

Formation aux méthodes d'évaluation d'impact scientifiques

Introduction

Florent Bédécarrats, Thomas Thivillon



Le contenu de cette formation est largement tiré de Bédécarrats, Bouvier, HOUNGbedji; de Montalembert, Ferry. 2022. Impact des aires protégées sur la déforestation : guide de formation pratique, Tany Vao, IRD. [En ligne](#).

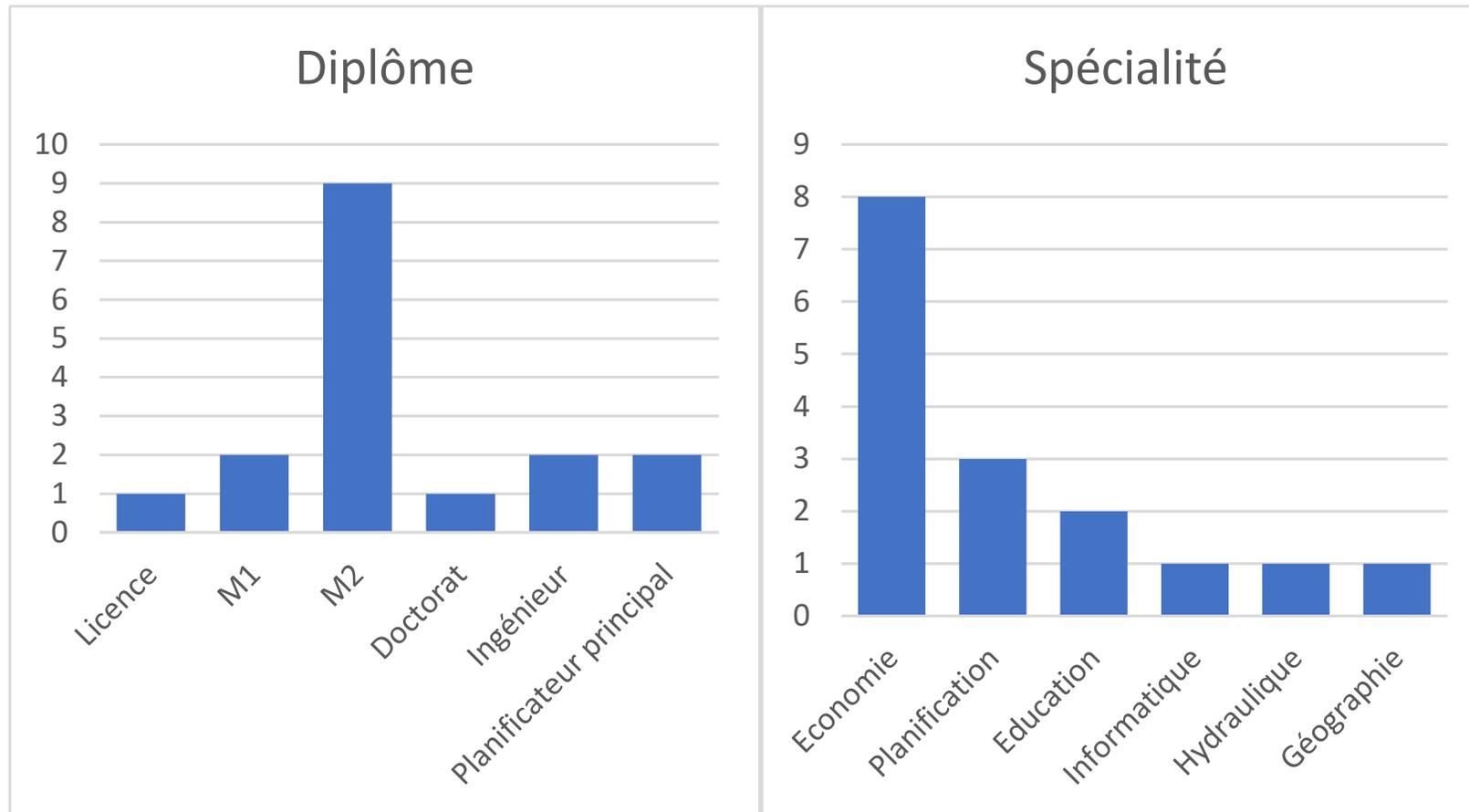
Tour de table

Présentation de chaque participant

Attendus de la formation

Niveau d'expérience préalable avec R, l'évaluation d'impact et les données spatiales

