

Evaluation des politiques publiques : apports et pratiques

Pauline Givord

INSEE - DMS

31 Janvier 2013

Introduction

- ▶ Objectif : évaluer l'effet causal d'une politique, d'un programme, d'un médicament...
Exemple : loi antitabac dans les cafés sur la santé des personnels exposées, niveaux des taxes sur la consommation d'alcool, suivi prénatal sur la mortalité infantile...
- ▶ difficile car présence d'effets de sélection : a priori, les bénéficiaires ne le sont pas par hasard...
- ▶ on souhaiterait savoir ce qu'aurait été la situation des bénéficiaires en l'absence de cette mesure.
- ▶ on parle d'un « contrefactuel », donc de quelque chose que l'on n'observe pas.

Le problème de l'évaluation

- ▶ pour chaque personne, l'effet causal du traitement correspond à la différence de sa situation lorsqu'elle bénéficie du traitement Y_1 de celle lorsqu'elle n'en bénéficie pas Y_0
- ▶ en pratique, on observe soit l'un, soit l'autre
- ▶ on veut estimer l'effet causal du traitement pour les traités $\Delta^{ATT} = E(Y_1 - Y_0 | T = 1)$
- ▶ on observe $E(Y_1 | T = 1)$ mais pas le contrefactuel (ie. qu'on n'observe pas) $E(Y_0 | T = 1)$

Les biais de sélection

- ▶ la comparaison des personnes traitées et non traitées est en général biaisée :

$$E(Y_1|T=1) - E(Y_0|T=0) = \underbrace{E(Y_1|T=1) - E(Y_0|T=1)}_{\Delta^{ATT}} + \underbrace{E(Y_0|T=0) - E(Y_0|T=1)}_{B^{ATT}}$$

- ▶ Δ^{ATT} est l'effet causal qu'on souhaite estimer
- ▶ B^{ATT} est un biais de sélection, lié aux différences de composition de la population des traités et des non traités (état de santé, habitudes alimentaires, conditions de travail...)

Les effets de sélection

- ▶ il n'est pas possible de se contenter de suivre les personnes "traitées" car elles l'ont été probablement pour des raisons qui peuvent aussi avoir un effet sur le résultat
- ▶ exemple : évaluation de l'efficacité d'un traitement médicamenteux d'aide au sevrage tabagique
- ▶ les fumeurs auxquels il est prescrit ne le sont pas par hasard : il peut s'agir des plus motivés pour arrêter...ou de ceux qui ne le sont pas assez pour le faire sans cette aide
- ▶ la décision d'arrêt peut dépendre avant tout de cette motivation et non du traitement : un simple suivi des fumeurs ne permet pas d'évaluer l'efficacité propre (effet causal) du traitement

Les expérimentations aléatoires contrôlées

- ▶ rendre les deux populations (traitées et non traitées) comparables : randomisation
- ▶ on suit un échantillon de personnes, dont certaines sont traitées et d'autres non, ce choix étant *aléatoire*
- ▶ plus d'effet de population (en moyenne, les traités et les non traités sont identiques) ; quand c'est possible, procédure avec placebo, double aveugle...
- ▶ constituent la manière la plus crédible d'estimer l'effet causal qui nous intéresse

Les limites des expérimentations contrôlées

- ▶ parfois difficiles à mettre en place pour des questions éthiques (exemple : impact de la tabagie sur le cancer)
- ▶ se préparent souvent très en amont de la mise en œuvre souhaitée, ce qui ne correspond pas toujours au calendrier politique, alors qu'elles impliquent un coût élevé
- ▶ ont du mal à capter des effets faibles (taille de l'expérimentation) ou qui mettent du temps à apparaître (durée de l'expérimentation)
- ▶ ne sont pas toujours généralisables en population réelle (exemple : loi antitabagie, suivi des personnes âgées en risque de dépendance par leur généraliste...)

