

Évaluation d'impact : Revue de littérature théorique

OUAKIL HICHAM¹, LECHHEB HOUDA² & JOUILIL YOUNESS³

¹(Professeur Habilité, Département d'Economie et de Gestion, Faculté des Sciences Juridiques, Economiques et Sociales, Université Ibn Tofail, Kénitra, Maroc.)

²(Professeure Habilitée, Département d'Economie et de Gestion, Faculté des Sciences Juridiques, Economiques et Sociales, Université Ibn Tofail, Kénitra, Maroc.)

³(Doctorant, Département d'Economie et de Gestion, Faculté des Sciences Juridiques, Economiques et Sociales, Université Ibn Tofail, Kénitra, Maroc.)

Correspondance : JOUILIL YOUNESS, doctorant FSJESK, Maroc. y.jouilil@gmail.com

Abstrait :

L'objectif de ce papier est de présenter de façon exhaustive les méthodes d'évaluation d'impact communément utilisées pour examiner l'effet propre d'une politique donnée sur un groupe de bénéficiaires ou de traités. Ainsi, plusieurs questions se posent auxquelles il faut apporter des réponses claires et précises. Qu'entend-on par évaluation d'impact ? Quelles raisons pour mener une évaluation d'impact ? Quelles sont les méthodes fréquemment utilisées pour l'évaluation d'impact d'une politique donnée ? Comment choisir la méthode ou les méthodes les plus adéquates ?

Il s'avère d'après l'investigation menée dans le cadre de papier que, la réussite d'une évaluation d'impact impose l'estimation d'un scénario contrefactuel crédible et rigoureusement défini afin de contrôler, dans la mesure du possible, les facteurs autres que la mesure qui pourraient expliquer les changements observés. Autrement dit, l'évaluation d'impact nécessite de trouver un groupe témoin (de contrôle) pour estimer les résultats qu'auraient connus les traités (les bénéficiaires) par une mesure si ladite mesure n'avait pas existée.

Dans les faits, la construction d'un « bon » contrefactuel permet d'estimer par la suite les effets réels, propres, et directs des interventions puisqu'il fournit des estimations de l'ampleur et de la nature (positive ou négative) des impacts et des mesures statistiques de l'incertitude.

De plus, le choix de la méthode ou de la meilleure méthode à utiliser pour étudier l'effet causal d'une intervention donnée est souvent conditionné par les ressources disponibles et des contraintes existantes ou encore par l'information ou les informations dont on dispose pour y répondre. Parfois aucune n'est applicable !

Mots-clés : Évaluation de l'impact, double différence, appariement des scores de propension, variables instrumentales, régression sur discontinuités.

I. Introduction

L'évaluation d'impact des politiques et des actions publiques est un outil efficace et efficient pour le développement et l'amélioration de la gestion publique (Lechheb H., Ouakil H. & JOUILIL, Y., 2018). D'où l'importance cruciale du choix de ce sujet.

Alors, qu'entend-on par évaluation d'impact des politiques publiques ? Quelles sont les raisons pour mener une analyse d'impact ? Quels sont les problèmes rencontrés lors de la réalisation de l'analyse d'impact ? Quelles sont les différentes méthodes d'évaluation d'impact ? Et comment choisir la méthode ou les méthodes les plus adéquates ? Pour répondre à toutes ces questions, nous allons adopter le plan suivant.

Dans le premier paragraphe, on présentera ce qu'on entend par le mot « évaluation d'impact », les raisons de la conduite d'une évaluation d'impact et la difficulté de sa mise en œuvre. Dans le deuxième paragraphe, on passera en revue sur l'ensemble de techniques économétriques qui permettent, chacune à sa façon, de mesurer l'effet causal d'une politique ou d'un projet sur la population des bénéficiaires. Dans le dernier paragraphe, on présentera les critères sur lesquelles on se base pour choisir la méthode ou les méthodes appropriées pour la l'examen d'une intervention donnée.

