

Économétrie — TD 6

Endogénéité et Méthode des Variables Instrumentales

Pierre Beaucoral

```
library(knitr)
knit_hooks$set(optipng = hook_optipng)
```

Introduction

Objectif du TD

- Rappeler les hypothèses des **MCO** et la notion d'**exogénéité**.
- Identifier trois **sources d'endogénéité** (omission, causalité inverse, erreur de mesure).
- Introduire la **méthode des variables instrumentales** (2SLS/DMC) : pertinence, exogénéité des instruments, et tests associés.

Rappel MCO (BLUE)

Hypothèses clés (linéaire, MCO)

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

i Note

Si $\text{Cov}(X, \varepsilon) \neq 0$, l'estimateur MCO est **biaisé** et **non convergent** : il ne mesure pas l'effet causal de X sur Y .

Origines de l'endogénéité

(1) Omission d'une variable pertinente

Vrai modèle : $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$.

Mais on estime : $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$ avec X_1 corrélé à X_2 .

Sens du biais sur $\hat{\beta}_1$:

	$\text{corr}(X_1, X_2) > 0$	$\text{corr}(X_1, X_2) < 0$
$\beta_2 > 0$	+	(vers le haut)
$\beta_2 < 0$	-	(vers le bas)

L'omission « pousse » $\hat{\beta}_1$ dans le sens de la corrélation entre X_1 et la variable manquante X_2 .

(2) Causalité inverse

On estime $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ alors qu'en réalité $X_i = \gamma_0 + \gamma_1 Y_i + \gamma_2 Z_i + \nu_i$ (boucle de rétroaction).

Exemple : croissance du PIB \leftrightarrow dette publique.

Sens du biais (sur $\hat{\beta}_1$) :

	$\gamma_1 > 0$	$\gamma_1 < 0$
$\beta_1 > 0$	+	-
$\beta_1 < 0$	-	+

L'effet estimé « récupère » une partie du retour $Y \rightarrow X$.

(3) Erreur de mesure sur X

On souhaite $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, mais on observe $\tilde{X}_i = X_i + \nu_i$.

Alors : $Y_i = \beta_0 + \beta_1 \tilde{X}_i + (\varepsilon_i - \beta_1 \nu_i)$, d'où \tilde{X} corrélée à l'erreur composite.

Conséquence : biais d'atténuation (vers 0) sur $\hat{\beta}_1$.

Variables instrumentales (VI)

Principe (2SLS / DMC)

But : isoler la variation exogène de X avec un instrument Z .

1. 1 étape : régresser X sur Z (et autres contrôles), obtenir \hat{X} .
2. 2 étape : remplacer X par \hat{X} dans l'équation de Y .
3. Recalculer des écarts-types adaptés (2SLS).

Conditions pour Z :

- Pertinence : $\text{Cov}(Z, X) \neq 0$ (pouvoir explicatif).
 - Exogénéité exclue : Z n'affecte Y que via X ($\text{Cov}(Z, \varepsilon) = 0$).
-

Pertinence : instruments faibles

- Vérifier la 1 étape : test F des instruments.
 - Règle pratique : $F > 10 \Rightarrow$ pertinence acceptable.
 - Plusieurs instruments faibles aggravent le biais.
 - (EViews) View → IV Diagnostics and Tests → Weak Instrument Diagnostics.
-

Exogénéité de l'instrument

- L'instrument ne doit pas être corrélé à Y autrement que via X .
- Tests de sur-identification (si $q > p$) : **Sargan** (homoscédasticité) / **Hansen-J** (robuste).
- Statistique $\chi^2(k)$ avec $k = \text{nb de restrictions}$ (sur-id).
- (EViews) View → IV Diagnostics and Tests → Instrument Orthogonality Test.



Tip

En pratique, on dispose rarement de sur-identification « confortable »; la **justification théorique** de Z reste centrale.

Faut-il instrumenter ? (DWH)

- Perte de précision avec VI : vérifier si l'instrumentation est nécessaire.
 - **Durbin–Wu–Hausman** (a.k.a. Nakamura–Nakamura) :
 - H_0 : MCO non biaisé ($\beta^{\text{MCO}} \approx \beta^{\text{DMC}}$)
 - H_A : MCO biaisé ($\beta^{\text{MCO}} \neq \beta^{\text{DMC}}$)
 - Statistique $\chi^2(k)$ avec $k = \text{nb de variables endogènes}$.
 - (EViews) View → IV Diagnostics and Tests → Regressor Endogeneity Test.
-

En pratique (guidelines)