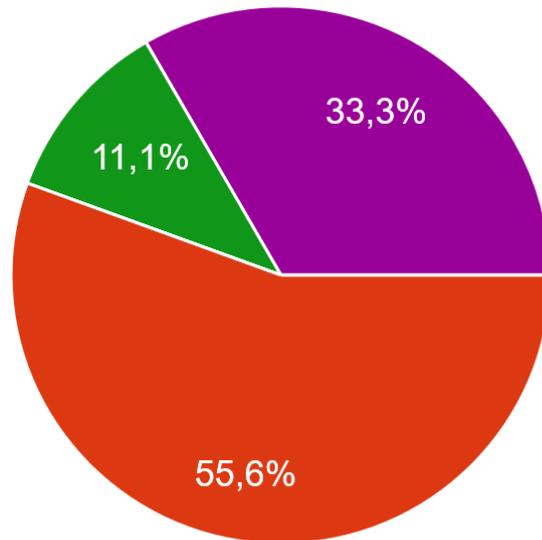
Profils



Expérience en analyse de données

Avez-vous déjà travaillé avec des données ?

18 réponses



- Jamais
- Oui, uniquement dans Excel ou un tableur
- Oui, dans un logiciel statistique
- Oui, pour des analyses descriptives
- Oui, pour des modélisations statistiques

Objectifs de la formation

- Formation aux méthodes d'évaluation d'impact
- Accent sur les données spatialisées et les traitements statistiques qui s'y rapportent
- Fil rouge de ces travaux est l'impact des aires protégées sur la déforestation
- Alternier sessions théoriques et pratiques pour favoriser la réutilisation des principes
- Se familiariser avec un outil (package mapme.biodiversity) qui facilite l'acquisition de données spatiales et leur transformation en indicateurs exploitables pour l'évaluation économétrique
- Formation largement fondée sur les [supports élaborés pour l'université Tany Vao 2022](#) avec Jeanne de Montalembert, Marin Ferry et Marc Bouvier

Programme de l'atelier

- **Jour 1 (1^{er} juillet) :**
 - Matin : Introduction et cadrage théorique.
 - Après-midi : Initiation à R
- **Jour 2 (2 juillet) :**
 - Matin : Principes de l'évaluation d'impact avec contrefactuel.
 - Après-midi : Initiation à R (suite)
- **Jour 3 (3 juillet) :**
 - Matin : Revue des types de données.
 - Après-midi : Initiation à R (fin)
- **Jour 4 (4 juillet) :**
 - Matin : Données spatiales (travail avec les données vectorielles et raster)
 - Après-midi : Données spatiales (opérations avec les packages sf et mapme.biodiversity)
- **Jour 5 (7 juillet) :**
 - Matin : Expérimentations aléatoires (théorie)
 - Après-midi : Expérimentations aléatoires (application)
- **Jour 6 (8 juillet) :**
 - Matin : Comparaisons avant-après ; Clarifications sur les unités d'analyse
 - Après-midi : Méthodes d'appariement (théorie)
- **Jour 7 (9 juillet) :**
 - Matin : Méthodes d'appariement (application)
 - Après-midi : Méthode des doubles différences (théorie)
- **Jour 8 (10 juillet) :**
 - Matin : Méthodes des doubles différences (application)
 - Après-midi : Projet de groupe.
- **Jour 9 (11 juillet) :**
 - Matin : Exercices pratiques dans R.
 - Après-midi : Restitution, évaluation et conclusion.

Modalités technique

- Utilisation d'environnements "cloud":
 - SSP Cloud
 - Liens et mot de passe disponible chaque jour
- Supports disponibles en ligne

- Version technique :

https://github.com/BETSAKA/statcap_impact_training

- Rendu visuel :

https://betsaka.github.io/statcap_impact_training



Attention à bien réserver l'usage des connexions wifi aux besoins du cours
SVP

Jour 1 : Fondements théoriques

Objectifs de cette session

- Spécificité des évaluations d'impact dans la « boîte à outils » évaluative
- Enjeux de l'évaluation d'impact
- Cadre d'analyse économétrique de la causalité : le cadre des résultats potentiels de Rubin
- Affiner sa compréhension des mécanismes de causalité : diagrammes acycliques orientés et combinaisons quanti-quali
- Tour d'horizon des principales méthodes d'évaluation d'impact

Spécificités de l'évaluation d'impact
scientifique

Attention aux différentes acceptions de l'évaluation

- Terme très usité:
 - Signification parfois floue
 - Même quand bien défini : différents référentiels
- Deux cadres bien définis qui ont des définitions différentes
 - Référentiel des politiques et projets de développement
 - Référentiel scientifique d'identification causale

Qu'est-ce que c'est que l'évaluation de politique ou de projets ?