II. Qu'entend-on par évaluation d'impact ?

1. Définition

L'évaluation d'impact étudie les changements causés par une intervention, une politique, un projet, ou une mesure sur le groupe témoin et d'analyser et mesurer l'ampleur de cet impact (effet causal). Une telle évaluation permet d'établir un lien de cause à effet entre un projet, une mesure, un traitement ou une politique et

un ou plusieurs résultats au niveau des outcomes. Cela implique de reconstituer un groupe témoin dite « contrefactuelle » ou encore groupe de comparaison (Bonnet et al. 2004).

Autrement dit, l'évaluation d'impact est une évaluation de la manière dont le traitement évalue les résultats. L'analyse appropriée de l'évaluation l'impact nécessite un contrefactuel (un groupe de comparaison ou de contrôle). C'est-à-dire de ce que ces résultats auraient été en l'absence du traitement en question.

Une évaluation d'impact évalue les changements dans le bien-être des personnes, des ménages ou des communautés pouvant être attribués à un programme ou une politique spécifique. La question centrale de l'évaluation de l'impact est de répondre à la question suivante. Quel serait arrivé à ceux qui reçoivent l'intervention s'ils n'avaient pas reçu le programme ? Comme nous ne pouvons pas observer simultanément l'état de ce groupe (avec et sans l'intervention), le principal défi est de développer un contrefactuel, c'est-à-dire un groupe aussi similaire que possible à ceux qui reçoivent l'intervention en question. Cette comparaison permet d'établir une causalité définitive en attribuant au programme les changements observés dans le bien-être, tout en éliminant les facteurs de confusion.

Il apparaît de ce qui précède que l'évaluation d'impact est une évaluation ex post destinée, surtout, à identifier si une mesure donnée est efficace ou non. Elle sert également à quantifier la part du changement observé dans la population étudiée et qui serait attribuée à la mesure évaluée.

Les évaluations d'impact sont un type particulier d'évaluation qui cherche surtout à répondre une question de cause à effet et à savoir quel est l'impact (ou l'effet causal) d'un programme sur un résultat d'intérêt ? Cette question de base intègre une importante dimension (dimension causale) ou l'accent est mis uniquement sur l'impact. C'est-à-dire sur les changements directement attribuables à un programme, à une modalité de programme ou à une conception.

Ainsi, l'objectif primordial de l'évaluation d'impact est de mesurer l'impact ou l'effet causal d'un programme spécifique sur un résultat donné ? En effet, ce type d'évaluation peut être appliqué à de nombreux contextes et domaines. Par exemple, quel est l'effet causal d'un dispositif d'aide à l'emploi sur la durée de chômage des bénéficiaires ? Quel est l'impact d'une bourse scolaire sur les résultats scolaires ? Quel est l'impact de la taille des classes sur les résultats scolaires des élèves ? La généralisation de l'électricité a-t-elle un impact sur la fécondité ? Les cours d'éducation sexuelle ont un impact sur les grossesses chez les adolescentes ? Qu'elle est l'effet causal du sexe de chef de ménage sur la pauvreté du ménage ?

2. *Quelles raisons pour mener une évaluation d'impact ?*

L'évaluation d'impact est un outil efficace de prise de décision puisqu'elle fournit des preuves fiables, rigoureuses, pertinentes et solides et qui ne peuvent pas être ignorées dans les responsables politiques (Lechheb H., Ouakil H. & JOUILIL, Y., 2018). En effet, les praticiens de l'évaluation d'impact permettent d'éclairer les décideurs lorsqu'ils veulent prendre plusieurs types de décisions entre autres l'interruption des interventions inefficaces, l'expansion des programmes éprouvés, l'ajustement des bénéfices, la sélection entre plusieurs modalités de conception de programmes.

3. *Problème de l'évaluation d'impact*

Comme mentionnée ci-dessous, l'évaluation d'impact s'attache à déterminer si les éventuelles améliorations observées dans la population des traitées sont dues par l'intervention entreprise (évaluée).

