

Variables instrumentales

Introduction aux méthodes d'évaluation d'impact
des politiques publiques

Philippe De Vreyer

Quand va-t-on employer cette
méthode ?

Conditions d'emploi de la méthode

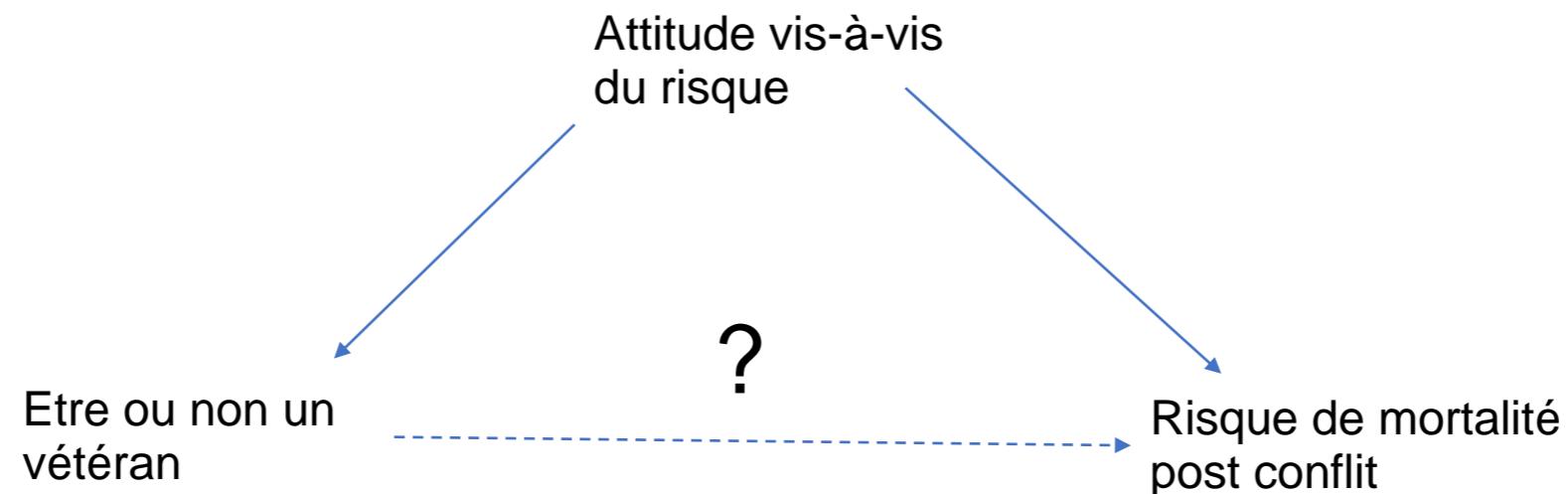
- Quand on suspecte la présence de caractères inobservés des individus qui expliquent à la fois la participation au programme et la valeur prise par la variable d'intérêt (indépendamment de la participation). On parle d'hétérogénéité inobservée.
- Quand cette hétérogénéité ne peut être retirée par une méthode de double différence.
- Quand on dispose *d'une ou de plusieurs sources exogènes de variation permettant d'expliquer la participation* (instruments)

Principe de la méthode

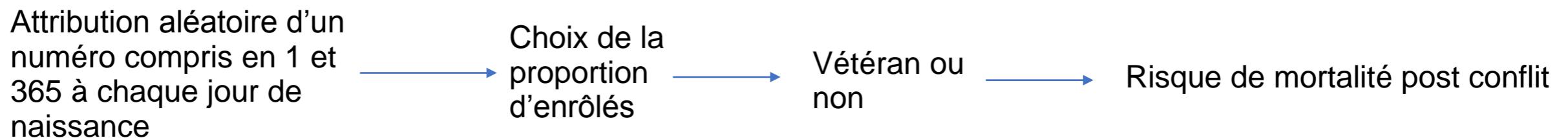
- Angrist et Imbens (1996): effet d'être un vétéran du Vietnam sur la mortalité une fois revenu à la vie civile.
- Principe: aux USA, comparer les risques de mortalité d'individus de caractéristiques semblables, à ceci près que certains ont pris part à la guerre du Vietnam et les autres pas.
- Est-ce que les vétérans ont, de ce fait, un risque de mortalité différent du reste de la population ?
- Problème: les vétérans peuvent s'être autosélectionnés.

