

# **ÉCONOMÉTRIE (UGA S2) PRÉSENTATION <sup>1</sup>**

Michal W. Urdanivia \*

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

1<sup>er</sup> mars 2023

# Contenu

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?
2. La démarche des économètres
3. Spécificités de l'économétrie (statistique)
4. Objectifs, organisation et plan du cours
5. Notations et rappels de statistique

# PLAN

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?

2. La démarche des économètres

3. Spécificités de l'économétrie (statistique)

4. Objectifs, organisation et plan du cours

5. Notations et rappels de statistique

## **1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?**

# **Introduction de l' « Introduction à l'économétrie »**

- 1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?**
- 2. La démarche des économètres**
- 3. Les spécificités de l'économétrie, en tant que « domaine » de la statistique**
- 4. Objectifs, organisation et plan du cours**
- 5. Notations et rappels de statistique (listés ici et rappelés par la suite en temps utile)**

# 1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?

**Econométrie** : Modélisation statistique des comportements économiques

- Choix individuels : **micro-économétrie**

- Consommateurs : demande des biens marchands (dépenses d'alimentation, de transport, ...), choix de marques, ...
- Firmes : investissement, main d'œuvre, localisation, ...
- Salariés : durée de chômage, déterminants des salaires, ...
- Ménages : épargne, portefeuille financier, ...
- *Données individuelles, panels (enquêtes répétées dans le temps)*

- Grands agrégats économiques : **macro-économétrie**

- PNB, importations, consommation, épargne, taux d'intérêt, taux de change, taux de salaire, ...
- Déterminants de la croissance, du taux de chômage, ...
- *Séries temporelles, un pays ou plusieurs pays, ...*

# 1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?

**Econométrie** : Modélisation statistique des comportements économiques

- Choix individuels : **micro-économétrie**

- Consommateurs : demande des biens marchands (dépenses d'alimentation, de transport, ...), choix de marques, ...
- Firmes : investissement, main d'œuvre, localisation, ...
- Salariés : durée de chômage, déterminants des salaires, ...
- Ménages : épargne, portefeuille financier, ...
- *Données individuelles, panels (enquêtes répétées dans le temps)*

- Grands agrégats économiques : **macro-économétrie**

- PNB, importations, consommation, épargne, taux d'intérêt, taux de change, taux de salaire, ...
- Déterminants de la croissance, du taux de chômage, ...
- *Séries temporelles, un pays ou plusieurs pays, ...*

## L'économétrie, c'est d'abord de l'économie

Rationalité des choix économiques



Modèles de comportement économique

- Modèles inspirés d'éléments de **théorie micro-économique**
  - Théorie du consommateur et du producteur
  - Equilibres de marché, concurrence  $\pm$  parfaite, ...
- Modèles inspirés d'éléments de **théorie macro-économique**
  - Courbes IS-LM, théorie Keynésienne, ...

**Ceci-dit, l'idée de « lois » de l'économie est à utiliser avec précaution.**



## Utilisation des techniques de la statistique

- **Estimer** les paramètres des modèles de comportement économique
- **Tester** (ou juger de) la validité des modèles de comportement économique
- **Exemple** : Modèle de consommation

$$\text{Dépenses} = f(\text{prix\_biens}, \text{revenu}, \text{description\_ménage}; \theta_0) + \text{erreur}$$

- **Données** : Enquête de consommation (SECODIP, IPSOS, ...)
- **Estimation** : Estimateur  $\hat{\theta}_N$  et données  $\Rightarrow$  estimation de  $\theta_0$  et mesure de sa « précision »
- **Tests** : Tests de validité interne (« taille » des erreurs, signes des éléments de  $\hat{\theta}_N$ ) et externe (capacité à prédire les dépenses hors échantillon)

## Etudes économétriques, pour analyser

- Les données disponibles décrivent les choix passés des agents économiques
- Elles permettent d'analyser les *déterminants/mécanismes* de ces choix

**Exemple.** Plusieurs enquêtes sur la consommation de tabac, années marquées par des taxes croissantes sur le tabac

- On veut mesurer l'*efficacité de la taxation*. Effets « purs », hors campagnes « anti-tabac », interdictions, ... (*ceteris paribus*)
- Globalement, les *augmentations de prix diminuent la consommation*
  - Effet significatif mais limité sur les quantités consommées des fumeurs
  - Effet significatif sur les décisions d'arrêter
  - *Effet majeur* : empêche les jeunes de commencer à fumer
- Etudes récentes (avec des toxicologues) : les fumeurs consomment un peu moins de tabac, mais pratiquement autant de nicotine.

## Etudes économétriques, pour simuler/prédire

- Les données disponibles permettent l'analyse quantifiée des *déterminants/mécanismes* des choix des agents économiques
- La modélisation économétrique permet de *simuler/prédire* les effets de changements des déterminants des choix

**Exemple** : Taxation *hypothétique* des pesticides sur les choix des agriculteurs

- A court terme :
  - Peu de changement sur les choix pour une culture donnée
  - Effets significatifs sur les choix de cultures (en priorité les moins utilisatrices de pesticides)
- A moyen terme :
  - Effet significatif sur les choix pour une culture donnée
  - Effets significatifs sur les choix de cultures
  - *Idee* : Réorganisation des systèmes de production moins dépendants des pesticides

## Qui utilise l'analyse économétrique ?

- **Les décideurs publics** : ministères et institutions internationales (UE, OCDE, Banque Mondiale, FMI, ...)
  - Analyse des effets des politiques économiques ou non économiques mises en œuvre
  - Simulation/prédiction des effets des politiques économiques envisagées
  - Calcul des indices de prix
- **Les (grandes) entreprises** :
  - Finance (banque/assurance) : choix d'investissements financiers et gestion des contrats d'assurance
  - Marketing quantitatif
  - Entreprises spécialisées (BIPE, ...)
  - (Consultation des études économétriques macro- ou micro-économétriques publiées, scientifiques ou non)

# PLAN

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?

