



Mesurer l'effet des politiques publiques : l'essor des évaluations aléatoires

Quel est l'effet des programmes d'accompagnement des demandeurs d'emploi ? Quel est l'impact des politiques de lutte contre la pauvreté dans les pays en développement ? L'objectif d'une évaluation d'impact est de répondre de manière rigoureuse à ce genre de questions. Les évaluations aléatoires, qui consistent à comparer la situation de deux groupes tirés au sort, un groupe «traitement» recevant le programme alors que l'autre groupe «contrôle» ne le reçoit pas, connaissent un essor important et deviennent une méthode phare. Elles permettent de mesurer précisément la valeur ajoutée des politiques publiques. Cette mesure est indispensable pour modifier et améliorer les politiques.

William Parienté*

Dans le cadre du «tax shift» belge, tous les employeurs qui engagent un premier travailleur depuis le 1^{er} janvier 2016 bénéficient d'une exonération des cotisations sociales patronales à vie sur ce premier emploi. De son côté, la France a démarré en 2012 le Contrat de génération qui a pour objectif de favoriser l'embauche à la fois de jeunes et de seniors au sein des entreprises. Au Mexique, le gouvernement a décidé au début des années 2000 de généraliser les allocations conditionnelles (à l'assiduité scolaire et à des visites médicales des enfants) à l'ensemble des ménages pauvres. Ces interventions visent, chacune dans leur contexte, à améliorer la situation des populations concernées (par le chômage ou la pauvreté). Si la motivation principale de ces politiques est souvent claire, se pose néanmoins la question de leur efficacité.

L'efficacité d'une politique se mesure essentiellement par la comparaison de deux paramètres : son coût et son impact. L'estimation du coût nécessite de disposer d'informations précises sur l'ensemble des composantes d'un programme et d'être capable de les valoriser monétairement. La mesure de l'impact quant à elle vise à identifier la valeur ajoutée de la politique par rapport à la situation sans la politique. Si l'estimation des coûts est possible dans un grand nombre de cas (en faisant certaines hypothèses sur la valorisation des composantes du programme), la mesure de l'impact est en général beaucoup plus difficile à obtenir car beaucoup plus complexe d'un point de vue méthodologique. De ce fait, la «simple» comparaison coût-impact n'est pas réalisable pour une grande partie des politiques existantes. Par exemple, un grand nombre d'interventions sont mises en œuvre pour lutter contre le chômage dans les pays développés. Les politiques actives sur le marché du travail représentaient plus de 13 milliards d'euros en France en 2012 (DARES, 2014) et près de 2,8 milliards d'euros en Belgique en

* Le présent article se base partiellement sur un autre article de l'auteur à paraître dans un numéro spécial de la revue *Reflets et perspectives de la vie économique* sur le thème de l'économie du développement. L'auteur remercie Muriel Dejemepe, Vincent Bodart et Bart Cockx pour leurs commentaires et suggestions.

2013 (OCDE, 2015). Dans un autre contexte, l'aide aux pays en développement s'élevait à 131 milliards de dollars en 2012 (OCDE, 2014). Dans les deux cas, il est légitime de se demander comment ces politiques contribuent à diminuer le chômage ou la pauvreté et ce en regard des moyens utilisés.

L'évaluation prend une place de plus en plus importante dans l'analyse des politiques publiques, que ce soient des politiques gouvernementales, des programmes provenant de bailleurs internationaux ou des interventions d'organisations non gouvernementales (ONG). Une évaluation d'impact vise à mesurer les effets de la politique sur les bénéficiaires directs et éventuellement indirects. En cela, l'évaluation d'impact se distingue des évaluations de processus ou d'audit qui visent à mesurer si une intervention s'est correctement mise en œuvre. Une évaluation d'impact revient à comparer la situation des individus qui ont participé à la politique avec leur situation s'ils n'avaient pas participé. Cette situation contrefactuelle n'existe évidemment pas dans la nature. Toute la difficulté d'une évaluation consiste à reconstruire cette situation manquante.

Les économistes se sont abondamment penchés sur ce problème, notamment sur l'évaluation des politiques de l'emploi (Heckman *et al.*, 1999). Il existe plusieurs méthodes pour évaluer l'impact d'une politique. On distingue les méthodes non ou quasi-expérimentales des méthodes expérimentales ou évaluations aléatoires (ou encore expériences contrôlées), que nous développerons dans ce numéro de *Regards économiques*.

Aujourd'hui, un grand nombre de politiques ont été ou sont évaluées avec la méthode aléatoire dans de nombreux domaines tels que les politiques actives de l'emploi, les politiques d'éducation, les politiques de développement. L'objectif de cet article est de présenter la méthode de l'évaluation aléatoire et son apport, les conditions de sa mise en œuvre ainsi que son application à différents domaines de l'économie et des politiques publiques. Nous discuterons également de certaines limites et présenterons les stratégies existantes pour les surmonter. Enfin nous concluons sur l'intérêt de la méthode pour l'amélioration des politiques publiques et de son apport à la science économique.

1. Qu'est-ce que l'évaluation aléatoire ?

L'évaluation aléatoire en sciences sociales est directement inspirée des essais cliniques en sciences médicales utilisés pour tester l'efficacité d'un médicament. Elle consiste à comparer la situation de deux groupes tirés au sort, un groupe «traitement» recevant le programme alors que l'autre groupe «contrôle» ne le reçoit pas. Ces deux groupes sont statistiquement identiques (c'est-à-dire qu'ils ont en moyenne les mêmes caractéristiques) au départ grâce au tirage au sort, à condition d'avoir un échantillon suffisamment large.

L'origine des évaluations aléatoires est attribuée à Ronald Fisher (1923) avec des premières expériences en agriculture qui visaient à mesurer l'impact de certains intrants sur la productivité des terres agricoles. Fisher a été le premier à introduire la notion de groupe contrefactuel (groupe qui ne reçoit pas le programme) comme élément fondamental pour identifier l'impact d'un traitement (Levitt et List, 2009).

