

# Économétrie — TD 6

## Endogénéité & Méthode des Variables Instrumentales

Pierre Beaucoral

### Introduction

Ce polycopié reprend les notions des slides (TD 6) et les développe avec des explications, exemples simulés et schémas produits en **R**. Objectifs:

- 1) Rappeler les hypothèses MCO et la notion d'**exogénéité**;
- 2) Identifier trois sources d'**endogénéité** (omission, causalité inverse, erreur de mesure);
- 3) Introduire la **méthode des variables instrumentales** (2SLS/DMC), les tests (faiblesse, Sargan/Hansen, DWH) et l'implémentation EViews.

#### Note

**Notation.** On note  $Y$  la variable expliquée,  $X$  la(les) variable(s) potentiellement endogène(s),  $Z$  l'instrument (ou le vecteur d'instruments),  $W$  les contrôles exogènes, et  $u$ /*uarepsilon* l'erreur.

### Rappel MCO (BLUE) et exogénéité

Hypothèses MCO usuelles (modèle linéaire) :

- $\mathbb{E}[u_i] = 0$
- $\text{Var}(u_i) = \sigma^2$  (**homoscédasticité**) ;
- $\text{Cov}(u_i, u_j) = 0$  (**pas d'autocorrélation**) ;
- $\text{Cov}(X, u) = 0$  (**exogénéité**).

**Conséquence clé :** si  $\text{Cov}(X, u) \neq 0$ , l'estimateur MCO de l'effet causal de  $X$  sur  $Y$  est **biaisé** (et non convergent).

## Effet du biais d'endogénéité sur l'estimation OLS

Nuage de points : données simulées avec  $X$  corrélé à l'erreur

Rouge : régression OLS biaisée | Noir pointillé : vraie pente (2)

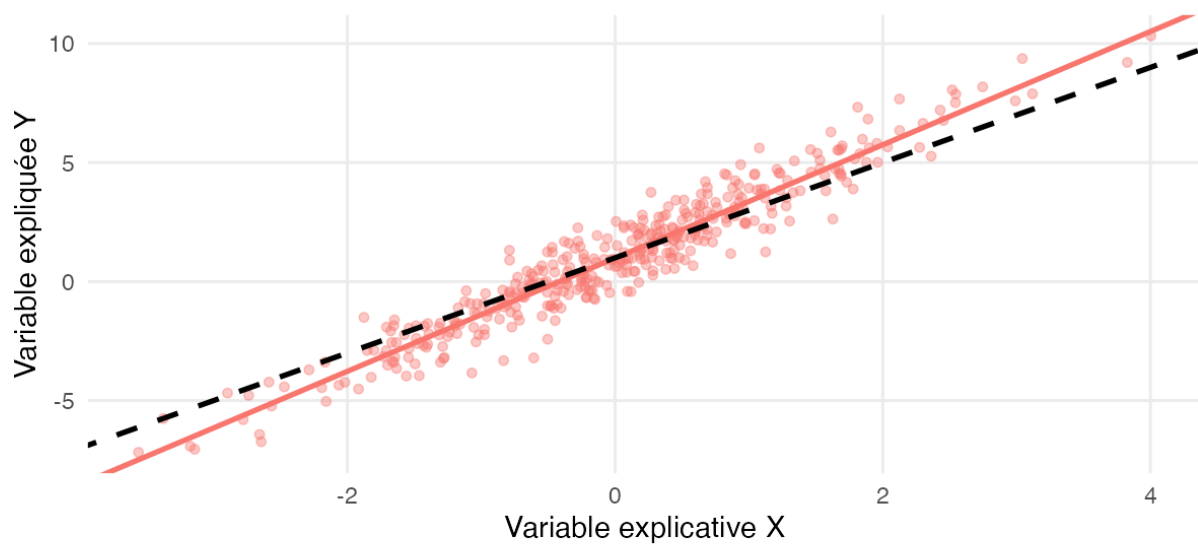


Illustration pédagogique : l'écart entre la droite rouge et la droite noire montre le biais de l'estimateur OLS lorsque  $\text{Cov}(X, u) \neq 0$ .

Figure 1: Intuition : si  $X$  est corrélé au choc  $u$ , l'OLS attribue à  $X$  ce qui vient de  $u \rightarrow$  biais.

## Origines de l'endogénéité (avec schémas)

### Omission d'une variable pertinente

Vrai modèle  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$ , mais on omet  $X_2$  et on estime  $Y = \beta_0 + \beta_1 X_1 + u$ . Si  $X_1$  est **corrélé** avec  $X_2$ , alors  $\text{Cov}(X_1, u) \neq 0$  et  $\hat{\beta}_1^{OLS}$  est **biaisé**.