2. La démarche des économètres

3. Spécificités de l'économétrie (statistique)

4. Objectifs, organisation et plan du cours

5. Notations et rappels de statistique

## **2. La démarche des économètres**

## 2. La démarche des économètres

**1a. Analyse de la question posée :** effets de la taxe sur le tabac

⇒ Construction d'un modèle avec effets du prix du tabac

**1b. Analyse des données disponibles :** enquêtes disponibles depuis 1990

⇒ Attention : campagnes anti-tabac, interdictions dans les lieux publics, ...

**2. Spécification d'un modèle mathématique** des choix liés au tabac :

⇒ Commencer, arrêter, quantité consommée, effets d'addiction

⇒ Effets des prix et du revenu, des campagnes et interdictions

**3. Utilisation des techniques de l'inférence statistique :**

⇒ Estimation des paramètres du modèle, choix de l'*estimateur approprié*

⇒ Tests de la validité du modèle, interprétation des résultats

**4. Réponse à la question posée :**

⇒ Décomposition des effets estimés, simulations/prédictions

# PLAN

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?
2. La démarche des économètres
- 3. Spécificités de l'économétrie (statistique)**
4. Objectifs, organisation et plan du cours
5. Notations et rappels de statistique



### **3. Spécificités de l'économétrie (statistique)**

### 3. Spécificités de l'économétrie (statistique)

- **Questions posées aux économètres : analyses « contre-factuelles »**
  - Que se serait-il passé si ... ?
  - Nécessité de spécifier des modèles mettant en évidence des *mécanismes causaux* ; pas spécifique à l'économétrie mais ...
- **Pas (ou très peu) de données expérimentales**
  - **Expérience** : On veut savoir ce qui se passerait pour des sujets si « Condition A » ou « Condition B ». On place des sujets en « Condition A » et des sujets en « Condition B » et on compare.
  - Très difficile en économie, seulement « Condition réelle »
  - **Les techniques usuelles** de la statistique sont bien adaptées à l'analyse de données expérimentales:
    - Calculs de moyennes conditionnelles (tris à plat) ; régression ; ...
  - ... mais elles sont à utiliser avec précaution en économétrie.

## Point de vue « technique »

- Comportement mesuré par  $y_i$  (salaire de  $i$ ), déterminants d'intérêt mesurés par  $x_i$  (niveau d'études de  $i$ ) et (modèle linéaire simple) :

$$y_i = \alpha_0 + b_0 x_i + u_i \quad \text{avec} \quad E[u_i] \equiv 0$$

- **Pour un économètre :**  $y_i$ ,  $u_i$  et  $x_i$  *sont des variables aléatoires*
  - **Salaire**  $y_i$  : résultat du **choix** du salarié  $i$  et de son employeur (ou de ses employeurs potentiels)
  - **Niveau d'études**  $x_i$  : résultat du **choix** ( $\pm$  contraint) de  $i$
  - **Terme d'erreur**  $u_i$  : contient tout ce qui explique  $y_i$  et n'est pas expliqué par  $\alpha_0 + b_0 x_i$
- Le niveau d'étude  $x_i$  n'est ni fixé par un expérimentateur, ni parfaitement aléatoire. Il a été **choisi** (sous  $\pm$  de contraintes) par  $i$

$$y_i = \alpha_0 + b_0 x_i + u_i \quad \text{avec} \quad E[u_i] \equiv 0$$

- **Analyse économétrique :**

▪ *L'estimateur des MCO de  $b_0$  est en général biaisé* car  $Cov[x_i, u_i] \neq 0$ .

▪ En général :  $Cov[x_i, u_i] > 0$

▪ ***Idée :***  $u_i$  contient les effets de nombreux facteurs expliquant  $x_i$

▪ ***Problème d'endogénéité*** de  $x_i$  par rapport à  $u_i$

⇒ Un autre modèle du salaire est nécessaire, lequel ?

⇒ Un autre estimateur que celui des MCO doit être utilisé, lequel ?

*Un des objectifs de ce cours : analyse de ces problèmes et de leurs solutions*

# PLAN

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?
2. La démarche des économètres
3. Spécificités de l'économétrie (statistique)
4. Objectifs, organisation et plan du cours
5. Notations et rappels de statistique

## **4. Objectifs, organisation et plan du cours**

## 4. Objectifs, organisation et plan du cours

### Objectifs du cours : introduction à l'analyse économétrique

- La *démarche* des économètres
- Les *principaux modèles* utilisés : « Pièges » à éviter
- Les *principales techniques d'inférence* : « Astuces » utilisées
- Micro-économétrie essentiellement

**Remarque :** J'ai utilisé pour mes travaux tout ce je vais présenter

### Organisation du cours : classique

- Cours : théorie et exemples, mais tous les résultats pas démontrés (intuition, démonstrations dans le poly)
- TD/TP : utilisation des concepts théoriques introduits et applications (SAS)

## **Plan du cours**

### **Partie A. Modèles (linéaires) de variables continues ( $y_i$ continue)**

**Mots clés** : Identification, exogénéité/endogénéité, variables instrumentales

**Inférence statistique** : Moindres carrés et Méthode des Moments (Généralisée)

### **Partie B. Modèles à variables latentes ( $y_i$ discrète, continue/discrète)**

**Mots clés** : Variable observée/latente, mécanisme d'observation, choix discret, variable censurée, échantillon tronqué

**Inférence statistique** : Maximum de Vraisemblance

### **Partie C. Modèles à régime et mesure des effets de traitement**

**Synthèse** : Mobilise des éléments des parties A et B

**Notations et rappels de statistique particulièrement utiles pour l'économétrie**



# PLAN

1. Qu'est-ce que l'économétrie ? A quoi (à qui) ça sert ?
2. La démarche des économètres
3. Spécificités de l'économétrie (statistique)
4. Objectifs, organisation et plan du cours
5. Notations et rappels de statistique

## **5. Notations et rappels de statistique**

## 5. Notations et rappels de statistique

Convention pour l'écriture des variables, paramètres ou fonctions

*scalaire*

**vecteur colonne**

**MATRICE**

## 5.1. Echantillon

1. *Echantillon aléatoire*:  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  pour  $i = 1, \dots, N$
2. Les  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  sont des *vecteurs de variables aléatoires*
3. Les  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  sont indépendants et équi-distribués pour  $i = 1, \dots, N$   
(pas « trop » dépendants entre eux et  
avec des distributions par « trop » différentes)
4.  $N$  est grand (pour justifier des approximations asymptotiques)
5. Le tirage de l'échantillon est aléatoire, *i.e.* les  $N$  individus de l'échantillon  
sont tirés aléatoirement dans la population d'intérêt des «  $i$  ».