En comparaison avec la méthode expérimentale, les méthodes non-expérimentales essaient de reconstruire la situation contrefactuelle (celle qu'on aurait observée en l'absence du programme) du groupe traité en sélectionnant non aléatoirement un groupe témoin qui a des caractéristiques similaires au groupe qui a reçu le programme. Dans la plupart des cas, le groupe contrôle est construit après la mise en œuvre du programme et après que le groupe traité a reçu le programme. La difficulté majeure de ce type de méthode se trouve dans la capacité de l'évaluateur de construire un groupe contrôle qui a les mêmes caractéristiques que le

... Qu'est-ce que l'évaluation aléatoire ?

groupe traité. En effet, il existe toute une série de caractéristiques inobservables qui peuvent à la fois déterminer l'accès au programme et affecter la variable d'impact d'intérêt. Dans ce cas la comparaison entre les deux groupes sera biaisée. Ce biais de sélection peut provenir des individus qui s'auto-sélectionnent dans un programme mais aussi des responsables du programme qui ciblent une population spécifique. Par exemple, les demandeurs d'emploi qui participent à une politique de formation volontaire ont certainement des caractéristiques différentes de ceux qui ne participent pas alors qu'ils sont aussi éligibles. Certaines méthodes cherchent à améliorer la comparaison entre ces deux groupes en prenant en compte un ensemble de caractéristiques observables tels que l'âge, l'expérience professionnelle, le lieu d'habitation, etc. (régressions ou méthodes d'appariement) ou en prenant en compte l'évolution de la situation (en termes de variable d'impact) des individus des deux groupes pendant une période donnée (méthodes de différence de différence). Il reste cependant très difficile de contrôler pour des caractéristiques inobservables (par exemple le niveau de motivation ou l'esprit d'initiative) qui peuvent affecter à la fois la décision de participer, l'évolution et le niveau de la variable d'impact.

L'avantage de la méthode aléatoire est qu'elle supprime le biais de sélection. La comparabilité des groupes est obtenue grâce à la randomisation et non par hypothèse contrairement à l'ensemble des méthodes non-expérimentales. Dans certains cas spécifiques, il peut exister une assignation quasi-aléatoire du programme survenant de manière naturelle sans intervention de l'évaluateur. Certaines règles régissant l'accès à un programme (par exemple l'âge des bénéficiaires) créent un accès différentiel au programme. Par exemple, les individus qui se trouvent en-dessous d'un certain seuil d'âge sont affectés par la politique alors que ceux qui sont un peu plus âgés ne le sont pas. Sur la partie de l'échantillon qui est proche du seuil de sélection, les conditions sont proches d'une expérience aléatoire. On parle alors d'expériences naturelles. Si ces cas fournissent des opportunités très intéressantes pour évaluer l'impact d'une politique, ils sont relativement rares et ne concernent en général qu'une partie de la population d'intérêt (dans notre exemple, celle qui se concentre autour du seuil d'âge). Les résultats sont donc difficilement généralisables à l'ensemble de la population.

Un certain nombre d'études (Lalonde, 1986; Smith et Todd, 2005; Glazerman *et al.*, 2003; Arceneaux *et al.*, 2006) comparant les évaluations expérimentales et non-expérimentales sur une même intervention montrent que les résultats obtenus sont significativement divergents. Ceci suggère que les méthodes non-expérimentales n'arrivent pas à réduire totalement les biais de sélection existants.

2. Développement des évaluations aléatoires en sciences sociales

Les évaluations aléatoires ont connu un essor important avec la multiplication des essais cliniques en sciences médicales qui cherchaient essentiellement à mesurer l'efficacité de certains traitements. En général, les essais cliniques ont un protocole en «double aveugle» dans lequel les groupes traités et contrôles ne sont connus ni des individus testés ni des évaluateurs. Les conditions d'évaluations sont celles d'un laboratoire où la plupart des paramètres du dispositif expérimental sont contrôlés. En sciences sociales, les conditions sont en général assez différentes (et il n'est naturellement pas possible d'expérimenter en double aveugle) en ce sens que le terrain d'expérimentation n'est pas entièrement contrôlable par l'évaluateur. Ceci peut avoir des conséquences sur la validité du dispositif expérimental (point abordé en section 5).

Une des premières expériences contrôlée à grande échelle en sciences sociales a été effectuée sur une politique de l'emploi mise en œuvre dans les Etats de Pennsylvanie et du New Jersey aux Etats-Unis (Ross, 1970; List et Rasul, 2011). La politique consistait en un crédit d'impôt (ou impôt négatif) et visait à accroître

... Développement des évaluations aléatoires en sciences sociales

L'offre de travail des individus. L'évaluation a couvert un échantillon de 1.300 ménages auxquels étaient assignés aléatoirement différents niveaux de crédits d'impôt. Une première série de résultats a démontré des effets plutôt positifs sur l'offre de travail des hommes. Leur validité a été cependant remise en cause ultérieurement car les conditions de l'expérience se sont éloignées du dispositif initial (List et Rasul, 2011).

Les évaluations aléatoires ont connu une progression importante par la suite aux Etats-Unis et ont été appliquées à de nombreux domaines des politiques sociales (santé, éducation, marché du travail, etc.). Une part significative des évaluations ont eu pour objectif direct d'informer les décideurs des politiques publiques, notamment celles menées par MDRC («Manpower Demonstration Research Corporation»). En parallèle, certaines expérimentations ont reçu un écho très important au sein de la recherche scientifique. Par exemple, les évaluations du projet «Moving to Opportunity» qui visait à sortir certaines familles pauvres de zones défavorisées (Katz *et al.*, 2001, 2003, 2007) ou encore les programmes préscolaires aux Etats-Unis comme le projet «Perry School» destiné aux enfants venant de milieux défavorisés (Heckman *et al.*, 2010).

L'expérimentation de Progres, programme d'allocations conditionnelles aux ménages pauvres au Mexique, menée par l'IFPRI (International Food Policy Research Institute) a été une des premières évaluations à grande échelle mise en œuvre dans un pays en développement. Ces dix dernières années, les évaluations aléatoires ont connu un développement très important avec le Abdulatif Jameel Poverty Action Lab (J-PAL) qui a été créé par trois économistes aux Etats-Unis, Esther Duflo, Abhijit Banerjee et Sendhil Mulainathan. Le J-PAL est un réseau de plus d'une centaine de chercheurs en économie qui a pour objectif d'évaluer les politiques publiques avec la méthodologie d'évaluation aléatoire et d'en diffuser les résultats. Les chercheurs du J-PAL sont impliqués dans près de 600 évaluations (terminées ou en cours) dans 61 pays portant sur un grand nombre de domaines, tels que l'éducation, la santé, l'accès aux services financiers, la gouvernance, etc.

Dans certains cas, les évaluations peuvent directement venir d'une initiative politique. En France, un grand nombre d'évaluations ont été mises en œuvre ces dernières années grâce à un programme gouvernemental d'expérimentations sociales (voir encadré 1).

