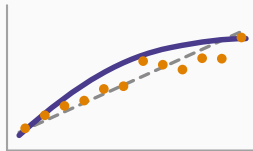


Économétrie avancée

Chapitre 15 : Variables instrumentales et doubles moindres carrés



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Avant de commencer

Le problème d'endogénéité

L'estimateur par variables instrumentales

Doubles moindres carrés

Tester la validité

Synthèse

Avant de commencer

Pourquoi ce chapitre ?

Le panel (chapitres 13-14) neutralise l'hétérogénéité inobservée **constante dans le temps**. Mais l'endogénéité a d'autres visages — variable omise variable dans le temps, **erreur de mesure**, **simultanéité** — contre lesquels le panel ne peut rien. La méthode des **variables instrumentales** (VI) offre une solution générale : elle exploite une variable extérieure, l'**instrument**, pour retrouver l'effet causal quand les MCO échouent.

Le problème d'endogénéité

Quand les MCO deviennent non convergents

Régresseur endogène

Un régresseur x est **endogène** s'il est corrélé au terme d'erreur : $\text{Cov}(x, u) \neq 0$. Les MCO sont alors **biaisés et non convergents** : ils attribuent à x une part de variation qui revient en réalité à des facteurs inobservés.

Trois sources classiques

- (1) **Variable omise** corrélée à x (ex. capacité inobservée dans l'équation de salaire);
- (2) **erreur de mesure** sur x (le régresseur observé diffère du vrai);
- (3) **simultanéité**, quand y et x se déterminent mutuellement (offre-demande). Dans les trois cas, il faut un levier externe.

L'estimateur par variables instrumentales

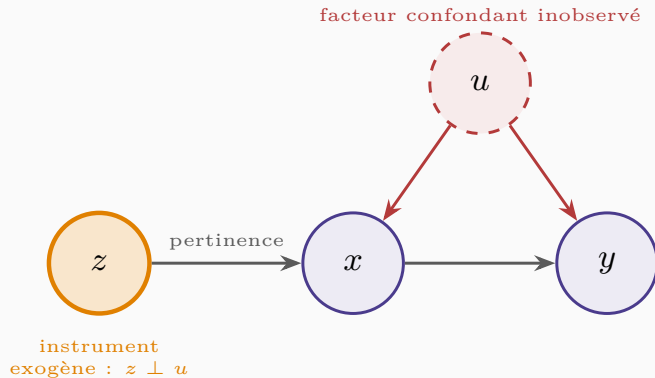
Pertinence et exogénéité

Une variable z est un **instrument valide** pour x si :

- **pertinence** — z est corrélé au régresseur endogène : $\text{Cov}(z, x) \neq 0$ (vérifiable et testable);
- **exogénéité** — z n'est pas corrélé à l'erreur : $\text{Cov}(z, u) = 0$ (hypothèse d'exclusion, non testable directement).

L'estimateur VI vaut alors $\hat{\beta}_1^{\text{VI}} = \frac{\text{Cov}(z, y)}{\text{Cov}(z, x)}$.

Les deux conditions d'un bon instrument (suite)



Exemple : rendement de l'éducation

Estimer l'effet de l'éducation sur le salaire par MCO est biaisé (la capacité, inobservée, gonfle les deux). On instrumente l'éducation par une variable qui l'influence **sans** agir directement sur le salaire — proximité d'une université, trimestre de naissance. L'instrument isole la part « exogène » de l'éducation.

Doubles moindres carrés

Estimateur 2SLS

Avec plusieurs instruments, la méthode des **doubles moindres carrés** (2SLS) procède en deux étapes : (1) régresser le régresseur endogène sur **tous** les instruments et les régresseurs exogènes, et récupérer la valeur ajustée $\hat{x}$; (2) régresser y sur $\hat{x}$ et les régresseurs exogènes. La partie « prédite » $\hat{x}$ n'utilise que la variation exogène, ce qui rétablit la convergence.

MCO comme cas particulier

Si le régresseur suspect est en fait exogène, il est son propre instrument et 2SLS **coïncide** avec les MCO. 2SLS est donc convergent mais **moins précis** que les MCO : on ne l'emploie que si l'endogénéité est réelle. Un instrument trop faiblement corrélé au régresseur (**instrument faible**) rend l'estimateur très imprécis et biaisé en petit échantillon.

Tester la validité

Trois diagnostics

- **Endogénéité** (Durbin-Wu-Hausman) : comparer MCO et 2SLS; une différence significative confirme qu'il fallait instrumenter.
- **Suridentification** (Sargan/Hansen) : quand il y a plus d'instruments que de régresseurs endogènes, on teste conjointement leur exogénéité.
- **Instruments faibles** : la statistique F de première étape doit être élevée (règle empirique $F > 10$) pour garantir la pertinence.

La discipline de l'instrument

Un bon travail empirique par VI se juge à la **crédibilité de l'exclusion** : pourquoi l'instrument n'agirait-il sur y que via x ? Cette justification est un argument économique, pas un test — d'où l'importance d'instruments bien motivés et, quand c'est possible, d'un test de suridentification à l'appui.

Synthèse

Synthèse du chapitre

Face à un régresseur **endogène** (variable omise, erreur de mesure, simultanéité), les MCO sont non convergents. Les **variables instrumentales** rétablissent la convergence à l'aide d'un instrument **pertinent** ($\text{Cov}(z, x) \neq 0$) et **exogène** ($\text{Cov}(z, u) = 0$); les **2SLS** généralisent à plusieurs instruments. On valide l'approche par les tests d'**endogénéité**, de **suridentification** et de **force des instruments**. La simultanéité, cas particulier majeur, est développée au chapitre 16 (modèles à équations simultanées).