- La distribution commune des  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  est celle de  $(y, \mathbf{x}, \mathbf{z})$  le vecteur décrivant la distribution des réalisations des  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  dans la population puisque le *tirage de l'échantillon est aléatoire* (notion de *représentativité de l'échantillon*).
- Le modèle commun aux  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  est celui de  $(y, \mathbf{x}, \mathbf{z})$ , nommé *modèle de population*.
- On peut donc « *inférer statistiquement* » les *relations entre les éléments de  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  pour chacun des  $i$*  à partir de l'*observation des réalisations de  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  pour l'ensemble des  $i$* .
- La notion d'équidistribution sous-tend celle de modèle : les  $(y_i, \mathbf{x}_i, \mathbf{z}_i)$  suivent le même modèle s'ils ont le même PGD.
  - Si on s'intéresse à la distribution conditionnelle des  $(y_i, \mathbf{x}_i)/\mathbf{z}_i$ , le PGD des  $\mathbf{z}_i$  importe peu.

## Vecteur $\mathbf{x}_i$

$$\mathbf{x}_i \equiv \begin{bmatrix} x_{1,i} \\ x_{2,i} \\ \vdots \\ x_{K,i} \end{bmatrix}_{K \times 1}$$

$x_{1i}$  est généralement la variable constante :  $x_{1i} = 1$

$$\mathbf{x}_i \equiv \begin{bmatrix} 1 \\ x_{2,i} \\ \vdots \\ x_{K,i} \end{bmatrix} = \begin{bmatrix} 1 \\ \tilde{\mathbf{x}}_i \end{bmatrix} \text{ avec } \tilde{\mathbf{x}}_i \equiv \begin{bmatrix} x_{2,i} \\ x_{3,i} \\ \vdots \\ x_{K,i} \end{bmatrix} \equiv \begin{bmatrix} \tilde{x}_{1,i} \\ \tilde{x}_{2,i} \\ \vdots \\ \tilde{x}_{K-1,i} \end{bmatrix}$$

## Vecteur $\mathbf{z}_i$

$$\mathbf{z}_i \equiv \begin{bmatrix} z_{1,i} \\ z_{2,i} \\ \vdots \\ z_{L,i} \end{bmatrix}_{L \times 1}$$

$z_{1i}$  est généralement la variable constante :  $z_{1i} = 1$

$$\mathbf{z}_i = \begin{bmatrix} 1 \\ z_{2,i} \\ \vdots \\ z_{L,i} \end{bmatrix} = \begin{bmatrix} 1 \\ \tilde{\mathbf{z}}_i \end{bmatrix} \text{ avec } \tilde{\mathbf{z}}_i \equiv \begin{bmatrix} z_{2,i} \\ z_{3,i} \\ \vdots \\ z_{L,i} \end{bmatrix} \equiv \begin{bmatrix} \tilde{z}_{1,i} \\ \tilde{z}_{2,i} \\ \vdots \\ \tilde{z}_{L-1,i} \end{bmatrix}$$

## 5.2. Modèle linéaire

### 1. Modèle linéaire de $y_i$ en $\mathbf{x}_i$

$$y_i = \mathbf{a}_0' \mathbf{x}_i + u_i \text{ avec } E[u_i] \equiv 0$$

### 2. Vecteur de paramètres à estimer : $\mathbf{a}_0$

$$\mathbf{a}_0 \equiv \begin{bmatrix} a_{1,0} \\ a_{2,0} \\ \vdots \\ a_{K,0} \end{bmatrix}, \text{ terme constant } a_{10} = \alpha_0 \text{ et } \mathbf{a}_0 \equiv \begin{bmatrix} \alpha_0 \\ \mathbf{b}_0 \end{bmatrix} \text{ avec } \mathbf{b}_0 \equiv \begin{bmatrix} a_{2,0} \\ a_{3,0} \\ \vdots \\ a_{K,0} \end{bmatrix} \equiv \begin{bmatrix} b_{1,0} \\ b_{2,0} \\ \vdots \\ b_{K-1,0} \end{bmatrix}$$



## Modèle linéaire

$$y_i = \mathbf{a}'_0 \mathbf{x}_i + u_i \text{ avec } E[u_i] \equiv 0$$

$$\Updownarrow$$

$$y_i = \sum_{k=1}^K a_{k,0} x_{ki} + u_i \text{ avec } E[u_i] \equiv 0$$

$$\Updownarrow$$

$$y_i = \mathbf{x}'_i \mathbf{a}_0 + u_i \text{ avec } E[u_i] \equiv 0$$

$$\Updownarrow$$

$$y_i = \alpha_0 + \tilde{\mathbf{x}}'_i \mathbf{b}_0 + u_i \text{ avec } E[u_i] \equiv 0$$

### 5.3a. Espérances, variances et covariances, cas scalaire

**Définition.** *Espérance de  $x_{ki}$*  :  $m_{k,0} \equiv E[x_{ki}] \equiv \int_{X_k} x f_{xk}(x) dx$

- $X_k$  est le domaine de variation (commun) des  $x_{ki}$
- $f_k(\cdot)$  la fonction de densité de la distribution (commune) des  $x_{ki}$
- L'intégrale définit une somme et  $f_k(\cdot)$  définit une probabilité dans le cas où les  $x_{ki}$  sont discrètes

**Définition.** *Variance de  $x_{ki}$*  :  $v_{kk,0} \equiv V[x_{ki}] \equiv \int_{X_k} (x - m_{k,0})^2 f_k(x) dx$

**Variance et espérances** :  $V[x_{ki}] = E[(x_{ki} - m_{k,0})^2] = E[x_{ki}^2] - E[x_{ki}]^2$

**Définition. Covariance de  $x_{ki}$  et  $x_{\ell i}$  :**

$$v_{k\ell,0} \equiv \text{Cov}[x_{ki}; x_{\ell i}] \equiv \int_{X_{(k,\ell)}} (e_k - m_{k0})(e_\ell - m_{\ell 0}) f_{(k,\ell)}(e_k, e_\ell) d(e_k, e_\ell)$$