L'évaluation d'impact se trouve confronter à des problèmes de sélection. Une intervention cible généralement des populations particulières (les chômeurs, les pauvres, les veuves, les enfants de la rue,...) et dont les caractéristiques socio-économiques et sociodémographiques auraient différentes de celle du reste de la société. A ce niveau on distingue entre deux types de problèmes de sélection :

Sélection liée au ciblage des bénéficiaires (*Program placement*) : Comment peut-on identifier les éligibles au programmes et qui remplissent les critères d'éligibilité ? Comment peut-on minimiser, dans la mesure du possible, les erreurs de ciblage (erreurs d'inclusions et d'exclusions) ?

Problème d'auto-sélection (*Self selection*) : Les personnes ont la liberté de participer ou non aux programmes qui leur sont proposés. A titre illustratif, on peut citer l'exemple du microcrédit ou de la micro finance.

III. Méthode d'évaluation d'impact

Ce paragraphe fait un état des lieux et une description détaillée des méthodes d'évaluation d'impact et donne des pistes d'utilisation raisonnée de ces techniques.

Certes, les techniques d'évaluation d'impact sont multiples mais il est important de signaler que leurs hypothèses de base et leurs spécificités conditionnement fortement leurs outputs.

Définie de façon simple, l'évaluation d'impact vise à établir la situation qu'aurait connue la société en l'absence de la mesure évaluée. Cette situation fictive, aussi appelée contrefactuelle, permet, en la comparant à la situation effectivement observée, de déduire une relation de causalité entre l'intervention publique et un indicateur jugé pertinent (la santé, l'habitat, l'emploi, la morbidité, l'éducation, etc.).

1. Les méthodes économétriques dans le cadre des expériences expérimentales.

Dites encore aléatoires ou encore contrôlées. Le défi majeur de l'évaluation est lié au fait que les individus ou les ménages ciblés par une intervention donnée ne sont pas tirés au hasard dans la population. Une méthode de surmonter ce problème est de réaliser une expérience aléatoire contrôlée dont on choisit aléatoirement « tirage au sort » un groupe d'individus qui se verra appliquer une politique, tandis qu'un autre groupe constituera ce qu'on appelle le « contrefactuel ».

Pour cela, pour examiner l'effet causal d'une intervention, on procède comme pour le test d'un médicament ou encore pour le test clinique. On sélectionne au hasard deux groupes, groupe témoin (ne bénéficie pas de du remède) et groupe de traitement (bénéficie du remède). Si, au cours du temps, on constate que leurs trajets (état de santé par exemple) diffèrent, on peut dire que c'est du forcément au programme.

L'avantage majeur de ces méthodes, résulte dans l'annulation du biais de sélection. En effet, la différence simple entre le groupe de contrôle et le groupe des traités permet d'aboutir à un estimateur sans biais. On parle à ce niveau de l'estimateur « naïf ».

Ces types d'expériences sont très transparents, ne requièrent que peu d'hypothèses identifiantes. Toutefois, ils sont peu voire rarement utilisables vu leurs complexités et leurs coûts « importants ». (Angrist, J et al. 2010)

2. Les méthodes économétriques dans le cadre des expériences quasi-expérimentales et non expérimentales

Les techniques économétriques dans le cadre des expériences quasi-expérimentales et non expérimentales sont multiples. On distingue généralement entre les variables instrumentales, la régression autour d'une discontinuité, la double différence et l'appariement.

1. Les variables instrumentales

Une variable instrumentale (*Instrumental Variable* ou *IV*) est une variable qui explique d'être bénéficiaires ou non et qui n'est pas corrélée avec les composantes inobservées des résultats.

