choix. Samuel Ferey, Yannick Gabuthy et Nicolas Jacquemet (2013) définissent ainsi ces nouvelles conceptions des politiques publiques :

En insistant sur les aspects cognitifs des choix individuels (biais, heuristiques, évaluation, faiblesse de la volonté), elles considèrent que l'efficacité d'une politique, ici entendue comme sa capacité à orienter les choix des acteurs, dépend moins des contraintes et des incitations mises en place que de la manière dont ces politiques construisent les contextes de choix des individus. (Ferey, Gabuthy et Jacquemet 2013, page 160)

Construire le contexte de choix des individus pour qu'ils prennent des décisions considérées comme rationnelles ne signifie cependant pas que l'on va pouvoir rendre les individus rationnels. Seuls leurs choix le seront, pas la façon dont ils auront pris leur décision. Considérons par exemple la proposition faite par Thaler et Sunstein (2008) pour augmenter l'épargne retraite des salariés aux Etats-Unis. Il s'agissait de demander aux entreprises de mettre en place par défaut un plan d'épargne pour leurs salariés, ceux-ci ayant la possibilité de le refuser. Si le résultat est effectivement une augmentation du taux d'épargne, ce n'est pas parce que les salariés auront pris une décision de façon rationnelle mais bien parce que le paternalisme libéral aura su exploiter leur biais cognitif.

En économie du développement, cette méthode est qualifiée par Esther Duflo de *modestement normative*. Lorsqu'elle déclare que « nous devons comprendre comment les pauvres raisonnent », ce peut être pour deux raisons. Il peut s'agir de comprendre comment les pauvres raisonnent pour optimiser les effets d'une politique de développement. Ainsi, lorsque l'on veut distribuer des moustiquaires pour lutter contre la malaria, on doit savoir quel est le moyen le plus efficace de les distribuer. Faut-il les vendre à un prix modeste ou les

donner⁹ ? Il peut s'agir aussi d'imaginer une politique de développement dont le seul objectif est de corriger le comportement des pauvres. Ainsi, en discutant cette fois-ci de mesures propres à améliorer les rendements dans l'agriculture, Esther Duflo pose l'hypothèse que les agriculteurs n'utilisent pas assez d'engrais par rapport à ce qui serait optimal, c'est-à-dire rationnel, parce qu'ils procrastinent :

Même lorsque les paysans pensent que les engrais sont efficaces, beaucoup d'entre eux ne les utilisent pas. Ils ont de l'argent après la récolte mais dépensent plus que prévu entre récolte et plantation et n'ont plus d'argent au moment de planter. Les comportements de procrastination sont communs. (Esther Duflo, discours au Collège de France)

A l'appui de cette déclaration, Esther Duflo cite Matthew Rabin et Ted O'Donoghue, qui ont proposé « un modèle d'incohérence temporelle qui explique que nous sommes impulsifs dans le présent, mais rationnels lorsque nous considérons le futur ». Elle explique alors que son équipe a imaginé une mesure, pour corriger cette tendance tout en respectant la liberté de choix des agriculteurs, qui consiste à proposer une livraison gratuite à ceux qui choisissent d'acheter les engrais juste après la récolte. Cette mesure est alors testée grâce à une expérience aléatoire contrôlée et semble accroître la proportion d'agriculteurs utilisant des engrais. Surtout, la mesure est comparée à une autre, qui consiste simplement à subventionner l'achat d'engrais. La livraison gratuite juste après la récolte apparaît plus efficace que la subvention, ce qui est interprété comme suit : les agriculteurs sont bien victimes de procrastination et pas seulement d'un manque d'argent.

A travers ces différents exemples nous voyons comment la théorie du choix rationnel, l'économie expérimentale et comportementale et les EAC sont utilisées conjointement pour la conception et l'évaluation des politiques publiques. Esther Duflo soutient aussi qu'inversement, certains résultats des méthodes d'évaluation peuvent contribuer à l'avancement de la théorie.

