

Concepts fondamentaux de l'évaluation quantitative de l'impact : Module 4

Aneta Bonikowska, PhD

Division de l'analyse sociale et de la modélisation

19 mars 2024



Statistique
Canada

Statistics
Canada

Canada

Module 4

Approches plus courantes de l'EQI

Contenu du module 4

- Auto-évaluation
- Estimateurs de différence
- Estimateurs de discontinuité
- Variables instrumentales
- Appariement
- Régression

Exemple :

- Une agence de développement régional (ADR) finance l'acquisition de machines et d'équipements spécifiques
- L'ADR s'intéresse aux ventes des entreprises. D'autres résultats intéressants pourraient être le nombre d'employés ou la propension à exporter. Il pourrait aussi s'agir de la croissance plutôt que des niveaux.
- L'ADR veut savoir si les bénéficiaires du programme (A) ont obtenu de meilleurs résultats que leurs homologues (B)

Auto-évaluation

- On demande aux participants au programme comment le programme les a touchés
- « À quel point vos ventes sont-ils plus élevées cette année grâce au programme ? »

Auto-évaluation

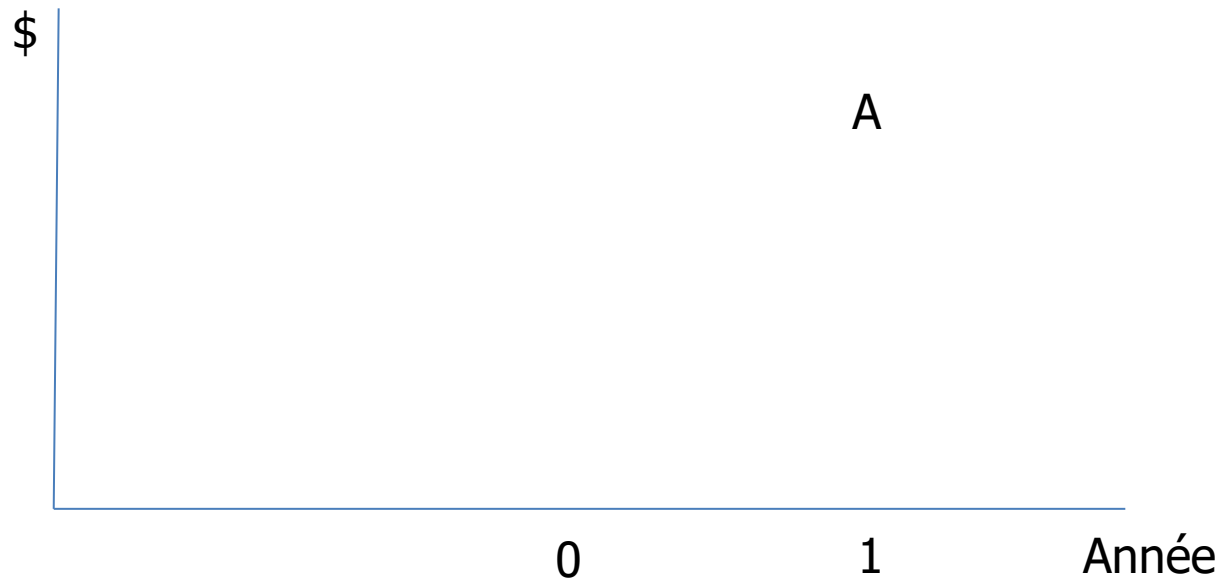
- Avantages :
 - Toujours disponible tant qu'une enquête peut être financée
 - Des approches plus rigoureuses ne sont pas toujours possibles
 - Plus facile à comprendre
- Inconvénients :
 - Les participants ne savent peut-être pas comment leurs ventes auraient évolué s'ils n'avaient pas suivi le programme
 - Les effets placebo sont bien documentés dans divers contextes

Estimateurs de différence

- Méthode n ° 1 : Aucune différence
- Il suffit de regarder les ventes des participants au programme après avoir suivi le programme

Estimateurs de différence

- Méthode n ° 1 : Aucune différence



Un an après la mise en œuvre du programme, les entreprises participantes ont réalisé des ventes de \$A valeur en moyenne.

\$A est le résultat réel des entreprises participantes. Résultat, pas impact.

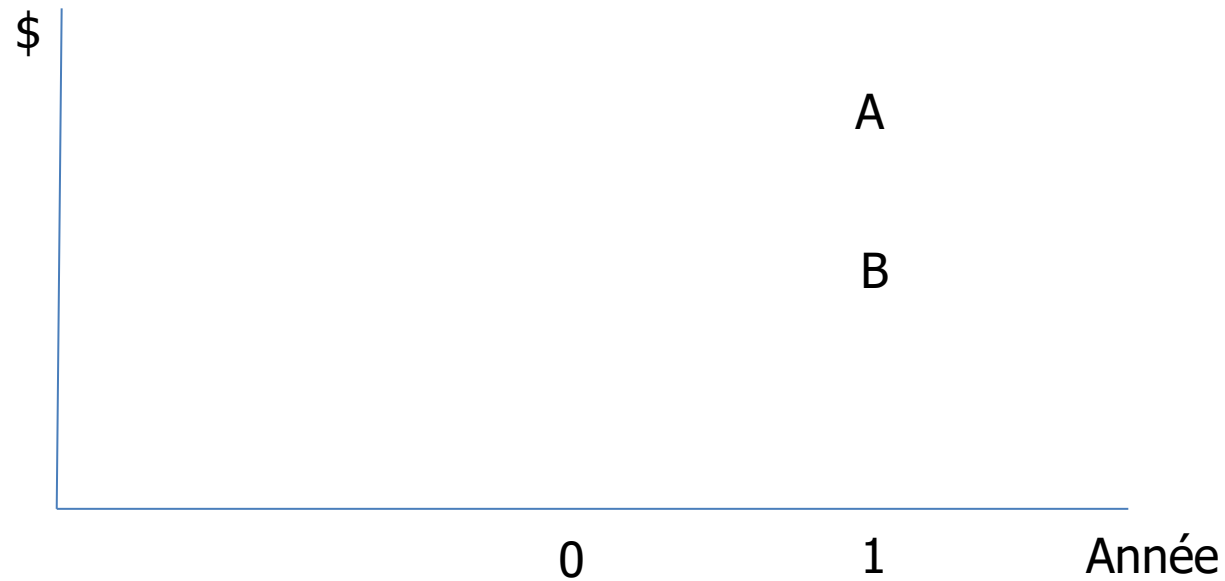
Aucune comparaison ne permet de dire si le programme a eu un impact.