La technique des variables instrumentales est une technique quasi expérimentale permettent, en théorique, de s'affranchir des problèmes de sélection. Toutefois, leurs défis majeurs se trouvent dans la difficulté de retrouver ou de disposer d'une ou de plusieurs variables instrumentales qui soient convaincantes, valides et pertinentes. (Pauline Givord, 2014)

2. La régression sur discontinuité

L'idée derrière la régression sur discontinuité (*Regression Discontinuity Design* ou *RDD*) est qu'autour de ces seuils, les unités statistiques vont avoir des chances différentes de bénéficier d'une intervention selon qu'elles se trouvent en dessus ou en dessous d'un seuil.

Figure 1: Régression autour de la discontinuité – Situation de départ-

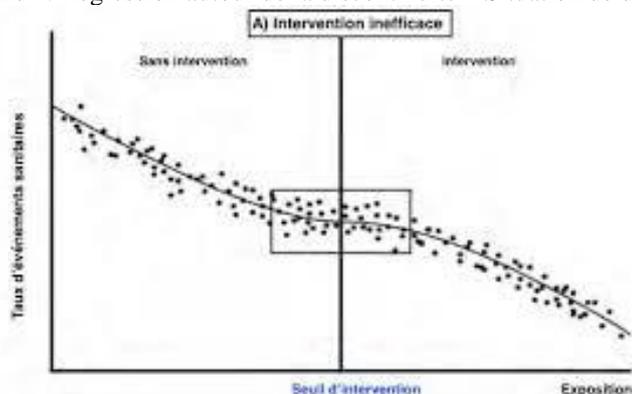
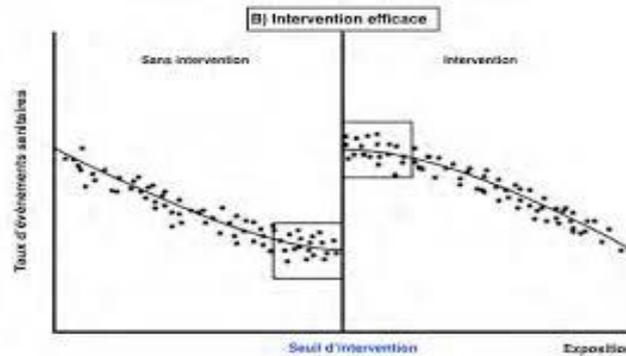


Figure 2: Régression autour de la discontinuité – après intervention-



Les personnes autour de ces seuils (valeurs critiques) auraient des caractéristiques similaires voire identiques. Ce qui implique que la probabilité d'être sélectionnée pour ses personnes serait indépendante des caractéristiques des unités

Elle est utilisée lorsqu'une intervention est survenue au-dessus ou au-dessous d'un seuil spécifique. L'application de cette technique requière la disposition des données sur les participants et les non-participants.

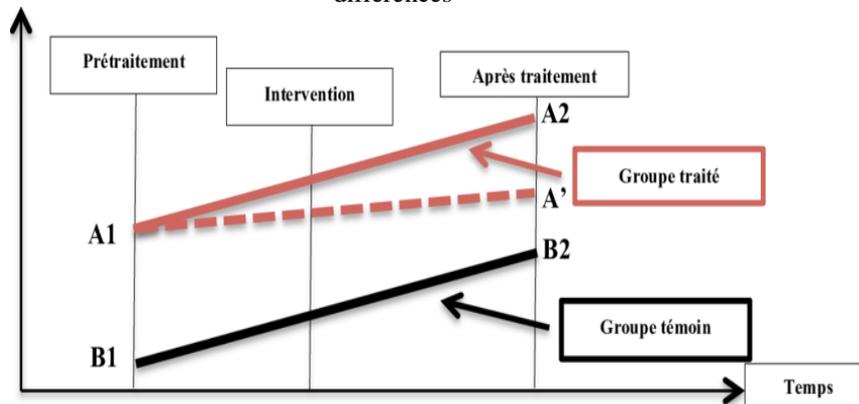
Toutefois, l'application de cette méthode nécessite à ce que la variable dépendante (la variable des résultats) soit continue au point de sélection de participation au programme.