- $X_{(k,\ell)}$  est le domaine de variation (commun) des  $(x_{ki}, x_{\ell i})$
- $f_{(k,\ell)}(.,.)$  la fonction de densité de la distribution (commune) des  $(x_{ki}, x_{\ell i})$

$$\textbf{Symétrie : } \text{Cov}[x_{ki}; x_{\ell i}] = \text{Cov}[x_{\ell i}; x_{ki}]$$

**Variance et espérances :**

$$\begin{aligned} \text{Cov}[x_{ki}; x_{\ell i}] &= E[(x_{ki} - m_{k0})(x_{\ell i} - m_{\ell 0})] = E[x_{ki}x_{\ell i}] - E[x_{ki}]E[x_{\ell i}] \\ &= E[(x_{ki} - m_{k0})x_{\ell i}] = E[x_{ki}(x_{\ell i} - m_{\ell 0})] \end{aligned}$$

$$\textbf{Variance et covariance : } \text{Cov}[x_{ki}; x_{ki}] = V[x_{ki}]$$

### 5.3b. Espérances, variances et covariances, cas vectoriel

**Définition. Espérance de  $\mathbf{x}_i$  :**

$$E[\mathbf{x}_i] \equiv \begin{bmatrix} E[x_{1i}] \\ E[x_{2i}] \\ \vdots \\ E[x_{Ki}] \end{bmatrix} \equiv \mathbf{m}_0 \equiv \begin{bmatrix} m_{1,0} \\ m_{2,0} \\ \vdots \\ m_{K,0} \end{bmatrix} \text{ et } m_{1,0} = 1 \text{ si } x_{1i} = 1$$

**Définition. Matrice de variance-covariance de  $\mathbf{x}_i$  :**

$$V[\mathbf{x}_i] \equiv \begin{bmatrix} V[x_{1i}] & Cov[x_{1i}, x_{2i}] & \cdots & Cov[x_{1i}, x_{Ki}] \\ Cov[x_{2i}, x_{1i}] & V[x_{2i}] & \cdots & Cov[x_{2i}, x_{Ki}] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[x_{Ki}, x_{1i}] & Cov[x_{Ki}, x_{2i}] & \cdots & V[x_{Ki}] \end{bmatrix}_{K \times K} \equiv \mathbf{C}_0$$

$$\textit{Transposition} : E[\mathbf{x}_i]' = E[\mathbf{x}_i']$$

$$\textit{Symétrie} : V[\mathbf{x}_i]' = V[\mathbf{x}_i]$$

$V[\mathbf{x}_i]$  est *semi-définie positive*, i.e.  $\mathbf{r}'V[\mathbf{x}_i]\mathbf{r} \geq 0$  pour tout  $\mathbf{r} \in \mathbb{R}^K$

$$\text{Variance et espérances : } V[\mathbf{x}_i] = E[\mathbf{x}_i \mathbf{x}_i'] - E[\mathbf{x}_i] E[\mathbf{x}_i']$$

où  $\mathbf{x}_i \mathbf{x}_i'$  est le *produit croisé* de  $\mathbf{x}_i$

$$\mathbf{x}_i \mathbf{x}_i' = \begin{bmatrix} x_{1i}^2 & x_{1i}x_{2i} & \cdots & x_{1i}x_{Ki} \\ x_{2i}x_{1i} & x_{2i}^2 & \cdots & x_{2i}x_{Ki} \\ \vdots & \vdots & \ddots & \vdots \\ x_{Ki}x_{1i} & x_{Ki}x_{2i} & \cdots & x_{Ki}^2 \end{bmatrix}_{K \times K}$$

et :

$$E[\mathbf{x}_i \mathbf{x}_i'] \equiv \begin{bmatrix} E[x_{1i}^2] & E[x_{1i}x_{2i}] & \cdots & E[x_{1i}x_{Ki}] \\ E[x_{2i}x_{1i}] & E[x_{2i}^2] & \cdots & E[x_{2i}x_{Ki}] \\ \vdots & \vdots & \ddots & \vdots \\ E[x_{Ki}x_{1i}] & E[x_{Ki}x_{2i}] & \cdots & E[x_{Ki}^2] \end{bmatrix}_{K \times K}$$

Avec  $x_{1i} = 1$  on a :

$$E[\mathbf{x}_i \mathbf{x}_i'] \equiv \begin{bmatrix} 1 & E[x_{2i}] & \cdots & E[x_{Ki}] \\ E[x_{2i}] & E[x_{2i}^2] & \cdots & E[x_{2i}x_{Ki}] \\ \vdots & \vdots & \ddots & \vdots \\ E[x_{Ki}] & E[x_{Ki}x_{2i}] & \cdots & E[x_{Ki}^2] \end{bmatrix} = \begin{bmatrix} 1 & E[\tilde{\mathbf{x}}_i'] \\ E[\tilde{\mathbf{x}}_i] & E[\tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i'] \end{bmatrix}$$

et :

$$V[\mathbf{x}_i] \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & V[x_{2i}] & \cdots & Cov[x_{2i}, x_{Ki}] \\ \vdots & \vdots & \ddots & \vdots \\ 0 & Cov[x_{Ki}, x_{2i}] & \cdots & V[x_{Ki}] \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & V[\tilde{\mathbf{x}}_i] \end{bmatrix}$$

**Définition. Matrice de covariance de  $\mathbf{x}_i$  et  $\mathbf{z}_i$ :**

$$Cov[\mathbf{x}_i; \mathbf{z}_i] \equiv \begin{bmatrix} Cov[x_{1i}; z_{1i}] & Cov[x_{1i}; z_{2i}] & \cdots & Cov[x_{1i}; z_{Li}] \\ Cov[x_{2i}; z_{1i}] & Cov[x_{2i}; z_{2i}] & \cdots & Cov[x_{2i}; z_{Li}] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[x_{Ki}; z_{1i}] & Cov[x_{Ki}; z_{2i}] & \cdots & Cov[x_{Ki}; z_{Li}] \end{bmatrix}_{K \times L}$$