Estimateurs de différence

- Méthode n ° 2 : Première différence après la mise en œuvre du programme
- Comparez les ventes des participants au programme et des non-participants (un peu) similaires après le lancement du programme

Estimateurs de différence

- Méthode no 2 : Première différence après la mise en œuvre du programme



Les entreprises non participantes avaient \$B valeur de ventes en moyenne

Impact estimé = $A - B$

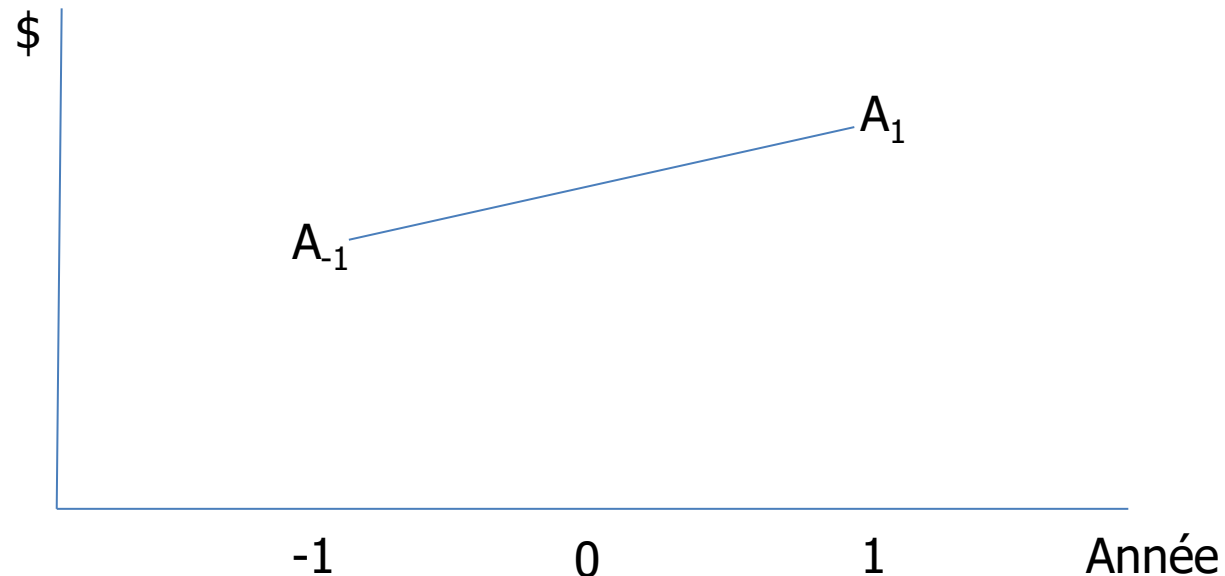
Que se passe-t-il si les ventes des participants au programme ont diminué, plutôt que d'augmenter, après la mise en œuvre ?

Estimateurs de différence

- Méthode n ° 3 : Première différence avant et après la mise en œuvre du programme (avant/après)
- Comparer les ventes des participants au programme avant et après le lancement du programme

Estimateurs de différence

- Méthode n ° 3 : Première différence avant et après la mise en œuvre du programme



$$\text{Impact estimé} = A_1 - A_{-1}$$

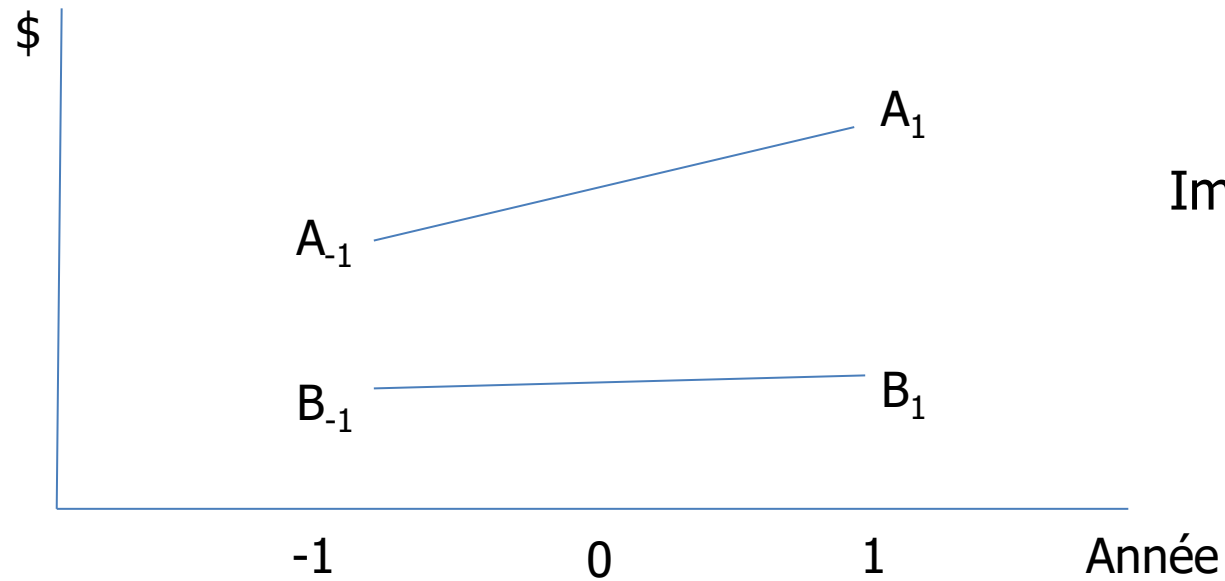
On dirait que le programme a fonctionné, mais cela pourrait être dû à des améliorations de l'économie qui pourraient également affecter les ventes des non-participants.