- Qui part à la guerre ?
 - Il y a les « risque tout » qui recherchent l'aventure et veulent se battre (« always takers »).
 - A l'opposé on trouve les « prudents et les pacifistes » qui font tout pour ne pas avoir à se battre (« never takers »).
 - Au milieu il y a les « conformistes » qui vont se battre s'ils y sont contraints et n'y vont pas sinon (« compliers »)
- Quel lien avec le risque de mortalité ?
 - Les « risque tout » prennent plus de risque et ont par conséquent un risque de mortalité plus élevé, même revenus à la vie civile
 - Les « prudents » sont...prudents et ont de ce fait un risque de mortalité moins élevé.

- Si seuls les « risque tout » vont se battre et si seuls les « prudents » n'y vont pas (autrement dit s'il n'y a pas de conformistes) il n'est pas possible d'évaluer l'impact du statut de vétéran sur le risque de mortalité, car celui-ci se confond avec l'attitude vis-à-vis du risque:



- Pour identifier **il faut que l'on observe des « conformistes »**, c'est-à-dire des individus dont la participation a une origine exogène, contre laquelle ils n'ont rien pu faire.
- Pour ces individus, toute différence entre la valeur moyenne du taux de mortalité observée entre les vétérans et les autres peut être attribuée au statut de vétéran lui-même.
- Angrist et Imbens utilisent le tirage au sort des conscrits comme origine exogène de variation dans le statut de vétéran:



Théorie

Notations

- Comme nous en avons maintenant l'habitude notons:
 - Y_i la valeur de la variable d'intérêt pour l'individu i
 - T_i une indicatrice prenant la valeur 1 si l'individu participe (est « traité ») et 0 sinon
 - X_i un ensemble de caractéristiques observables de l'individu (éducation, âge, sexe etc...)
- La relation entre la participation au traitement, les caractéristiques observables de l'individu et la valeur prise par la variable d'intérêt Y est donnée par l'équation:

$$Y_i = \alpha + \beta X_i + \gamma T_i + \varepsilon_i$$

- ε_i terme aléatoire qui mesure tout ce qui influe sur la valeur de Y_i et que l'on observe pas.
- γ effet du traitement sur l'individu
- **Problème:** on ne peut pas identifier la valeur du coefficient γ lorsque ε_i et T_i sont corrélés. On dit dans ce cas que la variable T_i est **endogène**.

Origines de l'endogénéité

- **Cause commune:** par exemple: attitude vis-à-vis du risque dans l'analyse de l'impact du statut de vétéran sur le risque de mortalité.
- **Causalité inverse:** la participation au traitement est déterminée par la valeur de la variable d'intérêt
- **Erreur de mesure** sur T_i

Remède à l'endogénéité

- Quelle que soit la source, le remède est le même: on peut arriver à identifier l'effet du traitement lorsque l'on dispose d'une **source exogène de variation qui influe sur la probabilité d'être traité sans relation directe avec la variable d'intérêt.**
- Une telle source est appelée *un instrument.*

Principe de l'instrumentation

- Notons donc Z_i la valeur prise par l'instrument Z pour l'individu i .
- Le principe est d'*extraire* des variations de T_i ce qui a une origine exogène, donc ce qui provient de Z_i .
- Puis d'utiliser cette partie exogène de T_i pour évaluer l'impact du traitement sur la variable d'intérêt.
- Cette méthode est aussi connue sous le nom de « méthode des doubles moindres carrés ».