- Définition du CAD OCDE (2000)

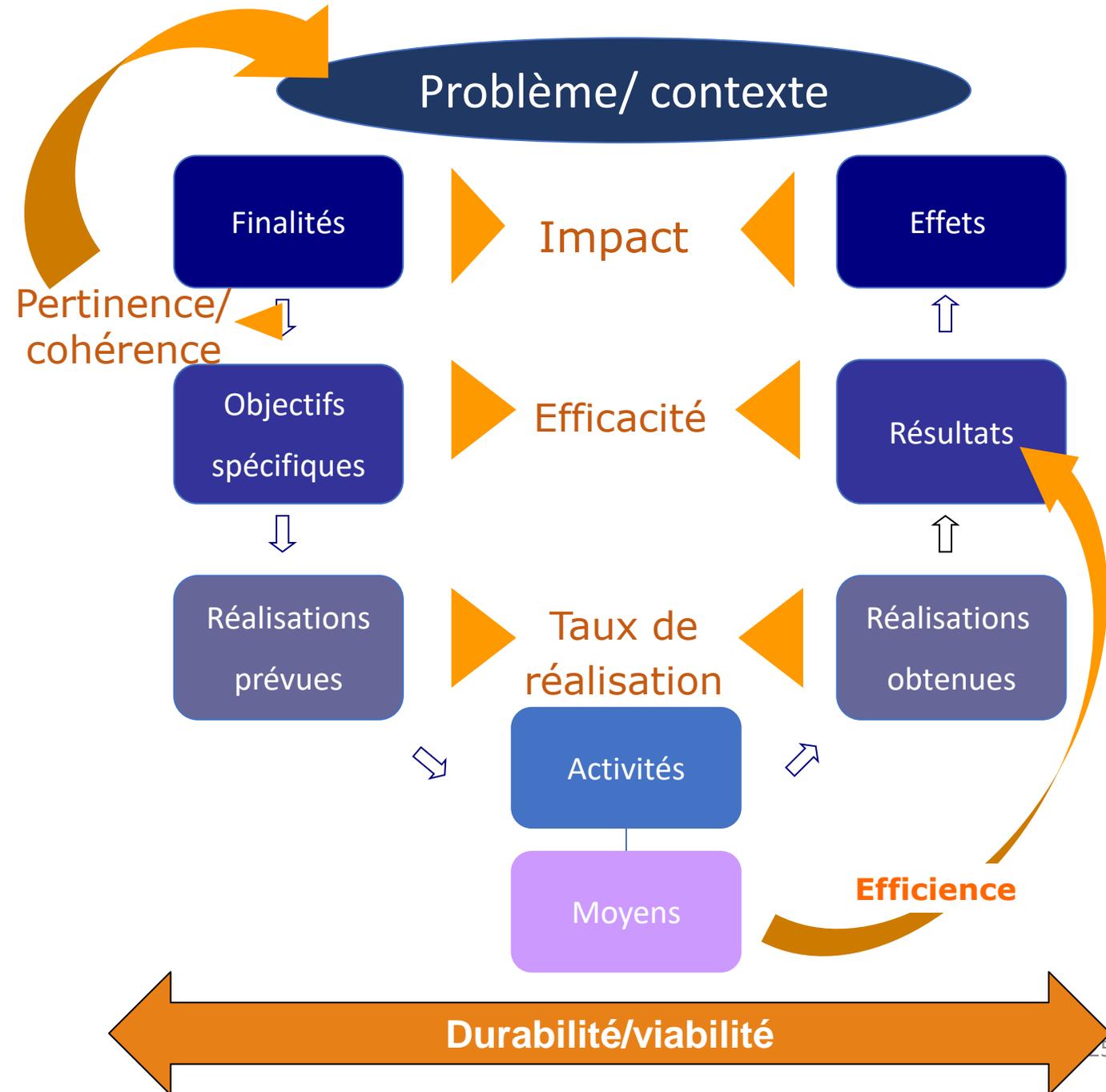
« **Appréciation** systématique et objective d'un projet, d'un programme ou d'une politique en cours ou achevé, de sa conception, de sa mise en œuvre et de ses résultats. Le but est de déterminer la pertinence, le niveau d'atteinte des objectifs, l'efficacité (...), l'impact et la durabilité».