3. L'évaluation aléatoire en pratique

3.1. Comment randomiser ?

Une des conditions primordiales des évaluations aléatoires est que le dispositif d'évaluation soit pensé et mis en œuvre avant que le programme ne démarre car celui-ci doit être assigné de manière aléatoire dès le début. Il n'est pas forcément nécessaire d'évaluer un nouveau programme, une nouvelle composante dans un programme existant pouvant aussi être l'objet de l'évaluation.

Les évaluations aléatoires en sciences sociales ne sont pas mises en œuvre dans les conditions de laboratoire comme pour un essai clinique¹. Cela implique forcément que les conditions expérimentales sont plus difficiles à conserver, notamment sur une longue période. Cela nécessite un effort important de suivi du dispositif. La façon de randomiser est aussi plus flexible que dans le cas d'essais

¹ Il existe cependant une série d'expérimentations en laboratoire en sciences économiques. Ces expériences visent généralement à tester le comportement des agents sous différents scénarios. Leur limite principale est d'avoir une validité externe relativement faible (les conditions de laboratoires et les populations étudiées sont souvent très spécifiques).

ENCADRÉ 1

Le Fonds d'expérimentation pour la jeunesse

Le Fonds d'expérimentation pour la jeunesse (FEJ) a été créé en France en 2008 sous l'impulsion de Martin Hirsch alors Haut-Commissaire aux Solidarités Actives. Les objectifs de cette initiative sont d'expérimenter de nouvelles interventions visant à favoriser l'insertion des jeunes et d'évaluer leur impact de manière rigoureuse. En termes de politiques publiques, l'objectif est de généraliser (d'essaimer) les interventions dont les évaluations auront démontré leur efficacité. Ces interventions ont été sélectionnées par le biais de nombreux appels à projet (16) portant sur plusieurs thématiques telles que l'insertion professionnelle, le décrochage scolaire et universitaire, la lutte contre les discriminations, la santé, etc. En termes de méthodologie d'évaluation, le FEJ encourageait la démarche expérimentale et si cela n'était pas faisable une méthode non expérimentale qui comprenait un groupe témoin. Chaque porteur de projet devait donc intégrer un dispositif d'évaluation dans la mise en place du programme, le plus souvent en collaboration avec une équipe de recherche. Au total, en juillet 2013, 554 projets avaient été soutenus lors de 16 appels à projets pour un montant de 163 millions d'euros. L'ensemble de ces expérimentations donnera lieu à 295 rapports d'évaluation (Rapport d'activité FEJ, 07/2013). Ces projets ont permis de générer des collaborations fructueuses entre un porteur de projet et une équipe de recherche. Par exemple, le projet «la Malette des Parents» (Avisati *et al.*, 2014) mené en collaboration avec le rectorat de Créteil visait à améliorer, à l'aide d'un ensemble d'outils, le dialogue avec les parents d'élèves des quartiers défavorisés afin d'accroître leur implication dans la scolarité de leurs enfants. L'évaluation aléatoire a montré que le programme avait des effets positifs sur les résultats scolaires, réduisait l'absentéisme et les problèmes de comportement.

... Comment randomiser ?

cliniques. Il s'avère qu'il existe en fait différentes possibilités d'insérer une composante aléatoire dans la mise en œuvre d'un programme. En général, les programmes parviennent rarement à toucher toute la population éligible pour des raisons budgétaires ou à toucher tout le monde au début du programme pour des raisons opérationnelles. Ce rationnement naturel offre une opportunité intéressante pour assigner aléatoirement soit le programme soit le moment auquel le programme est proposé.

En termes pratiques, la randomisation peut se faire au niveau individuel (par exemple des demandeurs d'emploi) ou à un niveau plus agrégé (des agences d'emploi). Le niveau de randomisation a des implications sur la précision statistique (en général la puissance statistique diminue avec le niveau d'agrégation de la randomisation, cf. infra) et peut avoir des effets sur la validité de l'estimation. En effet, en cas d'externalités potentielles entre les bénéficiaires et non bénéficiaires de la politique, il est préférable de choisir une unité de randomisation plus agrégée qui limite ces externalités. Ensuite, il n'est pas nécessaire que l'ensemble des individus assignés au traitement participent et que l'ensemble des individus assignés au contrôle ne participent pas. La proportion de participation doit cependant être substantiellement plus élevée dans le groupe traitement que dans le groupe contrôle.

Enfin, l'expérimentation peut poser des questions d'ordre éthique car elle génère une inégalité d'accès au programme. En effet, seule une partie de la population éligible à un programme y aura accès (au moins dans un premier temps). Cette différence d'accès au programme ne résulte d'ailleurs pas de la randomisation mais de l'expérimentation (c'est-à-dire de la mise en œuvre du programme sur une partie seulement de la population éligible). Par ailleurs, étant donné que l'impact du programme n'est pas connu, fournir un programme dont on ne connaît pas l'efficacité peut également poser un problème éthique. D'autre part, un principe important des évaluations est le consentement éclairé des participants : les indi-

... Comment randomiser ?

vidus doivent avoir le choix d'accepter ou non de participer à l'expérimentation. La plupart des protocoles expérimentaux des évaluations sont d'ailleurs avalisés par un comité d'éthique².

En fin de compte, il semble qu'un grand nombre de programmes qui se mettent en œuvre peuvent être évalués avec la méthode d'évaluation aléatoire. La diversité des programmes qui ont été évalués ces dernières années en atteste.

3.2. Conditions nécessaires et analyse

Si le programme a pu être assigné aléatoirement, le biais de sélection est égal à zéro par construction. Cependant il faut disposer d'un minimum d'unités (individus ou groupes d'individus) tirées au sort pour que le dispositif ait une puissance statistique suffisante, c'est-à-dire pour détecter un effet du programme évalué si cet effet existe. La puissance statistique dépend d'une série de paramètres dont les plus importants sont la taille de l'échantillon, le degré de variabilité des variables d'impact mesurées et la différence de participation dans les groupes traités et contrôles. Il est également nécessaire de pouvoir recueillir des informations sur les populations des groupes traités et contrôles. Les évaluations reposent sur des systèmes de collecte de données souvent exigeants, avec des données administratives ou des données d'enquêtes, où il est nécessaire de suivre l'ensemble de la population des deux groupes dans le temps.

L'analyse des données expérimentales est relativement simple dans le cas d'un scénario standard où tous les individus assignés au traitement participent et ceux au contrôle ne participent pas (conformité parfaite). Il s'agit en fait de comparer les moyennes des variables d'intérêt dans les groupes traitement et contrôles, la différence étant due à la participation au programme. En pratique, l'analyse peut s'avérer plus complexe lorsque la conformité n'est pas parfaite ou lorsque les conditions de l'expérimentation s'écartent du dispositif initial. En effet, il est possible qu'une partie de l'échantillon ne soit plus observable dans les données (on parle d'attrition).