Méthode instrumentale ou des doubles moindres carrés

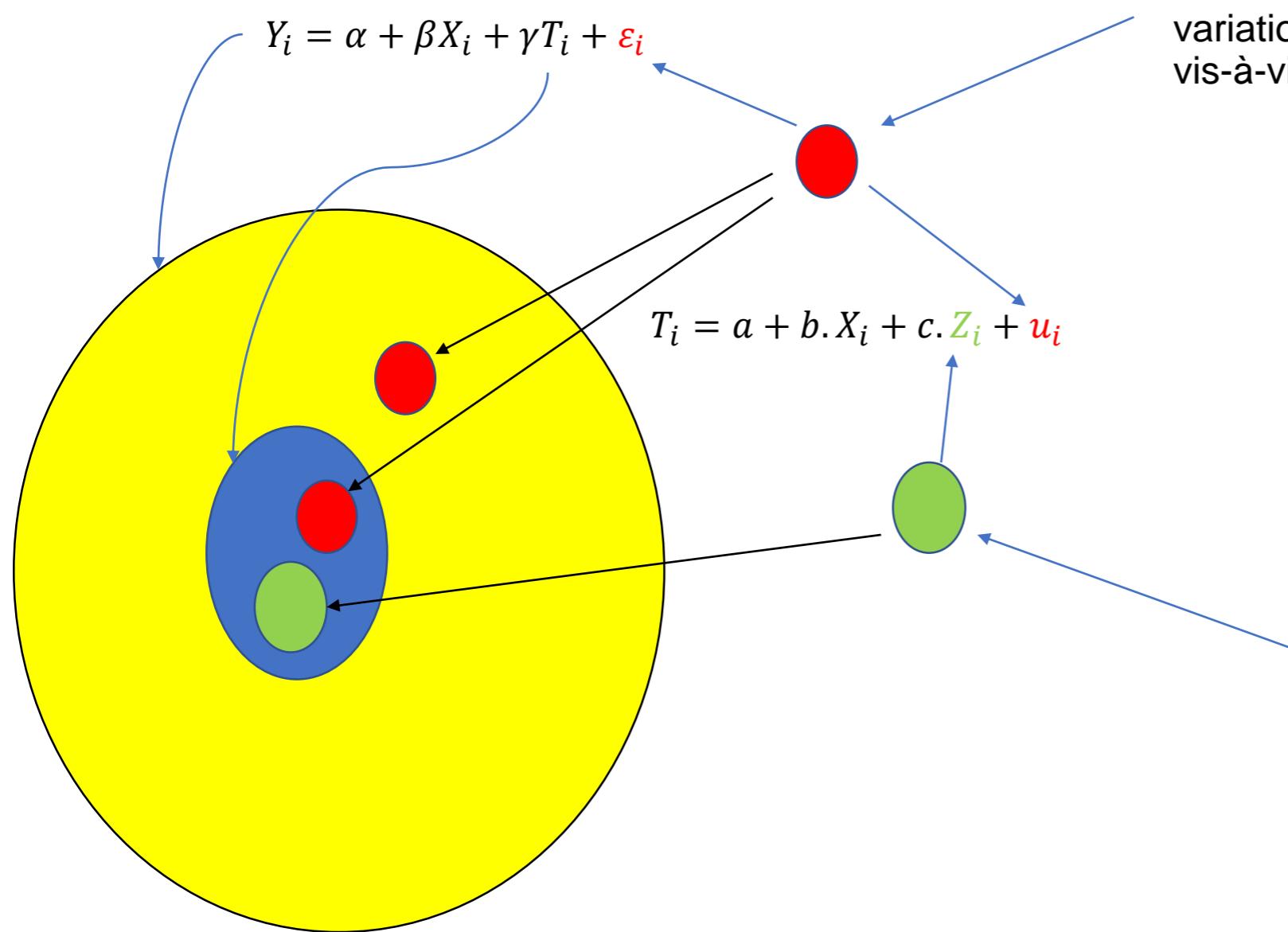
$$T_i = a + b \cdot X_i + c \cdot Z_i + u_i$$

Effet de l'instrument Z_i sur la sélection dans le traitement, net de celui des autres variables exogènes