Notons que le problème est cependant différent dans les situations d'interaction comme le jeu de contribution volontaire au bien public où les expériences de laboratoire ont montré que les sujets sont plus efficaces que ce que prédirait la théorie. Le modèle de rationalité ne peut donc être considéré, dans ces cas-là, comme une référence normative. L'objectif est alors de comprendre pourquoi les sujets en interaction font mieux que ce que la théorie prédit à

⁹ Référence à la controverse qui a opposé Jeffrey Sachs et William Easterly au milieu des années 2000.

l'équilibre de Nash. La littérature expérimentale sur les jeux de contribution au bien public en est un bon exemple. On peut enfin se demander s'il existe un contexte sans interaction. En effet, même lorsqu'un sujet doit prendre une décision par lui-même, sans aucun lien apparent avec les autres, nous savons que la façon dont il pense est largement influencée par les interactions passées et / ou futures.

Conclusion

Les développements de l'économie expérimentale d'une part et des EAC d'autre part ont certainement des origines distinctes et obéissent à des logiques différentes. Elles ont cependant toutes été développées en économie pour accroître la légitimité de la discipline à des moments critiques où les méthodes plus traditionnelles de modélisation et de validation empirique ont présenté des failles. Nous avons attribué l'origine de l'économie expérimentale à une remise en question de la cohérence interne de la théorie microéconomique et celle des EAC à une remise en question de la validité empirique des modèles macroéconomiques. La logique de développement de chacune de ces méthodes à partir de ces origines spécifiques était aussi différente. L'économie expérimentale s'est construite au départ sur la base d'un critère de scientificité de type réfutationniste alors que les EAC sont apparues comme un outil confirmationniste.

Ces histoires aux origines différentes n'ont pas empêché une certaine convergence dans la façon dont ces méthodes ont influencé la discipline. Si l'on compare l'économie aujourd'hui avec la description que Daniel Hausman brosse, au début des années 90, d'une science inexacte et séparée, on voit apparaître des différences fondamentales que l'on peut clairement attribuer au développement des pratiques expérimentales. Tout d'abord, l'ambition est moins aujourd'hui de défendre l'existence de relations de causalité conçues comme des lois générales que de tenter de comprendre la réalité économique dans ses détails et les spécificités des contextes. On l'observe autant en économie expérimentale que dans l'évaluation des politiques publiques. L'autre changement important, par rapport à la description de Hausman, c'est une remise en question de l'idée d'une science séparée et un bouleversement dans la conception du domaine de l'économie. Cependant, à partir d'un même constat d'une définition du domaine de l'économie insatisfaisante à plusieurs égards

nous voyons s'ouvrir des directions de recherche en partie incompatibles. Une partie de la littérature semble vouloir préciser les conditions dans lesquelles la théorie du choix rationnel s'applique quand une autre tente de modifier cette théorie pour qu'elle puisse s'appliquer à davantage de contextes. Bien sûr, la question de l'interdisciplinarité est au centre de ces bouleversements. Malgré toutes ces remises en cause la théorie du choix rationnel semble garder un rôle parfois central, aussi bien en économie comportementale que dans des EAC, en servant de référence normative. Ce qui mériterait d'ailleurs certainement d'être questionné.

Les différentes méthodes expérimentales ont aussi beaucoup évolué et apparaissent aujourd'hui plus complémentaires que concurrentes. Elles ont en un sens convergé. D'une part l'économie expérimentale a accepté de perdre une partie de ce qu'elle considérait comme un contrôle parfait du protocole en sortant du laboratoire au bénéfice des expériences de terrain. D'autre part les EAC ne voient plus autant d'intérêt à revendiquer un pragmatisme a-théorique. Ces évolutions pourraient rendre un dialogue possible qui tirerait partie des objectifs restés complémentaires des deux approches : l'étude du comportement et le lien avec la théorie dans un cas et l'évaluation des politiques publiques de l'autre.

Références

Banerjee Abhijit, Chandrasekhar Arun G., Duflo Esther, and Jackson Matthew O. (2013), « The Diffusion of Microfinance », *Science*, Vol 341, Issue 6144.

Banerjee Abhijit V., Duflo E. « L'approche expérimentale en économie du développement », *Revue d'économie politique* 2009/5 (Vol. 119), p. 691-726.