Estimateurs de différence

- Méthode n ° 4 : Différence dans les différences
- Comparez le changement dans les ventes (avant et après le lancement du programme) des participants au programme et des non-participants

Estimateurs de différence

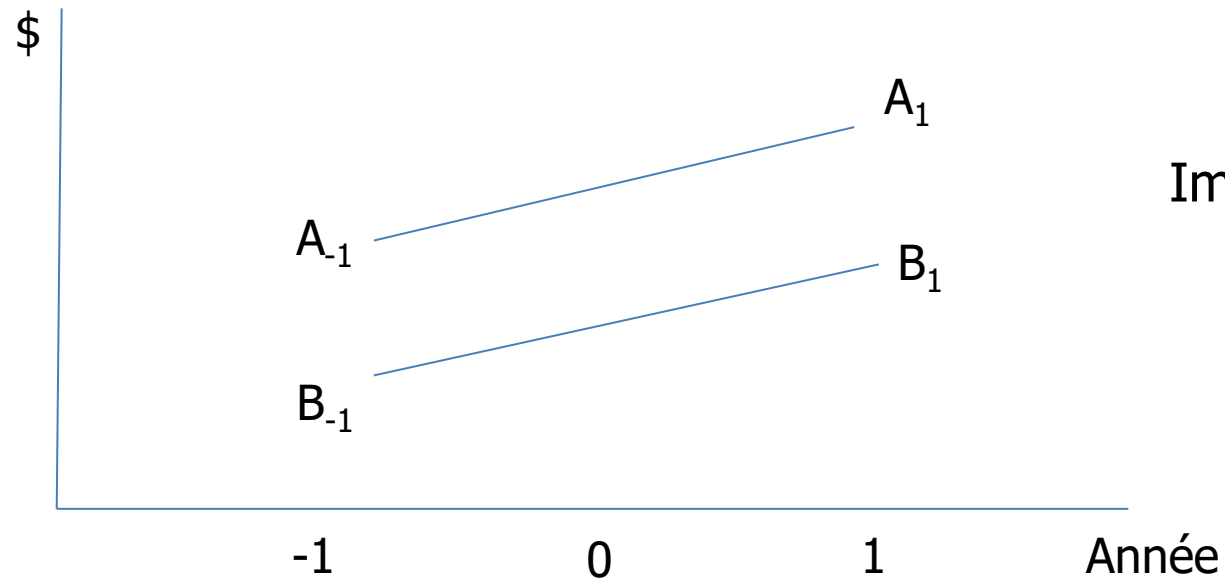
- Méthode n ° 4 : Différence dans les différences



$$\text{Impact estimé} = (A_1 - A_{-1}) - (B_1 - B_{-1})$$

Estimateurs de différence

- Méthode n ° 4 : Différence dans les différences



$$\text{Impact estimé} = (A_1 - A_{-1}) - (B_1 - B_{-1})$$

Estimateurs de différence

- Méthode n ° 4 : Différence dans les différences
- Point de départ pour les évaluations quantitatives de l'impact les plus sérieuses
- Cependant, il y a de nombreuses raisons pour lesquelles cela peut encore ne pas être suffisant

Estimateurs de différence

- Problème n ° 1 avec différences dans les différences :

Facteurs de confusion

- D'autres facteurs peuvent avoir une incidence sur le groupe de programmes en particulier
- Habituellement, il s'agit d'un problème avec l'analyse provinciale ou d'un autre groupe (p. ex. une nouvelle politique provinciale est entrée en vigueur en même temps que la mise en œuvre du programme)
- Peut également être un problème lors du suivi des résultats individuels (par exemple, les non-participants peuvent chercher un autre traitement)

Estimateurs de différence

- Problème n ° 1 avec différences dans les différences :
 - La solution au problème des facteurs de confusion est d'effectuer une analyse de l'environnement pour s'assurer qu'il n'y en a pas

Estimateurs de différence

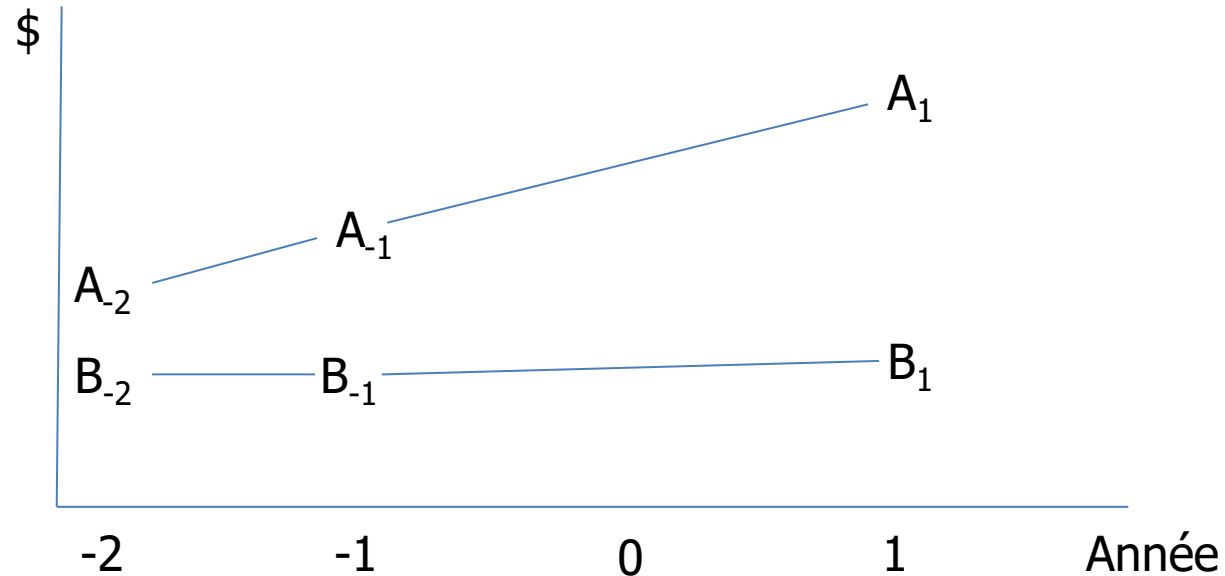
- Problème n ° 2 avec la différence dans les différences :

Hypothèse des tendances communes

- Les groupes de programme et de comparaison peuvent avoir été sur des chemins différents avant la mise en œuvre du programme