« évaluation » désigne également un processus aussi systématique et objectif que possible par lequel on détermine la valeur et la portée d'une action de développement

<http://www.oecd.org/dac/dac-glossary.htm#Evaluation>

Relation entre logique d'intervention et critères d'évaluation

- **Pertinence:** Les objectifs de l'évaluation correspondent-ils aux problèmes identifiés?
- **Cohérence:** L'intervention est-elle compatible avec le contexte général?
- **Efficacité:** L'intervention a-t-elle atteint ses objectifs prévus?
- **Effizienz:** Les ressources allouées à l'intervention ont-elles été bien utilisées?
- **Impact:** L'intervention a-t-elle conduit à un changement réel et mesurable pour les bénéficiaires?
- **Durabilité/viabilité:** Les bénéfices de l'intervention peuvent-ils perdurer dans le temps?



Différentes acceptions de l'impact

Approche classique de l'évaluation du développement

- **Définition** : « Mesure dans laquelle l'intervention a produit, ou devrait produire, des effets importants et de vaste portée, positifs ou négatifs, intentionnels ou non. » (CAD-OCDE, 2019)
- **Spécificités** :
 - Adopte une perspective **holistique**, englobant une gamme variée d'effets.
 - Est orientée vers des effets **durables** et de **long terme**.
 - Reconnaît l'existence d'effets **non intentionnels**, qu'ils soient positifs ou négatifs.

Approches scientifiques (en particulier économétriques)

- **Définition** : « La différence entre le résultat observé avec l'intervention et le résultat qui aurait été observé sans elle. » (Angrist et Pischke, 2009)
- **Spécificités** :
 - Se focalise sur des effets à **court terme** : relèvent souvent de l'efficacité.
 - Adopte une rigueur méthodologique pour isoler l'effet de l'intervention des autres facteurs.
 - Priorise la **causalité**, cherchant à déterminer les effets directs et mesurables d'une action.

Discussion

Que pensez-vous de ces définitions de l'impact ?

Qu'est-ce qui manque selon vous ?

Les enjeux de l'évaluation d'impact

Le problème d'identification et les corrélations fallacieuses

Le problème d'identification (exemple 1)

- Exemple de Angrist et Pischke (2009, p. 10-12) tiré de l'enquête nationale sur la santé aux Etats-Unis en 2005
- Relation entre :
 - états de santé : de 1 (excellente santé) à 5 (santé très dégradée)
 - le fait d'être allé à l'hôpital au cours des 12 derniers mois ?

Groupe	Taille de l'échantillon	Statut de santé moyen	Erreur standard
Hôpital	7774	2.79	0.014
Sans hôpital	90049	2.07	0.003