$$\hat{T}_i = \hat{a} + \hat{b} \cdot X_i + \hat{c} \cdot Z_i$$

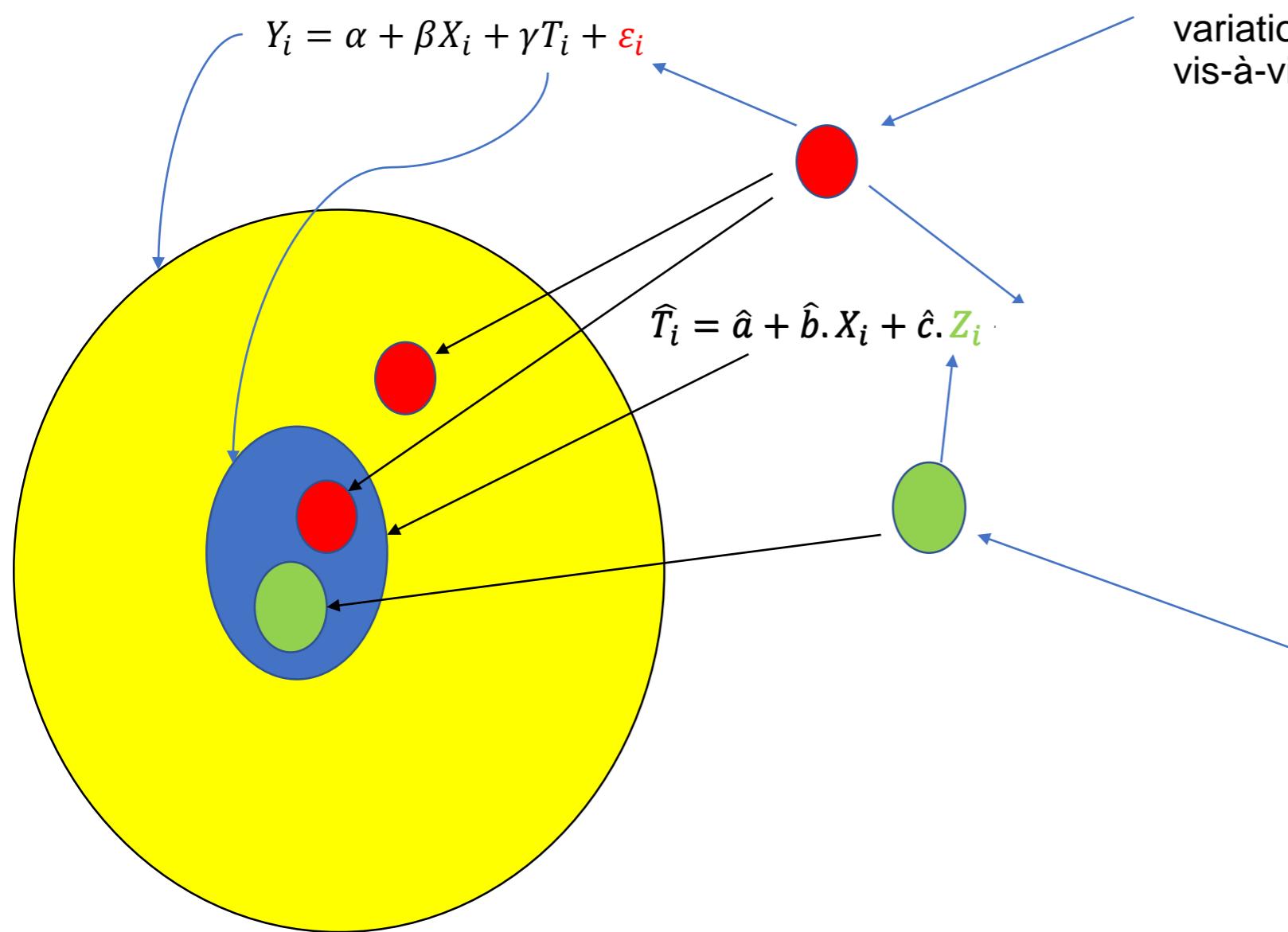
Seule la « partie exogène » de T_i est conservée: l'impact de T_i sur Y_i est estimé sans biais.

$$Y_i = \alpha + \beta X_i + \gamma \hat{T}_i + \varepsilon_i$$



Source endogène de
variation (ex. attitude
vis-à-vis du risque)

Source exogène de
variation (ex.
affectation aléatoire
par tirage au sort)



Source endogène de variation (ex. attitude vis-à-vis du risque)

Source exogène de variation (ex. affectation aléatoire par tirage au sort)

Compléments et limites

Qualité des instruments

- Deux propriétés essentielles:
 - Exogénéité
 - Puissance

Exogénéité

$$Y_i = \alpha + \beta X_i + \gamma T_i + \varepsilon_i$$

$$Z_i$$

- Permet de reproduire les conditions du tirage au sort
- Garantit l'absence de biais dans l'estimation de l'effet du traitement

Sources d'instruments

- Les expériences naturelles constituent les meilleurs instruments:
 - Sexe des deux premiers nés pour instrumenter la probabilité d'avoir un troisième enfant lorsqu'on en a déjà deux
 - Trimestre de naissance pour instrumenter le nombre d'années d'éducation dans une régression visant à évaluer les rendements de l'éducation sur le marché du travail
 - Etc.
- Des instruments différents conduisent à des résultats différents

Puissance

- L'instrument doit contribuer de façon importante à expliquer le traitement



Si: et

l'instrument n'est pas parfaitement exogène

l'instrument est peu puissant

Biais **plus important** avec la méthode instrumentale

A blue bracket groups the two conditions above it, and another blue bracket groups these with the resulting bias statement. A blue arrow points from the warning sign to the word 'Si:'.

Que faire si plus d'un instrument est disponible ?

- Tous les utiliser (et non pas choisir)
- Principe de la méthode et conditions d'exogénéité et de puissance identiques
- Ne conserver que les meilleurs si certains s'avèrent peu puissants

Limites

- Grands échantillons nécessaires
- Portée limitée de l'estimateur:
 - Estimateur LATE (Local Average Treatment Effect): l'évaluation ne vaut que pour ceux qui sont traités en raison de l'origine exogène représentée par Z_i .
 - Problème moins aigu quand plus d'un instrument est disponible.