Estimateurs de différence

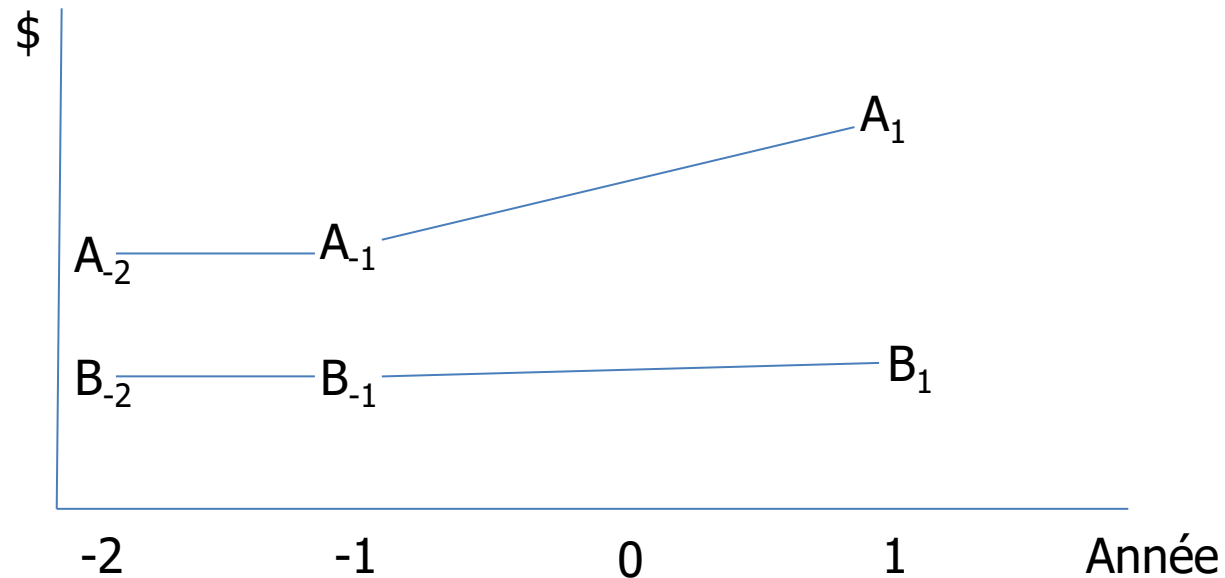
- Le programme a-t-il fonctionné ?



Cela n'en a pas l'air puisque les tendances n'ont pas changé après la mise en œuvre du programme

Estimateurs de différence

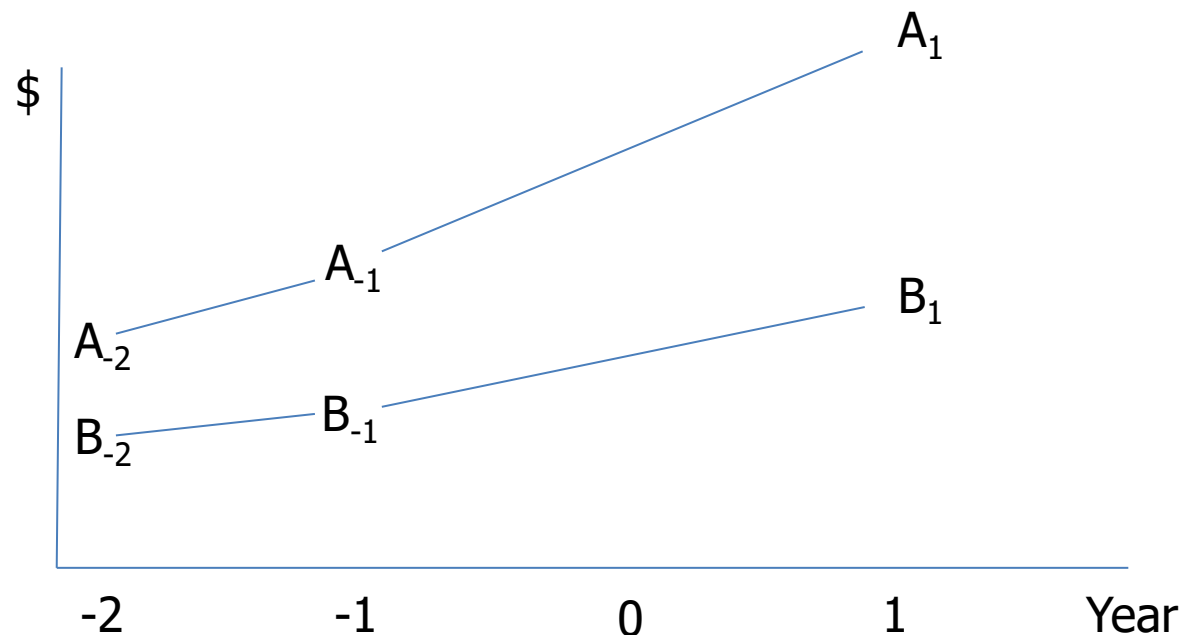
- Le programme a-t-il fonctionné ?



Il semble que ce soit le cas puisque les tendances étaient similaires avant la mise en œuvre du programme

Estimateurs de différence

- Le programme a-t-il fonctionné ?