3. La double différence

Comme son nom l'indique, la double différence ou la différence de différences (*Difference in Difference ou DID*) est souvent utilisée pour mesurer les changements dans la quantité de différence entre les groupes de traitement et le contrefactuel au fil du temps. Cela se compare les différences survenues lors de l'intervention pour le groupe de traitement et le groupe de comparaison (contrefactuel).

Autrement dit, elle consiste à quantifier la différence entre la moyenne de résultats entre les traités et les non traités avant et après l'intervention évaluée. D'où son nom. (Donald Rubin, 1974)

Figure 3 : Evolution prévue des résultats du contrefactuel et du traité dans un cas de différence de différences



D'après le graphique ci-dessus, on constate qu'en l'absence de l'intervention, l'évolution du résultat dans le groupe des traités entre les périodes aurait été identique à celle observée dans le groupe témoin (entre B_1 et B_2). Et grâce à la mise en place de l'intervention, le résultat a progressé de A_1 à A_2 . Ce constat est biaisé, car les unités des traités diffèrent des non traités par la distance A_1B_1 . L'impact réel du programme en question correspond à l'écart entre les deux points A_2 et A' .

En fait, la première étape permet à différencier les variables selon la dimension temporelle alors que la deuxième utilise les écarts à la moyenne temporelle. Ce qui permet de neutraliser les effets fixes β_j . Et en utilisant les données panel, on peut estimer α l'effet du traitement sur les bénéficiaires par la méthode des moindres carrés ordinaires ou par des estimateurs à effets fixes (within).

L'application de cette méthode repose sur l'hypothèse identifiante suivante est qu'en absence de l'intervention en question, le résultat potentiel aurait évolué de la même manière pour le contrefactuel que pour les bénéficiaires.

4. L'appariement

L'appariement (*Propensity Score Matching ou PSM*) est une technique non expérimentale commune introduite par le statisticien Rosenbaum et Rubin (1974). Elle est utilisée pour estimer l'effet d'un programme ou d'une intervention lorsque nous n'avons pas d'attribution aléatoire.

Un autre avantage cette méthode empirique est l'élimination des sources de biais en recherchant des groupes traités et non traités, qui présentent des caractéristiques similaires voire identiques.

3. Comment choisir les méthodes d'évaluation d'impact ?

Pour les décideurs publics, l'existence d'un large éventail de méthodes d'évaluation soulève toutefois un certain nombre de questions notamment par rapport le choix de la méthode à utiliser. De façon précise :

Sur la base de quels critères une méthode d'évaluation devrait-elle être choisie ? Sur quels fondements scientifiques reposent les résultats de nombreuses évaluations ? Comment peut la fiabilité et validité empirique être testé de la méthode choisie ?

Il existe énormément de débats sur la question du choix de la/les meilleure(s) méthode(s) à préconiser pour l'évaluation d'impact d'une intervention donnée. Ces discussions reflètent des visions différentes de ce qui constitue une donnée probante crédible, rigoureuse et utile et de qui devrait mener et contrôler les évaluations.

Il y a de nombreux méthodologies et approches pour évaluer l'impact d'une mesure ou d'un programme. Le choix d'une méthode se fait notamment en fonction des objectifs, des moyens, contraintes, et de la question évaluative. Chaque démarche nécessitera des adaptations. Avoir cadré sa démarche au préalable est incontournable pour choisir et construire sa méthode.

Stern et al. (2012) stipule que les trois questions suivantes permettent de définir quel est le type d'outil qui conviendra le mieux, en tenant compte du contexte et des besoins spécifiques: Quelle est la finalité de l'évaluation? Qui fera l'évaluation ? Quelles sont les ressources à disposition ?