**Partition de  $Cov[\mathbf{x}_i; \mathbf{z}_i]$  :**

$$Cov[\mathbf{x}_i; \mathbf{z}_i] = \begin{bmatrix} Cov[\mathbf{x}_i; z_{1i}] & Cov[\mathbf{x}_i; z_{2i}] & \cdots & Cov[\mathbf{x}_i; z_{Li}] \end{bmatrix} = \begin{bmatrix} Cov[x_{1i}; \mathbf{z}_i] \\ Cov[x_{2i}; \mathbf{z}_i] \\ \vdots \\ Cov[x_{Ki}; \mathbf{z}_i] \end{bmatrix}$$



$$\textit{Covariance et espérances : } \text{Cov}[\mathbf{x}_i; \mathbf{z}_i] = E[\mathbf{x}_i \mathbf{z}_i'] - E[\mathbf{x}_i] E[\mathbf{z}_i']$$

*Covariance et transposition :*

$$\text{Cov}[\mathbf{x}_i; \mathbf{z}_i]' = \text{Cov}[\mathbf{z}_i; \mathbf{x}_i] = E[\mathbf{z}_i \mathbf{x}_i'] - E[\mathbf{z}_i] E[\mathbf{x}_i']$$

- Avec  $x_{1i} = 1$  et  $z_{1i} = 1$  on a :

$$Cov[\mathbf{x}_i; \mathbf{z}_i] \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & Cov[x_{1i}; z_{2i}] & \cdots & Cov[x_{2i}; z_{Li}] \\ \vdots & \vdots & \ddots & \vdots \\ 0 & Cov[x_{Ki}; z_{2i}] & \cdots & Cov[x_{Li}; z_{Li}] \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & Cov[\tilde{\mathbf{x}}_i; \tilde{\mathbf{z}}_i] \end{bmatrix}$$

et :

$$E[\mathbf{x}_i \mathbf{z}_i'] = \begin{bmatrix} 1 & E[z_{2i}] & \cdots & E[z_{Li}] \\ E[x_{2i}] & Cov[x_{1i}; z_{2i}] & \cdots & Cov[x_{2i}; z_{Li}] \\ \vdots & \vdots & \ddots & \vdots \\ E[x_{Ki}] & Cov[x_{Ki}; z_{2i}] & \cdots & Cov[x_{Li}; z_{Li}] \end{bmatrix} = \begin{bmatrix} 1 & E[\tilde{\mathbf{z}}_i'] \\ E[\tilde{\mathbf{x}}_i] & E[\tilde{\mathbf{x}}_i \tilde{\mathbf{z}}_i'] \end{bmatrix}$$

## 5.4. Résultats essentiels de la statistique asymptotique

- La plupart des estimateurs présentés dans le cours n'ont pas de propriétés connues à distance finie, *i.e.* pour  $N$  fixe (beaucoup d'entre eux sont même biaisés à distance finie).
- On étudie leurs propriétés asymptotiques, *i.e.* pour  $N \rightarrow +\infty$ , et on approxime les propriétés de ces estimateurs en considérant que «  $N$  est grand mais (tout de même) pas infini ».
- Un estimateur est une fonction (explicite ou non, compliquée ou non) de moyennes, de variances et de covariances empiriques de variables aléatoires.
  - *Lois des Grands Nombres* (LGN)  $\Rightarrow$  Convergence des estimateurs
  - *Théorème Central Limite* (TCL)  $\Rightarrow$  Distribution as. des estimateurs

**Propriété 1a. Loi Forte des Grands Nombres**  
(Convergence presque sûre)

Soit  $\{\mathbf{w}_i; i = 1, 2, \dots\}$  une suite de vecteurs aléatoires tels que les  $\mathbf{w}_i$  sont iid pour  $i = 1, 2, \dots$  avec  $E[\mathbf{w}_i] = \boldsymbol{\mu}_0 < +\infty$  et  $V[\mathbf{w}_i] = \boldsymbol{\Omega}_0 < +\infty$ . On a :

$$N^{-1} \sum_{i=1}^N \mathbf{w}_i \xrightarrow[N \rightarrow +\infty]{p.s.} E[\mathbf{w}_i] = \boldsymbol{\mu}_0.$$

Si, de plus,  $V[\text{vech}(\mathbf{w}_i \mathbf{w}_i')] = \boldsymbol{\Psi}_0 < +\infty$ , alors :

$$N^{-1} \sum_{i=1}^N \mathbf{w}_i \mathbf{w}_i' \xrightarrow[N \rightarrow +\infty]{p.s.} E[\mathbf{w}_i \mathbf{w}_i'] = \boldsymbol{\Omega}_0 + \boldsymbol{\mu}_0 \boldsymbol{\mu}_0'.$$

**Fonction**  $\text{vech}(\mathbf{M})$  :  $\mathbf{M}$  est une matrice symétrique,  $\text{vech}(\mathbf{M})$  renvoie le vecteur des éléments non redondants de  $\mathbf{M}$ .

**Propriété 1b. Loi faible des Grands Nombres**  
(Convergence en probabilité)

Soit  $\{\mathbf{w}_i; i = 1, 2, \dots\}$  une suite de vecteurs aléatoires tels que les  $\mathbf{w}_i$  sont iid pour  $i = 1, 2, \dots$  avec  $E[\mathbf{w}_i] = \boldsymbol{\mu}_0 < +\infty$  et  $V[\mathbf{w}_i] < +\infty$ . On a :

$$N^{-1} \sum_{i=1}^N \mathbf{w}_i \xrightarrow[N \rightarrow +\infty]{p.} E[\mathbf{w}_i] = \boldsymbol{\mu}_0.$$

Si, de plus,  $V[\mathbf{w}_i] = \boldsymbol{\Omega}_0 < +\infty$  et  $V[\text{vech}(\mathbf{w}_i \mathbf{w}_i')] < +\infty$ , alors :

$$N^{-1} \sum_{i=1}^N \mathbf{w}_i \mathbf{w}_i' \xrightarrow[N \rightarrow +\infty]{p.} E[\mathbf{w}_i \mathbf{w}_i'] = \boldsymbol{\Omega}_0 + \boldsymbol{\mu}_0 \boldsymbol{\mu}_0'.$$

**Rmq.** Loi Forte  $\Rightarrow$  loi faible

**Rmq.** Les conditions de régularité, celles qui portent sur la variance des  $\mathbf{w}_i$ , indiquent que les LGN ne s'appliquent qu'à des variables aléatoires à variation « limitée », *i.e.* pas trop « explosives ».

