# ÉCONOMÉTRIE S2
# (UGA S2)
## CHAPITRE 4:
## SYSTÈMES D'ÉQUATIONS SIMULTANÉES


Michal W. Urdanivia*

*Université de Grenoble Alpes, Faculté d'Économie, GAEL,
e-mail: michal.wong-urdanivia@univ-grenoble-alpes.fr


27 février 2022

# Contenu

# Motivation et exemple

# 1. Motivation et exemple

## 1.1. Motivation

- Historiquement, les premiers problèmes d'***endogénéité*** mis en évidence par les économètres sont les problèmes de ***simultanéité*** des prix et des quantités échangées dans le cadre de ***marchés en équilibre***.
- Ces problèmes d'estimation ont tout d'abord été traités dans le cadre dit des « ***systèmes d'équations simultanées*** ».
- En fait, on trouve ici les ***fondements de l'économétrie*** en tant que branche spécifique de la statistique.
- Tout problème d'endogénéité peut se traiter dans ce cadre.
- L'objectif est ici :
    - De présenter les systèmes d'équations simultanées et leur estimation
    - De présenter les différences et similarités entre l'approche présentée jusqu'à présent pour traiter les problèmes d'endogénéité et celle dans les systèmes d'équations simultanées.

## 1.2. Exemple : délinquance économique et chômage

***Exemple inspiré*** d'une étude de Fougère, Kramarz et Pouget (2007, INSEE).

**Objectifs de l'étude**. Mesurer l'effet causal du chômage sur la délinquance économique (cambriolages, vols et trafic de stupéfiants).

Proposer des arguments en faveur/défaveur de l'efficacité de politiques visant à réduire les taux de délinquance.

**Données utilisées**. Différentes mesures de la délinquance, de la structure du tissu économique, du chômage, … pour les 95 départements français métropolitains suivies sur la période 1990-2000.

**Rmq**. Dans la suite, l'indice $i$ désigne un couple (*département* , *année*)

## Délinquance économique et chômage : effets causaux et simultanéité

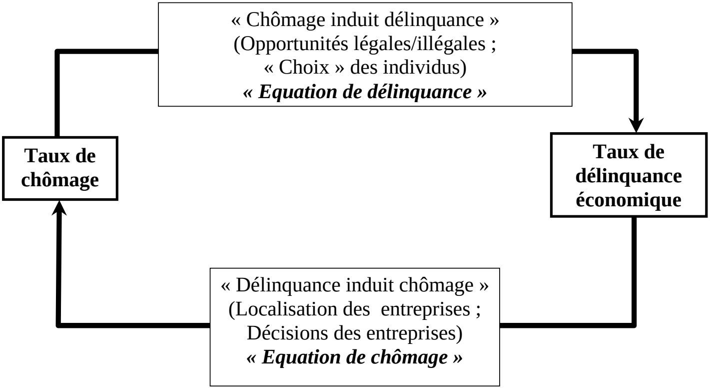
***Chômage « cause » délinquance économique***

- L'absence d'***opportunités légales*** de revenu tend à favoriser le choix d'***opportunités illégales***.
    - ***Effets atténués*** par la présence policière, les aides aux chômeurs, la densité du tissu économique (opportunités d'emploi) et par l'éducation (un diplômé peut espérer sortir vite plus du chômage).
    - ***Effets accentués*** par la densité de population, les inégalités de revenu et la criminalité organisée implantée.

***Délinquance économique « cause » chômage***

- Localement et à moyen terme, la délinquance tend à défavoriser l'***implantation d'entreprises***.
    - ***Effets atténués*** par des aides à l'implantation d'entreprises (zones franches, …), l'éducation et la densité de population (bassin de main d'œuvre).

# Schéma causal « basique »



## Formalisation du schéma causal « de base » mis en évidence

***Causalité dans les deux sens = simultanéité*** :

$$\begin{cases} d\acute{e}li_i = a_0 + b_0 \times chom_i + e_i \\ chom_i = \alpha_0 + \beta_0 \times d\acute{e}li_i + \varepsilon_i \end{cases} \text{ avec } \begin{cases} E[e_i] \equiv 0 \\ E[\varepsilon_i] \equiv 0 \end{cases}$$

et :

$$\begin{cases} Cov[chom_i; e_i] \neq 0 \quad \text{("effet investissement des entreprises")} \\ Cov[d\acute{e}li_i; \varepsilon_i] \neq 0 \quad \text{("effet opportunités légales/illégales")} \end{cases}$$

**Définition.** Un ***système d'équations simultanées*** est un système : (*i*) avec au moins une variable explicative endogène dans au moins une des équations et (*ii*) où chaque variable endogène a « son » équation.

**Définition.** Un ***système de régressions empilées*** est un système composé de modèles de régression.

Dans le système d'équations simultanées :

$$\begin{cases} déli_i = a_0 + b_0 \times chom_i + e_i \\ chom_i = \alpha_0 + \beta_0 \times déli_i + \varepsilon_i \end{cases} \text{ avec } \begin{cases} E[e_i] \equiv 0 \\ E[\varepsilon_i] \equiv 0 \end{cases}$$

et :

$$\begin{cases} Cov[chom_i ; e_i] \neq 0 \\ Cov[déli_i ; \varepsilon_i] \neq 0 \end{cases}$$

on s'attend aux signes suivants pour les effets causaux :

| | | |
|---|---|---|
| $b_0 > 0$ | et | $\beta_0 > 0$ |
| (chômage cause délinquance) | | (délinquance cause chômage) |

***On retrouve empiriquement*** $Cov[chom_i ; déli_i] > 0$ mais ***on ne sait pas trop ce que mesure cette covariance***

On a en fait deux types de problème, intimement liés l'un à l'autre :

- ***Rien ne distingue réellement les équations de « délinquance » et de « chômage »***. Par exemple :

$$déli_i = a_0 + b_0 \times chom_i + e_i \quad \Rightarrow \quad chom_i = -b_0^{-1} a_0 + b_0^{-1} \times déli_i - b_0^{-1} \times e_i$$

***mais en général on a :***

$$\alpha_0 \neq -b_0^{-1} a_0, \quad \beta_0 \neq b_0^{-1} \quad \text{et} \quad \varepsilon_i \neq -b_0^{-1} \times e_i .$$

  - L'équation de « délinquance » décrit le ***comportement des individus***
  - L'équation de « chômage » décrit le ***comportement des entreprises et des salariés***

- ***Les deux équations du système ont des variables explicatives endogènes :***

$$Cov\left[chom_i ; e_i\right] \neq 0 \text{ et } Cov\left[déli_i ; \varepsilon_i\right] \neq 0$$

Il est impossible d'estimer $\beta_0$ ou $b_0$ à partir des seuls $(déli_i , chom_i)$.

Or : **Problème d’identification = déficit d’information**

Jusqu’à présent, le problème de l’***endogénéité*** a été géré par :

- des ***techniques de variables instrumentales***

et/ou :

- l’utilisation de ***variables de contrôle*** (de l’hétérogénéité).

On va voir ici qu’on utilise implicitement ces deux solutions (et conjointement) lorsqu’on cherche à **compléter les modèles de comportement** décrits par les équations de « chômage » et de « délinquance », mais en suivant une approche différente.

Cette ***approche*** est dite ***en information complète*** car on construit ***un modèle « causal » pour chacune des variables endogènes*** du système.

***Equation de « délinquance »***

$$déli_i = a_0 + b_0 \times chom_i + \mathbf{d}_0' \begin{bmatrix} dens\_pop_i \\ \textbf{éduc}_i \\ \textbf{tissu\_éco}_i \end{bmatrix} + \mathbf{r}_0' \begin{bmatrix} \textbf{aid\_chom}_i \\ \textbf{inég}_i \\ \textbf{crim\_org}_i \end{bmatrix} + u_{d,i}$$

***Equation de « chômage »***

$$chom_i = \alpha_0 + \beta_0 \times déli_i + \boldsymbol{\delta}_0' \begin{bmatrix} dens\_pop_i \\ \textbf{éduc}_i \\ \textbf{tissu\_éco}_i \end{bmatrix} + \boldsymbol{\rho}_0' \textbf{prév\_crois}_i + u_{c,i}$$

Pour l’observation *i*, *i.e.* un couple (*département* , *année*), les vecteurs introduits sont ***des mesures*** de :

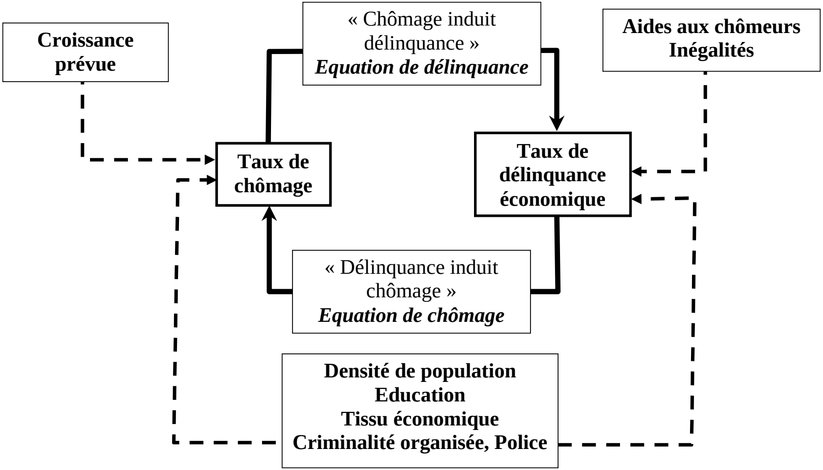
$$
\mathbf{q}_i \quad \begin{cases} \textbf{éduc}_i : \text{ éducation des jeunes} \\ \textbf{tissu\_éco}_i : \text{ description du tissu économique local} \\ \textbf{crim\_org}_i : \text{ description de l'implantation du crime organisé} \\ \textbf{police}_i : \text{ description de la présence policière} \end{cases}
$$

$$
\mathbf{q}_{d,i} \quad \begin{cases} \textbf{aid\_chom}_i : \text{ description des aides au chomeurs (}\textit{temps}\text{)} \\ \textbf{inég}_i : \text{ description des inégalités de revenu} \end{cases}
$$

$$
\mathbf{q}_{c,i} \quad \left\{ \textbf{prév\_crois}_i : \text{ prévisions de la croissance économique locale} \right.
$$

Ces variables explicatives sont ***prédéterminées par rapport à*** $chom_i$ et $déli_i$, elles sont donc ***exogènes dans le système d’équations*** considéré, *i.e.* ***par rapport à*** $u_{d,i}$ ***<u>et</u> à*** $u_{c,i}$.
**Rmq**. « Limite » pour la police.