Les expériences naturelles et les quasi-expériences

- ▶ à défaut d'une expérience aléatoire contrôlée, on utilise des dispositifs "quasi-expérimentaux", ou des expériences "naturelles"
- ▶ on utilise des interventions qui n'ont pas été créées dans un but de recherche, mais qui peuvent être utilisées comme telles ; elles se traduisent par des différences d'exposition qui peuvent être utilisées pour évaluer leur effet
- ▶ la présentation aujourd'hui va présenter les principales méthodes qui peuvent être utilisées en s'appuyant sur des exemples concrets

Plan

Introduction

Exemples de quasi expériences ou d'expériences naturelles

- Les méthodes d'appariement (matching)

- Différences de Différences

- Variables Instrumentales

- Régressions sur discontinuités

Discussion

Les méthodes d'appariement (matching)

- ▶ on a un échantillon de personnes exposées au traitement, et d'autres non exposées
- ▶ différences de composition : peuvent avoir des conséquences sur les variables d'intérêt
- ▶ on a une connaissance suffisamment fine des personnes pour pouvoir "contrôler" des différences
- ▶ on va comparer chaque personne qui bénéficie de l'intervention avec un non bénéficiaire identique
- ▶ conditionnellement à ces caractéristiques X , le fait de bénéficier du traitement peut être considéré comme aléatoire :
$$E(Y_0|T = 1, X) = E(Y_0|T = 0, X)$$
- ▶ difficile de comparer sur de très nombreuses dimensions : utilisation du "score de propension"

Exemple

Lim, Dandona, Hoisington, James, Hogan et Gakidou (2010), India's Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities : an impact evaluation, The Lancet.

- ▶ évaluation d'un programme d'allocations conditionnelles à des accouchements médicalement assistés en Inde
- ▶ quel impact sur les risques de mortalité péri et néonatale ?
- ▶ les femmes qui en bénéficient peuvent avoir des caractéristiques différentes des autres
- ▶ comparaison avec celles identiques en termes socio-démographiques

Les méthodes d'appariement (matching)

- ▶ on a un échantillon de personnes exposées au traitement, et d'autres non exposées
- ▶ différences de composition : peuvent avoir des conséquences sur les variables d'intérêt
- ▶ on a une connaissance suffisamment fine des personnes pour pouvoir "contrôler" des différences
- ▶ on va comparer chaque personne qui bénéficie de l'intervention avec un non bénéficiaire identique
- ▶ difficile de comparer sur de très nombreuses dimensions : utilisation du "score de propension"

Limites

il peut rester de la sélection liée à des composantes inobservées par les chercheurs

Estimateur Avant - Après

- ▶ principe : on évalue l'évolution de la variable d'intérêt avant et après la mise en place d'une politique par exemple généralisée à la population entière

Exemple

Friedman, Powell, Hutwagner, Graham et Teague (2001), Impact of Changes in Transportation and Commuting Behaviors During the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma, JAMA

- ▶ évaluation de la pollution atmosphérique lié au trafic routier sur l'asthme infantile
- ▶ durant les jeux olympiques d'Atlanta, des politiques volontaristes de réduction du trafic automobile
- ▶ chute spectaculaire de la pollution, concomitante avec une baisse significative des crises d'asthme infantile aiguës.

Estimateur Avant - Après

- ▶ principe : on évalue l'évolution de la variable d'intérêt avant et après la mise en place d'une politique par exemple généralisée à la population entière

Limites

d'autres facteurs peuvent aussi avoir des effets sur la variable d'intérêt (par exemple dans le cas de l'asthme, conditions météo...). Il faut être certain qu'on ne mesure pas ces effets.

Différences de Différences

- ▶ Principe : combine ces deux types d'estimateurs :
 1. l'estimateur Participants/non Participants est biaisé par des effets de composition
 2. l'estimateur Avant/Après risque de capter les effets de la conjoncture, ou de la dynamique des revenus.
- ▶ L'estimateur de Différence de Différence consiste à comparer les **évolutions** des revenus des bénéficiaires avant et après, avec celles de non bénéficiaires

Remarques

- ▶ suppose de disposer de données (éventuellement agrégées) sur le groupe des bénéficiaires et le groupe des non bénéficiaires avant et après
- ▶ la première différence permet d'éliminer les effets de composition
la deuxième différence permet d'éliminer les effets temporels.
- ▶ donc la démarche n'est valide que si les évolutions auraient été les mêmes dans les deux groupes.