Il peut aussi être intéressant d'évaluer l'impact du programme sur des sous-groupes de la population tirée au sort car les effets peuvent être hétérogènes (par exemple, il s'agirait d'évaluer, pour une politique de l'emploi, l'effet sur les demandeurs d'emploi les plus éloignés ou les moins éloignés du marché du travail). Cela suppose cependant que les sous-groupes soient suffisamment larges pour avoir le niveau de puissance statistique nécessaire.

Enfin, de nombreux travaux cherchent à mesurer l'impact de différents traitements (et éventuellement leur interaction) ou différentes intensités de traitement afin de tester différentes hypothèses. Par exemple, Cohen et Dupas (2010) évaluent l'effet du prix sur l'adoption des moustiquaires imprégnées par les femmes enceintes au Kenya. Pour ce faire, elles randomisent le niveau des prix des moustiquaires fournies dans vingt cliniques et montrent que faire payer les moustiquaires, à des niveaux de prix même très faibles, a des effets négatifs très importants sur la demande. L'effet négatif du prix sur la demande de biens de santé préventif a été identifié dans plusieurs autres évaluations aléatoires sur d'autres types de produits (chlorine, vaccination, déparasitage, etc.) (J-PAL, 2011³). Ces résultats ont des implications importantes en termes de santé publique, notamment lorsque les comportements individuels peuvent affecter la situation de l'ensemble de la communauté (présence d'externalités).

² La plupart des universités américaines disposent de comités d'éthique appelés «Institutional Review Board» (IRB) auxquels sont soumis les protocoles de recherche impliquant des êtres humains. L'IRB vérifie notamment que l'expérimentation minimise les risques pour les participants et que ceux-ci sont correctement informés.

³ J-PAL Bulletin, April 2011 (<http://www.povertyactionlab.org/publication/the-price-is-wrong>).

4. Focus sur les politiques d'emploi et de lutte contre la pauvreté

Un nombre très important de politiques et d'interventions ont été ou sont évaluées avec la méthode expérimentale. Ces évaluations se focalisent sur les politiques de l'emploi et d'éducation dans les pays occidentaux et sur un grand nombre d'interventions dans les pays en développement sur les thèmes de la santé, éducation, gouvernance, accès aux services financiers, entrepreneuriat etc. On se focalisera ici sur quelques exemples d'évaluations de politiques actives sur le marché du travail et de politiques de lutte contre la pauvreté et d'éducation dans les pays en développement.

4.1. Evaluation des politiques actives sur le marché du travail

Comme mentionné en introduction, les politiques de l'emploi dans les pays occidentaux et particulièrement en Europe mobilisent des ressources importantes. Les politiques actives sur le marché du travail ont pour objectif d'améliorer le retour à l'emploi des populations au chômage. Ces questions ont été à l'origine d'une littérature théorique et empirique abondante en économie du travail et d'un nombre important d'évaluations aléatoires aux Etats-Unis, au Royaume-Uni, en Europe du Nord et plus récemment en France.

Il existe une série de mesures – reflétant différentes hypothèses sur les barrières au retour à l'emploi – telles que le renforcement du contrôle et de l'application des sanctions, l'accompagnement plus ou moins intensif des demandeurs d'emploi, la formation, des subventions à l'embauche, etc.

La plupart des évaluations aléatoires étudient l'impact joint de plusieurs de ces interventions, car les programmes sont souvent développés de cette manière. Une des premières évaluations aléatoires sur l'accompagnement à la recherche d'emploi a été menée dans plusieurs Etats aux Etats-Unis dans les années 1980 (Meyer, 1995). L'accompagnement consistait en plusieurs interventions conjointes : donner une meilleure information sur les offres d'emploi, fournir un meilleur accompagnement ainsi qu'un contrôle renforcé des efforts dans la recherche d'emploi. Ces programmes ont eu un impact positif sur la baisse du recours aux allocations de chômage et sur les revenus. En Europe, plusieurs travaux évaluent également l'effet joint de l'accompagnement et du contrôle. Au Royaume-Uni, Dolton et O'Neill (2002) évaluent les effets à long terme du programme Restart qui combine un contrôle accru de l'activité de recherche des demandeurs d'emploi et un meilleur accompagnement. L'effet à long terme du programme est positif pour les hommes avec une réduction du taux de chômage de 6 points de pourcentage comparé au groupe contrôle. Il est cependant nul pour les femmes (car le programme aurait, en partie, eu pour effet de diminuer leur recours à l'assurance chômage à cause du contrôle accru). Aux Pays-Bas, Van den Berg et Van der Klaauw (2006) évaluent par la méthode expérimentale l'effet d'un programme d'accompagnement et de contrôle des demandeurs d'emploi qui reçoivent des allocations de chômage. Ils montrent que l'effet sur le retour à l'emploi est faible, la cause principale étant la faible intensité du programme.

La façon dont l'ensemble de ces programmes étaient proposés – avec plusieurs interventions jointes – ne permet pas de déterminer l'impact respectif de chacune de ces interventions. Ashenfelter *et al.* (1999) s'intéressent quant à eux à l'effet spécifique du contrôle renforcé des demandeurs d'emploi. Ils montrent, dans une expérimentation conduite dans quatre Etats aux Etats-Unis, qu'une plus forte vérification de l'effort de recherche d'emploi et une application plus stricte des sanctions n'ont pas d'effet sur le recours aux allocations de chômage.

Au Danemark, Graversen et Van Ours (2008) évaluent l'effet d'un programme d'activation obligatoire qui combine des composantes de formations, d'accompagnement et de contrôle pour retrouver un emploi. L'effet global du programme est très positif avec une augmentation du taux de retour à l'emploi de 30 % par rapport au groupe contrôle. L'analyse des données suggère que l'impact vient plus de l'accompagnement à la recherche d'emploi que des formations.

... Evaluation des politiques actives sur le marché du travail

Plus récemment, plusieurs évaluations de grandes envergures ont été menées en France, en collaboration avec Pôle Emploi (service public de l'emploi en France). Deux politiques majeures sur l'accompagnement renforcé des demandeurs d'emploi ont été évaluées sur deux segments relativement différents : les demandeurs d'emploi présentant un risque de chômage de longue durée (supérieure à un an) et les demandeurs d'emploi jeunes diplômés. L'encadré 2 présente la mise en œuvre et les résultats de l'expérimentation sur les demandeurs d'emploi de longue durée.

