

Evaluation de projets de développement

Les variables instrumentales

Björn Nilsson

`bjorn.nilsson@universite-paris-saclay.fr`

Faculté Jean Monnet
Université Paris-Saclay

M2 GPDS, 1er semestre
Année universitaire 2024-2025

université
PARIS-SACLAY

Présentation à faire pour janvier

- Vous aurez à faire une présentation en binôme pour le mois de janvier.
 - Il s'agira de présenter de manière critique un article de recherche.
 - L'article utilisera une méthode d'évaluation d'impact pour répondre à une question liée à la lutte contre la pauvreté.
- 20 minutes (+ 10 minutes de discussion)
- Ce qui est attendu:
 - Présentation de la motivation, de la question de recherche qui est posée, et de la stratégie mise en oeuvre pour y répondre.
 - Veillez particulièrement à identifier les hypothèses derrière le résultat obtenu et de discuter leur crédibilité.
 - Discutez les résultats (portée et limites)

Plan du cours

- Séance 1 : Introduction
- Séance 2 : Mesurer la pauvreté
- Séance 3 : Lire des travaux d'évaluation d'impact
- Séance 4 : La différence de différence
- Séance 5 : La régression par discontinuité
- **Séance 6 : Les variables instrumentales**
- Séance 7 : Les essais aléatoires randomisés
- Séance 8 : Présentations d'articles
- Séance 9 : Présentations d'articles
- Séance 10 : Examen

Plan du chapitre

1. Avant-propos
2. Introduction IV
3. Hypothèses et estimation
4. Exemples

C'est quoi une variable instrumentale ?

Rappel : le problème fondamental de l'évaluation

- Les variables instrumentales ont deux utilisations en évaluation d'impact.
 - Contrôler pour des différences entre assignation à un traitement, et le fait d'être traité (ex Fuzzy RDD, RCT avec des non-compliers).
 - Constituer une source de variation exogène entre traité et non-traité en l'absence de randomisation ou de discontinuité.
- Le fond du problème : $Cor[T_i | \epsilon_i] = 0$ pas toujours vérifié.
- Ex: $Y_i = \alpha + \beta T_i + \epsilon_i$
 - Imaginons que Y est le revenu, T ([0,1]) est une formation.
 - **Possibilité:** Les gens qui assistent à la formation ont un niveau de motivation plus élevé que ceux qui n'assistent pas.
 - Cette motivation n'est pas mesurée (c'est difficile).
 - La motivation est une **variable omise**.

C'est quoi une variable instrumentale ?

Une variable omise est-ce grave ?

- La réponse dépend de la variable.
- Si c'est quelque chose qui est corrélé à la fois à ϵ_i et à T, **alors oui**.
- Si c'est quelque chose qui est corrélé à l'un mais pas à l'autre, **alors non**.
 - Il y a toujours plein de variables omises. La plupart du temps, elles ne sont pas inquiétantes.
 - Ex: dans l'exemple précédent, on n'introduit pas les goûts musicaux pour expliquer le revenu.
 - ▶ Pas grave, cela ne devrait pas être lié à ce qui nous intéresse.
- Revenons à la motivation.

C'est quoi une variable instrumentale ?

$$Y_i = \alpha + \beta T_i + \epsilon_i$$

- Dans l'équation ci-dessus, je pense que T est lié à la motivation.
 - Celle-ci n'étant pas mesurée, elle va être capturée par le terme d'erreur ϵ .
 - Quand T est élevé, ϵ est donc élevé.
 - Mais dans mes données, je n'ai que Y et T: l'effet de la motivation va être confondu avec l'effet de T (le traitement).
 - Autrement dit, on n'a aucun moyen de distinguer l'effet de la formation et celui de la motivation.
 - **Biais d'endogénéité** dû à un *problème de sélection*.
 - \Rightarrow C'est le problème fondamental de l'évaluation.
 - Il peut nous faire conclure à un effet quand il n'y en a pas, ou en l'absence d'effet quand il y en a un.

C'est quoi une variable instrumentale ?