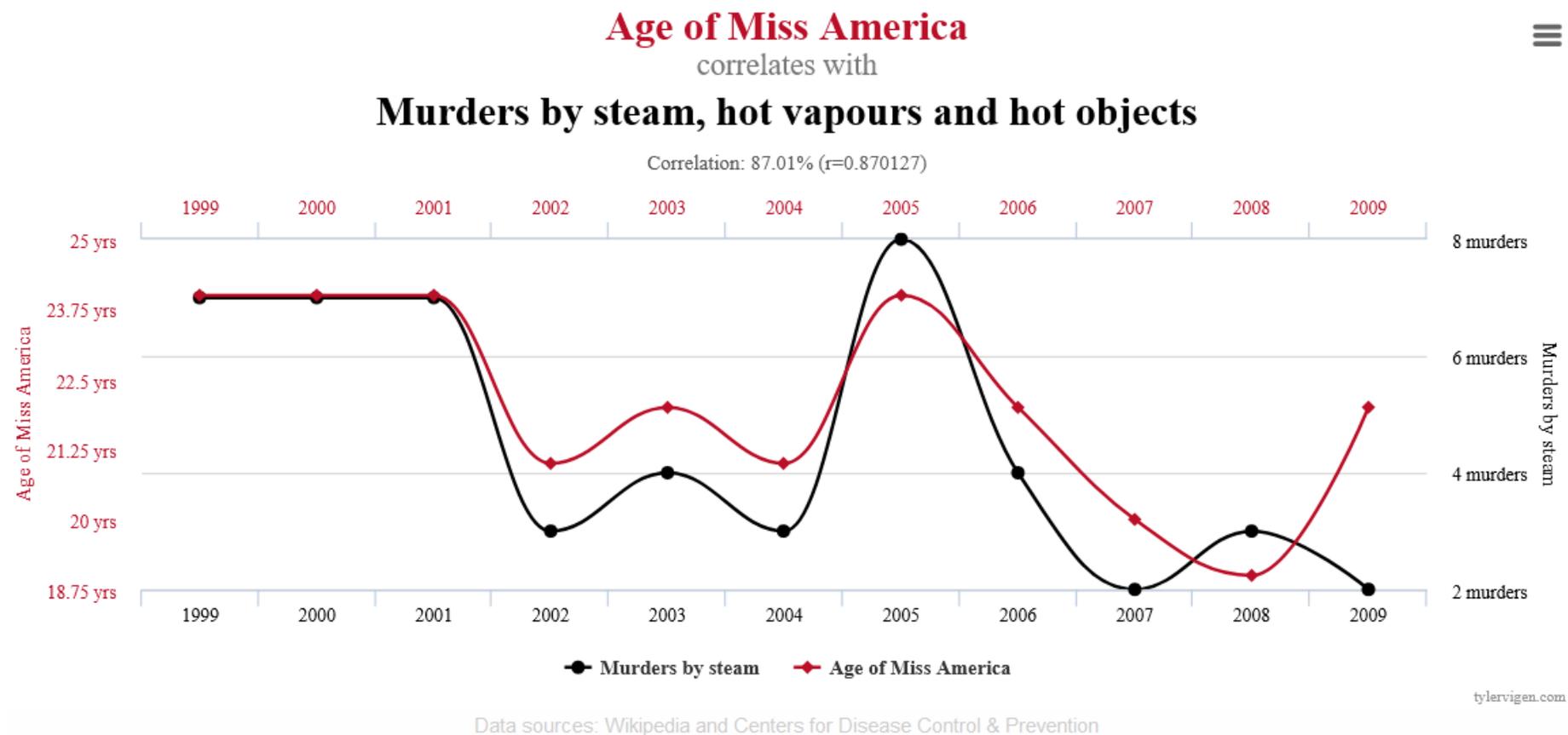
Quelle conclusion pouvez-vous en tirer sur l'impact de l'hôpital sur la santé ?

Le problème d'identification (exemple 2)

- A New York et à Chicago, on observe une étroite correspondance entre :
 - Le niveau de ventes de glaces (crèmes glacées)
 - Le nombre de meurtres

Y a-t-il un impact de la consommation de glaces sur les meurtres ?

Le problème d'identification (exemple 3)



Source: <https://tylervigen.com/spurious-correlations>

Risques de biais d'identification causale si on se contente de corrélations naïves

Variable confondante : Variable Z influence à la fois X et Y. Si on omet de tenir compte de Z, on conclut à une relation causale erronée.

- Exemple : La relation entre les ventes de glaces et les taux d'homicides est confondue par la température.

Causalité simultanée ou inverse : Lorsqu'on présuppose que c'est X qui influence B, alors que c'est B qui cause A, ou que X et Y s'influencent mutuellement.

- Exemple : La relation entre l'hospitalisation et l'état de santé

Coïncidence : X et Y peuvent être corrélés purement par hasard, en particulier si la taille de l'échantillon est petite ou qu'on teste un grand nombre de variables.

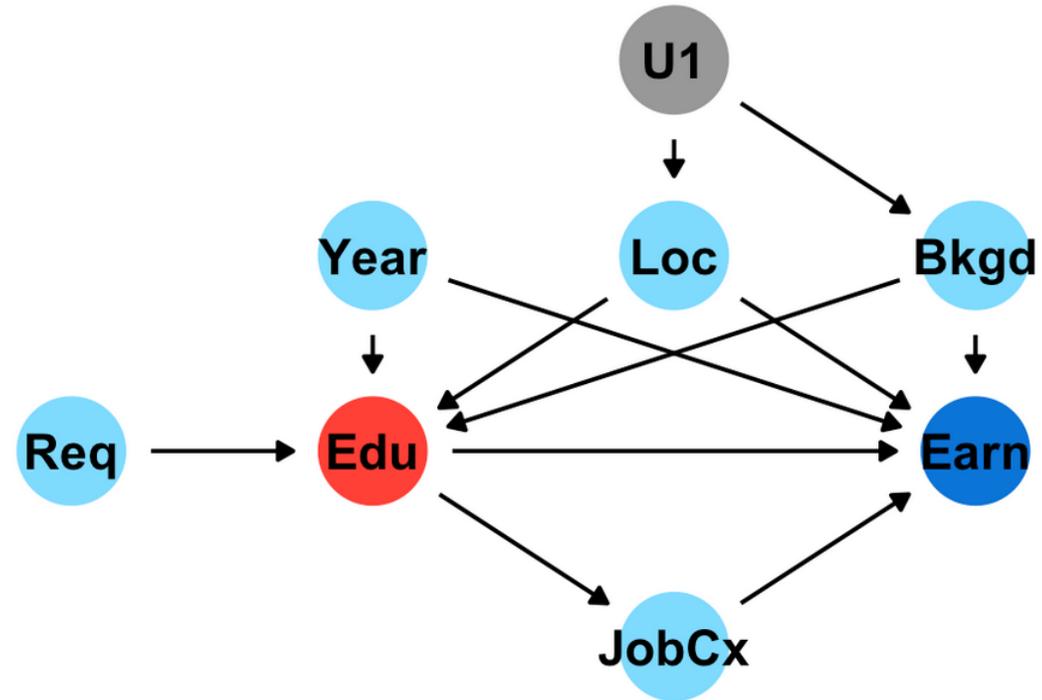
- Exemple : La relation entre l'âge des Miss America et les meurtres par objets brulants