ENCADRÉ 2

**L'expérimentation
OPP/CVE**

Cette expérimentation mise en œuvre par l'ANPE et l'Unédic (qui ont fusionné pour devenir Pôle Emploi en 2009) visait à évaluer l'efficacité d'un accompagnement renforcé des demandeurs d'emploi ayant un risque de chômage de longue durée. Une caractéristique intéressante est que ce programme a été mis en œuvre par deux organisations : l'une publique, l'ANPE (futur Pôle Emploi) et l'autre privée, les Opérateurs Privés de Placement (OPP) auxquels l'ANPE a sous-traité l'accompagnement d'une partie des demandeurs d'emploi. Le programme visait à concentrer plus d'efforts sur les demandeurs d'emploi qui en avaient le plus besoin en faisant passer le nombre de demandeurs d'emploi par conseiller d'environ 120 à 40 (Behaghel *et al.*, 2013)⁴. Les hypothèses principales justifiant ce programme sont que les demandeurs d'emploi n'ont pas suffisamment d'informations sur les emplois disponibles et que l'encadrement des conseillers pour les aider dans leur recherche n'est pas assez intensif.

Le dispositif expérimental d'évaluation a consisté à construire trois groupes, sélectionnés aléatoirement parmi la population éligible au programme : un groupe de demandeurs d'emploi recevant l'accompagnement renforcé par l'ANPE, un groupe de demandeurs d'emploi recevant l'accompagnement renforcé par OPP et un dernier groupe (contrôle) recevant le parcours classique d'accompagnement (de l'ANPE). Près de 44.000 demandeurs d'emploi sont entrés dans le dispositif expérimental⁵. La sélection aléatoire a été effectuée par un système informatique qui indiquait aux conseillers à quel groupe appartenait le demandeur d'emploi qu'ils suivaient.

En résumé, le dispositif permettait d'étudier l'impact des deux programmes d'accompagnement renforcé (public et privé) et aussi de comparer leur efficacité lorsqu'ils étaient déployés dans les mêmes endroits auprès d'un public de demandeur d'emploi similaire⁶. Grâce aux données administratives les évaluateurs ont pu ensuite suivre mois par mois l'évolution des taux de retour à l'emploi (et une série d'autres indicateurs) des populations de l'échantillon. L'impact du programme est obtenu en comparant le taux d'emploi dans le groupe traité (accompagnement public ou privé) avec celui dans le groupe contrôle (accompagnement classique). Les résultats montrent que les deux dispositifs améliorent significativement le retour à l'emploi. En outre, dans les endroits où les deux programmes peuvent être comparés, l'effet est plus fort pour le programme public que privé. En effet le programme public augmente le taux d'emploi de 10,2 points de pourcentage six mois après l'assignation aléatoire alors que OPP l'augmente seulement de 4,2, la référence du groupe contrôle étant un taux de sortie d'environ 20 pourcents (Behaghel *et al.*, 2013). Selon les auteurs, plusieurs hypothèses peuvent expliquer cette différence de résultats : les conseillers de l'accompagnement public auraient une plus grande expérience, la structure publique aurait un meilleur accès aux offres d'emploi ou encore les contrats avec les partenaires privés (avec une part fixe importante payée pour chaque demandeur d'emploi suivi quel que soit le résultat) ne les inciteraient pas à fournir un effort maximal.

⁴ Au FOREM, le service public de l'emploi de la Région wallonne (Belgique), ce ratio était d'environ un conseiller pour 245 demandeurs d'emploi en 2012 (Maystadt *et al.*, 2013, p. 297).

⁵ 4.565 ont été assignés au parcours classique (contrôle), 3.385 à l'accompagnement renforcé public et 36.027 à l'accompagnement renforcé privé (Behaghel *et al.*, 2014).

⁶ Les deux programmes n'ont pas été développés conjointement dans toutes les régions et n'ont pas toujours le même public cible. Dans ces cas comparer leur efficacité n'a pas de sens.

4.2. Evaluation des politiques de lutte contre la pauvreté dans les pays en développement

Dans un autre domaine, un grand nombre d'interventions sont mises en œuvre dans les pays en développement pour lutter contre la pauvreté. Il existe tout une série d'hypothèses pouvant expliquer que la pauvreté perdure. Existe-t-il une trappe à la pauvreté liée à un déficit nutritionnel ? Les plus pauvres n'arrivent pas à sortir de la trappe à la pauvreté car ils n'ont pas accès au capital qui leur permettrait d'investir dans des activités profitables ? Manquent-ils de capital humain pour être capable de saisir ces opportunités ? De nombreuses interventions cherchent à surmonter ces contraintes avec pour objectif de réduire la pauvreté des populations touchées. Certaines de ces interventions ont été évaluées avec la méthode expérimentale. Un projet de développement emblématique évalué avec la méthode expérimentale est le programme Progesa au Mexique. Progesa visait à lutter contre la pauvreté au moyen d'un programme de transfert d'allocations aux ménages pauvres conditionnellement à la participation des enfants à l'école et à la visite des mères et des enfants à des centres de santé. La motivation principale de ce type de programme vient du fait que pour les ménages les plus pauvres l'investissement en capital humain représente des coûts directs ou d'opportunité trop importants (surtout en l'absence d'accès au crédit). La phase pilote du programme a été évaluée par l'IFPRI. Le programme a été déployé tout d'abord dans une centaine de communes rurales au Mexique tirées au sort parmi plus de 200 qui étaient éligibles pour participer. L'impact était donc mesuré en comparant les

ENCADRÉ 3

Evaluation d'une intervention en éducation au Kenya

En économie de l'éducation dans les pays en développement, plusieurs travaux évaluent l'impact d'interventions visant à améliorer le fonctionnement des écoles et des classes. En effet plusieurs paramètres rentrent dans la «fonction de production» scolaire, notamment certains inputs tels que les manuels scolaires ou le type d'enseignants (contractuels ou fonctionnaires), l'organisation des classes – leur taille, l'existence de classes de niveaux – ou encore la redevabilité des écoles auprès de la communauté. Ces questions sont très importantes à l'heure où l'éducation est devenue gratuite dans un grand nombre de pays en développement et où le nombre d'enfants inscrits a beaucoup plus augmenté que les moyens disponibles. Il en résulte que la taille des classes a augmenté et que les niveaux au sein des classes sont très hétérogènes.

Pour tenter de répondre à ces questions, Duflo *et al.* (2011) évaluent un programme dans 210 écoles au Kenya consistant à augmenter le nombre d'enseignants en ayant recours à des enseignants contractuels, à renforcer le contrôle des parents et à mettre en œuvre des classes de niveaux. Le dispositif expérimental développé par les chercheurs permet de répondre à ces trois questions.