## Schéma causal « complet »



Le ***système d'équations simultanées*** « ***complété*** » est donné par la :

**Forme structurelle du système d'équations simultanées**

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0'\mathbf{q}_i + \mathbf{r}_0'\mathbf{q}_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0'\mathbf{q}_i + \boldsymbol{\rho}_0'\mathbf{q}_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{u}_i / \mathbf{z}_i\right] = E\left[\mathbf{u}_i\right] \equiv \mathbf{0}$$

où :

$$\mathbf{u}_i \equiv (u_{d,i}, u_{c,i}) \quad \text{et} \quad \mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i}).$$

La forme ***structurelle du modèle décrit les relations causales d'intérêt*** dans leur forme « naturelle », *i.e.* les paramètres d'intérêt, $b_0$ (surtout) et $\beta_0$ ici, sont des paramètres de la forme structurelle du modèle.

En notant :

$$\mathbf{g}_0 \equiv \begin{bmatrix} a_0 \\ b_0 \\ \mathbf{d}_0 \\ \mathbf{r}_0 \end{bmatrix},\ \mathbf{x}_{d,i} \equiv \begin{bmatrix} 1 \\ y_{c,i} \\ \mathbf{q}_i \\ \mathbf{q}_{d,i} \end{bmatrix},\ \boldsymbol{\gamma}_0 \equiv \begin{bmatrix} \alpha_0 \\ \beta_0 \\ \boldsymbol{\delta}_0 \\ \boldsymbol{\rho}_0 \end{bmatrix} \text{ et } \mathbf{x}_{c,i} \equiv \begin{bmatrix} 1 \\ y_{d,i} \\ \mathbf{q}_i \\ \mathbf{q}_{c,i} \end{bmatrix}, \text{ et } \mathbf{z}_i \equiv \begin{bmatrix} 1 \\ \mathbf{q}_i \\ \mathbf{q}_{d,i} \\ \mathbf{q}_{c,i} \end{bmatrix} \text{ et } \mathbf{u}_i \equiv \begin{bmatrix} u_{d,i} \\ u_{c,i} \end{bmatrix},$$

ce modèle peut également être écrit sous la forme plus compacte :

**Forme structurelle du système d’équations simultanées**

$$\begin{cases} y_{d,i} = \mathbf{g}_0' \mathbf{x}_{d,i} + u_{d,i} \\ y_{c,i} = \boldsymbol{\gamma}_0' \mathbf{x}_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{u}_i / \mathbf{z}_i\right] = E\left[\mathbf{u}_i\right] \equiv \mathbf{0}$$

où :

$$\mathbf{u}_i \equiv (u_{d,i}, u_{c,i}) \text{ et } \mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i}).$$

# Formes structurelle et réduite

## 2. Formes structurelle et réduite

On considère ici le système d'équations simultanées « général » suivant :

**Forme structurelle du modèle**

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0'\mathbf{q}_i + \mathbf{r}_0'\mathbf{q}_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0'\mathbf{q}_i + \boldsymbol{\rho}_0'\mathbf{q}_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{u}_i / \mathbf{z}_i\right] = E\left[\mathbf{u}_i\right] \equiv \mathbf{0}$$

où :

$$\mathbf{u}_i \equiv (u_{d,i}, u_{c,i}) \ \text{ et } \ \mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i}).$$

Les ***deux équations*** de ce système contiennent ***des variables explicatives endogènes*** (les variables à expliquer de l'autre équation).

Compte-tenu de ce qui a été vu dans le chapitre précédent, ***on sait qu'il est a priori possible d'estimer sans biais (as.) les paramètres de ce système*** par les ***2MC***, pour peu qu'on dispose de VI …

Une ***approche alternative*** est présentée ici, celle dite par les ***MCI***.

- Elle est peu utilisée pour des systèmes de ce type, …
- … mais s’avère parfois utile dans le cadre de problèmes plus complexes.
- Elle est en tous cas ***très instructive*** quant aux fondements de la stratégie d’identification par les techniques de VI.

En fait cette approche est peu utilisée pour plusieurs raisons, la principale étant qu’elle n’est (simplement) utilisable que dans des cas particuliers, les cas où les ***paramètres de la forme du modèle sont juste-identifiés*** …

… notion qu’on définit par la suite.

# Cas juste-identifié : Estimateur des MCI

## 3. Cas juste-identifié : Estimateur des MCI

### 3.1. Définition de l'estimateur des MCI

On considère ici une version simplifiée du modèle :

**Forme structurelle du modèle**

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + d_0 q_i + r_0 q_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i + \rho_0 q_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E[\mathbf{u}_i / \mathbf{z}_i] = E[\mathbf{u}_i] \equiv \mathbf{0}$$

où :

$$\mathbf{u}_i \equiv (u_{d,i}, u_{c,i}) \text{ et } \mathbf{z}_i \equiv (1, q_i, q_{d,i}, q_{c,i}).$$

On verra dans la suite que les paramètre de ce modèle sont juste-identifiés.

**Principe de l’approche par les MCI**

Utiliser la ***structure du système d’équations*** pour :

1. Définir les ***variables à expliquer***, $y_{d,i}$ et $y_{c,i}$, ***sous la forme de modèles de régression***, *i.e.* de modèles ne contenant que ***des variables explicatives exogènes***. On obtient ainsi ce qu’on appelle la ***forme réduite*** du système d’équations structurelles.

2. Définir ***les relations entre les paramètres*** de la ***forme structurelle*** et de la ***forme réduite*** du modèle considéré.

3. Calculer des ***estimateurs des paramètres de la forme réduite*** du modèle (simple : cette forme réduite est un système de régressions empilées).

4. Calculer des ***estimateurs des paramètres de la forme structurelle*** du modèle en utilisant :
   (*a*) Les estimateurs des paramètres de la forme réduite du modèle (3)
   (*b*) Les relations entre les paramètres de forme réduite et de la forme structurelle du modèle (2).

***1. Détermination de la forme réduite :***

On veut exprimer $y_{d,i}$ et $y_{c,i}$ ***en fonction des variables exogènes*** du système, *i.e.* des éléments de $\mathbf{z}_i \equiv (1, q_i, q_{d,i}, q_{c,i})$.

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + d_0 q_i + r_0 q_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i + \rho_0 q_{c,i} + u_{c,i} \end{cases}$$

Il suffit ici d'utiliser un des modèles, pour ***remplacer*** la variable à expliquer correspondante dans l'autre modèle, *e.g.* :

$$y_{d,i} = a_0 + b_0(\alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i + \rho_0 q_{c,i} + u_{c,i}) + d_0 q_i + r_0 q_{d,i} + u_{d,i}$$

$$\Downarrow$$

$$y_{d,i} = (1 - b_0\beta_0)^{-1} \times \left[ (a_0 + b_0\alpha_0) + (d_0 + b_0\delta_0) q_i + r_0 q_{d,i} + b_0\rho_0 q_{c,i} + (u_{d,i} + b_0 u_{c,i}) \right]$$

$$\Downarrow$$

$$y_{d,i} = \pi_{d,1} + \pi_{d,2} q_i + \pi_{d,3} q_{d,i} + \pi_{d,4} q_{c,i} + v_{d,i}$$

On obtient alors la ***forme réduite*** du système d’équations simultanées

$$\begin{cases} y_{d,i} = \pi_{d,1} + \pi_{d,2} q_i + \pi_{d,3} q_{d,i} + \pi_{d,4} q_{c,i} + v_{d,i} \\ y_{c,i} = \pi_{c,1} + \pi_{c,2} q_i + \pi_{c,3} q_{d,i} + \pi_{c,4} q_{c,i} + v_{c,i} \end{cases}$$

avec :

$$\boldsymbol{\pi}_d = \begin{bmatrix} \pi_{d,1} \\ \pi_{d,2} \\ \pi_{d,3} \\ \pi_{d,4} \end{bmatrix} \equiv \Delta^{-1} \times \begin{bmatrix} a_0 + b_0 \alpha_0 \\ d_0 + b_0 \delta_0 \\ r_0 \\ b_0 \rho_0 \end{bmatrix} \text{ et } \boldsymbol{\pi}_c = \begin{bmatrix} \pi_{c,1} \\ \pi_{c,2} \\ \pi_{c,3} \\ \pi_{c,4} \end{bmatrix} \equiv \Delta^{-1} \times \begin{bmatrix} \alpha_0 + \beta_0 a_0 \\ \delta_0 + \beta_0 d_0 \\ \beta_0 r_0 \\ \rho_0 \end{bmatrix}$$

$$\text{où } \Delta \equiv 1 - b_0 \beta_0$$

$$v_{d,i} \equiv \Delta^{-1} \times (u_{d,i} + b_0 u_{c,i}) \text{ et } v_{c,i} \equiv \Delta^{-1} \times (u_{c,i} + \beta_0 u_{d,i})$$

et :

$$E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0} \text{ où } \mathbf{v}_i \equiv (v_{d,i}, v_{c,i}).$$

Le vecteur $\mathbf{z}_i \equiv (1, q_i, q_{d,i}, q_{c,i})$ est celui des ***variables exogènes du système***