Exemples

Costello, Erkanli, Copeland et Angold (2010) : Association of family income supplements in adolescence with development of psychiatric and substance use disorders in adulthood among an American Indian population, JAMA

- ▶ équipe de chercheurs menaient une étude sur les déterminants de la santé mentale dans les Appalaches : suivi sur dix ans de cohortes âgées de 9, 11 et 13 ans en 1993, dont une partie d'indiens cherokee
- ▶ en 1996, décision d'implanter un casino dans la réserve : source de revenu important pour les populations indiennes (revenu moyen est passé de 500\$ à 9000\$)... sans équivalent pour les familles non indiennes
- ▶ réduction significative des problèmes mentaux chez les enfants indiens même revenus à l'âge adulte

Variable instrumentale

Principe :

- ▶ utiliser une caractéristique qui influence le fait d'avoir accès, de participer à un dispositif
- ▶ mais pas directement le résultat auquel on s'intéresse i.e. n'est pas corrélée avec les autres déterminants du revenu (en particulier inobservés).

Exemples

De “bons” instruments peuvent être fournis par :

- ▶ des différences législatives

Etude de la tabagie maternelle sur le poids de naissance :
instrument = différences dans les législations anti-tabac dans les différents états américains

- ▶ des caractéristiques géographiques

McClellan, McNeil et Newhouse (1993) Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables, JAMA.

Efficacité de traitements intensifs pour les accidents cardiaques :

instrument = distance des patients traités à un hôpital mettant en œuvre ces protocoles

Exemples

- ▶ des “événements”
Evans et Lien (2005), The benefits of prenatal care : evidence from the PAT bus strike, Journal of Econometrics
Efficacité des visites prénatales :
instrument = grève des transports publics dans un état américain

Variable instrumentale

Principe :

- ▶ utiliser une caractéristique qui influence le fait d'avoir accès, de participer à un dispositif
- ▶ mais pas directement le résultat auquel on s'intéresse
i.e. n'est pas corrélée avec les autres déterminants du revenu (en particulier inobservés).

Limites

Les instruments sont rares... et ne fournissent des informations que sur les personnes "marginales"

Régression sur discontinuités

Principe :

- ▶ de nombreux dispositifs comprennent des seuils : la probabilité d'en bénéficier est discontinue.

Exemple : seuils de revenus pour la perception d'une allocation, conditions d'âge pour accéder à un dispositif...

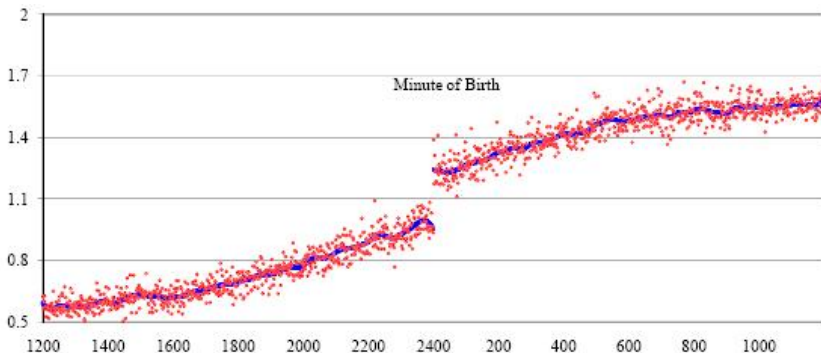
- ▶ les personnes qui sont juste en dessous ou juste au dessus ont par ailleurs des caractéristiques très proches
On peut donc les comparer sans introduire des effets de compositions trop importants.
- ▶ remarque : en terme de mise en œuvre, suppose d'avoir des données suffisamment précises pour avoir suffisamment d'observations au dessus et en dessous du seuil.