Solution: les variables instrumentales

- Elles ont un rôle *instrumental* : elles ne nous intéressent pas directement, c'est leur capacité à résoudre le problème d'endogénéité qui est intéressante pour nous.
- **Principe** : la variable instrumentale va nous permettre d'utiliser une partie de la variation de T seulement, et cette partie ne sera pas corrélée à des variables omises qui menacent l'identification d'un effet.
- Comme on l'a déjà vu pour le Fuzzy design en RDD, la variable instrumentale permet en général de mesurer un effet moyen du traitement **local** = le **LATE** (Local Average Treatment Effect).
 - L'effet estimé vaut pour les personnes qui passent du groupe non-traité au groupe traité à **cause d'une variation de l'instrument**.

C'est quoi une variable instrumentale ?

Comment ça fonctionne ?

- On explique **T** par une autre variable **Z** (l'instrument).
- Dans l'exemple précédent, il est essentiel que **Z explique T**, mais qu'il ne soit pas lié à la motivation.
- On utilise la variation en **T expliquée par Z** dans la régression principale.
 - Si **Z** n'est pas lié à la motivation, alors la variation de notre nouvelle **T**, \hat{T} , n'est pas non plus corrélée à la motivation, et ainsi n'est pas corrélée à ϵ .
 - \Rightarrow Problem solved!

Hypothèses de la méthode

Hypothèse de pertinence

- La variable instrumentale qu'on propose doit bien expliquer la participation au groupe traité.
 - $Cor(X, Z) \neq 0$
- Exemple de la formation: la couleur des cheveux est probablement un mauvais instrument, car *non-pertinent*.
- L'instrument doit en outre avoir un certain pouvoir explicatif
 - S'il n'est que trop faiblement corrélé au traitement, cela crée un biais (le coefficient se rapproche du coefficient naïf en MCO).
 - On peut mesurer le pouvoir explicatif de l'instrument pour expliquer T. C'est la statistique F. Une règle heuristique veut que la F-stat soit au moins proche de 10.

Hypothèses de la méthode

Tableau avec F-stat

Table 4. 2SLS models of health outcomes

| | Model I | | Model II | |
|--|-----------------------|----------------------|-----------------------|---------------------------|
| | 1st stage: Ln(Income) | 2nd stage: Ln(BMI) | 1st stage: Ln(Income) | 2nd stage: No. conditions |
| Ln(rainfall) | 2.83*** (0.36) | – | 2.84*** (0.35) | – |
| Ln(income) | – | 0.01* (0.003) | – | –0.15*** (0.05) |
| Age | –0.02 (0.02) | –0.002 (0.002) | 0.02 (0.02) | 0.01 (0.03) |
| Age squared | 0.0002 (0.0002) | –0.00004** (0.00) | –0.0002 (0.0002) | 0.00001 (0.0002) |
| Age*Female | 0.02* (0.01) | 0.002* (0.001) | 0.02* (0.01) | 0.03** (0.02) |
| HH size | 0.02 (0.02) | –0.001 (0.001) | 0.02 (0.02) | 0.01 (0.01) |
| No. of Observations | | 16,083 | | 16,105 |
| No. of individuals | | 4,845 | | 4,848 |
| Kleibergen-p F-stat ($H_0 =$ weak IV) | | 63.72 | | 64.05 |

Note: Std. errors in () clustered by villages. All models include year-cluster interaction effects, a month linear time trend, and individual fixed effects.

*** $p < 0.01$.

** $p < 0.05$.

* $p < 0.10$.

Source: Fichera & Savage (2014)

Hypothèses de la méthode

Critère d'exclusion

- On dit que l'instrument doit être exogène.
- C'est-à-dire qu'on veut qu'il explique T , mais absolument pas ϵ .
 - $Cor(Z, \epsilon) = 0$.
- Rappel : dans ϵ , on a les effets de l'ensemble des variables omises.
 - Exemple formation : on considère l'âge comme un instrument, et on voit que l'âge explique bien le fait d'avoir participé.
 - **L'âge est-il un bon instrument ?**

2SLS—Doubles moindres carrés ordinaires

- Une façon d'estimer un effet avec une variable instrumentale est les doubles moindres carrés ordinaires, ou 2SLS.
- On s'intéresse toujours à $Y_i = \alpha + \beta T_i + \epsilon_i$.
- On procède en deux temps:
 - $T_i = \delta + \gamma_1 Z_i + \omega_i$.
- Cette régression nous donne donc δ et γ .
- On crée $\hat{T}_i = \delta + \gamma Z_i$.
- Que contient \hat{T} ?