***Propriété 2. Théorème Central Limite***

(Convergence en loi ou en distribution après  $\times\sqrt{N}$ )

Soit  $\{\mathbf{w}_i; i=1,2,\dots\}$  une suite de vecteurs aléatoires tels que les  $\mathbf{w}_i$  sont iid pour  $i=1,2,\dots$  avec  $E[\mathbf{w}_i] = \boldsymbol{\mu}_0 < +\infty$  et  $V[\mathbf{w}_i] = \boldsymbol{\Omega}_0 < +\infty$ . On a :

$$\sqrt{N} \left( N^{-1} \sum_{i=1}^N \mathbf{w}_i - \boldsymbol{\mu}_0 \right) \xrightarrow[N \rightarrow +\infty]{L} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_0).$$

**Rmq.** Si  $\{\mathbf{w}_i; i=1,2,\dots\}$  est une suite de vecteurs aléatoires tels que les  $\mathbf{w}_i$  sont iid pour  $i=1,2,\dots$ , alors pour toute fonction  $\mathbf{g}(\cdot)$  des  $\mathbf{w}_i$  on a :

$\{\mathbf{g}(\mathbf{w}_i); i=1,2,\dots\}$  est une suite de vecteurs aléatoires tels que les  $\mathbf{g}(\mathbf{w}_i)$  sont iid pour  $i=1,2,\dots$

Ce résultat s'applique en particulier pour tout sous-vecteur de  $\mathbf{w}_i$ .

**Rmq.** Si on a  $E[\mathbf{w}_i] = \boldsymbol{\mu}_0 < +\infty$ , on n'a pas toujours  $E[\mathbf{g}(\mathbf{w}_i)] = \boldsymbol{\gamma}_0 < +\infty$ .

## 5.5. Propriétés des estimateurs

Un estimateur de  $\mathbf{a}_0$ ,  $\hat{\mathbf{a}}_N$ , est construit à partir des  $\mathbf{w}_i \equiv (y_i, \mathbf{x}_i, \mathbf{z}_i)$ :

$$\hat{\mathbf{a}}_N = \mathbf{f}(\mathbf{w}_i; i = 1, \dots, N).$$

C'est une variable aléatoire puisque  $\mathbf{w}_i$  contient des termes aléatoires.

**A distance finie** ( $N$  fixe et  $< +\infty$ )

**Définition.** L'estimateur  $\hat{\mathbf{a}}_N$  est sans biais si :

$$E[\hat{\mathbf{a}}_N] = \mathbf{a}_0 = E_{(\mathbf{w}_i; i=1, \dots, N)}[\mathbf{f}(\mathbf{w}_i; i = 1, \dots, N)]$$

La distribution de  $\hat{\mathbf{a}}_N$  est celle de  $\mathbf{f}(\mathbf{w}_i; i = 1, \dots, N)$

Nécessite : 1) des hypothèses sur la distribution de  $(\mathbf{w}_i; i = 1, \dots, N)$

2) une forme simple et explicite de  $\mathbf{f}(\mathbf{w}_i; i = 1, \dots, N)$

## Point de vue asymptotique ( $N \rightarrow +\infty$ )

**Définition.**  $\hat{\mathbf{a}}_N$  est convergent si :

$$\hat{\mathbf{a}}_N \xrightarrow[N \rightarrow +\infty]{p.} \mathbf{a}_0 \text{ ou } p \lim_{N \rightarrow +\infty} \hat{\mathbf{a}}_N = \mathbf{a}_0$$

**Définition.**  $\hat{\mathbf{a}}_N$  est fortement convergent si  $\hat{\mathbf{a}}_N \xrightarrow[N \rightarrow +\infty]{p.s.} \mathbf{a}_0$

**Définition.**  $\hat{\mathbf{a}}_N$  est as. normal (convergent en  $\sqrt{N}$ ) si :

$$\sqrt{N}(\hat{\mathbf{a}}_N - \mathbf{a}_0) \xrightarrow[N \rightarrow +\infty]{L} \mathcal{N}(\mathbf{0}, \Sigma_0)$$

et :

$\mathcal{N}(\mathbf{0}, \Sigma_0)$  est la distribution asymptotique de  $\hat{\mathbf{a}}_N$  et  $\Sigma_0$  est sa matrice de variance-covariance asymptotique.



- Dans le cas de modèles linéaires, on utilise le fait que  $\hat{\mathbf{a}}_N$  est une fonction d'éléments de :

$$N^{-1} \sum_{i=1}^N \mathbf{w}_i, N^{-1} \sum_{i=1}^N \mathbf{w}_i \mathbf{w}_i', \dots$$

- La normalité as. de  $\hat{\mathbf{a}}_N$  ne repose pas sur la normalité des  $\mathbf{w}_i$ , c'est une conséquence du TCL.

### **Interprétation des notions de convergence de $\hat{\mathbf{a}}_N$ vers $\mathbf{a}_0$ :**

- Si  $\hat{\mathbf{a}}_N$  est convergent pour  $\mathbf{a}_0$  alors l'évènement  $\hat{\mathbf{a}}_N = \mathbf{a}_0$  survient avec une probabilité approchant 1 lorsque  $N \rightarrow +\infty$ .
- Si  $\hat{\mathbf{a}}_N$  est fortement convergent pour  $\mathbf{a}_0$  alors l'évènement  $\hat{\mathbf{a}}_N = \mathbf{a}_0$  survient presque sûrement lorsque  $N \rightarrow +\infty$ .