Sur les 210 écoles, 140 sélectionnées aléatoirement ont reçu un enseignant supplémentaire alors que les 70 restantes constituaient le groupe de contrôle. Parmi les 140 écoles sélectionnées, pour la moitié d'entre elles, sélectionnées à nouveau aléatoirement, la composition des classes était déterminée au hasard alors que pour l'autre moitié elles étaient composées par niveau (avec des classes de niveau faible et des classes de niveau élevé). Enfin pour la moitié des écoles du programme (incluant des écoles avec et sans classe de niveau), sélectionnées à nouveau de manière aléatoire, les parents des élèves ont reçu une formation visant à renforcer leur implication dans l'école. Ils avaient aussi la possibilité de contrôler plus étroitement la performance des enseignants.

L'apprentissage des élèves est la variable centrale de l'évaluation. Il est mesuré à l'aide de tests standardisés en mathématique et en langue.

Les résultats montrent que le recours à des enseignants contractuels a un effet positif sur l'apprentissage des élèves. Cet effet est plus important et persiste plus longtemps dans les écoles où les parents ont reçu une formation et dans les écoles où il existe des classes de niveaux (indépendamment du niveau initial des étudiants). Ces résultats montrent que la façon de mettre en œuvre la politique de recrutement d'enseignants contractuels supplémentaires est importante.

... *Evaluation des politiques de lutte contre la pauvreté dans les pays en développement*

ménages éligibles des communes assignées au programme avec ceux des communes qui n'étaient pas assignées au programme. Les résultats positifs (Skoufias, 2005) de l'évaluation (une nette augmentation du niveau d'éducation et de santé des enfants) ainsi que la rigueur de la méthode ont convaincu les responsables politiques de l'utilité de ce programme qui a été généralisé dans le pays. Le programme a été répliqué et évalué dans un grand nombre de pays en développement depuis.

Cette expérience novatrice (qui constitue une des premières expériences dans un pays en développement) a été suivie par de nombreuses évaluations (l'une d'entre elles est détaillée dans l'encadré 3), notamment sous l'impulsion du J-PAL, depuis les années 2000.

5. Limites méthodologiques et réponses

L'avantage principal des méthodes d'évaluations aléatoires est de résoudre le problème de biais de sélection grâce à la construction par assignation aléatoire des groupes traitement et contrôle. Les évaluations aléatoires peuvent comporter certaines limites dont certaines sont d'ailleurs communes à l'ensemble des évaluations. Ces limites sont de deux ordres (cette section s'intéressera surtout au deuxième type de problème).

Une première limite a trait aux conditions expérimentales elles-mêmes qui peuvent s'écarter du dispositif initial. En effet, il est possible que la composition des groupes traitement et contrôle changent au cours de l'évaluation. Cela peut provenir de l'attrition (perte d'individus) de l'échantillon enquêté au départ (si l'attrition est différentielle entre les deux groupes, elle peut être source de biais) ou de la mobilité des individus entre les groupes. Ces problèmes peuvent être en général limités par un contact régulier avec les individus enquêtés et par une supervision étroite de la mise en œuvre du dispositif expérimental. Des corrections d'ordre statistiques sont également possibles pour prendre en compte l'attrition différentielle. La mobilité entre groupes traités et contrôles a un effet limité si la différence de proportion de participants dans les groupes traités et contrôle initiaux reste importante. Les conditions expérimentales en elles-mêmes peuvent également être source de biais lorsque les participants ont conscience de faire partie d'un dispositif d'évaluation. En effet, cela peut avoir un effet sur leur comportement. En général une gestion appropriée de la façon dont l'information est diffusée peut suffire pour réduire ce genre de problèmes⁷. Enfin, la validité de l'effet mesuré peut être remise en cause si le groupe contrôle est lui aussi affecté par la mise en œuvre du programme auprès du groupe traitement. Cette situation peut subvenir lorsque les individus des deux groupes interagissent ou en présence d'externalités d'un groupe à l'autre.

Un deuxième type de limite, plus complexe, concerne la généralisation des résultats obtenus avec une évaluation aléatoire. Il existe plusieurs problèmes de généralisation appelant à des réponses différentes.

En premier lieu, les programmes sont évalués dans un contexte donné, avec une organisation, une population et un programme spécifique. Les résultats sont donc applicables à ce contexte. L'impact d'un programme scolaire en Inde n'est pas généralisable au Kenya, ou un programme en milieu rural en Inde n'aura pas le même effet qu'en milieu urbain, etc. Les critères de sélection des populations faisant partie de l'évaluation peuvent aussi poser un problème de représentativité. En effet, un biais de population peut exister si la randomisation est effectuée sur une

⁷ A nouveau, dans le cas des essais cliniques, ce problème est résolu par la méthodologie du double aveugle qui n'est bien entendu pas transposable à l'évaluation en sciences sociales.

... Limites méthodologiques et réponses

population spécifique sur laquelle il est par exemple plus facile de randomiser. Ce biais peut aussi provenir du type d'organisations elles-mêmes, les organisations ou interventions qui acceptent d'être évaluées et de se soumettre au protocole expérimental sont susceptibles d'être spécifiques.

Même si ce type de critique s'applique à tout type d'évaluation empirique, elle n'en demeure pas moins valide. Une réponse au problème de généralisation consiste à évaluer une politique similaire dans des contextes différents. Un exemple intéressant est une évaluation d'un programme de lutte contre la pauvreté appelé «Targeting the Ultra-Poor» (Banerjee *et al.*, 2015) mis en œuvre presque simultanément dans six pays en développement (Ethiopie, Inde, Pakistan, Pérou, Honduras, Ghana) par des organisations différentes. Le cœur de l'intervention est le transfert d'un actif productif à des ménages identifiés comme très pauvres en combinaison avec des formations, un transfert monétaire limité et un appui en santé. La motivation principale du programme vient de l'idée que les plus pauvres n'ont pas accès au capital, et s'ils y avaient accès ne seraient pas dans les conditions de pouvoir rembourser un crédit avec un taux d'intérêt, et sont sujets à des chocs importants pouvant affecter leur niveau de revenu. Le programme a été évalué dans l'ensemble des pays avec la méthode expérimentale. Les résultats montrent globalement que le programme a un effet positif et durable sur les conditions de vie dans cinq pays sur six et l'ampleur des effets est relativement stable en fonction des contextes.