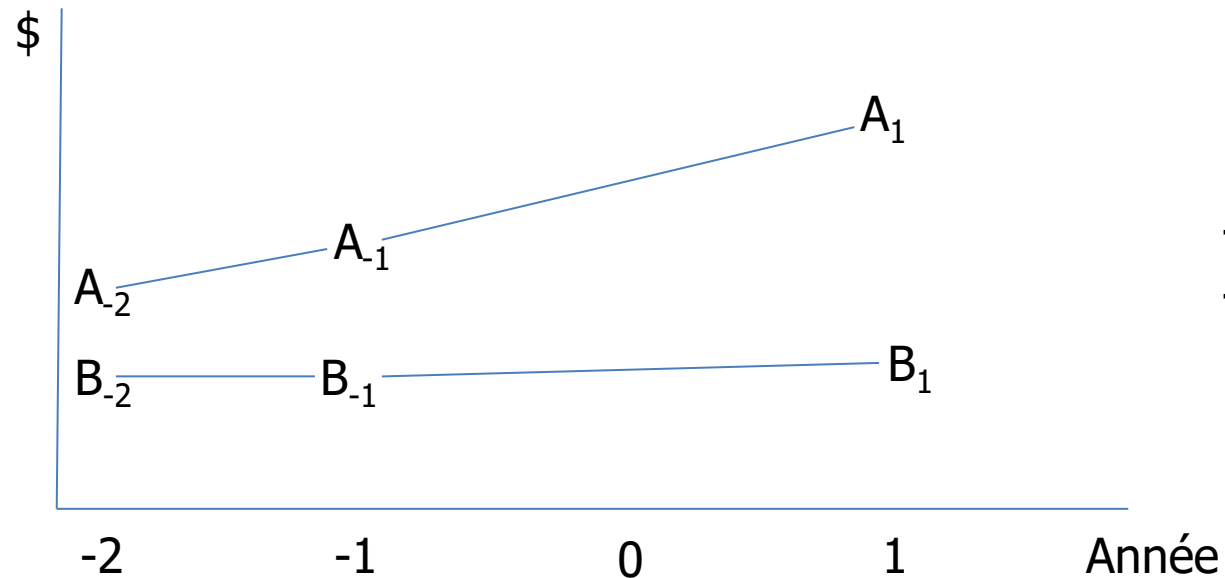
Il semble que ce soit le cas puisque les ventes ont augmenté plus rapidement pour A

- après la mise en œuvre du programme par rapport à avant la mise en œuvre et
- par rapport à la croissance de B après la mise en œuvre du programme

Estimateurs de différence

- Problème n ° 2 avec la différence dans les différences :
 - La solution à l'hypothèse des tendances communes est de ne pas faire cette hypothèse
 - Au lieu de cela, prenez en considération les tendances préexistantes

Estimateurs de différence



Estimation de l'incidence du programme : comparer les différences dans les tendances avant et après la mise en œuvre du programme

Tendance après : $(A_1 - A_{-1}) - (B_1 - B_{-1})$
Tendance avant : $(A_{-1} - A_{-2}) - (B_{-1} - B_{-2})$
La différence entre ces deux chiffres est l'impact estimé de notre programme

Estimateurs de différence

- Problème n ° 3 avec la différence dans les différences :

Sélection sur le traitement

- Le groupe du programme a peut-être choisi de suivre le programme parce qu'il savait que cela lui serait bénéfique
- Le groupe témoin a peut-être choisi de ne pas suivre le programme parce qu'il savait que cela ne lui serait pas bénéfique
- Problème majeur dans l'EQI

Estimateurs de différence

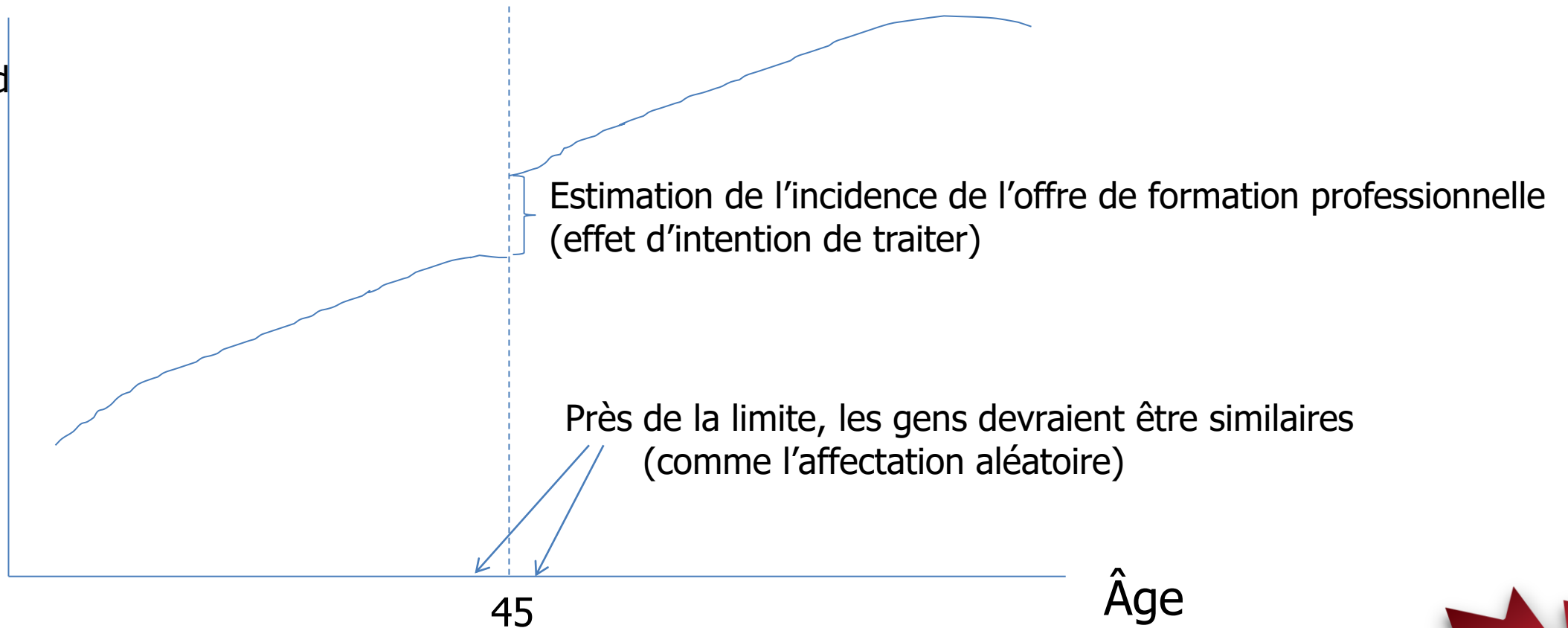
- Problème n ° 3 avec la différence dans les différences :
 - Solutions pour la sélection sur le traitement, par ordre décroissant d'efficacité :
 - Éliminer le choix (p. ex. affectation aléatoire)
 - Trouver des situations dans le monde réel où le choix est supprimé (discontinuités et variables instrumentales)
 - S'assurer que les programmes et les groupes témoins sont aussi semblables que possible au moyen d'ajustements mathématiques (appariement et régression)