**Rmq**. ***Le terme*** $\Delta$ est le déterminant de la matrice :

$$\begin{bmatrix} 1 & -b_0 \\ -\beta_0 & 1 \end{bmatrix} \text{ dont l'inverse est } \Delta^{-1} \times \begin{bmatrix} 1 & b_0 \\ \beta_0 & 1 \end{bmatrix}.$$

En fait le modèle structurel peut également s'écrire sous **forme matricielle** :

$$\begin{bmatrix} 1 & -b_0 \\ -\beta_0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} y_{d,i} \\ y_{d,i} \end{bmatrix}}_{\mathbf{y}_i} = \begin{bmatrix} a_0 & d_0 & r_0 & 0 \\ \alpha_0 & \delta_0 & 0 & \rho_0 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ q_i \\ q_{d,i} \\ q_{c,i} \end{bmatrix}}_{\mathbf{z}_i} + \underbrace{\begin{bmatrix} u_{d,i} \\ u_{d,i} \end{bmatrix}}_{\mathbf{u}_i}$$

et on a :

$$\begin{bmatrix} y_{d,i} \\ y_{d,i} \end{bmatrix} = \begin{bmatrix} 1 & -b_0 \\ -\beta_0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a_0 & d_0 & r_0 & 0 \\ \alpha_0 & \delta_0 & 0 & \rho_0 \end{bmatrix} \mathbf{z}_i + \begin{bmatrix} 1 & -b_0 \\ -\beta_0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_{d,i} \\ u_{d,i} \end{bmatrix}.$$

**Rmq**. Si $\Delta = 0$, le système n'a pas d'équilibre. On suppose ici que $\Delta \neq 0$.

La forme réduite du modèle (écrite ici sous forme compacte) :

**Forme réduite du modèle**

$$\begin{cases} y_{d,i} = \boldsymbol{\pi}'_d \mathbf{z}_i + v_{d,i} \\ y_{c,i} = \boldsymbol{\pi}'_c \mathbf{z}_i + v_{c,i} \end{cases} \text{ avec } \mathbf{v}_i \equiv \begin{bmatrix} v_{d,i} \\ v_{c,i} \end{bmatrix} \text{ et } E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0}$$

est un ***système de régressions empilées***, *i.e.* un système d'équations de régression.

Les conditions d'exogénéité de $\mathbf{z}_i$ par rapport à $\mathbf{v}_i$ :

$$E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0}$$

sont obtenues par construction, *e.g.* avec $v_{d,i} \equiv (1 - b_0\beta_0)^{-1} \times (u_{d,i} + b_0 u_{c,i})$ on a :

$$E\left[v_{d,i} / \mathbf{z}_i\right] = (1 - b_0\beta_0)^{-1} \times \left(E\left[u_{d,i} / \mathbf{z}_i\right] + b_0 E\left[u_{c,i} / \mathbf{z}_i\right]\right) = 0 = E\left[v_{d,i}\right].$$

Bien entendu, il est aisé d'estimer les paramètres, $\boldsymbol{\pi}_d$ et $\boldsymbol{\pi}_d$, de la forme réduite du modèle :

$$\begin{cases} y_{d,i} = \boldsymbol{\pi}'_d \mathbf{z}_i + v_{d,i} \\ y_{c,i} = \boldsymbol{\pi}'_c \mathbf{z}_i + v_{c,i} \end{cases} \quad \text{avec } E[\mathbf{v}_i / \mathbf{z}_i] = E[\mathbf{v}_i] = \mathbf{0},$$

il suffit d'utiliser l'estimateur des MCO pour chaque équation (***cette estimation est en outre efficace***).

***2. Déterminer les paramètres « structurels » en fonction des paramètres « réduits »***

Ces estimations ne sont cependant utiles que si on est capable de calculer les paramètres de la forme structurelle du modèle, ou ***paramètres structurels***, ***à partir*** des paramètres de la forme réduite, ou ***paramètres réduits***.

C'est possible dans la version simplifiée du modèle, car on a autant de ***paramètres réduits*** que de ***paramètres structurels (8)***.

On dit alors que les ***paramètres structurels sont juste-identifiés dans le modèle*** (par les paramètres réduits).

Le système de 8 ***équations*** (nombre de paramètres réduits) à 8 ***inconnues*** (nombre de paramètres structurels) :

$$\begin{cases} \pi_{d,1} = \Delta^{-1} \times (a_0 + b_0\alpha_0) \\ \pi_{d,2} = \Delta^{-1} \times (d_0 + b_0\delta_0) \\ \pi_{d,3} = \Delta^{-1} \times r_0 \\ \pi_{d,4} = \Delta^{-1} \times b_0\rho_0 \end{cases} \quad \text{et} \quad \begin{cases} \pi_{c,1} = \Delta^{-1} \times (\alpha_0 + \beta_0 a_0) \\ \pi_{c,2} = \Delta^{-1} \times (\delta_0 + \beta_0 d_0) \\ \pi_{c,3} = \Delta^{-1} \times \beta_0 r_0 \\ \pi_{c,4} = \Delta^{-1} \times \rho_0 \end{cases}$$

admet une solution unique :

$$\begin{cases} a_0 = \pi_{d,1} - \dfrac{\pi_{d,4}}{\pi_{c,4}}\pi_{c,1}, \quad b_0 = \dfrac{\pi_{d,4}}{\pi_{c,4}}, \quad d_0 = \pi_{d,2} - \dfrac{\pi_{d,4}}{\pi_{c,4}}\pi_{c,2} \ \text{ et } \ r_0 = \pi_{d,3} - \pi_{c,3}\dfrac{\pi_{d,4}}{\pi_{c,4}} \\ \alpha_0 = \pi_{c,1} - \dfrac{\pi_{c,3}}{\pi_{d,3}}\pi_{d,1}, \quad \beta_0 = \dfrac{\pi_{c,3}}{\pi_{d,3}}, \quad \delta_0 = \pi_{c,2} - \pi_{d,2}\dfrac{\pi_{c,3}}{\pi_{d,3}} \ \text{ et } \ \rho_0 = \pi_{c,4} - \pi_{d,4}\dfrac{\pi_{c,3}}{\pi_{d,3}} \end{cases}$$

(sous certaines conditions analysées ci-après).

**Pour résumer ce qu'on vient d'obtenir**

Le ***schéma causal*** considéré est décrite par la ***forme structurelle*** du modèle :

$$\begin{cases} y_{d,i} = \mathbf{g}_0'\mathbf{x}_{d,i} + u_{d,i} \\ y_{c,i} = \boldsymbol{\gamma}_0'\mathbf{x}_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{u}_i / \mathbf{z}_i\right] = E\left[\mathbf{u}_i\right] \equiv \mathbf{0}$$

qui est un ***système d'équations simultanées***. Il est possible d'écrire les variables endogènes, $y_{d,i}$ et $y_{c,i}$, du modèle considéré comme des fonctions des variables exogènes du système, $\mathbf{z}_i$, ce qui donne la ***forme réduite*** du modèle :

$$\begin{cases} y_{d,i} = \boldsymbol{\pi}_d'\mathbf{z}_i + v_{d,i} \\ y_{c,i} = \boldsymbol{\pi}_c'\mathbf{z}_i + v_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0}$$

qui est un ***système de régressions empilées***.

Dans le ***cas juste-identifié***, les paramètres structurels et réduits sont liés par une bijection notée sous la forme :

$$\mathbf{g}_0 = \mathbf{g}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c) \quad \text{et} \quad \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c).$$

***3. Estimation des paramètres réduits :***

Puisque :

$$\begin{cases} y_{d,i} = \boldsymbol{\pi}'_d \mathbf{z}_i + v_{d,i} \\ y_{c,i} = \boldsymbol{\pi}'_c \mathbf{z}_i + v_{c,i} \end{cases} \quad \text{avec } E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0}$$

on sait que :

$$\boxed{\hat{\boldsymbol{\pi}}_{d,N}^{MCO} \xrightarrow[N \to +\infty]{p} \boldsymbol{\pi}_d \quad \text{et} \quad \hat{\boldsymbol{\pi}}_{c,N}^{MCO} \xrightarrow[N \to +\infty]{p} \boldsymbol{\pi}_c ,}$$

*i.e.* les équations de la forme réduite du modèle sont des équations de régression.

| |
|---|
| **Rmq.** On peut en outre montrer que cet estimateur de $\boldsymbol{\pi}$ est as. efficace. |

***4. Estimation des paramètres structurels (fin) :***

En utilisant les équations :

$$\mathbf{g}_0 = \mathbf{g}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c) \text{ et } \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c),$$

la convergence de $\hat{\boldsymbol{\pi}}_{d,N}^{MCO}$ et $\hat{\boldsymbol{\pi}}_{c,N}^{MCO}$, et le ***principe d'analogie***, il est aisé de construire des estimateurs convergents de $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$.