### Interprétation de la normalité as. de $\hat{\mathbf{a}}_N$ :

Avec  $\sqrt{N}(\hat{\mathbf{a}}_N - \mathbf{a}_0) \xrightarrow[N \rightarrow +\infty]{L} \mathcal{N}(\mathbf{0}, \Sigma_0)$ , lorsque «  $N$  est grand mais pas infini » on a :

$$\hat{\mathbf{a}}_N \sim_{app} \mathcal{N}(\mathbf{a}_0, N^{-1}\Sigma_0)$$

**Rmq.**  $\hat{\mathbf{a}}_N$  est as. normal  $\Rightarrow \hat{\mathbf{a}}_N$  est convergent.

**Rmq.** La précision *approchée* de  $\hat{\mathbf{a}}_N$ , mesurée par  $N^{-1}\Sigma_0$ , croît « mécaniquement » en  $N$ .

**Rmq.** L'efficacité as. (la précision lorsque  $N$  est grand) de  $\hat{\mathbf{a}}_N$  est d'autant plus élevée que  $\Sigma_0$  est petite dans le pré-ordre des matrices semi-définies positive

**Shématiquement** : précision = éléments de la diagonale de  $\Sigma_0$  « petits »

**Attention.**  $\Sigma_0$  est la variance as. de l'estimateur  $\hat{\mathbf{a}}_N$ ,  $N^{-1}\Sigma_0$  est sa variance approchée.

**Rmq.** On a besoin d'un estimateur convergent de  $\Sigma_0$  :  $\hat{\Sigma}_N \xrightarrow[N \rightarrow +\infty]{p} \Sigma_0$   
pour calculer des statistiques de test ou des intervalles de confiance.

**Utilisation de la normalité as. de  $\hat{\mathbf{a}}_N$  :**

$$\text{On a : } \Sigma_0 \equiv \begin{bmatrix} \sigma_{1,0}^2 & c_{12,0} & \cdots & c_{1K,0} \\ c_{12,0} & \sigma_{2,0}^2 & \cdots & c_{2K,0} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1K,0} & c_{2K,0} & \cdots & \sigma_{K,0}^2 \end{bmatrix} \text{ et } \hat{\Sigma}_N \equiv \begin{bmatrix} \hat{\sigma}_{1,N}^2 & \hat{c}_{12,N} & \cdots & \hat{c}_{1K,N} \\ \hat{c}_{12,N} & \hat{\sigma}_{2,N}^2 & \cdots & \hat{c}_{2K,N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}_{1K,N} & \hat{c}_{2K,N} & \cdots & \hat{\sigma}_{K,N}^2 \end{bmatrix}$$

**Sortie typique de logiciel :**

<i>Paramètre</i>	<i>Estimation</i>	<i>Estimation de l'écart- -type de l'estimateur</i> $\hat{\mathbf{a}}_N$	<i>Statistiques de test,</i> ...
$a_{1,0}$	$\hat{a}_{1,est}$	$\hat{\sigma}_{1,est} / \sqrt{N}$	...
$a_{2,0}$	$\hat{a}_{2,est}$	$\hat{\sigma}_{2,est} / \sqrt{N}$	...
$\vdots$	$\vdots$	$\vdots$	...
$a_{K,0}$	$\hat{a}_{K,est}$	$\hat{\sigma}_{K,est} / \sqrt{N}$	...

**Intervalles de confiance** de  $\hat{a}_{k,est}$ . Avec  $\hat{a}_{k,N} \sim_{app} \mathcal{N}(a_{k,0}, N^{-1}\sigma_{k,0}^2)$  on a :

$$\left[ \hat{a}_{k,est} - 1,96 \times \hat{\sigma}_{k,est} / \sqrt{N}; \hat{a}_{k,est} + 1,96 \times \hat{\sigma}_{k,est} / \sqrt{N} \right] \text{ à } 5\%$$

et :

$$\left[ \hat{a}_{k,est} - 2,58 \times \hat{\sigma}_{k,est} / \sqrt{N}; \hat{a}_{k,est} + 2,58 \times \hat{\sigma}_{k,est} / \sqrt{N} \right] \text{ à } 1\%$$

**Attention.** Significativité statistique (/0) de  $\hat{a}_{k,est} \neq$  importance dans le modèle de  $a_{k,0}$

- Les  $\hat{\sigma}_{k,est} / \sqrt{N}$  mesurent la capacité des données à fournir des estimations précises des  $a_{k,0}$  dans le modèle considéré.
- Un paramètre important d'un point de vue économique peut être « non différent de 0 statistiquement » parce qu'il est mal mesuré :  $N$  petit, trop peu de variation de variable explicative associée, ...
- Lorsque  $N$  est très très grand, tout est « différent de 0 statistiquement »

## 5.6. La notion d'espérance conditionnelle

- La notion d'espérance conditionnelle est essentielle dans toute la statistique, elle l'est également en économétrie

**Définition.** *Espérance* de  $y_i$  conditionnelle à (sachant)  $\mathbf{x}_i = \mathbf{x}$  où  $\mathbf{x} \in X$  :

$$E[y_i / \mathbf{x}_i = \mathbf{x}] \equiv E_{y_i / \mathbf{x}_i = \mathbf{x}}[y_i] \equiv \int_Y y f(y; \mathbf{x}) dy \equiv \mu(\mathbf{x})$$

- $X$  est le domaine de variation (commun) des  $\mathbf{x}_i$
  - $Y$  est le domaine de variation (commun) des  $y_i$
  - $f(., \mathbf{x})$  la fonction de densité de la distribution (commune) des  $y_i$  conditionnelle à  $\mathbf{x}_i = \mathbf{x}$ .
- Le terme  $\mu(\mathbf{x}) = E[y_i / \mathbf{x}_i = \mathbf{x}]$  est un réel.

- On utilisera également souvent la notion d'espérance conditionnelle :

$$E[y_i/\mathbf{x}_i] \equiv E_{y_i/\mathbf{x}_i}[y_i] \equiv \mu(\mathbf{x}_i),$$

*i.e.* « sans choisir » de valeur pour la réalisation de  $\mathbf{x}_i$ .

Cette espérance conditionnelle est une **variable aléatoire**, puisque c'est une fonction de la variable aléatoire  $\mathbf{x}_i$ .

