

Econométrie des données de Panel

Insa DIOP
Ingénieur Statisticien Economiste

December 26, 2024

- **Introduction**
- **Chapitre 1: Généralités**
- **Chapitre 2: Modèle statique en panel**
- **Chapitre 3: Modèle dynamique en panel**
- **Chapitre 4: Extensions**

Introduction

Overview

- Rappels
- Intérêt
- Background

Rappels

- **Econometrie**
- **Modele:**
- **Modele lineaire:** Une relation lineaire entre une variable de reponse et d'autres variables explicatives.
- **Hypotheses de Gauss Markov:** bonne specification, exogeneite, homoscedasticite, non-autocorrelation, normalite
- **Endogeneite:**
- **Variable instrumentale:**
- **Stationnarite:**
- **Types d'enquete et d'echantillon:**

Interet

- *Explicatif: Transformer des modèles économiques théoriques en outils pratiques pour la prise de décision.*
- *Prédicatif: Fournir des prédictions qui permettent d'anticiper des perturbations de l'économie (baisse de la production, forte Inflation...)*

Outline

- *Background*
- *Definitions*
- *Notations et Illustrations*
- *Avantages et Inconvénients*

Background

- *Modelisation de l'heterogeneite*
- *ANOVA: Une analyse de la variance (ANOVA) permet de vérifier s'il existe des différences statistiquement significatives entre plus de deux échantillons. Pour ce faire, les moyennes et les variances des groupes respectifs sont comparées entre elles*
- *Premieres etudes: Astroonomie, Agronomie, statistiques*
- *Mundlak, Y. (1961). Empirical Production Function Free of Management Bias. Journal of Farm Economics, 43(1), 44.
doi:10.2307/1235460*
- *Nerlove (1966): Balestra, P., and M. Nerlove (1966), "Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," Econometrica, 34: pp. 585–612.*

Definitions

- *Type de données*
 - *Données transversales*
 - *Séries temporelles*
 - *Données empilées*
 - *Données longitudinales ou données de panel*
- *Données de panel: un ensemble qui suit un échantillon donné d'individus dans le temps, et fournit donc des observations multiples (au minimum 2 périodes) sur chaque individu de l'échantillon (Hsiao 2003, page 2).*
 - *Micro Panel: lorsque $T \ll N$ (problème d'hétérogénéité)*
 - *Macro panel: lorsque $T \gg N$ (problème de non-stationnarité)*
 - *Panel cylindre: lorsque toutes les unités statistiques sont observées durant la même période considérée*
 - *Panel non-cylindre: lorsque toutes les unités statistiques ne sont pas observées durant la même période considérée (selection bias)*

Notations

- *Linear Model*

$$Y_{it} = \alpha_{it} + \beta_{it} * X_{it} + \varepsilon_{it}$$

- $\forall i = 1, \dots, N$ (*Nombre d'unites*) $\forall t = 1, \dots, T$ (*Nombre de periodes*)
- α_{it} est un parameter qui peut varier en fonction de i ou t
- $\beta_{it} = (\beta_{1it}, \beta_{2it}, \dots, \beta_{kit})'$ represente un vecteur de K parametres qui peut varier en fonction de i ou t
- $X_{it} = (X_{1it}, X_{2it}, \dots, X_{kit})'$ represente un vecteur de K variables exogenes
- ε_{it} represente le terme d'erreur

Illustrations

- Comment le rendement de l'éducation et l'écart salarial entre les hommes et les femmes ont évolué entre 1978 et 1985

$$\log(wage) = \beta_0 + \delta_0 * y85 + \beta_1 * educ + \delta_1 * y85 * educ + \beta_2 * exper + \beta_3 * exper^2 + \beta_4 * union + \beta_5 * female + \delta_5 * y85 * female$$

$$\begin{aligned} \log(wage) = & .459 + .118 y85 + .0747 educ + .0185 y85 \cdot educ \\ & (.093) (.124) (.0067) (.0094) \end{aligned}$$

$$\begin{aligned} & + .0296 exper - .00040 exper^2 + .202 union \\ & (.0036) (.00008) (.030) \end{aligned}$$

$$\begin{aligned} & - .317 female + .085 y85 \cdot female \\ & (.037) (.051) \end{aligned}$$

$$n = 1,084, R^2 = .426, \bar{R}^2 = .422.$$

Source: Wooldridge (2012)

Illustrations

- National Opinion Research Center's General Social Survey (de 1972 à 1984): Comment on peut expliquer la fécondité des femmes au fil du temps?