Estimateurs de discontinuité

- Exemple : Un programme de formation professionnelle unique est offert aux travailleurs « âgés », définis comme étant âgés de 45 ans et plus un jour précis (un jour de moins supprime l'admissibilité)

Estimateurs de discontinuité

Gains
5 ans plus tard



Estimateurs de discontinuité

- **Une discontinuité floue** est une discontinuité où la probabilité de recevoir le traitement est affectée par la discontinuité (comme le programme de formation professionnelle ci-dessus)
 - Comme l'intention de traiter
- **Une discontinuité nette** est une discontinuité où le traitement sera déterminé par la discontinuité

Estimateurs de discontinuité

- Limites de l'utilisation des discontinuités dans l'admissibilité :
 - Rare (règles d'admissibilité souvent plus graduelles)
 - Besoin d'un très grand échantillon très proche de la coupure
 - Effets de manipulation - les gens peuvent modifier (jeu) ou mentir sur leurs critères d'admissibilité pour obtenir un traitement (sélection sur le traitement)
 - Discontinuités qui se chevauchent (facteurs de confusion)
 - EMTL (les effets ne s'appliquent que près de la discontinuité)

Estimateurs de variables instrumentales

- Exemple : Nouveau programme de formation professionnelle, mais au lieu d'être fondé sur l'âge, le programme s'adresse aux parents seuls ayant au moins trois enfants
- Impossible d'appliquer un estimateur de discontinuité puisque deux et trois enfants sont des choix très différents probablement faits par des personnes très différentes, c'est-à-dire que les personnes avec deux enfants pourraient être très différentes des personnes avec trois enfants et donc ne pas être un groupe témoin approprié pour les personnes avec trois enfants
- Au lieu de cela, pouvons-nous penser à une situation dans laquelle il y a un élément de chance dans le nombre d'enfants et donc l'admissibilité au programme ?

Estimateurs de variables instrumentales

- De nombreux parents visent deux enfants, de préférence un de chaque sexe
- Deux cas où les parents peuvent se retrouver avec trois, même s'ils n'en espéraient que deux, d'intérêt pour les chercheurs¹:
 1. Naissance multiple à la deuxième naissance
 2. Deux enfants du même sexe lors des deux premières naissances (fille-fille ou garçon-garçon)
- Nous nous référons à 1. et 2. en tant que variables instrumentales
- Dans une large mesure, la chance détermine les naissances multiples et le sexe des enfants
- La chance est comme l'affectation aléatoire (c'est-à-dire qu'elle enlève le choix, ce qui est une bonne chose pour l'EQI)

¹ Source: Frenette, Marc. (2011). How does the stork delegate work? Childbearing and the gender division of paid and unpaid labour. *Journal of Population Economics*. 24. 895-910. 10.1007/s00148-010-0307-y.

Estimateurs de variables instrumentales

Naissance multiple à la deuxième naissance, exemple

NAISSANCE	GROUPE 1	GROUPE 2
	Nombre d'enfants nés	
1 ^{er}	1	1
2 ^e	1	2

- Le groupe 2 devient admissible au programme par chance
- Le fait d'avoir des jumeaux à la deuxième naissance est un candidat potentiel pour une variable instrumentale pour l'admissibilité au programme

Estimateurs de variables instrumentales

2 enfants de même sexe lors des 2 premières naissances, exemple

NAISSANCE	GROUPE 1	GROUPE 2
	Sex of child	
1 ^{er}	Garçon	Garçon
2 ^e	Fille	Garçon
3 ^e	N/A	Fille

- Les parents avec 2 enfants du même sexe lors des 2 premières naissances sont empiriquement plus susceptibles d'avoir plus de 2 enfants au total !
- 2 enfants du même sexe sur les 2 premières naissances est (généralement) due à la chance
- 2 enfants du même sexe lors des 2 premières naissances est un « prédicteur » d'avoir plus de 2 enfants au total
- Par conséquent, le fait d'avoir 2 enfants du même sexe lors des 2 premières naissances est un candidat potentiel pour une variable instrumentale pour l'admissibilité au programme

Estimateurs de variables instrumentales

- La mécanique réelle des estimateurs de variables instrumentales est assez complexe et mieux étudiée dans une classe de méthodes avancées
- Notez que vous aurez besoin de données sur l'admissibilité au programme (dans cet exemple, le nombre d'enfants) et la ou les variables instrumentales choisies
- Une variable instrumentale crédible a deux caractéristiques :
 - Force
 - Validité

Estimateurs de variables instrumentales

Force :

- Il devrait être fortement corrélé avec les critères d'admissibilité (le fait d'avoir une naissance multiple à la deuxième naissance, ou d'avoir deux enfants de même sexe lors des deux premières naissances, doit être fortement corrélé avec l'admissibilité au programme)
 - Facile à vérifier avec les données

Estimateurs de variables instrumentales

Validité :

- Il ne doit être lié au résultat (les gains dans notre exemple) que par son incidence sur l'admissibilité au programme.
- Remarque : Tout ce qui ressemble à de la chance est toujours le meilleur
- En général, très difficile à établir autrement que pour des raisons conceptuelles

Estimateurs de variables instrumentales

- Limites de l'utilisation des variables instrumentales :
 - Rare (à la fois conceptuellement et en termes de disponibilité dans les ensembles de données)
 - La validité est difficile à établir
 - Les instruments faibles posent leurs propres problèmes
 - EMTL (les effets ne s'appliquent qu'à ceux qui modifieraient leur décision de formation en raison de la valeur de la variable instrumentale, mais cela ne dit rien sur l'impact du programme de formation sur ceux qui prendraient ou non une formation, quelle que soit la valeur de la variable instrumentale)

Estimateurs d'appariement

- Un outil mathématique pour rendre les groupes de programme et de comparaison plus similaires
- Peut être utilisé seul dans une EQI, ou en conjonction avec d'autres approches que nous avons vues pour les améliorer
- Ils peuvent même améliorer les expériences contrôlées randomisées puisque le traitement et les groupes témoins ne sont probablement pas identiques

Estimateurs d'appariement

- On nous donne les données suivantes sur les participants et les non-participants au programme de formation professionnelle :