**Définition.** ***Estimateurs des MCI***

Les estimateurs des paramètres structurels, $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$, obtenus comme des fonctions des paramètres réduits sont les estimateurs des ***Moindres Carrés Indirects*** (***MCI***) :

$$\hat{\mathbf{g}}_N^{MCI} \equiv \mathbf{g}(\hat{\boldsymbol{\pi}}_{d,N}^{MCO}, \hat{\boldsymbol{\pi}}_{c,N}^{MCO}) \text{ et } \hat{\boldsymbol{\gamma}}_N^{MCI} \equiv \boldsymbol{\gamma}(\hat{\boldsymbol{\pi}}_{d,N}^{MCO}, \hat{\boldsymbol{\pi}}_{c,N}^{MCO}).$$

**Propriété 41.** ***Convergence des estimateurs des MCI***

$$\hat{\mathbf{g}}_N^{MCI} \equiv \mathbf{g}(\hat{\boldsymbol{\pi}}_{d,N}^{MCO}, \hat{\boldsymbol{\pi}}_{c,N}^{MCO}) \xrightarrow[N\to+\infty]{p} \mathbf{g}_0 = \mathbf{g}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c)$$

et

$$\hat{\boldsymbol{\gamma}}_N^{MCI} \equiv \boldsymbol{\gamma}(\hat{\boldsymbol{\pi}}_{d,N}^{MCO}, \hat{\boldsymbol{\pi}}_{c,N}^{MCO}) \xrightarrow[N\to+\infty]{p} \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}(\boldsymbol{\pi}_d, \boldsymbol{\pi}_c).$$

Ceci provient simplement de ce que :

$$\hat{\boldsymbol{\pi}}_{d,N}^{MCO} \xrightarrow[N\to+\infty]{p} \boldsymbol{\pi}_d \quad \text{et} \quad \hat{\boldsymbol{\pi}}_{c,N}^{MCO} \xrightarrow[N\to+\infty]{p} \boldsymbol{\pi}_c .$$

### 3.2. Identification dans le cas juste-identifié

La solution précédente peut être réécrite sous la forme :

$$\begin{cases} b_0 = \dfrac{\pi_{d,4}}{\pi_{c,4}}, \quad a_0 = \pi_{d,1} - \pi_{c,1} b_0, \quad d_0 = \pi_{d,2} - \pi_{c,2} b_0 \quad \text{et} \quad r_0 = \pi_{d,3}\Delta \\ \\ \beta_0 = \dfrac{\pi_{c,3}}{\pi_{d,3}}, \quad \alpha_0 = \pi_{c,1} - \pi_{d,1}\beta_0, \quad \delta_0 = \pi_{c,2} - \pi_{d,2}\beta_0 \quad \text{et} \quad \rho_0 = \pi_{c,4}\Delta \end{cases}$$

pour mettre en évidence que les ***paramètres structurels*** :

$$\boxed{a_0,\ b_0 \text{ et } d_0 \text{ d'une part, et } \alpha_0,\ \beta_0 \text{ et } \delta_0 \text{ d'autre part}}$$

ne peuvent calculés, *i.e.* ***ne sont identifiables***, que si :

$$\boxed{\pi_{c,4} \neq 0 \ (\Leftrightarrow \rho_0 \neq 0) \text{ d'une part et } \pi_{d,3} \neq 0 \ (\Leftrightarrow r_0 \neq 0) \text{ d'autre part}}$$

(étant donné que $\Delta \neq 0$).

**« Schéma d'identification »** des paramètres de la **forme structurelle** par ceux de la **forme réduite :**

***Rôle des variables exogènes « spécifiques »***

***Forme réduite***

$$\begin{cases} y_{d,i} = \pi_{d,1} + \pi_{d,2} q_i + \boxed{\pi_{d,3}} q_{d,i} + \pi_{d,4}\ q_{c,i} + v_{d,i} \\ y_{c,i} = \pi_{c,1} + \pi_{c,2} q_i + \pi_{c,3} q_{d,i} + \boxed{\pi_{c,4}}\ q_{c,i} + v_{c,i} \end{cases}$$

$$\boxed{\pi_{d,4} \neq 0}$$

$$\boxed{\pi_{c,3} \neq 0}$$

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + d_0 q_i + \boxed{r_0}\ q_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i + \boxed{\rho_0}\ q_{c,i} + u_{c,i} \end{cases}$$

***Forme structurelle***

Les paramètres structurels ne sont identifiables qu'à certaines conditions :

- $\mathbf{g}_0$ n'est identifiable que si $\pi_{c,4} \neq 0$ ($\Leftrightarrow \rho_0 \neq 0$).

et :

- $\boldsymbol{\gamma}_0$ n'est identifiable que si $\pi_{d,3} \neq 0$ ($\Leftrightarrow r_0 \neq 0$).

L'examen de ces conditions va en fait permettre de mettre en évidence deux points importants :

- Les ***conditions sur les paramètres réduits du modèle***, $\pi_{c,4} \neq 0$ et $\pi_{d,3} \neq 0$, sont ***analogues aux conditions de rang relatives à l'identification des paramètres des modèles à VI à partir de l'estimateur des 2MC ou des VI***.
- Les estimateurs des MCI de $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$ décrits ici sont en fait égaux aux estimateurs des VI (ou des 2MC dans le juste-identifié) de ces paramètres dans leurs équations respectives.

## 3.3. Estimation des paramètres structurels par les techniques de VI « équation par équation »

La forme structurelle du modèle permet d'examiner l'intérêt des conditions $\pi_{d,3} \neq 0$ et $\pi_{c,4} \neq 0$ :

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + d_0 q_i + r_0 q_{d,i} \qquad + u_{d,i} \text{ avec } E\left[u_{d,i} / \mathbf{z}_i\right] = 0 \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i \qquad + \rho_0 q_{c,i} + u_{c,i} \text{ avec } E\left[u_{c,i} / \mathbf{z}_i\right] = 0 \end{cases}$$

- Sous cette forme on voit que la forme structurelle du modèle est un ***empilement de modèles à VI avec*** $\mathbf{z}_i \equiv (1, q_i, q_{d,i}, q_{c,i})$ ***pour vecteur de VI pour les deux modèles***.
- Si on souhaite ***estimer*** $\mathbf{g}_0$ ***et*** $\boldsymbol{\gamma}_0$ ***dans leurs équations respectives***, il faut en fait utiliser ***des estimateurs des 2MC ou des VI***.

Les variables $q_i$, $q_{d,i}$ et $q_{c,i}$ étant exogènes, il faut instrumenter $y_{c,i}$ dans l'équation de $y_{d,i}$ et $y_{d,i}$ dans l'équation de $y_{c,i}$.

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + d_0 q_i + r_0 q_{d,i} \qquad\qquad + u_{d,i} \text{ avec } E\left[u_{d,i} / \mathbf{z}_i\right] = 0 \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \delta_0 q_i \qquad\qquad + \rho_0 q_{c,i} + u_{c,i} \text{ avec } E\left[u_{c,i} / \mathbf{z}_i\right] = 0 \end{cases}$$

- Or ce système d'équations donne ***des modèles de*** $y_{d,i}$ ***et*** $y_{c,i}$.
    - Dans un système d'équations simultanées, ***toutes les variables endogènes sont modélisées***.
    - On parle alors de ***d'inférence en information complète***.
    - ***Il n'existe pas de VI « externe » au système***.

- Les VI potentielles de $y_{d,i}$ et $y_{c,i}$ sont à trouver ***dans*** $\mathbf{z}_i \equiv (1, q_i, q_{d,i}, q_{c,i})$.

$$\begin{cases} y_{d,i} = a_0 + b_0 \boxed{y_{c,i}} + d_0 q_i + r_0 \boxed{q_{d,i}} + u_{d,i} \quad \text{avec } E\left[u_{d,i}/\mathbf{z}_i\right] = 0 \\ y_{c,i} = \alpha_0 + \beta_0 \boxed{y_{d,i}} + \delta_0 q_i + \rho_0 \boxed{q_{c,i}} + u_{c,i} \quad \text{avec } E\left[u_{c,i}/\mathbf{z}_i\right] = 0 \end{cases}$$

La variable $q_{c,i}$ ***est la seule VI potentielle de*** $y_{c,i}$ dans l'équation de $y_{d,i}$. C'est :

(*i*) le seul élément de $\mathbf{z}_i$, donc exogène

(*ii*) qui n'est pas explicative dans le modèle de $y_{d,i}$.

Mais $q_{c,i}$ ***n'est une VI valide que si*** $\rho_0 \neq 0 \Leftrightarrow \pi_{c,4} \neq 0$ car :

$$y_{c,i} = \underbrace{\pi_{c,1} + \pi_{c,2} q_i + \pi_{c,3} q_{d,i} + \pi_{c,4} q_{c,i}}_{\text{Projection de } y_{c,i} \text{ sur } \mathbf{z}_i} + v_{c,i}.$$

- L'analyse est analogue pour $y_{d,i}$ dans l'équation de $y_{c,i}$ avec $\pi_{d,3} \neq 0$.

Pour l'équation :

$$y_{d,i} = \mathbf{g}_0' \mathbf{x}_{d,i} + u_{d,i} \quad \text{avec} \quad E\left[ u_{d,i} / \mathbf{z}_i \right] = 0$$

on a :

$$\mathbf{x}_{d,i} \equiv \begin{bmatrix} 1 \\ y_{c,i} \\ q_i \\ q_{d,i} \end{bmatrix} \begin{matrix} \}\text{exog} \\ \}\textbf{endo} \\ \}\text{exog} \\ \}\text{exog} \end{matrix} \quad \text{et} \quad \mathbf{z}_i \equiv \begin{bmatrix} 1 \\ q_i \\ q_{d,i} \\ q_{c,i} \end{bmatrix} \begin{matrix} \\ \\ \\ \}\textbf{VI de } y_{c,i} \end{matrix} \quad ,$$

$\mathbf{g}_0$ est donc ***juste-identifié*** par $\mathbf{z}_i$ et :

$$\hat{\mathbf{g}}_N^{VI} = \left[ N^{-1} \sum\nolimits_{i=1}^{N} \mathbf{z}_i \mathbf{x}_{d,i}' \right] N^{-1} \sum\nolimits_{i=1}^{N} \mathbf{z}_i y_{d,i} \,.$$

## Egalité des estimateurs des VI des MCI, intuitions

Il n'y a pas de manière simple de montrer que, ***dans un système d'équations simultanées juste-identifié*** :

**Propriété 42.** ***Egalité de l'estimateur des MCI et des VI***

$$\hat{\mathbf{g}}_N^{MCI} = \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{x}'_{d,i} \right] N^{-1} \sum_{i=1}^{N} \mathbf{z}_i y_{d,i} = \hat{\mathbf{g}}_N^{VI}$$

et :

$$\hat{\boldsymbol{\gamma}}_N^{MCI} = \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{x}'_{c,i} \right] N^{-1} \sum_{i=1}^{N} \mathbf{z}_i y_{c,i} = \hat{\boldsymbol{\gamma}}_N^{VI} .$$

On montre également que, dans un système d'équations simultanées juste-identifié, les ***estimateurs des MCI*** sont ***asymptotiquement efficaces pour l'estimation*** de $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$.