TABLE 13.1 Determinants of Women's Fertility

Dependent Variable: *kids*

Independent Variables	Coefficients	Standard Errors
<i>educ</i>	-.128	.018
<i>age</i>	.532	.138
<i>age</i> ²	-.0058	.0016
<i>black</i>	1.076	.174
<i>east</i>	.217	.133
<i>northcen</i>	.363	.121
<i>west</i>	.198	.167
<i>farm</i>	-.053	.147
<i>othrural</i>	-.163	.175
<i>town</i>	.084	.124
<i>smcity</i>	.212	.160
<i>y74</i>	.268	.173
<i>y76</i>	-.097	.179
<i>y78</i>	-.069	.182
<i>y80</i>	-.071	.183
<i>y82</i>	-.522	.172
<i>y84</i>	-.545	.175
<i>constant</i>	-7.742	3.052
<i>n</i> = 1,129		
<i>R</i> ² = .1295		
<i>R</i> ² = .1162		

Source: Wooldridge (2012)

Illustrations

- Comment la présence des étudiants dans une région affecte la rente de loyers? De 1980 à 1990.

$$\log(rent_{it}) = \beta_0 + \delta_0 * y90_t + \beta_1 * \log(pop_{it}) + \beta_2 * \log(avginc_{it}) + \beta_3 * pctstu_{it} + a_i + u_{it}$$

Avantages et Inconvénients

- Avantages

- Taille d'échantillon plus importante: Permet d'améliorer la qualité des estimateurs
- Modélisation de l'hétérogénéité
- Modélisation des relations dynamiques

- Inconvénients

- Biais de sélection
- Assurer une large couverture des données
- Présence de données atypiques

To sum up

- Les données de panel offrent plus d'opportunités d'analyse
 - Jusque là, la régression multiple a permis d'analyser les données transversales et les séries temporelles.
 - Ces données permettent de résoudre d'importantes questions
 - Mais ces données ne permettent pas d'analyser l'évolution des dynamiques temporelles et spatiales
- Les données de panel améliorent fortement l'analyse des politiques publiques et d'autres dynamiques (géographique, économique, physique...) temporelles
 - Elles fournissent plus de robustesse (efficience et convergence des estimateurs) et de diversité dans les analyses
 - L'utilisation de ces données permet de neutraliser plus facilement les caractéristiques individuelles ou temporelles dans l'analyse économétrique
 - Le défi majeur reste la fiabilité des données au cours du temps

Chapitre 2: Modèle statique en données de panel

Chapitre 2: Modèle statique en données de panel

Outline

- Introduction
- Tests de specification: HSIAO (2003)
- Modèles à effets fixes
- Modèles à effets aléatoires
- Conclusion

Introduction

- Modèles à effets fixes: $Y_{it} = \alpha_i + \gamma_t + \beta * X_{it} + \varepsilon_{it}$
- Modèles à effets fixes individuels: $Y_{it} = \alpha_i + \beta * X_{it} + \varepsilon_{it}$
- Modèles à effets fixes temporels: $Y_{it} = \gamma_t + \beta * X_{it} + \varepsilon_{it}$
- Effets fixes ? Tests de spécification: HSIAO (2003)
- Effets fixes; Effets aléatoires; Effets inobservés? Wooldridge (2010)
- Effets fixes or Effets aléatoires? Hausman test

Effets fixes or Effets aleatoires?