2SLS—Doubles moindres carrés ordinaires

- Une fois qu'on a \hat{T} , on l'utilise simplement pour estimer l'effet du traitement.
- NB: $T_i = \hat{T}_i + \omega_i$
- $\Rightarrow Y_i = \alpha + \beta T_i + \epsilon_i = \alpha + \beta(\hat{T}_i + \omega_i) + \epsilon_i$.
- $Y_i = \alpha + \beta \hat{T}_i + (\beta \omega_i + \epsilon_i)$
- Autrement dit, quand on régresse *la valeur prédite* de T sur Y on obtient quand même le vrai effet de T sur Y.

Estimation

Implémentation avec données fictives

| Y | T | Z |
|----|---|----|
| 11 | 1 | 3 |
| 9 | 0 | 8 |
| 16 | 1 | 5 |
| 8 | 0 | 11 |

- Avec **Y** le revenu, **T** le fait de suivre la formation et **Z**, la distance du logement au centre de formation.
- Avec T et Y on obtient $\beta = 5$ pour l'effet de T sur Y.
- Mais on a un doute sur cet effet: on mobilise une variable instrumentale : la distance du logement au centre de formation.
 - Cet instrument est-il bon au regard des deux hypothèses ?

Estimation

Implémentation avec données fictives

- On regarde d'abord si l'instrument prédit bien le traitement.

First-stage regressions

Number of obs = 4
F(1, 2) = 9.31
Prob > F = 0.0927
R-squared = 0.8231
Adj R-squared = 0.7347
Root MSE = 0.2974

| T | Coefficient | Std. err. | t | P> t | [95% conf. interval] | |
|-------|-------------|-----------|-------|-------|----------------------|----------|
| Z | -.1496599 | .0490551 | -3.05 | 0.093 | -.360727 | .0614073 |
| _cons | 1.510204 | .3629747 | 4.16 | 0.053 | -.0515502 | 3.071958 |

Implémentation avec données fictives

- Le premier critère semble donc plutôt validé.
- Quid du second? Rappel: $Cor(Z_i, \epsilon) = 0$
- Le critère d'exclusion est impossible à tester—il faut convaincre le lecteur qu'il n'y a pas de raison de le remettre en cause.
- Exemples de points d'interrogation ici:
 - Lien entre distance et urbain/rural ?
 - Lien entre distance et la ville dans laquelle les individus habitent... un problème?
 - Lien entre distance et composition sociodémographique au sens général.

Estimation

Implémentation avec données fictives

- Ensuite on régresse Y sur \hat{T} et on récupère le vrai effet du traitement sur Y .

```
Instrumental variables 2SLS regression
```

| | | |
|---------------|---|--------|
| Number of obs | = | 4 |
| Wald chi2(1) | = | 4.68 |
| Prob > chi2 | = | 0.0306 |
| R-squared | = | 0.6472 |
| Root MSE | = | 1.8306 |

| Y | Coefficient | Std. err. | z | P> z | [95% conf. interval] |
|-------|-------------|-----------|------|-------|----------------------|
| T | 4.363636 | 2.017756 | 2.16 | 0.031 | .4089066 8.318366 |
| _cons | 8.818182 | 1.362221 | 6.47 | 0.000 | 6.148278 11.48809 |

Instrumented: T

Instruments: Z

Implémentation avec données fictives

- Le vrai effet, une fois une variable instrumentale utilisée, est égale à 4.4.
- On a trouvé 5 initialement.
 - Biais positif du coefficient initial.
 - Il y a bien quelque chose qui est positivement corrélée au revenu qui explique aussi le traitement.
- Ici utilisé pour résoudre un problème de **variable omise**.
 - Fonctionne aussi comme remède aux problèmes de causalité inverse et d'erreur de mesure.
 - **Causalité inverse** si le revenu explique le recours à la formation.
 - Important alors de vérifier que la distance du lieu de formation n'est pas corrélée au revenu.