La formation professionnelle ?	Résultat (gains)	Âge
Yes	100 000	55
Yes	50 000	35
Yes	30 000	30
No	40 000	36
No	25 000	34
No	25 000	29
No	10 000	21

Résultat moyen des groupes de programmes
 $= (100\ 000 + 50\ 000 + 30\ 000) / 3 = 60\ 000$

Résultat moyen du groupe témoin =
 $(40\ 000 + 25\ 000 + 25\ 000 + 10\ 000) / 4 = 25\ 000$

Incidence estimée du programme = $60\ 000 - 25\ 000 = 35\ 000$

Estimateurs d'appariement

- À première vue, il semble que le programme ait amélioré les revenus de 35 000 \$ par personne
- Mais les participants au programme et les non-participants sont très différents
- Plus précisément, les participants sont plus âgés en moyenne
- Comparons des pommes à des pommes en trouvant un contrefactuel approprié pour chaque individu

Estimateurs d'appariement

- Mis à part le résultat et l'état du traitement, nous ne connaissons que l'âge des individus
- Trouvons le meilleur contrefactuel pour tout le monde en fonction de la proximité de l'âge
- De plus, assurons-nous que seules les personnes proches en âge sont comparées (disons, pas plus de cinq ans d'intervalle)

Estimateurs d'appariement

- On nous donne les données suivantes sur les participants et les non-participants au programme de formation professionnelle :

La formation professionnelle ?	Résultat (gains)	Âge	Résultat contrefactuel	Incidence du programme
Oui	100 000	55	--	
Oui	50 000	35	$(40\ 000 + 25\ 000)/2 = 32\ 500$	17 500
Oui	30 000	30	25 000	5 000
Non	40 000	36	50 000	10 000
Non	25 000	34	50 000	25 000
Non	25 000	29	30 000	5 000
Non	10 000	21	--	

Incidence estimée du programme = $(17\ 500 + 5\ 000 + 10\ 000 + 25\ 000 + 5\ 000)/5 = 12\ 500$

Estimateurs d'appariement

- Cet exemple a mis en évidence deux autres aspects de l'EQI :
 - L'impact peut être différent pour différentes personnes (impacts hétérogènes)
 - Nous pouvons estimer un impact pour ceux qui n'ont pas suivi le programme, en regardant les résultats de ceux qui l'ont fait (c'est-à-dire que le contrefactuel de ne pas suivre le programme est ... suivre le programme)

Estimateurs d'appariement

- Nous pouvons prendre plusieurs correspondances (par exemple, 10 contrefactuels les plus proches)
- Danger d'accepter trop de matchs - certains ne sont pas si similaires
- Essayer différents critères pour voir comment les résultats changent est une bonne approche (test de robustesse)

Estimateurs d'appariement

- Lorsque des estimateurs de correspondance ont été élaborés, ils ont été utilisés dans le contexte d'expériences contrôlées randomisées
- Les chercheurs avaient tendance à projeter les qualités des expériences sur des estimateurs appariés (ils ont été mélangés ensemble)
- En réalité, les estimateurs de correspondance ne sont aussi bons que les données avec lesquelles nous devons faire correspondre
- Il peut y avoir des facteurs non observés qui importent (problème de « correspondance sur les observables »)

Estimateurs d'appariement

- Différents types d'estimateurs d'appariement :
 - Appariement des voisins les plus proches (ce que nous venons de voir)
 - Correspondance exacte (imposer que les correspondances soient identiques)
 - Appariement du score de propension (populaire, mais plus technique)
 - Appariement exact grossier (plus récent)

Régressions

- Les régressions sont très similaires aux estimateurs de correspondance dans leur objectif :
 - Les deux approches rendent le programme et les groupes de comparaison plus similaires grâce à des techniques mathématiques
 - Les estimateurs appariants le font directement en comparant les membres des deux groupes qui sont similaires
 - Mais que se passe-t-il si les échantillons sont trop petits ? Il n'est peut-être pas possible de comparer des individus similaires
 - Les régressions rendent également les groupes de programme et de comparaison similaires, mais en prédisant ce qui se passerait si les deux groupes avaient les mêmes caractéristiques

Régressions

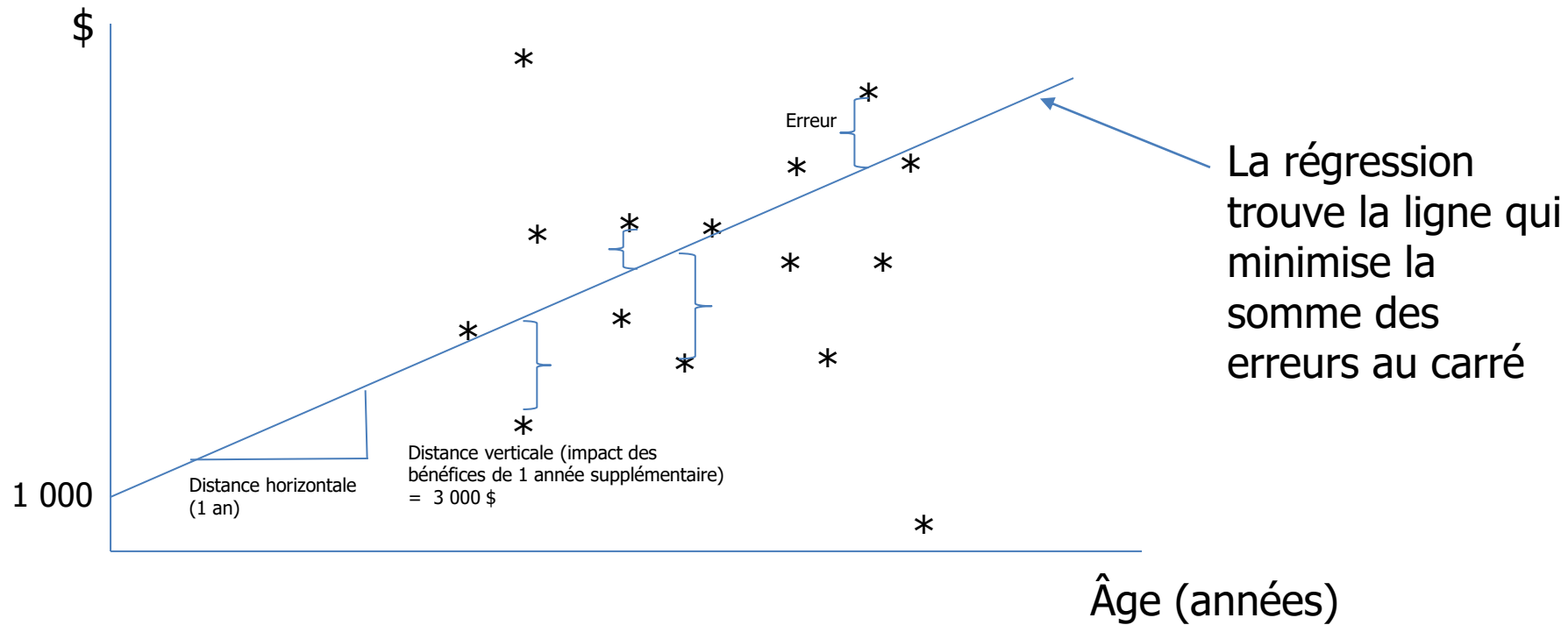
- Retour à notre exemple de formation professionnelle
- Nous avons des données sur les gains et l'âge pour le programme et les groupes témoins
- Le groupe de programme est beaucoup plus âgé et comme les travailleurs âgés ont généralement plus d'expérience, ils gagnent généralement plus