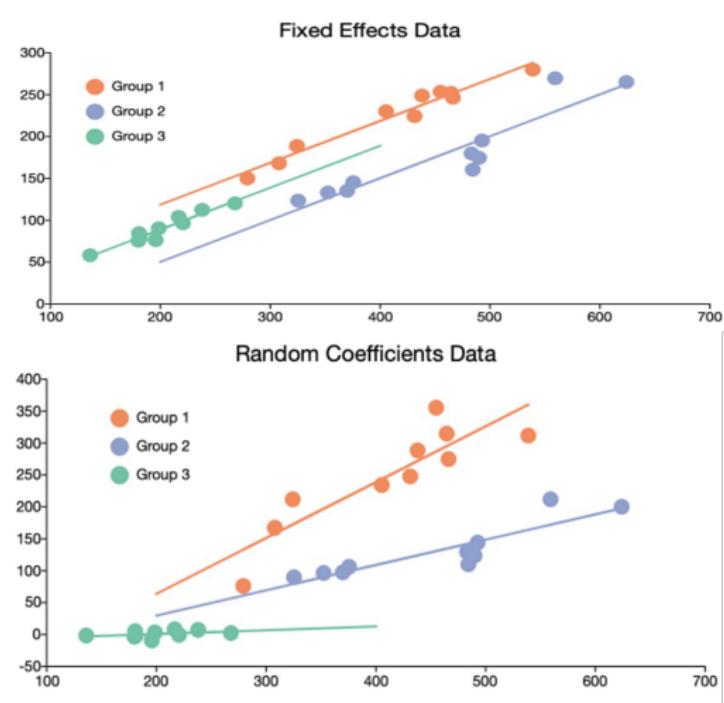


Figure: Effets fixes or Effets aleatoires

Tests de spécification: HSIAO (2003)

- *Model:*

$$Y_{it} = \alpha_i + \beta_i * X_{it} + \epsilon_{it}$$

- *Homogénéité des pentes:*

$$Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$$

- *Homogénéité des effets fixes individuels:*

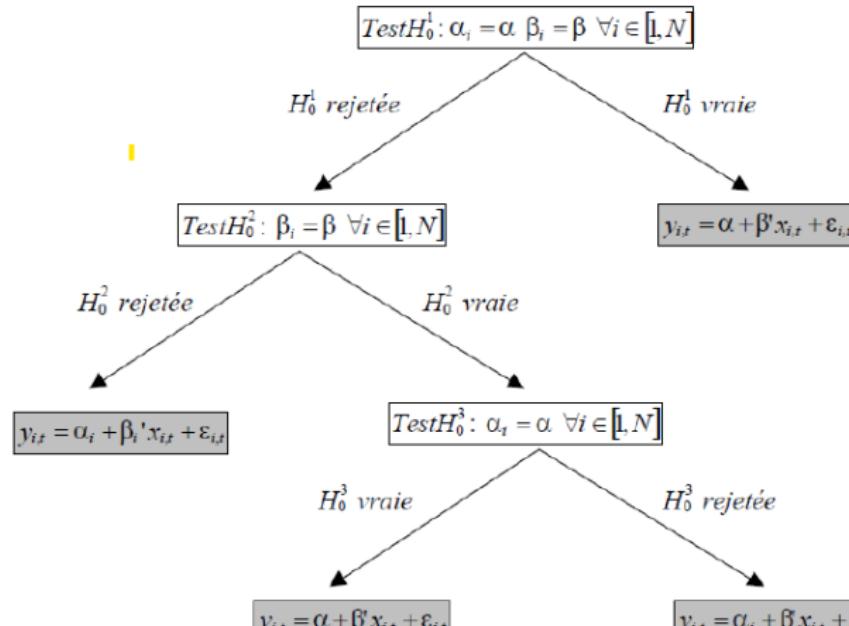
$$Y_{it} = \alpha + \beta_i * X_{it} + \epsilon_{it}$$

- *Homogénéité des pentes et effets fixes individuels (Pooled Panel):*

$$Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$$

Tests de spécification: HSIAO (2003)

- Comment choisir la meilleure specification
 - Justification économique ou empirique
 - Test économétrique: HSIAO (2003)



Tests de spécification: HSIAO (2003)

- Première étape: Homogénéité des pentes Vs Homogénéité des pentes et effets fixes individuels (Pooled Panel)