Exercice

Essayez de trouver un exemple de corrélation fallacieuse que vous avez rencontré dans votre vie professionnelle ou personnelle.

Clarification sur les manières de formaliser des liens causaux

- Exemple de modes de notation différentes :
- Diagrammes acycliques orienté (DAGs)
- Équations



$$Earn_i = \alpha + \beta \times Edu_i + \gamma_1 \times Year_i + \gamma_2 \times Loc_i + \gamma_3 \times Bkg_i + \gamma_4 \times JobCx_i + \varepsilon_i$$

Les ingrédients d'une équation simple

Variables : représentées par des lettres, elles sont soit dépendantes (expliquées par d'autres) soit indépendantes (elles expliquent)

- Exemple: Y pourrait représenter le revenu et X pourrait représenter des années d'éducation

Coefficients : valeurs constantes qui déterminent la relation entre les variables

- Souvent représentés par des lettres grecques comme α ou β

Équations: relient les variables et les coefficients.

- Exemple: $Y = \alpha + \beta X$

Quelques compléments

Indices : indiquent des éléments spécifiques dans un ensemble.

- Exemple: Y_i représente le revenu de la i -ème personne.

Deltas : Le symbole Δ représente un changement ou une différence.

- Exemple: ΔY fait référence au changement dans Y

• **Chapeau** : indique une estimation ou prédiction.

- Exemple: \hat{Y} est une prédiction de Y à partir d'un modèle

• **Macron** : représenter la moyenne (ou l'espérance) d'une variable

- Exemple : \bar{X} est la moyenne de X

• **Erreurs** (ou résidus) : Souvent notées ε représentent l'écart entre les prédictions d'un modèle et les valeurs réelles.

Les « DAG »

Qu'est-ce qu'un DAG?

- Graphique acyclique dirigé pour représenter des liens causaux.
- Les flèches indiquent la direction de la causalité entre variables.
- Pas de cycles; la causalité se déplace dans une seule direction
- L'absence de flèche indique l'absence de lien causal

Intérêts principaux :

- Interface entre la modélisation et la théorie économique, l'expérience de terrain, l'intuition, la littérature, le quali.
- Clarifie la question d'intérêt, les variables pertinentes et les hypothèses
- Formalisme en graphe permet une résolution logique pour le choix des variables à incorporer dans un modèle (et des tests à effectuer pour valider le modèle)
- Etablit un pont entre différentes écoles empiriques

Exercice

Imaginez un schéma causal qui représente les facteurs associant déforestation et conservation

Le cadre d'analyse des résultats potentiels

Le fondement de l'étude de la causalité en économie

Le modèle causal de Rubin

- Nous voulons mesurer l'impact d'un "traitement" (intervention, programme) T sur une variable de résultat Y pour une unité i
- Statut de traitement (T_i) et de résultat (Y_i) de l'unité i
 - Si une unité reçoit le traitement : $T_i = 1$ et $Y1_i$
 - Si une unité ne reçoit pas le traitement: $T_i = 0$ et $Y0_i$
- Effet du traitement pour l'unité i : $\Delta Y_i = Y1_i - Y0_i$
- **Problème** : un seul état est observé pour i , on ne peut résolument pas observer les deux états pour la même unité, soit elle est traitée
Autrement dit : $Y0_i$ n'existe pas pour $T_i = 1$

Le modèle causal de Rubin

- Le problème fondamental de l'identification causale : seulement un des deux résultats potentiels (Y_{1i} ou Y_{0i}) peut être observé pour l'individu i
- Idéalement, il faudrait pouvoir comparer le résultat de l'individu traité au résultat qu'il aurait obtenu en l'absence de traitement :
 - C'est la définition du **contrefactuel**
- Comment constituer ce contrefactuel ? Une solution hypothétique : le clonage ? Des mondes parallèles ?