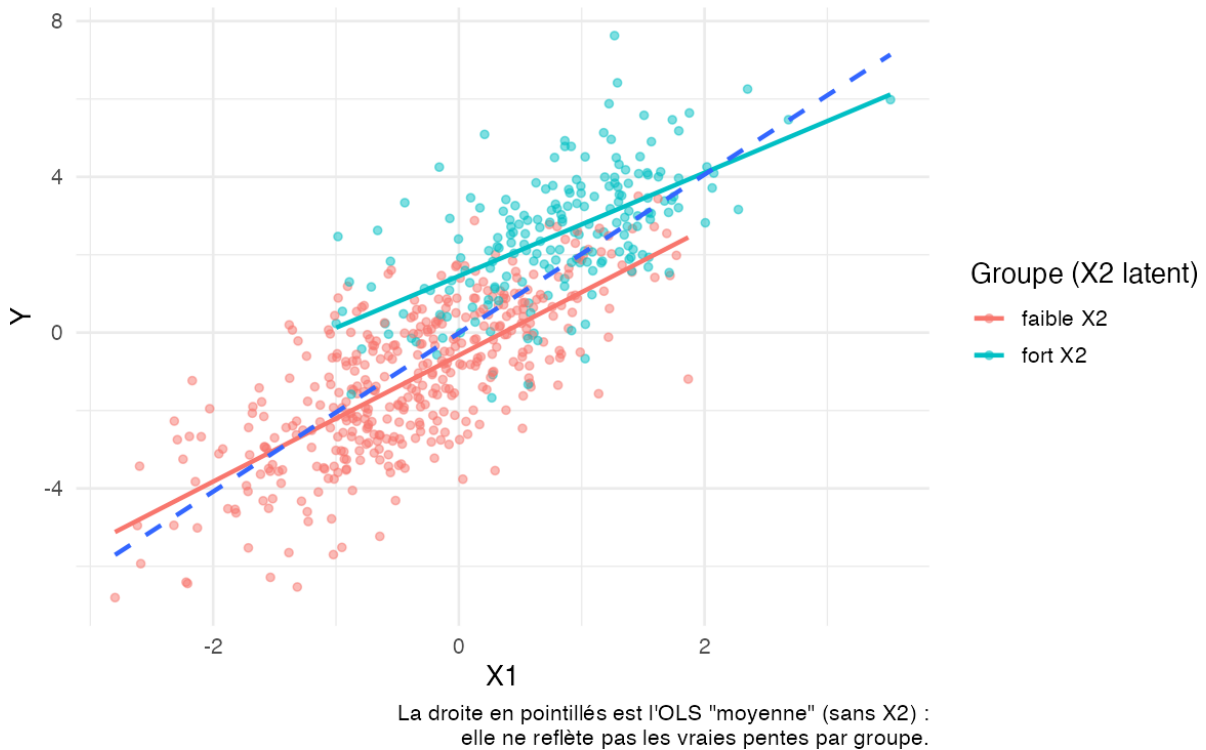


Figure 2: Omission d'une variable: régressions séparées par un facteur latent  $\rightarrow$  pente OLS "moyenne" biaisée.

**Sens du biais** (mémo) :

	$\text{corr}(X_1, X_2) > 0$	$\text{corr}(X_1, X_2) < 0$
$\beta_2 > 0$	biais <b>positif</b>	biais <b>négatif</b>
$\beta_2 < 0$	biais <b>négatif</b>	biais <b>positif</b>

### Causalité inverse

Boucle de rétroaction :  $Y \rightarrow X$  et  $X \rightarrow Y$ .

Ex. dette publique croissance.

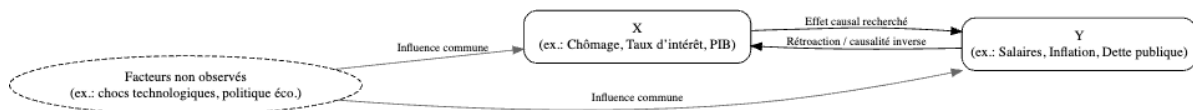


Figure 3: Causalité inverse générique  $X \rightarrow Y$  avec exemples économiques

**Effet sur  $\hat{\beta}_1^{OLS}$**  : il « récupère » une partie du retour  $Y \rightarrow X$ .

### Erreur de mesure sur $X$

On observe  $\tilde{X} = X + \nu$ . L'OLS de  $Y \sim \tilde{X}$  subit un **biais d'atténuation** (vers 0).