Exemples

Almond and Doyle (2010), "After Midnight : A Regression Discontinuity Design in Length of Postpartum Hospital Stays", American Economic Journal

- ▶ durée minimale de séjour pour les naissances : deux jours
- ▶ quel impact sur les taux de réadmissions / mortalité à un an ?
- ▶ système de remboursement des hôpitaux : nombres de jours (mesurés à minuit) par patients
- ▶ introduit donc une discontinuité sur la durée réelle d'hospitalisation : un enfant né à 00h05 restera plus longtemps qu'un enfant né 10mn plus tôt.

nombre de nuits supplémentaires suivant l'heure de naissance



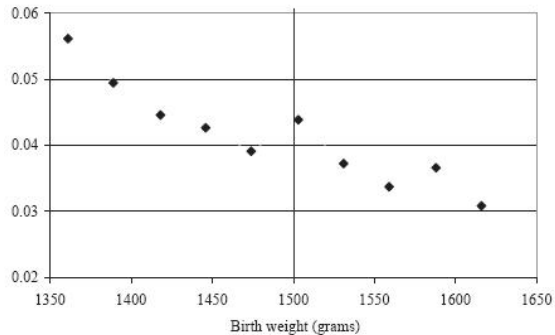
Exemple 2

Chen and van der Klaauw (2008), "The work disincentive effects of the disability insurance program in the 1990s. How do extended benefits affect unemployment duration? A regression discontinuity approach", Journal of Econometrics

- ▶ Définition "très petit poids de naissance" : 1500g
- ▶ Mais toujours arbitraire des seuils : les enfants de 1505g ne sont pas beaucoup moins exposés que ceux de 1495g... mais bénéficient d'un suivi moindre
- ▶ Avec des conséquences sur leur état de santé

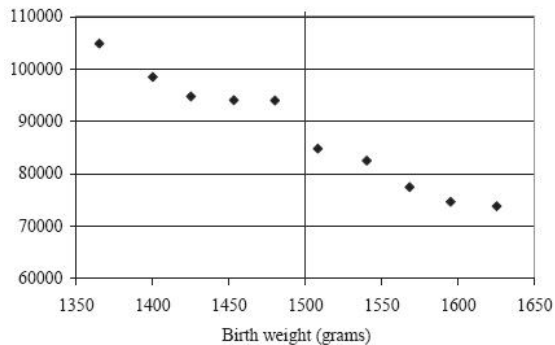
Mortalité à 28 jours

B: 28-Day mortality



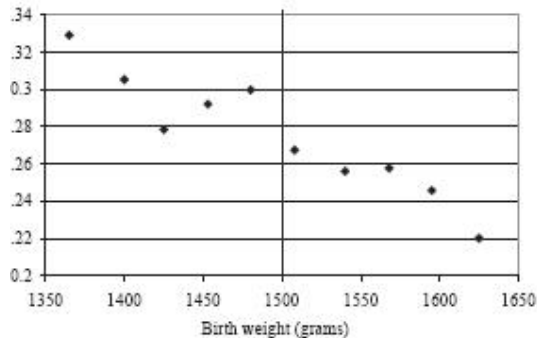
Dépenses par nouveau né

A. Hospital charges



Probabilité d'opération cardiaque

D. Operations on the heart



Régression sur discontinuités

Principe :

- ▶ de nombreux dispositifs comprennent des seuils : la probabilité d'en bénéficier est discontinue.
Exemple : seuils de revenus pour la perception d'une allocation, conditions d'âge pour accéder à un dispositif...
- ▶ les personnes qui sont juste en dessous ou juste au dessus ont par ailleurs des caractéristiques très proches
On peut donc les comparer sans introduire des effets de compositions trop importants.
- ▶ remarque : en terme de mise en œuvre, suppose d'avoir des données suffisamment précises pour avoir suffisamment d'observations au dessus et en dessous du seuil.

Limites

Là encore, ne fournit des résultats qu'"à la marge"

Expérience naturelle vs contrôlée, validité interne vs externe

- ▶ utiliser une expérience naturelle peut permettre de capter les effets d'une intervention qui peuvent être longs à apparaître, ou de faible ampleur
- ▶ les résultats sont souvent plus facilement généralisables que ceux d'une expérience naturelle
- ▶ ils sont aussi plus fragiles

Quand peut-on utiliser des expériences naturelles ?

Les conditions pour mettre en place une évaluation utilisant une expérience naturelle (*guidance for producers and users of evidence, MRC*)

- ▶ il existe des raisons rationnelles de l'existence de l'effet d'une politique mais une incertitude scientifique sur son ampleur
- ▶ pour des raisons éthiques, politiques ou pratiques, il n'est pas possible de faire une expérimentation contrôlée
- ▶ il est possible d'obtenir des données sur la population exposée et la population non exposée
- ▶ il s'agit d'une intervention répliquable