- Identification d'abord : quelles sources d'endogénéité ?
Quel sens du biais attendu ?
 - Choix de l'instrument :
 - Pourquoi est-il **corrélé** à X (pertinence) ?
 - Pourquoi est-il **exogène** (validité) ?
 - Les tests (faiblesse, Hansen/Sargan, DWH) aident, mais **ne remplacent pas** l'argument **économique**.
-

Annexes

Schéma 2SLS (rappel)

1. $X = \pi_0 + \pi_1 Z + W'\pi + v$ (1 ère étape)
 2. $Y = \beta_0 + \beta_1 \hat{X} + W'\gamma + u$ (2 ère étape)
- Z : instruments ; W : contrôles exogènes.
 - Conditions : $\text{rank}([Z, W])$ suffisant ; $\text{Cov}(Z, u) = 0$.
-

Tables « sens du biais » (récapitulatif)

Omission d'une variable

	$\text{corr}(X_1, X_2) > 0$	$\text{corr}(X_1, X_2) < 0$
$\beta_2 > 0$	+	-
$\beta_2 < 0$	-	+

Causalité inverse

	$\gamma_1 > 0$	$\gamma_1 < 0$
$\beta_1 > 0$	+	-
$\beta_1 < 0$	-	+

Références « dans EViews »

- Weak instruments : View → IV Diagnostics and Tests → Weak Instrument Diagnostics
 - Orthogonality (Sargan/Hansen) : View → IV Diagnostics and Tests → Instrument Orthogonality Test
 - Endogeneity (DWH) : View → IV Diagnostics and Tests → Regressor Endogeneity Test
-

Questions TD

Q1 — Charger le workfile

Énoncé

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

Afficher la réponse

- **File → Open → Workfile** puis sélectionner *Marshall.wf**.
 - Vérifier les séries dans l’arborescence (double-clic pour aperçu).
 - (Option) **View → Descriptive Statistics → Histogram and stats** pour un coup d’œil rapide.
-

Q2 — Estimations MCO

Énoncé

Estimez les fonctions d’offre par MCO :

pour $i = 1 \text{ à } 3$: $\text{offre}_i = + \text{Pi} +$
et $\text{offre}_4 = + \text{P4} + \text{W} +$, avec W exogène.

Afficher la réponse

- **Quick → Estimate Equation**, puis entrer la spécification :
 - `offre1 c p1`
 - `offre2 c p2`
 - `offre3 c p3`
 - `offre4 c p4 W`
- Noter : $\hat{\beta}$ (signe, magnitude), R^2 , **p-values**, et résidus.



Caution

Rappel : si $\text{Cov}(P, \varepsilon) \neq 0$ (ex. `offre_i p_i`), MCO est biaisé. On vérifiera ensuite avec des tests d’exogénéité.

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

Énoncé

Indiquez, à l'aide du test de **Nakamura & Nakamura**, le caractère exogène des variables de **prix**, en prenant comme **instrument** le **revenu Y**.

Afficher la réponse

Intuition

- H_0 : **exogénéité** du prix dans l'équation d'offre (MCO non biaisé).
- H_A : **endogénéité** → préférer VI (2SLS).

Mise en œuvre (EViews)

- Estime l'équation en **IV/2SLS** (voir Q4) pour récupérer les résidus nécessaires, ou utilise directement :
 - Quick → Estimate Equation → Method: TSLS/IV,
 - onglet View → IV Diagnostics and Tests → Regressor Endogeneity Test (Durbin–Wu–Hausman / Nakamura–Nakamura).
 - **Décision** : si la stat. χ^2 (ou F) est **significative**, rejeter H_0 → le prix est **endogène**.
-

Q4 — Estimations en Variables Instrumentales (2SLS)

Énoncé

Estimez, si nécessaire, les fonctions d'offre à l'aide de la méthode des **variables instrumentales**, en utilisant **Y** comme instrument.

Afficher la réponse

Rappel 2SLS (schéma)

1. 1^{re} étape : $P_i = \pi_0 + \pi_1 Y + v$ obtenir \widehat{P}_i .
2. 2^{re} étape : $\text{offre}_i = \alpha + \beta \widehat{P}_i + u$.
3. Écarts-types **robustes** si hétéroscédasticité (option *White*).

Guide (EViews)

- Quick → Estimate Equation → *Method*: TSLS – Two Stage Least Squares.
 - List of endogenous: p1 (ou p2/p3/p4).
 - Instrument list: Y (ajouter W pour offre4).
 - View → IV Diagnostics and Tests :
 - Weak Instrument Diagnostics (F 1 étape > 10 souhaitable),
 - Instrument Orthogonality Test (Hansen-J/Sargan), si sur-id.
-