### Erreur de mesure : comparaison des deux régressions

Vert : régression avec  $X$  vrai (pente  $\approx 2$ )

Orange : régression avec  $X$  observé (biais d'atténuation)

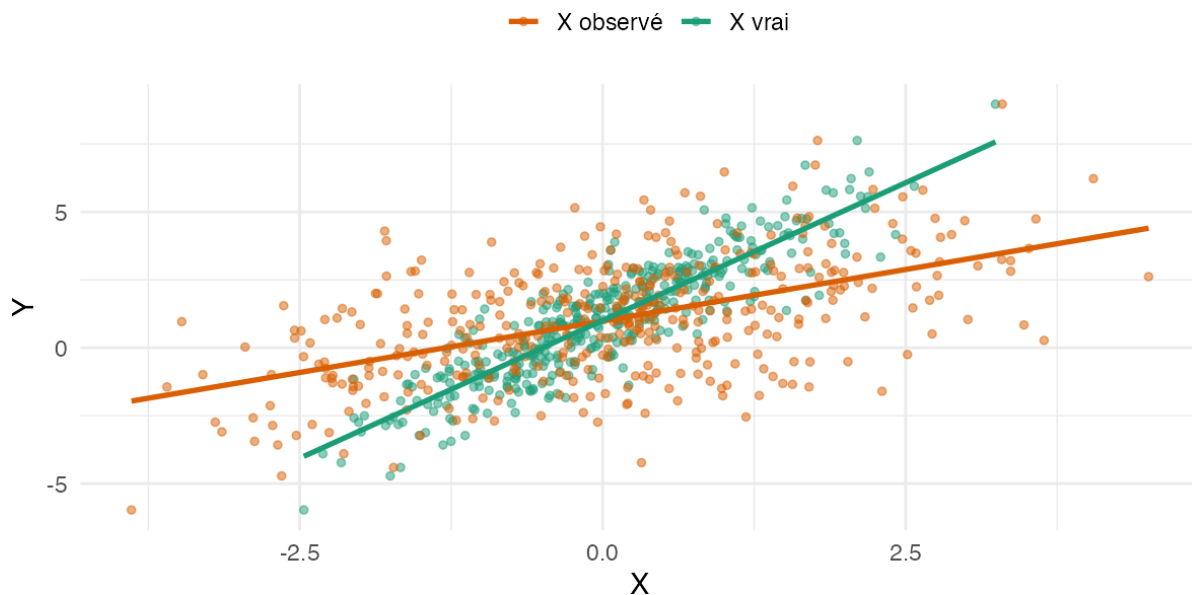


Figure 4: Erreur de mesure sur  $X$  : la pente OLS avec  $X$  bruité est plus faible (biais d'atténuation).

## Variables instrumentales (VI) et 2SLS

### Principe

But : utiliser un instrument  $Z$  pour isoler la **variation exogène** de  $X$ .

- **Pertinence** :  $\text{Cov}(Z, X) \neq 0$  (1 étape explicative).
- **Exogénéité exclue** :  $\text{Cov}(Z, u) = 0$  (Z n'affecte Y que via X).

#### Procédure 2SLS :

- 1)  $X = \pi_0 + \pi_1 Z + W' \pi + v \rightarrow$  obtenir  $\hat{X}$  ;
- 2)  $Y = \beta_0 + \beta_1 \hat{X} + W' \gamma + u$ .

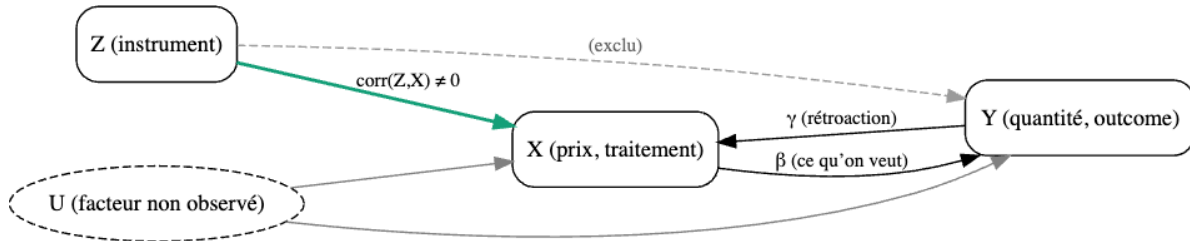


Figure 5: Causalité inverse :  $X \rightarrow Y$  et  $Y \rightarrow X$ . Z est un instrument valide s'il n'affecte Y que via X.

**Lecture attendue** :  $\hat{\beta}^{OLS}$  sur X est **biaisé** (ici vers le haut, car X corrélé à  $u$ ). L'estimateur **2SLS** se rapproche de la vraie valeur (1.5) si Z est suffisamment **pertinent** et **valide**.