On donne ici simplement quelques intuitions :

Les estimateurs des VI et des MCI ***utilisent la même information*** mais sous des ***formes différentes*** :

- Pour les estimateurs des MCI :

$$\mathbf{g}_0 = \mathbf{g}(\boldsymbol{\pi}) \;\text{ et }\; \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}(\boldsymbol{\pi})$$
$$\boldsymbol{\pi}_d = E\left[\mathbf{z}_i\mathbf{z}_i'\right]^{-1}E\left[\mathbf{z}_i y_{d,i}\right] \;\text{ et }\; \boldsymbol{\pi}_c = E\left[\mathbf{z}_i\mathbf{z}_i'\right]^{-1}E\left[\mathbf{z}_i y_{c,i}\right]$$

- Pour les estimateurs des VI :

$$\mathbf{g}_0 = E\left[\mathbf{z}_i\mathbf{x}_{d,i}'\right]^{-1}E\left[\mathbf{z}_i y_{d,i}\right] \text{ et } \boldsymbol{\gamma}_0 = E\left[\mathbf{z}_i\mathbf{x}_{c,i}'\right]^{-1}E\left[\mathbf{z}_i y_{c,i}\right].$$

- Les matrices $E\left[\mathbf{z}_i\mathbf{x}_{d,i}'\right]$ et $E\left[\mathbf{z}_i\mathbf{x}_{c,i}'\right]$ combinent les informations de $\mathbf{g}_0 = \mathbf{g}(\boldsymbol{\pi})$, $\boldsymbol{\gamma}_0 = \boldsymbol{\gamma}(\boldsymbol{\pi})$ et $E\left[\mathbf{z}_i\mathbf{z}_i'\right]$.

**Rmq**. La ***linéarité des équations du modèle structurel*** est une condition essentielle à l'égalité des estimateurs des MCI et des VI.

- Les estimateurs des VI utilisent des conditions d'orthogonalité qui sont des mesures de liaisons linéaires qui contiennent la même information que la forme réduite du modèle qui est elle-même linéaire en raison de la linéarité de la forme structurelle du modèle.

# Estimateur des 2MC « équation par équation » et conditions d'identification

## 4. Estimateur des 2MC « équation par équation » et conditions d'identification

On utilise à nouveau le modèle sous sa forme générale, *i.e.* pas nécessairement juste-identifié :

**Forme structurelle du modèle**

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0' \mathbf{q}_i + \mathbf{r}_0' \mathbf{q}_{d,i} + u_{d,i} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0' \mathbf{q}_i + \boldsymbol{\rho}_0' \mathbf{q}_{c,i} + u_{c,i} \end{cases} \quad \text{avec } E[\mathbf{u}_i / \mathbf{z}_i] = E[\mathbf{u}_i] \equiv \mathbf{0}$$

où :

$$\mathbf{u}_i \equiv (u_{d,i}, u_{c,i}) \quad \text{et} \quad \mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i}).$$

En utilisant la même technique que pour la version simplifiée, il est possible de déterminer la forme réduite de ce modèle :

**Forme réduite du modèle**

$$\begin{cases} y_{d,i} = \pi_{d,1} + \boldsymbol{\pi}'_{d,2}\mathbf{q}_i + \boldsymbol{\pi}'_{d,3}\mathbf{q}_{d,i} + \boldsymbol{\pi}'_{d,4}\mathbf{q}_{c,i} + v_{d,i} \\ y_{c,i} = \pi_{c,1} + \boldsymbol{\pi}'_{c,2}\mathbf{q}_i + \boldsymbol{\pi}'_{c,3}\mathbf{q}_{d,i} + \boldsymbol{\pi}'_{c,4}\mathbf{q}_{c,i} + v_{c,i} \end{cases}$$

avec :

$$\boldsymbol{\pi}_d = \begin{bmatrix} \pi_{d,1} \\ \boldsymbol{\pi}_{d,2} \\ \boldsymbol{\pi}_{d,3} \\ \boldsymbol{\pi}_{d,4} \end{bmatrix} \equiv \Delta^{-1} \times \begin{bmatrix} a_0 + b_0\alpha_0 \\ \mathbf{d}_0 + b_0\boldsymbol{\delta}_0 \\ \mathbf{r}_0 \\ b_0\boldsymbol{\rho}_0 \end{bmatrix} \quad \text{et} \quad \boldsymbol{\pi}_c = \begin{bmatrix} \pi_{c,1} \\ \boldsymbol{\pi}_{c,2} \\ \boldsymbol{\pi}_{c,3} \\ \boldsymbol{\pi}_{c,4} \end{bmatrix} \equiv \Delta^{-1} \times \begin{bmatrix} \alpha_0 + \beta_0 a_0 \\ \boldsymbol{\delta}_0 + \beta_0\mathbf{d}_0 \\ \beta_0\mathbf{r}_0 \\ \boldsymbol{\rho}_0 \end{bmatrix}$$

$$\text{où } \Delta \equiv 1 - b_0\beta_0$$

$$v_{d,i} \equiv \Delta^{-1} \times (u_{d,i} + b_0 u_{c,i}) \quad \text{et} \quad v_{c,i} \equiv \Delta^{-1} \times (u_{c,i} + \beta_0 u_{d,i})$$

et :

$$E\left[\mathbf{v}_i / \mathbf{z}_i\right] = E\left[\mathbf{v}_i\right] = \mathbf{0} \quad \text{où} \quad \mathbf{v}_i \equiv (v_{d,i}, v_{c,i})$$

Cependant dans ce cas, les équations :

$$\begin{cases} \pi_{d,1} = (1-b_0\beta_0)^{-1}\times(a_0+b_0\alpha_0) \\ \boldsymbol{\pi}_{d,2} = (1-b_0\beta_0)^{-1}\times(\mathbf{d}_0+b_0\boldsymbol{\delta}_0) \\ \boldsymbol{\pi}_{d,3} = (1-b_0\beta_0)^{-1}\times\mathbf{r}_0 \\ \boldsymbol{\pi}_{d,4} = (1-b_0\beta_0)^{-1}\times b_0\boldsymbol{\rho}_0 \end{cases} \quad \text{et} \quad \begin{cases} \pi_{c,1} = (1-b_0\beta_0)^{-1}\times(\alpha_0+\beta_0 a_0) \\ \boldsymbol{\pi}_{c,2} = (1-b_0\beta_0)^{-1}\times(\boldsymbol{\delta}_0+\beta_0\mathbf{d}_0) \\ \boldsymbol{\pi}_{c,3} = (1-b_0\beta_0)^{-1}\times\beta_0\mathbf{r}_0 \\ \boldsymbol{\pi}_{c,4} = (1-b_0\beta_0)^{-1}\times\boldsymbol{\rho}_0 \end{cases}$$

définissent un système à :

$$2\times(1+\dim\mathbf{q}_i+\dim\mathbf{q}_{d,i}+\dim\mathbf{q}_{c,i}) \text{ équations}$$

pour

$$2\times(2+\dim\mathbf{q}_i)+\dim\mathbf{q}_{d,i}+\dim\mathbf{q}_{c,i} \text{ inconnues.}$$

Il y a donc potentiellement :

$(\dim\mathbf{q}_{d,i}+\dim\mathbf{q}_{c,i}-2)$ équations ***sur-identifiant*** les paramètres structurels.

Dans le ***cas sur-identifié***, il est ***impossible d'écrire*** les éléments des paramètres structurels $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$ à partir des éléments des paramètres réduits $\boldsymbol{\pi}_d$ et $\boldsymbol{\pi}_c$ (sauf à éliminer des équations, ce qui constituerait une perte d'information)

Ce qui implique que ***l'approche par les MCI n'est pas*** (directement) ***utilisable***.

En revanche, il est toujours possible d'estimer les paramètres $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$ avec ***des estimateurs des 2MC*** à partir des équations du modèle structurel,

sous certaines conditions d'identification.

C'est ce qui est examiné ici.

La forme structurelle du modèle :

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0' \mathbf{q}_i + \mathbf{r}_0' \mathbf{q}_{d,i} + u_{d,i} \quad \text{avec } E\left[ u_{d,i} / \mathbf{z}_i \right] = 0 \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0' \mathbf{q}_i + \boldsymbol{\rho}_0' \mathbf{q}_{c,i} + u_{c,i} \quad \text{avec } E\left[ u_{c,i} / \mathbf{z}_i \right] = 0 \end{cases}$$

est en fait constitué de ***deux modèles à VI***, le vecteur de VI commun aux deux équations étant $\mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i})$.

Si on souhaite ***estimer*** $\mathbf{g}_0$ ***et*** $\boldsymbol{\gamma}_0$ ***dans leurs équations respectives***, il faut en fait utiliser ***des estimateurs des 2MC*** dans une logique équation par équation.

- Les variables $\mathbf{q}_i$, $\mathbf{q}_{d,i}$ et $\mathbf{q}_{c,i}$ étant exogènes, il faut instrumenter $y_{c,i}$ dans l'équation de $y_{d,i}$ et $y_{d,i}$ dans l'équation de $y_{c,i}$.
- ***Il n'existe pas de VI « externe » au système***. Donc il faut trouver des VI pour instrumenter $y_{c,i}$ et $y_{d,i}$ ***dans*** $\mathbf{z}_i$.

$$
\begin{cases}
y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0' \mathbf{q}_i + \mathbf{r}_0' \mathbf{q}_{d,i} + u_{d,i} & \text{avec } E\left[u_{d,i} / \mathbf{z}_i\right] = 0 \\
y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0' \mathbf{q}_i + \boldsymbol{\rho}_0' \mathbf{q}_{c,i} + u_{c,i} & \text{avec } E\left[u_{c,i} / \mathbf{z}_i\right] = 0
\end{cases}
$$

- Le vecteur $\mathbf{q}_{c,i}$ ***contient les seules VI potentielles de*** $y_{c,i}$ dans l'équation de $y_{d,i}$.
- Mais un élément de ce vecteur n'est potentiellement une VI valide que si l'élément correspondant dans $\boldsymbol{\delta}_0$ est non nul. Le paramètre structurel $\mathbf{g}_0$ n'est identifié que si $\boldsymbol{\delta}_0 \neq \mathbf{0}$, *i.e.* si $\boldsymbol{\delta}_0$ a (au moins) un élément non nul.
- L'analyse est analogue pour $y_{d,i}$ dans l'équation de $y_{c,i}$ avec $\mathbf{q}_{d,i}$.