Estimation

Pour résumer

- Une variable instrumentale est donc une variable qui explique la participation dans le groupe traité.
- Mais *qui n'explique pas Y autrement qu'à travers la participation*.
- Cette dernière condition, dite critère d'exclusion, n'est pas testable.
 - Elle doit être défendue par le raisonnement.
- Si les conditions de validité de l'instrument sont respectées, on récupère le vrai effet du traitement sur la variable d'intérêt Y.
- Cet effet doit être compris comme un effet local (**LATE**), i.e. valable pour les individus dont T varie à cause de Z.
 - Autrement dit, on ne connaît pas l'effet sur quelqu'un qui suivrait ou non la formation peu importe la distance.

Estimation

Jugez les instruments suivants: #1

- Un chercheur propose d'étudier le lien entre l'éducation et le vote aux élections parlementaires.
- La variable Y est donc "a voté aux dernières élections parlementaires", et la variable X est "nombre d'années d'études post-bac".
- On soupçonne un problème de variable omise.
 - Les études expliquent la forme de notre réseau d'amis, qui explique le vote.
- Le chercheur propose la variable instrumentale "éducation des parents".
 - Arguant que cela explique bien l'éducation des individus sans expliquer par d'autres canaux le vote.
 - Qu'en pensez-vous ?

Estimation

Jugez les instruments suivants: #2

- Un chercheur propose d'étudier le lien entre l'activité sportive et la santé cardiovasculaire.
- La variable Y est la tension artérielle, et la variable X est "nombre d'heures d'activité physique par semaine".
- On soupçonne encore un problème de variable omise (**comme?**).
- Ou encore, un problème de causalité inverse (j'ai une tension élevée, alors j'ai peur de faire du sport).
- Le chercheur propose la variable instrumentale "densité d'équipement sportifs du quartier"
 - Qu'en pensez-vous ?

Exemple: AJR (2001)

Mortalité des colons et institutions

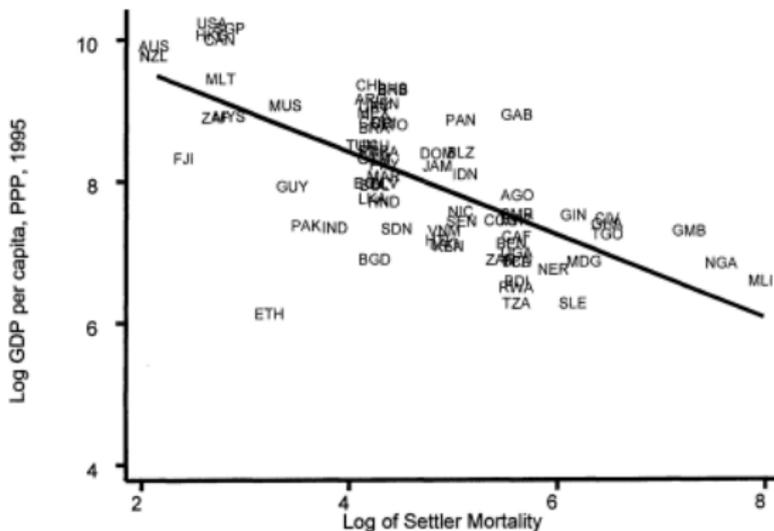
- Article important dans un courant de recherche essayant de comprendre la formation et l'importance des institutions pour le développement.
 - Prix 'Nobel' en 2024: Acemoglu, Johnson, Robinson.
- Probablement l'article le plus utilisé quand on présente les variables instrumentales.
- **Question** : Est-ce que de bonnes institutions sont la cause du développement ?
 - Ex : la protection du droit de propriété.
- Simplement régresser le développement (croissance) sur le droit de propriété n'est pas suffisant.
 - Risque de causalité inverse : c'est parce qu'un pays est riche qu'il peut mettre en place et protéger des institutions de qualité.