Régressions

- En d'autres termes, voici ce que fait une régression :
 - Estime la corrélation entre les gains et l'âge (disons qu'une année supplémentaire est associée à des gains de plus de 3 000 \$, en moyenne)
 - Rajuste l'écart entre les gains entre le programme et le groupe témoin en supposant que les deux ont le même âge
 - Par exemple, si l'écart réel dans les gains est de 15 000 \$ et que le groupe de programme a deux ans de plus, l'écart ajusté sera:
 - $15\ 000\ \$ - 2\ \text{ans} * 3\ 000\ \$ = 9\ 000\ \$$
 - Vous pouvez tenir compte des différences dans de nombreux facteurs entre ceux qui ont suivi la formation (programme) et ceux qui ne l'ont pas suivi, à condition d'avoir des données sur ces facteurs

Régressions

- Graphiquement, voici ce que fait une régression :



Régressions

- Mathématiquement, c'est ce que fait une régression (et ce que vous pourriez voir en tant que consommateur) :

Modèle simple de gains :

$$\text{Gain}_i = \alpha + \beta \hat{\text{âge}}_i + e_i$$

De combien les gains augmentent-ils avec 1 an de plus, en moyenne ?

$$\text{Gain} = 1\,000 + 3\,000 \times \hat{\text{âge}}$$

Les coefficients estimés vous permettent de prédire les gains moyens à différents âges

Regressions

- Mathématiquement, c'est ce que fait une régression (et ce que vous pourriez voir en tant que consommateur) :

$$\text{Gain}_i = \alpha + \beta I_{\text{programme}_i} + e_i$$

= gains moyens des personnes qui ont suivi le programme
– les gains moyens de ceux qui n'ont pas suivi le programme ;

$$\text{Gain}_i = \alpha + \beta_1 I_{\text{programme}_i} + \beta_2 \text{age}_i + e_i$$

= gains moyens des personnes qui ont suivi le programme
– les gains moyens de ceux qui n'ont pas suivi le programme
MAINTIEN DE L'ÂGE CONSTANT

Régressions

- Comme les estimateurs d'appariement, les régressions peuvent être utilisées seules dans l'EQI, ou en conjonction avec d'autres approches que nous avons vues pour les améliorer
- Avec de très grands échantillons, les estimateurs de correspondance sont potentiellement meilleurs car ils permettent des appariements plus directs
- Avec des échantillons plus petits, les régressions sont plus réalisables, mais elles nécessitent beaucoup de compétences pour les mettre en œuvre et les interpréter correctement
- Notez cependant que l'utilisation de l'analyse de régression seule avec des données d'observation (c'est-à-dire non des données obtenues par affectation randomisée) n'est pas suffisante pour estimer un impact causal

Principaux points à retenir du cours

- Dilemme : *l'amélioration des résultats peut-elle être attribuée à la participation au programme ou aux caractéristiques individuelles ou de l'entreprise de ceux qui ont choisi de participer, ou à d'autres facteurs ?*
- L'amélioration des résultats peut être corrélée à la participation au programme, mais pas nécessairement causée par la participation au programme
- Corrélation \neq causalité
- Pour estimer l'impact causal d'un programme, les résultats des participants au programme doivent être mis en contexte – doivent trouver un résultat contrefactuel crédible
- Un essai contrôlé randomisé est la meilleure façon de le faire parce que...

Principaux points à retenir du cours

- Dans la vraie vie, l'affectation aléatoire est rare. La plupart des politiques et des programmes ne seront pas mis à l'essai dans un cadre expérimental. Pourquoi cela complique-t-il les choses ?
 - Les personnes et les entreprises choisissent de participer aux programmes
 - Peut conduire à un biais d'autosélection
 - Ceux qui choisissent de participer peuvent différer systématiquement de ceux qui choisissent de ne pas participer en termes de caractéristiques qui pourraient influencer leurs résultats
 - D'autres facteurs (p. ex. les conditions économiques, les politiques gouvernementales, etc.) peuvent changer en même temps qu'un programme est mis en œuvre et avoir une incidence sur les résultats des participants au programme
- En l'absence d'affectation aléatoire, nous devons nous tourner vers d'autres approches (moins crédibles) qui tentent d'imiter l'affectation aléatoire (par exemple, la différence dans les différences, la discontinuité, les variables instrumentales, l'appariement)

Merci!



Aneta.Bonikowska@statcan.gc.ca



Restez branchés!

Application StatsCAN

Balado Hé-coutez bien!

StatsCAN Plus

Le Quotidien

Site Web

Enquêtes et programmes statistiques

Centre de service de données

Mon StatCan



Vous avez des questions? Contactez-nous : infostats@statcan.gc.ca



Statistique Canada – Votre organisme national de statistique



Éclairer grâce aux données, pour bâtir un Canada meilleur