Les variables explicatives exogènes n'apparaissant pas dans une équation sont dites ***exclues*** de cette équation. Ces ***relations d'exclusion*** sont fondamentales pour l'***identification des paramètres structurels***.

Pour l'équation :

$$y_{d,i} = \mathbf{g}_0'\mathbf{x}_{d,i} + u_{d,i} \quad \text{avec} \quad E\left[u_{d,i} / \mathbf{z}_i\right] = 0$$

on a :

$$\mathbf{x}_{d,i} \equiv \begin{bmatrix} 1 \\ y_{c,i} \\ \mathbf{q}_i \\ \mathbf{q}_{d,i} \end{bmatrix} \begin{matrix} \}\text{exog} \\ \}\textbf{endo} \\ \}\text{exog} \\ \}\text{exog} \end{matrix} \quad \text{et} \quad \mathbf{z}_i \equiv \begin{bmatrix} 1 \\ \mathbf{q}_i \\ \mathbf{q}_{d,i} \\ \mathbf{q}_{c,i} \end{bmatrix} \begin{matrix} \\ \\ \\ \}\textbf{vecteur de VI de } y_{c,i} \end{matrix} \quad ,$$

$\mathbf{g}_0$ est donc potentiellement ***sur-identifié*** par $\mathbf{z}_i$ et :

$$\begin{aligned} \hat{\mathbf{g}}_N^{2MC} = & \left\{ \left[ N^{-1} \sum_{i=1}^N \mathbf{x}_{d,i}\mathbf{z}_i' \right] \left[ N^{-1} \sum_{i=1}^N \mathbf{z}_i\mathbf{z}_i' \right]^{-1} \left[ N^{-1} \sum_{i=1}^N \mathbf{z}_i\mathbf{x}_{d,i}' \right] \right\}^{-1} \\ & \times \left[ N^{-1} \sum_{i=1}^N \mathbf{x}_{d,i}\mathbf{z}_i' \right] \left[ N^{-1} \sum_{i=1}^N \mathbf{z}_i\mathbf{z}_i' \right]^{-1} N^{-1} \sum_{i=1}^N \mathbf{z}_i y_{d,i} \end{aligned} .$$

Dans un système d'équations simultanées de forme générale :

$$\begin{cases} \vdots \\ y_{m,i} = a_{m,0} + \mathbf{b}'_{m,0}\mathbf{y}_{m,i} + \mathbf{d}'_{m,0}\mathbf{q}_{m,i} + \mathbf{r}'_0\mathbf{q}^s_{m,i} + u_{m,i} \quad \text{avec } E\left[u_{m,i} \middle/ \mathbf{z}_i\right] = 0 \\ \vdots \\ \text{pour } m = 1,\ldots,M \end{cases}$$

où :

$\mathbf{y}_i$ contient l'ensemble des var. endogènes du système

$\mathbf{z}_i$ contient l'ensemble des var. exogènes du système

$\mathbf{y}_{m,i}$ contient le vecteur des var. explicatives endogènes de $y_{m,i}$

$\mathbf{q}_{m,i}$ contient les var. explicatives exogènes communes à plusieurs équations

$\mathbf{q}^s_{m,i}$ contient les var. explicatives exogènes spécifiques de l'équation $m$

on examine ce qu'on appelle les ***conditions d'ordre*** pour chaque équation du système :

$$y_{m,i} = a_{m,0} + \mathbf{b}'_{m,0}\mathbf{y}_{m,i} + \mathbf{d}'_{m,0}\mathbf{q}_{m,i} + \mathbf{r}'_0\mathbf{q}^s_{m,i} + u_{m,i} \quad \text{avec} \quad E\left[u_{m,i} / \mathbf{z}_i\right] = 0$$

***Condition d'ordre pour l'équation*** *m*

La condition d'ordre est vérifiée pour l'équation *m* du système si :

$$\dim \mathbf{z}_i \geq 1 + \dim \mathbf{y}_{-m,i} + \dim \mathbf{q}^s_{m,i} + \dim \mathbf{q}_{m,i}$$

Nb de variables exog. du système ≥ Nb de variables explicatives de l'équation

« Nb de VI potentielles ≥ Nb de variables explicatives, pour l'équation »

Bien entendu :

- La satisfaction des conditions d'ordre n'est ***pas suffisante***. Elle ne garantit pas les conditions de rang.
- La satisfaction des conditions d'ordre n'est pas nécessaire si des contraintes entre les paramètres de différentes équations sont imposées.

***Condition** « **presque** » **suffisante***
***pour la condition de rang du système***

La condition de rang est « presque » vérifiée pour le système d'équations simultanées si :

$$\dim \mathbf{q}_{m,i}^{s} > 0 \text{ pour } m = 1,\ldots,M ,$$

*i.e.* si chaque variable endogène du système $y_{m,i}$ a des ***variables explicatives*** qui lui sont ***spécifiques***, c'est-à-dire qui ne sont pas des variables explicatives des autres variables endogènes du système.

- Cette condition n'est pas suffisante car elle ne garantit pas que $\mathbf{d}_{m,0} \neq \mathbf{0}$ pour $m = 1,\ldots,M$ , d'où le « presque ».
- Elle n'est pas nécessaire non plus. C'est par exemple le cas si $y_{\ell,i}$ n'est jamais variable explicative dans le système. Dans ce cas $\dim \mathbf{q}_{\ell,i} = 0$, *i.e.* « $\mathbf{q}_{\ell,i}$ n'existe pas », ne pose pas de problème d'identification.

## Exemple : une équation avec deux variables explicatives endogènes

***« Schéma » d'instrumentation*** :

$$y_{m,i} = a_{m,0} + b_{\ell,0}\boxed{y_{\ell,i}} + b_{n,0}\boxed{y_{n,i}} + \mathbf{d}'_{m,0}\mathbf{q}_{m,i} + \mathbf{r}'_0\mathbf{q}^s_{m,i} + u_{m,i}$$

$$y_{\ell,i} = \boldsymbol{\pi}'_\ell \mathbf{z}_i + v_{\ell,i} = \boldsymbol{\pi}_{-s,\ell}\mathbf{z}_{-s,\ell,i} + \boldsymbol{\pi}'_{s,\ell}\boxed{\mathbf{q}^s_{\ell,i}} + v_{\ell,i}$$

$$y_{n,i} = \boldsymbol{\pi}'_\ell \mathbf{z}_i + v_{n,i} = \boldsymbol{\pi}'_{-s,n}\mathbf{z}_{-s,n,i} + \boldsymbol{\pi}'_{s,n}\boxed{\mathbf{q}^s_{n,i}} + v_{n,i}$$

***2MC, forme réduite et projection*** (pour ce qui concerne $y_{\ell,i}$) :

- $\boldsymbol{\pi}'_\ell \mathbf{z}_i$ : Projection de $y_{\ell,i}$ sur $\mathbf{z}_i$, ***partie exogène*** de $y_{\ell,i}$.
- $\boldsymbol{\pi}'_{s,\ell}\mathbf{q}^s_{\ell,i}$ : Partie exogène « ***spécifique*** » de la projection de $y_{\ell,i}$ sur $\mathbf{z}_i$.
- $y_{\ell,i}$ est ***endogène*** dans l'équation de $y_{m,i}$ si $E\left[u_{m,i}v_{\ell,i}\right] \neq 0$ car $E\left[\mathbf{z}_i v_{\ell,i}\right] = \mathbf{0}$

## Estimation « en système » versus « équation par équation »

## 5. Estimation « en système » *versus* « équation par équation »

L'objectif de cette section est double :

- Il s'agit d'une part de discuter les différences (et avantages et inconvénients) de la ***logique d'estimation « en système »*** par opposition à la ***logique d'estimation « équation par équation »***.
    - Schématiquement :

La logique « en système » est ***plus efficace*** mais également ***plus risquée***, ce qui implique un ***arbitrage*** en pratique entre ***précision d'estimation*** et ***robustesse*** des résultats obtenus.

- Il s'agit d'autre part de discuter les intérêts et limites des systèmes d'équations simultanées, *i.e.* de discuter les intérêts et limites de ce que les économètres nomment la ***logique d'inférence en information complète*** et la ***logique d'inférence en information limitée***.

### 5.1. Estimateur en système *versus* estimateur équation par équation

- L'estimateur des MCI est le seul ***estimateur en système*** utilisé jusqu'à présent.
  - Il n'est applicable que lorsque les paramètres structurels sont juste-identifiés dans le système d'équations simultanées considérées.
  - Il exploite « à plein » la structure du système d'équations simultanées

- Pour le cas sur-identifié, l'estimation des paramètres du système a été considérée par les ***2MC*** dans une ***logique équation par équation***.

- Il existe un ***estimateur en système des paramètres d'un système d'équations simultanées sur-identifié***: l'***estimateur des Triples Moindres Carrés (3MC)***.
  - Cet estimateur est généralement ***plus efficace*** (***as.***) que « l'empilement » des estimateurs de 2MC de chaque équation.
  - Une exception est justement le cas juste-identifié.