Exemple: AJR (2001)

Mortalité des colons et institutions : hypothèses

1. Il y a des variétés de colonisation différentes :
 - Des colonies extractives (Congo Belge).
 - A l'opposé, des "Néo-Europe" (Etats-Unis, Canada, Australie, Nouvelle-Zélande)
2. La stratégie de colonisation est influencée par l'environnement sanitaire (la prévalence de maladies)
 - Pas de Néo-Europe dans des endroits où les colons européens faisant face à des taux de mortalité élevés.
3. Les institutions coloniales ont persisté à la fin de la colonisation.

Exemple: AJR (2001)



Source : Acemoglu, Johnson, Robinson (2001)

Exemple: AJR (2001)

- Le cheminement causal qu'ils ont en tête est le suivant : mortalité faible des colons -> colonie néo-europe -> institutions fortes établies -> institutions fortes à l'indépendance -> développement plus rapide.
- Est-ce que la mortalité des colons était bien connue et susceptible d'affecter la façon de coloniser ?
 - Afrique de l'ouest : pour les premiers colons, mortalité annuelle de l'ordre de 50%.
 - Sujet couvert par la presse à l'époque.
 - Le comité qui devait décider d'où envoyer les prisonniers britannique a tranché en faveur de l'Australie plutôt que la Gambie en raison de la mortalité.

Exemple: AJR (2001)

Schématiquement, les deux types de colonies :

- **Colonies d'implantation**

- Beaucoup d'européens
- Vie calquée sur la vie dans le pays d'origine
- Institutions favorisant le commerce et la propriété privée des colons

- **Colonies extractives**

- Objectif principal : obtenir des ressources naturelles (or, minéraux) ou des esclaves
- Peu de préoccupation quant à la structure sociétale et le pouvoir étatique
- Violence, destruction de structures politiques

Exemple: AJR (2001)

Pourquoi les institutions persistent-elles ?

- Coût de changement institutionnel (compenser les perdants)
- Gouvernance au jour le jour souvent délaissé aux élites locales, restées en place
- Investissements complémentaires à certaines institutions

Exemples de persistance

- Travail forcé en Amérique latine (plutôt colonies extractives)
 - Esclavage aboli au Brésil en 1886 (indépendance 1822).
 - Travail forcé réintroduit au Mexique jusqu'en 1910 (indépendance 1821), et au Guatemala jusqu'en 1945.
- Droit de propriétés forts préservés au Singapour, Hong-Kong (plutôt Néo-Europe).

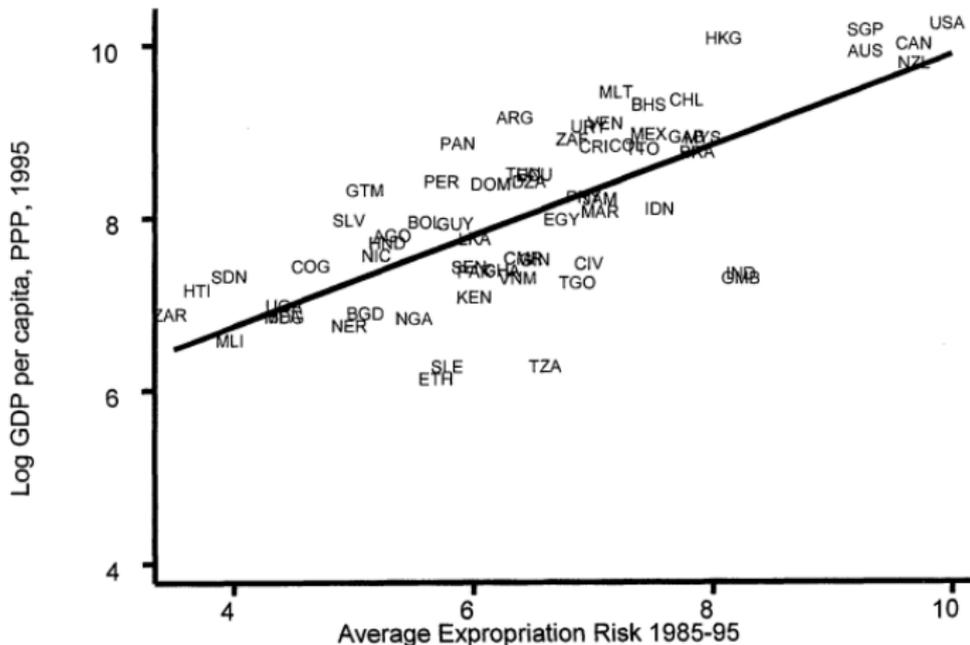
Exemple: AJR (2001)