Deux clones dans des mondes parallèles

$$\begin{array}{c} \text{Personne} \\ + \\ \$ \\ = \\ \text{Personne} \\ \times 8 \text{ ans} \end{array}$$

$$\begin{array}{c} \text{Personne} \\ + \\ \$ \\ = \\ \text{Personne} \\ \times 5 \text{ ans} \end{array}$$

Déduire l'impact de la bourse sur la durée d'éducation serait aisé
Mais cette situation n'existe pas

Le modèle causal de Rubin

- Le défi empirique devient alors de constituer un **contrefactuel pertinent et valide**, c.-à-d. qui se rapproche le plus possible de ce qu'aurait été la situation de l'unité traitée en l'absence de traitement
- Nécessiter de changer d'échelle dans l'approche contrefactuelle :
Niveau de l'unité \Rightarrow Niveau d'un groupe d'unités
- Comme l'estimation de l'effet du traitement individuel (D_{yi}) n'est pas possible, on se concentre sur l'**effet moyen du traitement** (average treatment effect : ATE)
- Impossible de trouver deux unités identiques, mais possible de trouver deux groupes statistiquement proches, voire identiques en moyenne

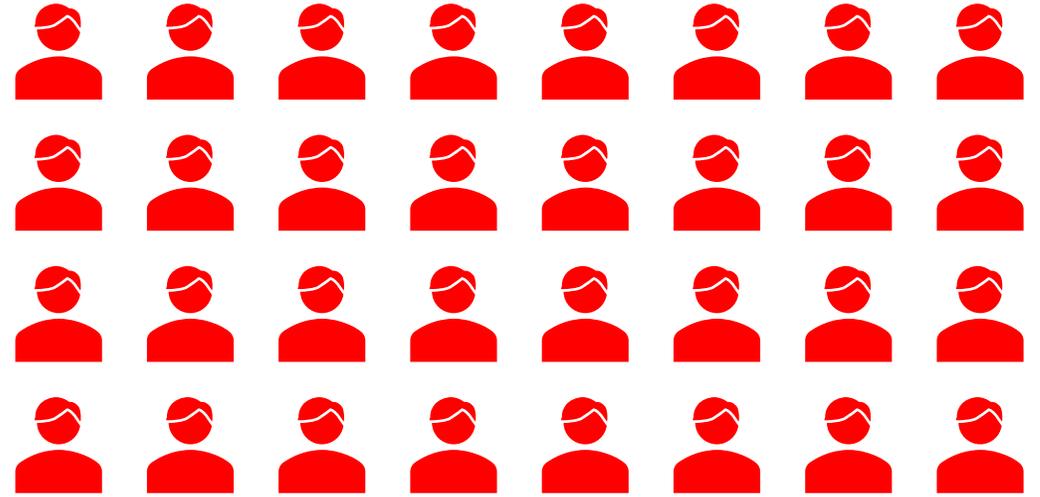
Le modèle causal de Rubin

Bénéficiaires
 $T_i = 1$



Résultat moyen des bénéficiaires
 $E(Y_i | T_i = 1) = E(Y1_i | T_i = 1)$

Non-bénéficiaires
 $T_i = 0$



Résultat moyen des non-bénéficiaires
 $E(Y_i | T_i = 0) = E(Y0_i | T_i = 0)$

Le modèle causal de Rubin

- Quel est l'impact du programme sur les bénéficiaires
- Effet moyen du traitement sur les traités (ATT)
- C'est la différence entre
 - Le résultat moyen des bénéficiaires:
 $E(Y1_i | T_i = 1)$...
 - ... et le résultat moyen des bénéficiaires, s'ils n'avaient pas été traités :
 $E(Y0_i | T_i = 1)$
- Estimation de l'effet sur les individus ciblés par le programme
- **Problème du contrefactuel** : $E(Y0_i | T_i = 1)$ n'est pas observable
- $E(Y0_i | T_i = 1) = E(Y0_i | T_i = 0)$
- **C'est toute la question !**