- L'***estimateur des 3MC*** est construit comme un estimateur des 2MC dans une ***logique en système***, ce qui consiste ici à tenir compte de ce que les éléments de $\mathbf{u}_i$ ***peuvent être corrélés.***
- L'***estimateur des 2MC équation par équation*** est en fait construit sans exploiter les corrélations entre les éléments de $\mathbf{u}_i$, ce qui explique qu'il n'est en général pas as. efficace.
    - L'estimateur des 3MC tient compte des relations entre les éléments de $\mathbf{u}_i$, *e.g.* $Cov\left[u_{d,i};u_{c,i}\right]\neq 0$.

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0'\mathbf{q}_i + \mathbf{r}_0'\mathbf{q}_{d,i} + \boxed{u_{d,i}} \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0'\mathbf{q}_i + \boldsymbol{\rho}_0'\mathbf{q}_{c,i} + \boxed{u_{c,i}} \end{cases}$$

    - C'est ce qui lui confère son efficacité mais également sa « fragilité » :

***Si une seule équation du système est mal spécifiée**, **l'estimateur des 3MC est généralement biaisé pour l'ensemble des paramètres du système**.*

## 5.2. Inférence en information complète *versus* en information limitée

- L'arbitrage entre ***estimation en système*** (plus efficace) et ***estimation équation par équation*** (plus robuste) se situe en fait au niveau même de ***la spécification du modèle***.

- Pour illustrer ce point, on reprend ici l'exemple du modèle relatif aux relations « chômage/délinquance » dont la forme structurelle est donnée par :

***Inférence en information complète (avec 3MC)***

$$\begin{cases} y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0' \mathbf{q}_i + \mathbf{r}_0' \mathbf{q}_{d,i} + u_{d,i} \quad \text{avec} \quad E\left[u_{d,i} / \mathbf{z}_i\right] = 0 \\ y_{c,i} = \alpha_0 + \beta_0 y_{d,i} + \boldsymbol{\delta}_0' \mathbf{q}_i + \boldsymbol{\rho}_0' \mathbf{q}_{c,i} + u_{c,i} \quad \text{avec} \quad E\left[u_{c,i} / \mathbf{z}_i\right] = 0 \end{cases}$$

avec $\mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i})$.

- On parle ici d’une ***logique d’inférence en information complète*** car ***chaque variable endogène du système est modélisée*** :

    - Le modèle décrivant les relations causales entre chômage et délinquance est en quelque sorte un ***modèle complet***.

- Les ***estimateurs des 3MC ou des MCI*** de $\mathbf{g}_0$ et $\boldsymbol{\gamma}_0$ exploitent l’ensemble de l’information contenue dans le modèle, ce qui leur donne leur efficacité asymptotique.

- En particulier, ils exploitent les hypothèses relatives à la spécification du modèle de la variable de chômage :

    (*i*) le modèle de $y_{c,i}$ est linéaire en $y_{d,i}$, $\mathbf{q}_i$ et $\mathbf{q}_{c,i}$

  et :

    (*ii*) les vecteurs $\mathbf{q}_i$ et $\mathbf{q}_{c,i}$ sont exogènes par rapport à $u_{c,i}$

- Or, et c'est là un point important :

***Si on ne s'intéresse qu'à l'effet causal de du chômage sur la délinquance*** (ce qui est en fait le cas de Fougère, Kramarz et Pouget), *i.e.* si on souhaite surtout estimer $\mathbf{g}_0$,
***il n'est alors pas nécessaire de spécifier un modèle pour la variable de délinquance***.

- L'idée essentielle est ici que l'estimation par les ***2MC*** de l'équation de chômage ne requiert que :

***Inférence en information limitée (avec 2MC)***

$$y_{d,i} = a_0 + b_0 y_{c,i} + \mathbf{d}_0' \mathbf{q}_i + \mathbf{r}_0' \mathbf{q}_{d,i} + u_{d,i} = \mathbf{g}_0' \mathbf{x}_{d,i} + u_{d,i} \quad \text{avec} \quad E\left[u_{d,i} / \mathbf{z}_i\right] = 0$$

$$\text{avec } \mathbf{z}_i \equiv (1, \mathbf{q}_i, \mathbf{q}_{d,i}, \mathbf{q}_{c,i}) \text{ et } rangE\left[\mathbf{x}_{d,i} \mathbf{z}_i'\right] = \dim \mathbf{x}_{d,i}$$

- En particulier, il n'est pas nécessaire :

  ($i$) que le modèle de $y_{c,i}$ soit linéaire en $y_{d,i}$, $\mathbf{q}_i$ et $\mathbf{q}_{c,i}$

  et :

  ($ii$) que les vecteurs $\mathbf{q}_i$ et $\mathbf{q}_{c,i}$ soient exogènes par rapport à $u_{c,i}$

  comme ***en information complète***.

- Ce qui est nécessaire ***en information limitée*** est que :

  ($a$) que le modèle de $y_{d,i}$ soit linéaire en $y_{d,i}$, $\mathbf{q}_i$ et $\mathbf{q}_{d,i}$,

  ($b$) que les vecteurs $\mathbf{q}_i$ et $\mathbf{q}_{d,i}$ soient exogènes par rapport à $u_{d,i}$

et :

  ($c$) que $\mathbf{q}_{c,i}$ soit exogène par rapport à $u_{d,i}$

  ($d$) $rangE\left[\mathbf{x}_{d,i}\mathbf{z}_i'\right] = \dim \mathbf{x}_{d,i}$.

- Les conditions (*a*) et (*b*) sont communes aux deux approches.

- Les conditions (*c*) et (*d*) assurent que $\mathbf{q}_{c,i}$ un vecteur de VI valide de $y_{c,i}$ dans l'équation de délinquance.
- Il faut que $\mathbf{q}_{c,i}$ permettent le ***calcul de bons prédicteurs*** de $y_{c,i}$ (sans nécessairement de sens causal)

    - Les conditions (*c*) et (*d*) sont impliquées par les conditions (*i*) et (*ii*) mais sont beaucoup moins restrictives que ces conditions.

    - C'est la raison pour laquelle ***l'inférence en information limitée est plus <u>robuste</u> que l'inférence en information complète***.

L'***arbitrage*** entre ***précision d'estimation*** (grâce à l'information complète) et ***robustesse des estimations*** (grâce à l'information limitée à apporter au modèle) est, ***en pratique***, généralement ***tranché en faveur de la robustesse***.

## Pour conclure

***Equation de « délinquance »***

$$déli_i = a_0 + b_0 \times chom_i + \mathbf{d}'_0 \begin{bmatrix} dens\_pop_i \\ \mathbf{éduc}_i \\ \mathbf{tissu\_éco}_i \end{bmatrix} + \mathbf{r}'_0 \begin{bmatrix} \mathbf{aid\_chom}_i \\ \mathbf{inég}_i \\ \mathbf{crim\_org}_i \end{bmatrix} + u_{d,i}$$

avec :

$$\mathbf{z}_i \equiv \begin{bmatrix} dens\_pop_i \\ \mathbf{éduc}_i \\ \mathbf{tissu\_éco}_i \\ \mathbf{prév\_crois}_i \end{bmatrix} \text{ et } \tilde{\mathbf{z}}_i^e \equiv \mathbf{prév\_crois}_i$$

Fougère, Kramarz et Pouget montrent que :

- Le vecteur d'instruments « externes » $\tilde{\mathbf{z}}_i^e \equiv \mathbf{prév\_crois}_i$ prédit bien la variable explicative $chom_i$.

- Le chômage des jeunes a un effet causal positif et statistiquement significatif sur la « petite » délinquance de type économique (vols, cambriolages, trafic de stupéfiants, …) mais pas beaucoup les crimes violents.

- L'effet du chômage des jeunes est atténué par l'éducation, la présence policière, …

- L'estimation par les MCO sous-estime très significativement $b_0$.
    - L'effet de $deli_i$ sur le chômage est pire plus les moins jeunes que pour les jeunes.
    - $chom_i$ mesure seulement le chômage des jeunes et $deli_i$ mesure l'ensemble de la délinquance (un autre intérêt de l'approche en information limitée ici).

## Calcul de la distribution as de $\hat{\theta}_N^{MCI} \equiv (\hat{g}_N^{MCI}, \hat{\gamma}_N^{MCI})$

## 6. Calcul de la distribution as de $\hat{\boldsymbol{\theta}}_N^{MCI} \equiv (\hat{\mathbf{g}}_N^{MCI}, \hat{\boldsymbol{\gamma}}_N^{MCI})$

Pour calculer la distribution as. de l'estimateur des MCI :

$$\hat{\boldsymbol{\theta}}_N^{MCI} \equiv \begin{bmatrix} \hat{\mathbf{g}}_N^{MCI} \\ \hat{\boldsymbol{\gamma}}_N^{MCI} \end{bmatrix} \text{ de } \boldsymbol{\theta}_0 \equiv \begin{bmatrix} \mathbf{g}_0 \\ \boldsymbol{\gamma}_0 \end{bmatrix}$$

à partir de celle de

$$\hat{\boldsymbol{\pi}}_N^{MCO} \equiv \begin{bmatrix} \hat{\boldsymbol{\pi}}_{d,N}^{MCO} \\ \hat{\boldsymbol{\pi}}_{c,N}^{MCO} \end{bmatrix} \text{ avec } \boldsymbol{\pi} \equiv \begin{bmatrix} \boldsymbol{\pi}_d \\ \boldsymbol{\pi}_c \end{bmatrix}$$

on utilise la « Méthode du delta » en remarquant que :

$$\boldsymbol{\theta}_0 = \mathbf{h}(\boldsymbol{\pi}) \text{ avec } \mathbf{h}(.) \equiv \begin{bmatrix} \mathbf{g}(.) \\ \boldsymbol{\gamma}(.) \end{bmatrix} \text{ et } \hat{\boldsymbol{\theta}}_N^{MCI} = \mathbf{h}(\hat{\boldsymbol{\pi}}_N^{MCO}).$$

**Propriété 42**. ***Méthode du delta et normalité asymptotique*.**

Si :

$$\sqrt{N}(\mathbf{a}_N - \mathbf{a}_0) \xrightarrow[N\to+\infty]{L} N(\mathbf{0}_{K\times 1}, \mathbf{\Psi}_0)$$

alors, pour toute fonction continûment différentiable, $\mathbf{g}(.): \mathbb{R}^K \to \mathbb{R}^Q$, on a :

$$\sqrt{N}\left(\mathbf{g}(\mathbf{a}_N) - \mathbf{g}(\mathbf{a}_0)\right) \xrightarrow[N\to+\infty]{L} N\left(\mathbf{0}_{K\times 1}, \frac{\partial \mathbf{g}(\mathbf{a}_0)}{\partial \mathbf{a}'} \mathbf{\Psi}_0 \frac{\partial \mathbf{g}(\mathbf{a}_0)'}{\partial \mathbf{a}}\right).$$

- Le nom « Méthode du delta » provient des « $\partial$ » de différentiation du calcul du Jacobien en **a** de $\mathbf{g}(\mathbf{a})$ en $\mathbf{a}_0$ : $\partial \mathbf{g}(\mathbf{a}_0) / \partial \mathbf{a}'$.