Toutes les méthodes d'évaluation d'impact portent sur une forme de question ou d'effet. L'approche de la causalité détermine la méthodologie à utiliser. Toutes ces méthodes doivent estimer ce qu'on appelle le contrefactuel. C'est-à-dire ce qui se serait passé si l'intervention n'avait pas eu lieu. En pratique, l'évaluation d'impact impose aux évaluateurs de trouver un groupe de comparaison (de contrôle, témoin, de référence) pour estimer ce qui serait arrivé aux participants au programme sans le programme, puis faire des comparaisons avec le groupe de traitement qui a reçu le programme en question. Il faut trouver des individus n'ont pas été traités par le programme et qui sont comparables et semblables en tout aux individus traités.

4. Conclusion

La réussite d'une évaluation d'impact impose l'estimation d'un scénario contrefactuel crédible et rigoureusement défini afin de contrôler, dans la mesure du possible, les facteurs autres que la mesure qui pourraient expliquer les changements observés. Autrement dit, l'évaluation d'impact nécessite de trouver un groupe témoin (de contrôle) pour estimer les résultats qu'auraient connus les traités (les bénéficiaires) par une mesure si ladite mesure n'avait pas existée.

Dans les faits, la construction d'un «bon» contrefactuel permet d'estimer par la suite les effets réels, propres, et directs des interventions. Il fournit des estimations de l'ampleur et de la nature (positive ou négative) des impacts et des mesures statistiques de l'incertitude. Il contribue également à la confirmation ou à exclusion du rapport causal présumé entre l'intervention et les résultats.

Finalement, le choix de la méthode ou de la meilleure méthode à utiliser pour étudier l'effet causal d'une intervention donnée est souvent conditionné par les ressources disponibles et des contraintes existantes ou encore par l'information ou les informations dont on dispose pour y répondre. Parfois aucune n'est applicable !

Références

- [1] Angrist, Joshua D., et Jörn-Steffen Pischke. (2010). "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives*, 24 (2) : 3-30.
- [2] Ashenfelter O. (1978). "Estimating the Effects of Training Programs on Earnings", *Review of Economic Studies*, vol. 60, n° 1, pp. 47-57.
- [3] Bonnet, B., M. Banzhaf, P.N. Giraud et M. Issa. (2004). *Analyse des impacts économiques, sociaux et environnementaux des projets d'hydraulique pastorale financés par l'AFD au Tchad*, AFD, Paris.

- [4] Hahn, J., Todd, P. et Van Der Klaauw W. . (1999). "Evaluating the Effect of an Antidiscrimination Law Using a Regression-Discontinuity Design." NBER cahier de recherche 7131.
- [5] Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. (2016). *Impact Evaluation in Practice*, second edition.
- [6] Lechheb, H., Ouakil, H. et Jouilil, Y. (2018). Evaluation des Politiques Publiques : Outil Efficace au Service du Développement. *Revue du Contrôle de la Comptabilité et de l'Audit*. n° 6 : pp. 905-908. URL: <https://revues.imist.ma/index.php?journal=RCCA&page=article&op=view&path%5B%5D=13727&path%5B%5D=7652>.
- [7] Morgan, S. L. and C. Winship. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. . New York: Cambridge University Press.
- [8] Pauline Givord. (2014). *Méthodes économétriques pour l'évaluation de politiques publiques*. *Economie & prévision*, 1-28.
- [9] Rubin, D. (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 56, 688-701.
- [10] Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. et B. Befani. (2012). "Broadening the range of designs and methods for impact evaluations. Report of a study commissioned by the Department for International Development", DFID Working Paper 38, April 2012, London.
- [11] Wooldridge, J. M. (1994b). "Estimation and Inference for Dependent Processes,". *Handbook of Econometrics*, volume 4, chapter 45, ed. R. F. Engle and D. L. McFadden, 2639–2738. Amsterdam: North-Holland.
- [12] Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

IOSR Journal of Business and Management (IOSR-JBM) is UGC approved Journal with SI. No. 4481, Journal no. 46879.

OUAKIL Hicham. " Évaluation d'impact : Revue de littérature." *IOSR Journal of Business and Management (IOSR-JBM)* 20.11 (2018): 13-17.