#### Instruments faibles (pertinence)

On regarde le **F** de 1 étape (régression de X sur Z et W) ; règle pratique usuelle :  $F > 10$ .

#### Validité de l'instrument (exogénéité)

Si l'on dispose de plus d'instruments que de variables endogènes (sur-id), on peut tester l'orthogonalité (Sargan/Hansen-J).

#### 🔥 Caution

Rappel : ces tests ne remplacent **jamais** l'argument **économique**.

#### Faut-il instrumenter ? (DWH / Nakamura–Nakamura)

Tester si l'estimation OLS est **cohérente** :

- $H_0$  : MCO non biaisé ( $\beta^{OLS} \approx \beta^{2SLS}$ )  $\rightarrow$  on préfère OLS (plus précis).
- $H_A$  : MCO biaisé  $\rightarrow$  on préfère 2SLS.

### 💡 Tip

**Décision** : si le résidu de 1<sup>ère</sup> étape est **significatif**, on **rejette**  $H_0$  (OLS biaisé) et l'on retient la spécification **2SLS**.

## Schémas/organigrammes utiles

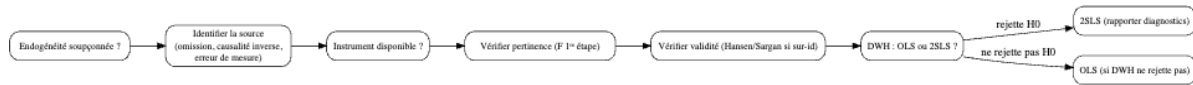


Figure 6: Organigramme (stabilité/ruptures & instruments)

## Mise en œuvre dans EViews (rappel)

- **2SLS/IV** : Quick → Estimate Equation → *Method*: TSLS.
- **Endogeneity test (DWH)** : View → IV Diagnostics and Tests → Regressor Endogeneity Test.
- **Weak instruments** : View → IV Diagnostics and Tests → Weak Instrument Diagnostics.
- **Instrument orthogonality** : View → IV Diagnostics and Tests → Instrument Orthogonality Test.

## Questions de TD

### Q1 — Charger le workfile

Chargez le workfile *Marshall* (contient offre1-offre4, p1-p4, Y, W).

Afficher la réponse

**EViews**. File → Open → Workfile puis sélectionnez Marshall.wf\*.

Vérifiez les séries (aperçu, stats descriptives : View → Descriptive Statistics).

## Q2 — Estimations MCO

Estimez  $\text{offre1} = c + p1$ ,  $\text{offre2} = c + p2$ ,  $\text{offre3} = c + p3$  et  $\text{offre4} = c + p4 + W$  (avec  $W$  exogène).

Afficher la réponse

Quick → Estimate Equation puis entrez :

$\text{offre1} \ c \ p1$  ;  $\text{offre2} \ c \ p2$  ;  $\text{offre3} \ c \ p3$  ;  $\text{offre4} \ c \ p4 \ W$ .

Notez signe/magnitude de  $\hat{\beta}$ , p-values,  $R^2$ , résidus.

**Rappel** : si  $\text{Cov}(P, \varepsilon) \neq 0$ , OLS biaisé.

## Q3 — Test d'exogénéité (Nakamura & Nakamura / DWH)

Avec  $Y$  comme instrument des prix, testez l'exogénéité des variables de prix.

Afficher la réponse

Dans EViews : View → IV Diagnostics and Tests → Regressor Endogeneity Test.

$H_0$  : exogénéité. Si stat. ( $t^2/F$ ) **significative** → **rejeter**  $H_0$  (prix endogène).

## Q4 — Estimation en VI (2SLS)

Si nécessaire, estimez les équations d'offre en **2SLS** avec  $Y$  instrument (et  $W$  exogène dans  $\text{offre4}$ ).

Afficher la réponse

Quick → Estimate Equation → *Method*: **TSLS**.

**Endogènes** :  $p1$  (ou  $p2/p3/p4$ ). **Instruments** :  $Y$  (+  $W$  en exogène).

Diagnostics : **F 1 étape** (faiblesse), **Hansen/Sargan** (sur-id), **DWH** (faut-il instrumenter ?).

## Points clés à retenir

- Identifier la **source** d'endogénéité oriente la solution (contrôles, interactions, VI...).
- **2SLS** corrige le biais si  $Z$  est **pertinent** et **valide**.
- **Diagnostics** : F 1 étape, Hansen/Sargan, **DWH**.
- L'argument **économique** reste central (au-delà des seuls tests).