En deuxième lieu, certaines évaluations sont effectuées sur des interventions qui se déploient à un niveau restreint qui correspond souvent à la phase pilote du programme. Les résultats d'évaluation d'un programme pilote ne sont pas forcément généralisables au cas où la politique serait appliquée à l'ensemble de la population éligible (par exemple un pays). On peut s'attendre à deux types d'effets. Premièrement des effets d'éviction peuvent survenir si le groupe traité améliore sa situation au détriment du groupe contrôle. Par exemple, une politique formation ou d'accompagnement des demandeurs d'emploi pourrait améliorer le taux de retour à l'emploi des individus traités mais pourrait avoir un effet négatif sur les individus non traités qui auraient moins de chance d'accéder aux emplois vacants surtout si ceux-ci sont constants dans le bassin d'emploi considéré. Une évaluation d'impact ne prenant pas en compte ce problème surestimerait l'effet du programme. Deuxièmement, une politique mise en œuvre à l'ensemble d'une économie peut avoir des effets d'équilibre. Par exemple, une politique qui viserait à augmenter le niveau d'éducation d'une part importante de la population pourrait avoir des répercussions sur le niveau général des qualifications et avoir des effets sur l'emploi et les salaires. Ces effets sont en général assez difficiles à identifier (Hsieh et Urquiola, 2006; Bettinger *et al.*, 2010). Il est donc vraisemblable que ces effets d'évictions et ces effets d'équilibre ne seront pas correctement mesurés au niveau du programme pilote.

Une réponse à ce type de problème serait d'évaluer un programme à une échelle suffisamment large pour capter ces effets d'équilibre mais cela est difficilement faisable étant donné les coûts qu'une telle évaluation représenterait. Une autre solution consiste à développer un dispositif expérimental qui permette de mesurer une partie de ces effets. Crépon *et al.* (2013) cherchent à identifier ces effets d'éviction lorsqu'une politique d'accompagnement des demandeurs d'emploi est mise en œuvre. La politique est similaire à celle décrite en section 4.1 (voir encadré 2) mais s'adresse spécifiquement aux jeunes diplômés. Pour mesurer ces effets d'éviction, le dispositif d'évaluation variait le niveau d'intensité du programme au sein des agences locales de l'emploi (ALE), le niveau d'intensité allant de 100 % des jeunes traités à 0 % de jeunes traités (groupe contrôle pur). Il s'agissait donc de randomiser à la fois le niveau d'intensité entre agences et les individus traités et contrôles au sein des agences. Les résultats montrent que les différences significatives de retour à l'emploi entre individus traités et contrôles constatées au sein

... *Limites méthodologiques et réponses*

des agences sont largement dues aux effets d'éviction. En effet, en comparant la situation des individus contrôles dans les agences où une proportion importante de jeunes est traitée avec celle dans les agences où aucun individu n'est traité, on observe que la situation des contrôles dans le premier cas est significativement dégradée. Ces résultats prouvent l'existence d'effets de déplacements ou d'éviction importants et ont une résonance importante en termes de politiques publiques. En période de faible croissance économique où le nombre d'emplois disponibles est constant, une politique active de l'emploi peut avoir un effet très limité étant donné les effets d'éviction.

6. Conclusion

La méthode d'évaluation aléatoire a un avantage important par rapport aux autres méthodes d'évaluation car elle permet de supprimer le biais de sélection par construction. L'effet de l'intervention évaluée est mesuré sans biais si le protocole expérimental est mis en œuvre correctement. Pour cela, l'évaluation doit être intégrée dans le processus de développement du programme (qui doit être assigné aléatoirement) et donc requiert une collaboration importante entre chercheurs et responsables du programme. Les échantillons doivent être suffisamment importants pour être capables de détecter des effets du dispositif si ceux-ci existent (comme pour toute évaluation) et le protocole expérimental doit être conservé pendant la durée de l'évaluation. Enfin, le design expérimental doit être capable de prendre en compte (et éventuellement mesurer) les effets d'externalités potentiels qui pourraient affecter la validité de la mesure d'impact. Les nombreuses évaluations mentionnées dans cet article montrent que ce type de méthode peut s'appliquer dans différents contextes et à différents types de politique.

La multiplication des évaluations aléatoires contribue à l'amélioration de la connaissance sur l'impact des politiques publiques. En effet, pour certaines interventions il existe aujourd'hui une taille critique de preuves empiriques rigoureuses. Par exemple, un grand nombre d'expérimentations se sont intéressées à l'effet du prix sur l'adoption des biens de santé préventive (vaccins, moustiquaires, traitements pour rendre l'eau potable, etc.) dans les pays en développement et montrent qu'il est préférable en termes de santé publique de subventionner totalement certains de ces biens plutôt que de faire payer un coût même très faible (Dupas, 2014). On dispose aujourd'hui d'une solide base empirique qui démontre que les transferts d'allocations conditionnels ont des effets positifs, le programme Progresa ayant été répliqué (et dans certains cas évalué) dans plus de 30 pays. En termes de politiques publiques, la mesure des effets d'un programme est d'autant plus importante qu'elle peut être comparée à son coût ou aux bénéfices d'une utilisation alternative des fonds.

Enfin, au-delà de l'intérêt direct en termes de politiques publiques, les évaluations aléatoires peuvent contribuer à la connaissance scientifique en économie. Les résultats des travaux basés sur des évaluations randomisées permettent de mieux comprendre certains comportements des individus. Les travaux sur les biens de santé préventive et sur l'adoption de certaines technologies montrent par exemple l'importance des préférences temporelles (les individus valorisent peu les bénéfices futurs potentiels associés à l'adoption de ces biens ou technologies en comparaison avec leur coût immédiat). Par ailleurs, de plus en plus d'évaluations aléatoires ont pour objectif de tester explicitement des paramètres provenant de modèles théoriques à l'aide de protocoles expérimentaux originaux ou «smart experiment».

William Parienté est professeur d'économie à l'UCL, chercheur à l'IRES (UCL) et membre de J-PAL.