Bédécarrats F., Guérin I. et Roubaud F. (2017), « L'étalon-or des évaluations randomisées : économie politique des expérimentations aléatoires dans le domaine du développement, *Papiers de Recherche AFD*, n° 2017-44, Mai.

Cappelen A.W., A. D. Hole, E. Ø. Sørensen and B. Tungodden, (2007). « The Pluralism of Fairness Ideals: An Experimental Approach », *American Economic Review*, 97(3): 818- 827.

Chalmers, Alan (1987), *Qu'est-ce qu'une science ? Popper, Kuhn, Feyerabend*, le Livre de Poche.

Cot Annie L, Ferey Samuel (2016), « La construction de « faits » économiques d'un nouveau type: éléments d'histoire de l'économie expérimentale », *L'actualité économique*, Volume 92, no 1-2.

Cozic, Mikaël (2014), « La méthodologie économique du déductivisme Millien au néo-positivisme », *Revue d'économie politique*, Vol. 124, pp. 23-73.

Crépon B., Duflo E., Huillery E., Parienté W., Seban J. (2014) « Les effets du dispositif d'accompagnement à la création d'entreprise CréaJeunes : résultats d'une expérience contrôlée », *Rapport d'évaluation fait par l'Ecole d'Economie de Paris pour le FEJ*

Deaton, A., Cartwright, Nancy (2018), « Understanding and misunderstanding randomized controlled trials », *Social Science & Medicine*, Volume 210, August 2018, Pages 2-21.

Devaux-Spartakis, Agathe, (2014). « l'expérimentation « telle qu'elle se fait » : leçons de trois expérimentations par assignation aléatoire », *Formation-Emploi n°126*.

De Vroey (2002) « La mathématisation de la théorie économique. Le point de vue de l'histoire des théories économiques », *Reflets et Perspectives de la Vie économique*, vol. XLI, pp.9-20.

De Vroey Michel, Malgrange Pierre (2007), « Théorie et modélisation macro-économiques, d'hier à aujourd'hui », *Revue française d'économie*, volume 21, n°3, pp. 3-38.

Duflo, Esther. (2009) *Expérience, science et lutte contre la pauvreté*, Paris, Fayard.

Du Parquet Loïc, Emmanuel Duguet, Yannick L'Horty, Pascale Petit, Florent Sari (2011), *Mobilité et accès à l'emploi, une expérimentation*, *Revue française d'économie*, 2011/4 volume XXVI, pages 33 à 56

Engelmann, Dirk et Martin Strobel, 2004. « Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments », *American Economic Review* 94 (4), 857-869.

Erkel-Rousse, H. (2014). « Introduction générale », *Economie et prévision*, 204-205,(1), I-XII.

Fehr, Ernst et Klaus M. Schmidt (1999), « A Theory of Fairness, Competition, and Cooperation », *The Quarterly Journal of Economics*, 114(3): 817-868.

Ferey S., Gabuthy Y., et Jacquemet, N. (2013), « L'apport de l'économie expérimentale dans l'élaboration des politiques publiques », *Revue française d'économie*, volume xxviii,(2), 155-194.

Friedman, Milton, (1953), « La méthodologie de l'économie positive », extrait d'*Essais d'économie positive*, Paris : LITEC, 1995. p.7, p.11.

Gerardo Infante, Guilhem Lecouteux and Robert Sugden (2016), « Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics », *Journal of Economic Methodology*, 23:1, 1-25

Guala, Francesco (2005), « Economics in the lab: Completeness vs. testability », *Journal of Economic Methodology*, 12:2, 185-196.

Harrison Glenn et John List (2004), « Field Experiments », *Journal of Economic Literature*, Vol. XLII, 1009-1055.

Hausman, Daniel (1992), *Essays on Philosophy and Economic Methodology*, Cambridge: Cambridge University Press.

Hausman, Daniel (2000), « Revealed Preference, Belief, and Game Theory », *Economics and Philosophy*, 16, p. 99-115.

Hausman, Daniel (ed.) (2008), *The Philosophy of Economics. An Anthology*, Cambridge: Cambridge University Press, 3ième édition.