- Ces notions d'espérance conditionnelle se généralisent directement au cas de vecteurs (et de matrices) aléatoires.
- Les espérances conditionnelles ont plusieurs propriétés importantes.

### ***Propriété 3. Espérance conditionnelle et prédiction***

Soit  $(y, \mathbf{x})$  un vecteur de variables aléatoires sur  $\mathbb{R}^{1+K}$  tel que  $V[y] < +\infty$  et  $\mu(\mathbf{x}) : \mathbb{R}^K \rightarrow \mathbb{R}$  est une fonction telle que  $E[y/\mathbf{x}] = \mu(\mathbf{x})$ , alors :

$$\mu(\mathbf{x}) = \min_{m(\cdot) \in F_K} E[(y - m(\mathbf{x}))^2]$$

où  $F_K$  est l'ensemble des fonctions  $m(\mathbf{x}) : \mathbb{R}^K \rightarrow \mathbb{R}$  telles que  $V[m(\mathbf{x})] < +\infty$ .

**Interprétation.** L'espérance de  $y$  conditionnelle en  $\mathbf{x}$  est une fonction de  $\mathbf{x}$ ,  $\mu(\mathbf{x})$  ici, qui prédit  $y$  au mieux au sens de l'*erreur quadratique moyenne*, i.e. au sens des *Moindres Carrés*.

**Rmq.** La condition  $V[y] < +\infty$  assure l'existence de  $\mu(\mathbf{x})$ , ce qui a été supposé jusqu'à présent dans les définitions des espérances, variances, ...



### ***Propriété 5. Espérance conditionnelle et résidu***

Soit  $(y, \mathbf{x})$  un vecteur de variables aléatoires réelles, on a :

$$E[y/\mathbf{x}] = \mu(\mathbf{x}) \Leftrightarrow y = \mu(\mathbf{x}) + \varepsilon \text{ avec } E[\varepsilon/\mathbf{x}] = 0$$

puisque  $\varepsilon \equiv y - \mu(\mathbf{x}) = y - E[y/\mathbf{x}]$ .

#### **Interprétation et utilisation.**

- Toute variable aléatoire peut être décomposée en la somme de son espérance (conditionnelle ou non) et d'un terme d'erreur d'espérance (conditionnelle ou non) nulle.
- Cette décomposition permet *parfois* d'écrire des modèles. Par exemple :

$$y_i = \mathbf{a}'_0 \mathbf{x}_i + u_i \text{ avec } E[u_i/\mathbf{x}_i] = 0 \Leftrightarrow E[y_i/\mathbf{x}_i] = \mathbf{a}'_0 \mathbf{x}_i$$

***Propriété 6. Caractérisation de  $E[\varepsilon/\mathbf{x}] = 0$***

Soit  $(\varepsilon, \mathbf{x})$  un vecteur de variables aléatoires réelles, on a :

$$E[\varepsilon/\mathbf{x}] = 0$$



$$E[\mathbf{g}(\mathbf{x})\varepsilon] = \mathbf{0} \text{ pour toute fonction } \mathbf{g}(\cdot) \text{ telle que } E[\mathbf{g}(\mathbf{x})\varepsilon] \text{ existe}$$

**Interprétation et utilisation.**

- $E[\varepsilon/\mathbf{x}] = 0$  ssi  $\varepsilon$  n'est corrélée avec aucune fonction de  $\mathbf{x}$ .
- En corollaire on a :

$$E[\varepsilon/\mathbf{x}] = 0 \Rightarrow E[\mathbf{x}\varepsilon] = \mathbf{0}$$

mais la réciproque n'est pas nécessairement vraie.

### ***Propriété 7. Loi des conditionnements successifs***

Soit  $(y, \mathbf{x}, \mathbf{q})$  un vecteur de variables aléatoires réelles tel que  $\mathbf{x} = \mathbf{g}(\mathbf{q})$ , on a alors :

$$E[y/\mathbf{x}] = E[E[y/\mathbf{q}]/\mathbf{x}] = E[E[y/\mathbf{x}]/\mathbf{q}].$$

**Interprétation** (Astuce : chercher à prédire  $y$ ).

- $\mathbf{x} = \mathbf{g}(\mathbf{q})$  implique que l'information contenue dans  $\mathbf{x}$  l'est déjà dans  $\mathbf{q}$ , *i.e.* l'information apportée par  $\mathbf{x}$  pour prédire  $y$  est entièrement contenue dans  $\mathbf{q}$ . La variable  $\mathbf{x}$  n'est qu'une transformation de  $\mathbf{q}$  et  $\mathbf{x} = \mathbf{g}(\cdot)$  n'est pas nécessairement bijective.
- Cette propriété indique que « c'est l'ensemble d'information le plus petit » qui domine après une succession de conditionnements. Ici l'information apportée par  $\mathbf{q}$  non contenue dans  $\mathbf{x}$  est « perdue » pour prédire  $y$ .
- Deux exemples importants d'application de cette propriété :

***Espérance et espérance conditionnelle.***

On a :

$$E[y_i] = E[E[y_i/\mathbf{x}_i]]$$

car (avec  $\mu_y(\mathbf{x}_i) \equiv E[y_i/\mathbf{x}_i]$ )

$$E[y_i] = E_{y_i}[y_i] = E_{(y_i, \mathbf{x}_i)}[y_i] = E_{\mathbf{x}_i}[E_{y_i/\mathbf{x}_i}[y_i]] = E_{\mathbf{x}_i}[\mu_y(\mathbf{x}_i)].$$

***Variance et variance conditionnelle, cas d'un terme d'erreur.***

Si  $E[u_i] = 0$  on a :

$$V[u_i] = E[V[u_i/\mathbf{x}_i]]$$

car :

$$V[u_i] = E_{u_i}[u_i^2] = E_{(u_i, \mathbf{x}_i)}[u_i^2] = E_{\mathbf{x}_i}[E_{u_i/\mathbf{x}_i}[u_i^2]] = E_{\mathbf{x}_i}[V_{u_i/\mathbf{x}_i}[u_i]].$$

**Rmq.** Cette propriété est utile pour exploiter des conditions d'homoscédasticité de termes d'erreurs, ce qu'on fera souvent dans la suite.