Mnc: $Y_{it} = \alpha_i + \beta_i * X_{it} + \epsilon_{it}$ Vs **Mc:** $Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$

$$H_0^1 : \beta_i = \beta \quad \alpha_i = \alpha \quad \forall i \in \{1, \dots, n\}$$

$$H_1^1 : \exists (i, j) \in \{1, \dots, n\}^2 / \beta_i \neq \beta \text{ or } \alpha_i \neq \alpha$$

- $F_1 = \frac{(SCR_{1,c} - SCR_1) / (n - 1)(K + 1)}{SCR_1 / (nT - n(K + 1))}$
- SCR_1 représente la somme des carrés des résidus du modèle non-contraint (**Mnc**)
- $SCR_{1,c}$ représente la somme des carrés des résidus du modèle contraint (**Mc**) $Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$
- Veuillez calculer SCR_1 et $SCR_{1,c}$!

Tests de spécification: HSIAO (2003)

- Deuxième étape: Effets hétérogènes Vs effets inobservés

Mnc: $Y_{it} = \alpha_i + \beta_i * X_{it} + \epsilon_{it}$ Vs **Mc:** $Y_{it} = \alpha_i + \beta * X_{it} + \epsilon_{it}$

$$H_0^2 : \beta_i = \beta \quad \forall i \in \{1, \dots, n\}$$

$$H_1^2 : \exists (i, j) \in \{1, \dots, n\}^2 / \beta_i \neq \beta_j$$

- $F_2 = \frac{(SCR_{1,c'} - SCR_1) / (n - 1)K}{SCR_1 / (nT - n(K + 1))}$

- SCR_1 représente la somme des carrés des résidus du modèle non-contraint (**Mnc**)

- $SCR_{1,c'}$ représente la somme des carrés des résidus du modèle contraint (**Mc**) $Y_{it} = \alpha_i + \beta * X_{it} + \epsilon_{it}$

- Veuillez calculer SCR_1 et $SCR_{1,c'}$!

- Troisième étape: Homogénéité des effets fixes individuels/s.c
homogénéité des pentes ($\beta_i = \beta$)

Mnc: $Y_{it} = \alpha_i + \beta * X_{it} + \epsilon_{it}$ Vs **Mc:** $Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$

$$H_0^3 : \alpha_i = \alpha \quad \forall i \in \{1, \dots, n\}$$

$$H_1^3 : \exists (i, j) \in \{1, \dots, n\}^2 \quad \alpha_i \neq \alpha_j$$

$$F_3 = \frac{(SCR_{1,c''} - SCR_1) / (n - 1)}{SCR_1 / (n(T - 1) - K)}$$

SCR_1 représente la somme des carres des résidus du modèle non-contraint (**Mnc**)

$SCR_{1,c''}$ représente la somme des carres des résidus du modèle contraint (**Mc**) $Y_{it} = \alpha + \beta * X_{it} + \epsilon_{it}$

Veuillez calculer SCR_1 et $SCR_{1,c''}$!

Tests de spécification: HSIAO (2003)

APPLICATION

Hypotheses

- ① Modèles à effets fixes individuels: $Y_{it} = \alpha_i + \beta * X_{it} + \varepsilon_{it}$
- ② ε_{it} est un processus "sans mémoire"
 - $\mathbb{E}(\varepsilon_{it}) = 0$: l'espérance des résidus du modèle est nulle
 - $\mathbb{E}(\varepsilon_i \varepsilon_j) = \sigma_\varepsilon^2 * I_T$: où I_T désigne la matrice identité ($T; T$)
- ③ Exogénéité des variables (X) explicatives
 - $\mathbb{E}(\varepsilon_{it}/X_j) = 0$
 - $\mathbb{E}(\varepsilon_{it}\varepsilon_{up}/X_j) = \delta_{iu}\delta_{tq}\sigma_\varepsilon^2$
- ④ $\mathbb{E}(\varepsilon_{it}\varepsilon_{jt}) = 0 \forall i \neq j$: Aucune corrélation pour deux individus distincts et cela quelle que soit la date considérée.