William Parienté
william.pariante@uclouvain.be

Références

- Arceneaux K., A. Gerber, & D. Green (2006), "Comparing Experimental and Matching Methods using a Large-Scale Field Experiment on Voter Mobilization", *Political Analysis*, 14(1): 37-62.
- Ashenfelter O., D. Ashmore & O. Deschenes (1999), "Do Unemployment Insurance Recipients Actively Seek Work? Randomized Trials in Four U.S. States", NBER Working Papers 6982, National Bureau of Economic Research, Inc.
- Avvisati F., M. Gurgand, N. Guyon & E. Maurin (2014), "Getting Parents Involved: a Field Experiment in Deprived Schools", *Review of Economic Studies*, 81(1): 57-83.
- Banerjee A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert & C. Udry (2015), "A multifaceted program causes lasting progress for the very poor: Evidence from six countries", *Science*, 348(6236): 1-16.
- Behaghel L., B. Crépon, M. Gurgand, L. Lequien, R. Rathelot & P. Zamora (2013), "L'accompagnement personnalisé des demandeurs d'emploi : les enseignements de trois expériences contrôlées menées en France", *Revue Française d'Economie*, 28(1): 123-158.
- Behaghel L., B. Crépon & M. Gurgand (2014), "Private and Public Provision of Counseling to Job-Seekers: Evidence from a Large Controlled Experiment", *American Economic Journal: Applied Economics*, 6(4): 142-74.
- Bettinger E., M. Kremer, and J. Saavedra (2010), "How Do Vouchers Work? Evidence from Colombia", *Economic Journal*, 546(08): F204-F228.
- Card D., J. Kluve, & A. Weber (2010), "Active Labour Market Policy Evaluations: A Meta-Analysis", *Economic Journal*, 120(548): F452-F477.
- Cohen J. & P. Dupas (2010), "Free Distribution or Cost Sharing? Evidence from a Malaria Prevention Experiment in Kenya", *Quarterly Journal of Economics*, 125(1): 1-45.
- Crépon B., E. Duflo, M. Gurgand, R. Rathelot & P. Zamora (2013), "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment", *Quarterly Journal of Economics*, 128(2): 531-580.
- DARES (2014), "Les dépenses en faveur de l'emploi et du marché du travail" (<http://travail-emploi.gouv.fr/>).
- Dolton P. & D. O'Neill (2002), "The Long-Run Effects of Unemployment Monitoring and Work-Search Programs: Some Experimental Evidence from the U.K.", *Journal of Labor Economics*, 20(2): 381-403.
- Duflo E., P. Dupas & M. Kremer (2011), "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya", *American Economic Review*, 101(5): 1739-1774.
- Dupas P. (2014), "Getting Essential Health Products to Their End Users: Subsidize, but How Much?", *Science*, 345(6202): 1279-1281.
- Fonds d'Expérimentation pour la Jeunesse (2013), «Rapport d'Activité pour 2013», Ministère de la Jeunesse, de la Ville et des Sports.
- Glazerman S., D. M. Levy, & D. Myers (2003), "Nonexperimental versus Experimental Estimates of Earnings Impacts", *Annals of the American Academy of Political and Social Science*, 589(1): 63-93.
- Glennerster R. & K. Takavarasha (2013), *Running Randomized Evaluations: A practical Guide*, Princeton University Press.
- Graversen Brian K. & Jan C. Van Ours (2008), "How to help unemployed find jobs quickly: Experimental evidence from a mandatory activation program", *Journal of Public Economics*, 92(10-11): 2020-2035.
- Graversen J., R. Lalonde & J. Smith (1999), "The economics and econometrics of active labor market programs", in: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, edition 1, volume 3, chapter 31, pp. 1865-2097, Elsevier.
- Heckman J. J., Hyeok Moon S., Pinto R., Savellyev P. A. & Yavitz A. (2010), "The Rate of Return to the High/Scope Perry Preschool Program", *Journal of Public Economics*, 94(1-2): 114-128.

- Hsieh C. & M. Urquiola. (2006), "The effects of generalized school choice on achievement and stratification: evidence from Chile's voucher program", *Journal of Public Economics*, 90(8-9): 1477-503.
- Katz L.F., Kling J.R. & Liebman J.B. (2001), "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment", *Quarterly Journal of Economics*, 116(2): 607-654.
- Katz L.F., Kling J.R. & Liebman J.B. (2003), "Boston Site Findings: The Early Impacts of Moving to Opportunity", in Goering J. & Feins J., *Choosing A Better Life: Evaluating the Moving to Opportunity Social Experiment*, Urban Institute Press, pp. 177-211.
- Katz L.F., Kling J.R. & Liebman J.B. (2007), "Experimental Analysis of Neighborhood Effects" *Econometrica*, 75(1): 83-119.
- Lalonde R. (1986), "Evaluating the Econometric Evaluations of Training Programs", *American Economic Review*, 76(4): 604-620.
- Levitt Steven D. & List John A. (2009), "Field experiments in economics: The past, the present, and the future", *European Economic Review*, 53(1): 1-18.
- List John A. & Rasul Imran (2011), "Field Experiments in Labor Economics", Chapter 2 in *Handbook of Labor Economics*, Volume 4a, O. Ashenfelter and D. Card (editors), Elsevier, pp. 104-228
- Maystadt Ph., E. Cantillon, L. Denayer, P. Pestieau, B. Van der Linden & M. Cattelain (2013), *Le modèle social belge : quel avenir ?*, Presses Interuniversitaires de Charleroi.
- Meyer B. (1995), "Lessons from the U.S. Unemployment Insurance Experiments", *Journal of Economic Literature*, 33(1): 91-131.
- OCDE (2015), *Répartition Géographique des Ressources Financières Allouées aux Pays en Développement*.
- OCDE (2015), *Dépenses publiques et nombre de participants PMT (Programmes Marché du Travail)*.
<http://stats.oecd.org/Index.aspx?lang=fr&SubSessionId=30e5d500-9d12-472f-becc-fa04bbd54566&themetreeid=13>
- Ross H.L. (1970), "An Experimental Study of the Negative Income Tax, Child Welfare", Massachusetts Institute of Technology.
- Skoufias E. (2005), «PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico», Research Report n°135, International Food Policy Research Institute, Washington DC.
- Smith J. & Todd P. (2005), "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators", *Journal of Econometrics*, 125(1-2): 305-353.
- Van Den Berg G. & B. Van Der Klaauw (2006), "Counselling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment", *International Economic Review*, 47(3): 895-936.

Directeur de la publication :
Vincent Bodart

Rédactrice en chef :
Muriel Dejemeppe

Comité de rédaction : Paul Belleflamme,
Vincent Bodart, Muriel Dejemeppe, Frédéric
Docquier, Jean Hindriks, Marthe Nyssens,
William Parienté, Mikael Petitjean

Secrétariat & logistique : Anne Davister
Graphiste : Dominos

Regards Économiques IRES-UCL

Place Montesquieu, 3
B1348 Louvain-la-Neuve

www.regards-economiques.be

regard-ires@uclouvain.be

tél. 010/47 34 26

ISSN 2033-3013