Modèle causal de Rubin

- En résumé, on va toujours comparer des unités non bénéficiaires en faisant comme si elles avaient été bénéficiaires
- Mais la question restera de déterminer dans quelle mesure les non-bénéficiaires sont identiques à des bénéficiaires qui n'auraient pas reçu de traitement
- Toute la boîte à outil de l'évaluation d'impact vise à constituer ce contrefactuel et à tester sa fiabilité

Discussion

Cette approche de contrefactuel vous semble-t-elle pertinente pour vos travaux ? Avez-vous déjà rencontré ou utilisé des méthodes s'appuyant sur cette approche contrefactuelle

Tour d'horizon rapide des méthodes

Méthodes contrefactuelle

Économétrie spatiale

Principales méthodes d'identification causale

Celles qu'on verra le plus en détail :

- Essais randomisés contrôlés (ERC, RCT en anglais)
 - Allocation aléatoire des individus dans les groupes traitement et contrôle.
- Appariement (Matching)
 - Jumelage d'individus traités avec des individus non traités ayant des caractéristiques observables similaires.
- Différence dans la différence (Diff-in-Diff)
 - Comparaison des écarts entre les groupes traitement et contrôle avant et après une intervention.
 - Aussi appelé Before-After-Control-Impact (BACI).

Principales méthodes d'identification causale

Celles qu'on abordera plus succinctement :

- **Discontinuité de la régression (Regression Discontinuity Design)**
 - Exploite un seuil prédéterminé d'attribution d'un traitement, permettant une comparaison juste au-dessus et juste en dessous de ce seuil.
- **Variables instrumentales (IVs)**
 - Utilise une variable (instrument) qui est corrélée avec la variable explicative d'intérêt, mais pas directement avec la variable dépendante, pour isoler la relation causale.
- **Contrôles synthétiques**
 - Constitue un "jumeau virtuel" à partir de plusieurs contrôles ("donor pool"). Exemple de la réunification allemande.

Économétrie spatiale : Une hybridation entre économie et géographie

- **Définition** : L'économétrie spatiale étudie la manière dont les données et les relations économiques sont influencées par la proximité géographique et spatiale des unités observées.
- Pourquoi est-ce important ? Dans de nombreux phénomènes économiques et sociaux, la localisation et les interactions spatiales jouent un rôle crucial.
- Première loi de la géographie selon Tobler : *"Tout interagit avec tout, mais deux objets voisins ont plus de chance d'entrer en interaction que deux objets éloignés."*

L'essence de la spatialité

- **Dépendance spatiale** : Lorsque la valeur d'une observation est influencée par les valeurs de ses voisines.
 - Exemple : La déforestation d'une zone accroît la probabilité de déforestation de ses voisines.
- **Hétérogénéité spatiale** : Chaque lieu peut avoir des caractéristiques spécifiques qui affectent le comportement des variables étudiées.
 - Exemple : Les différences climatiques entre régions influençant la production agricole.
- **Effet de bordure** : Considère comment les limites d'une zone peuvent influencer les résultats.
 - Exemple : des mécanismes de "fuite" de la déforestation à proximité immédiate d'une aire protégée.
- **Effet d'agrégation** : Se réfère à la manière dont les résultats peuvent changer selon l'échelle d'observation ou le découpage des zones.
 - Exemple : Analyser les taux de déforestation par région ou par pays peut donner des résultats différents

Discussion

Que vous inspirent ces approches ?

Travaux de groupe

Revue des articles

Revue croisée de deux articles

- Individuellement, lisez l'article qui vous a été assigné, parmi :
 - Desbureaux, et al. 2025. « Collaborative management partnerships strongly decreased deforestation in the most at-risk protected areas in Africa since 2000 ». *Proceedings of the National Academy of Sciences* 122 (1): e2411348121. <https://doi.org/10.1073/pnas.2411348121> .
 - Devenish, Katie, Sébastien Desbureaux, Simon Willcock, et Julia P. G. Jones. 2022. « On Track to Achieve No Net Loss of Forest at Madagascar's Biggest Mine ». *Nature Sustainability* 5 (6): 498-508. <https://doi.org/10.1038/s41893-022-00850-7>.
- Préparez des commentaires et questions :
 - L'article vous a-t-il paru facilement compréhensible ?
 - À quelle question d'évaluation l'article cherche-t-il à répondre ?
 - Avez-vous pu identifier les méthodes utilisées ?
 - Qu'avez-vous retenu des résultats présentés ?
 - Avez-vous identifié des limites de l'article ?