Heckman J. (1974), « Shadow Prices, Market Wages, and Labor Supply, *Econometrica* Vol. 42, No. 4, pp. 679-694.

Heckman, J.J. (1992). « Randomization and social policy evaluation », in: C Manski, I Garfinkel (Eds.), *Evaluating Welfare and Training Programs*, Cambridge, MA : Harvard University Press, p. 201-230.

Heckman, James J. and Jeffrey A. Smith. (1995), « Assessing the Case for Social Experiments », *Journal of Economic Perspectives*, 9 (2): 85-110.

Heckman J. J., Smith J., Clements N. (1997), « Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts », *Review of Economic Studies*, 64(4), p. 487-535.

Hodgson G., Lamberg J., (2018) « The Past and Future of Evolutionary Economics: Some Reflections Based on New Bibliometric Evidence », *Evolutionary and Institutional Economics Review*, 15(1), pp. 167-187.

Infante, Gerardo, Guilhem Lecouteux, Robert Sugden. (2016), « Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics », *Journal of Economic Methodology*, Taylor & Francis (Routledge), 23 (1), pp.1-25.

Kahneman, Daniel, Knetsch Jack L., and Richard H. Thaler. (1990), « Experimental Tests of the Endowment Effect and the Coase Theorem », *Journal of Political Economy* 98, no. 6.

Kahneman D. and Tversky A. (1979). « Prospect Theory: An Analysis of Decision under Risk », *Econometrica*, 47(2), 263-291.

Lalonde R. (1986), « Evaluating the Econometric Evaluations of Training Programs with Experimental Data », *American Economic Review*, vol. 76, issue 4, 604-20

Levitt, Steven D., and John A. List. (2007) « What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? », *The Journal of Economic Perspectives* 21, no. 2 : 153-74.

L'Horty, Y. et Petit, P. (2011). « Evaluation aléatoire et expérimentations sociales », *Revue française d'économie*, volume xxvi,(1), 13-48.

Lucas, R. (1976), « Econometric policy evaluation: A critique », *Carnegie-Rochester Conference Series on Public Policy*, vol. 1, issue 1, 19-46.

Marshall, Alfred (1920). *Principles of Economics* (Revised ed.). London: Macmillan

Mill, J. S. (1843). *A System of Logic*. London: Longmans, Green & Co., 1949.

O'Donoghue Ted et Rabin Matthew (1999), « Doing It Now or Later », *The American Economic Review*, vol. 89, no 1, p. 103-124.

Petit, E. (2015). *Économie des émotions*. Paris: La Découverte.

Petit, P., Bunel, M., Ene, E. Et L'Horty, Y. (2016), « Effets de quartier, effet de département : discrimination liée au lieu de résidence et accès à l'emploi », *Revue économique*, vol. 67, (3), 525-550.

Plott C. R. (1991). « Will Economics Become an Experimental Science? », *Southern Economic Journal*, 57(4), 901-919.

Plott, C. R. (1996) *Rational Individual Behavior in Markets and Social Choice Processes: the Discovered Preference Hypothesis*. In: *The Rational Foundations of Economic Behaviour*. McMillan, London, pp. 225-250.

Popper, K. (1963), *Conjectures and Refutations*, London: Routledge.

Roux, S. (2015). « Approches structurelles et non structurelles en micro-économétrie de l'évaluation des politiques publiques », *Revue française d'économie*, volume xxx, (1), 13-65.

Samuelson, Paul (1938), « A Note on the Pure Theory of Consumer's Behaviour », *Economica*, 5(17), p. 61-71.

Sen, Amartya (1973) « Behaviour and the Concept of Preference », *Economica*, 40, p. 241-259.

Smith, V. (1976), « Experimental Economics: Induced Value Theory », *The American Economic Review*, Vol. 66, No. 2

Thaler, Richard and Cass Sunstein (2008), *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

Titmuss, Richard, (1970), *The Gift Relationship: From Human Blood to Social Policy*.

Winter, Eyal, (2014), *Feeling Smart. Why Our Emotions Are More Rational Than We Think*. PublicAffairs.