Les données

- PIB par tête en 1995
- Indice de protection contre l'expropriation (1985-1995).
- Polity III
- Mesure d'institutions tôt
 - Soit à l'indépendance, soit en 1900.
- Fraction de population d'ascendance européenne.
- Mortalité des colons

Exemple: AJR (2001)

Lien entre droit de propriété et richesse



Exemple: AJR (2001)

Lien causal ?

- Risque de causalité inverse et de variables omises
- Instrument pour le risque d'expropriation : la mortalité des colons
- Est-ce un bon instrument ?
 - $Cor(X, Z) \neq 0$?
 - $Cor(Z, u) = 0$?
 - Encore une fois, le premier critère est visible dans la régression de première étape, tandis que la deuxième ne peut pas être formellement testé.
 - ▶ Voyons ce qui pourrait menacer la validité de cette deuxième hypothèse
 - ▶ Il ne faut pas que l'instrument influence la variable Y **autrement** que via X.

Exemple: AJR (2001)

Mortalité des colons affecte la croissance directement ?

- Si environnement avec plus de maladies, moins de croissance ?
 - AJR expliquent que 80% des décès de colons sont dûs à la fièvre jaune et au paludisme, et que les taux de mortalité des locaux pour ces maladies sont très faibles.
 - La malaria n'est par ailleurs pas corrélée avec des mesures géographiques simples, mais dépend beaucoup du microclimat, de la température et l'humidité.
- Mortalité infantile élevée -> liée à la croissance plus tard...?

Exemple: AJR (2001)

Le résultat principal

| | Base sample (1) | Base sample (2) | Base sample without Neo-Europes (3) | Base sample without Neo-Europes (4) | Base sample without Africa (5) | Base sample without Africa (6) | Base sample with continent dummies (7) | Base sample with continent dummies (8) | Base sample, dependent variable is log output per worker (9) |
|---|-----------------|-----------------|-------------------------------------|-------------------------------------|--------------------------------|--------------------------------|--|--|--|
| Panel A: Two-Stage Least Squares | | | | | | | | | |
| Average protection against expropriation risk 1985–1995 | 0.94 (0.16) | 1.00 (0.22) | 1.28 (0.36) | 1.21 (0.35) | 0.58 (0.10) | 0.58 (0.12) | 0.98 (0.30) | 1.10 (0.46) | 0.98 (0.17) |
| Latitude | | -0.65 (1.34) | | 0.94 (1.46) | | 0.04 (0.84) | | -1.20 (1.8) | |
| Asia dummy | | | | | | | -0.92 (0.40) | -1.10 (0.52) | |
| Africa dummy | | | | | | | -0.46 (0.36) | -0.44 (0.42) | |
| “Other” continent dummy | | | | | | | -0.94 (0.85) | -0.99 (1.0) | |

Source : Acemoglu, Johnson, Robinson (2001)

Exemple: AJR (2001)

Résumé et critiques

- Le résultat suggère donc qu'une bonne protection contre l'expropriation assure plus de croissance à long terme.
- **Beaucoup de critiques de ce papier**
- Sur les données (intrapolations de mortalité, adéquation de "l'indice d'expropriabilité")
- Sur l'agrégation des histoires coloniales (Amérique Latine et Afrique, deux périodes très différentes).
- **Sur la validité de l'instrument :**
- Le même exercice peut être mené pour le capital humain au lieu des institutions: les colons sont venus avec des institutions, mais aussi avec du capital humain et des connaissances.
- L'instrument mortalité des colons explique aussi l'environnement sanitaire aujourd'hui. ⇒ lié à l'espérance de vie, à l'investissement, à l'attractivité économique, etc.

Wooclap