- Cette propriété est très utile pour déterminer la distribution as. :
    - de fonctions d'estimateurs as. normaux
    - de statistiques de test (de Wald) en particulier.

Afin de simplifier la suite on utilise les produits de Kronecker ou produits tensoriels, et leurs propriétés.

**Définition**. ***Produit de Kronecker*** de **A** par **B**

Avec :

$$\underbrace{\mathbf{A}}_{P\times K} \equiv \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{21} \\ \vdots & \vdots & \ddots & \vdots \\ a_{P1} & a_{P2} & \cdots & a_{PK} \end{bmatrix},$$

on a :

$$\underbrace{\mathbf{A}}_{P\times K} \otimes \underbrace{\mathbf{B}}_{M\times N} \equiv \underbrace{\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1K}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{21}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{P1}\mathbf{B} & a_{P2}\mathbf{B} & \cdots & a_{PK}\mathbf{B} \end{bmatrix}}_{PM\times KN}.$$

**Propriété 41**. ***Principales propriétés des produits de Kronecker***.

(*i*) Pour des matrices de tailles conformes on a (*coagulation*) :

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD}).$$

(*ii*) Pour des matrices inversibles on a (*inversion*) :

$$(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}.$$

(*iii*) Pour toutes matrices on a (*transposition*) :

$$(\mathbf{A}\otimes\mathbf{B})' = \mathbf{A}'\otimes\mathbf{B}'.$$

Dans la suite, la matrice $\mathbf{I}_H$ est la matrice identité de dimension *H*.

Pour calculer la distribution as. de $\hat{\boldsymbol{\pi}}_N^{MCO} \equiv (\hat{\boldsymbol{\pi}}_{d,N}^{MCO}, \hat{\boldsymbol{\pi}}_{c,N}^{MCO})$ on utilise :

$$\hat{\boldsymbol{\pi}}_{d,N}^{MCO} = \boldsymbol{\pi}_d + \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1} N^{-1} \sum_{i=1}^{N} \mathbf{z}_i v_{d,i}$$

et :

$$\hat{\boldsymbol{\pi}}_{c,N}^{MCO} = \boldsymbol{\pi}_c + \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1} N^{-1} \sum_{i=1}^{N} \mathbf{z}_i v_{c,i}$$

pour écrire $\hat{\boldsymbol{\pi}}_N^{MCO}$ sous la forme :

$$\hat{\boldsymbol{\pi}}_N^{MCO} = \boldsymbol{\pi} + \begin{bmatrix} N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' & \mathbf{0} \\ \mathbf{0} & N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \end{bmatrix}^{-1} \begin{bmatrix} N^{-1} \sum_{i=1}^{N} \mathbf{z}_i v_{d,i} \\ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i v_{c,i} \end{bmatrix}$$

ou encore :

$$\hat{\boldsymbol{\pi}}_N^{MCO} = \boldsymbol{\pi} + \left( \mathbf{I}_2 \otimes \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1} \right) \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \otimes \mathbf{v}_i \right] \text{ avec } \mathbf{v}_i \equiv \begin{bmatrix} v_{d,i} \\ v_{c,i} \end{bmatrix}$$

Ceci implique que :

$$\sqrt{N}\,(\hat{\boldsymbol{\pi}}_N^{MCO} - \boldsymbol{\pi}) \xrightarrow[N\to+\infty]{L} N(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

avec :

$$\boldsymbol{\Sigma}_0 \equiv \begin{bmatrix} E\left[\mathbf{z}_i\mathbf{z}_i'\right] & \mathbf{0} \\ \mathbf{0} & E\left[\mathbf{z}_i\mathbf{z}_i'\right] \end{bmatrix}^{-1} \begin{bmatrix} E\left[v_{d,i}^2\mathbf{z}_i\mathbf{z}_i'\right] & E\left[v_{d,i}v_{c,i}\mathbf{z}_i\mathbf{z}_i'\right] \\ E\left[v_{d,i}v_{c,i}\mathbf{z}_i\mathbf{z}_i'\right] & E\left[v_{c,i}^2\mathbf{z}_i\mathbf{z}_i'\right] \end{bmatrix} \begin{bmatrix} E\left[\mathbf{z}_i\mathbf{z}_i'\right] & \mathbf{0} \\ \mathbf{0} & E\left[\mathbf{z}_i\mathbf{z}_i'\right] \end{bmatrix}^{-1}$$

ou encore :

$$\boldsymbol{\Sigma}_0 = \left\{\mathbf{I}_2 \otimes E\left[\mathbf{z}_i\mathbf{z}_i'\right]^{-1}\right\} E\left[(\mathbf{v}_i\mathbf{v}_i') \otimes (\mathbf{z}_i\mathbf{z}_i')\right] \left\{\mathbf{I}_2 \otimes E\left[\mathbf{z}_i\mathbf{z}_i'\right]^{-1}\right\} .$$

Dans le cas homoscédastique, *i.e.* si $V\left[\mathbf{v}_i/\mathbf{z}_i\right] = E\left[\mathbf{v}_i\mathbf{v}_i'/\mathbf{z}_i\right] \equiv \boldsymbol{\Omega}_0$, on a :

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Omega}_0 \otimes E\left[\mathbf{z}_i\mathbf{z}_i'\right]^{-1}.$$

Appliquée ici, la Méthode du delta permet de montrer que :

$$\sqrt{N}(\hat{\boldsymbol{\pi}}_N^{MCO} - \boldsymbol{\pi}) \xrightarrow[N\to+\infty]{L} N(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

implique :

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N^{MCI} - \boldsymbol{\theta}_0) \xrightarrow[N\to+\infty]{L} N\left(\mathbf{0}_{K\times 1}, \frac{\partial \mathbf{h}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}'} \boldsymbol{\Sigma}_0 \frac{\partial \mathbf{h}(\boldsymbol{\pi})'}{\partial \boldsymbol{\pi}}\right),$$

*i.e.* que $\hat{\boldsymbol{\theta}}_N^{MCI}$ est as. normal, puisque :

$$\boldsymbol{\theta}_0 = \mathbf{h}(\boldsymbol{\pi}) \text{ et } \hat{\boldsymbol{\theta}}_N^{MCI} = \mathbf{h}(\hat{\boldsymbol{\pi}}_N^{MCO}).$$

Bien entendu on a :

$$\frac{\partial \mathbf{h}(\hat{\boldsymbol{\pi}}_N^{MCO})}{\partial \boldsymbol{\pi}} \hat{\boldsymbol{\Sigma}}_N \frac{\partial \mathbf{h}(\hat{\boldsymbol{\pi}}_N^{MCO})'}{\partial \boldsymbol{\pi}} \xrightarrow[N\to+\infty]{p} \frac{\partial \mathbf{h}(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}} \boldsymbol{\Sigma}_0 \frac{\partial \mathbf{h}(\boldsymbol{\pi})'}{\partial \boldsymbol{\pi}} \text{ si } \hat{\boldsymbol{\Sigma}}_N \xrightarrow[N\to+\infty]{p} \boldsymbol{\Sigma}_0.$$

L'estimateur, $\hat{\boldsymbol{\Sigma}}_N$, de $\boldsymbol{\Sigma}_0$ peut être construit comme la contre-partie empirique de $\boldsymbol{\Sigma}_0$ en $\hat{\boldsymbol{\pi}}_N^{MCO}$.

Avec :

$$\hat{v}_{d,i,N} \equiv y_{d,i} - \hat{\boldsymbol{\pi}}^{MCO}_{d,N}\,,\;\; \hat{v}_{d,i,N} \equiv y_{c,i} - \hat{\boldsymbol{\pi}}^{MCO}_{c,N} \;\text{ et }\; \hat{\mathbf{v}}_{d,i,N} \equiv \begin{bmatrix} \hat{v}_{d,i,N} \\ \hat{v}_{c,i,N} \end{bmatrix}$$

on a :

***Cas hétéroscédastique***

$$\hat{\boldsymbol{\Sigma}}_N = \left\{ \mathbf{I}_2 \otimes \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1} \right\} \times \left[ N^{-1} \sum_{i=1}^{N} (\hat{\mathbf{v}}_{i,N} \hat{\mathbf{v}}'_{i,N}) \otimes (\mathbf{z}_i \mathbf{z}_i') \right] \times \left\{ \mathbf{I}_2 \otimes \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1} \right\}$$

***Cas homoscédastique***

$$\hat{\boldsymbol{\Sigma}}_N = \left[ N^{-1} \sum_{i=1}^{N} (\hat{\mathbf{v}}_{i,N} \hat{\mathbf{v}}'_{i,N}) \right] \otimes \left[ N^{-1} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i' \right]^{-1}$$