Estimateurs: LSDV or Within

- Par MCO simple: Modèle à Variables Muettes (MVM) « Least Squares Dummy Variables » LSDV

- $Y_i = \sum_{j=1}^N \alpha_j * d_j + \beta * X_i + \varepsilon_i$
- $\tilde{Y}_{it} = \beta * \tilde{X}_{it} + \tilde{\varepsilon}_{it}$ Avec
$$\begin{cases} \tilde{Y}_{it} &= Y_{it} - \bar{Y}_{it} \\ \tilde{X}_{it} &= X_{it} - \bar{X}_{it} \\ \tilde{\varepsilon}_{it} &= \varepsilon_{it} - \bar{\varepsilon}_{it} \end{cases}$$
- $\hat{\beta}_{LSDV} = (\tilde{X}' \tilde{X})^{-1} \tilde{X}' \tilde{Y}$
- $\hat{\alpha}_i = \bar{Y}_i - \hat{\beta}_{LSDV} * \bar{X}_i$

- Estimateur within: Application du théorème de Frish-Waugh (1993)

- $Y_{it} - \bar{Y}_{it} = \beta * (X_{it} - \bar{X}_i) + \varepsilon_{it} - \bar{\varepsilon}_i$
- $\mathbb{W}Y = \mathbb{W}X\beta + \mathbb{W}\varepsilon$
- Avec $\mathbb{W} = I_{NT} - (I_N \otimes J_T/T) \otimes$ est le produit de Kronecker
- $\hat{\beta}_{Within} = (X' \mathbb{W}X)^{-1} (X' \mathbb{W}Y)$
- $Var(\hat{\beta}_{Within}) = \sigma_\varepsilon^2 (X' \mathbb{W}X)^{-1}$
- L'estimateur de cette variance σ_ε^2 est donné par: $\hat{\sigma}_\varepsilon^2 = \frac{\hat{\varepsilon}_w' \hat{\varepsilon}_w}{NT - N - K}$

Background

- Problème du modèle à effets fixes: Le coefficient α_i est vraiment propre à l'individu i et uniquement à cet individu.
- Conséquences: Impossibilité d'extrapoler les résultats
- Quel intérêt y a-t-il de se focaliser sur la constante spécifique? Pour les pays, oui. Et les individus?
- Modèle à effets aléatoires: une autre façon de modéliser l'hétérogénéité des individus à l'intérieur d'un échantillon en données de panel: Le terme d'erreur est considéré comme la somme de deux composantes d'où la terminologie utilisée de Modèle à Composantes d'Erreur (MEC)
$$Y_{it} = \beta * X_{it} + \alpha_i + \varepsilon_{it}$$

Hypotheses

① Modèles à effets aléatoires:

② $Y_{it} = \beta * X_{it} + \alpha_i + \varepsilon_{it}$ with $\eta_{it} = \alpha_i + \varepsilon_{it}$

③ ε_{it} est un processus "sans mémoire"

- $\mathbb{E}(\varepsilon_{it}) = 0$: l'espérance des résidus du modèle est nulle
- $\mathbb{E}(\varepsilon_i \varepsilon_i') = \sigma_\varepsilon^2 * I_T$: où I_T désigne la matrice identité ($T; T$)

④ Exogénéité des variables (X) explicatives

- $\mathbb{E}(\alpha_i / X) = \mathbb{E}(\varepsilon_{it} / X_j) = 0$

⑤ $\mathbb{E}(\varepsilon_{it} \varepsilon_{up} / X_j) = \delta_{iu} \delta_{tq} \sigma_\varepsilon^2$: Les erreurs spécifiques ne sont ni autocorrelées et ni corrélées d'un individu à un autre

⑥ $\mathbb{E}(\varepsilon_{it} \varepsilon_{jt}) = 0 \quad \forall i \neq j$: Aucune corrélation pour deux individus distincts et cela quelle que soit la date considérée.

Hypotheses

- ① $\mathbb{E}(\alpha_i^2/X) = \sigma_\alpha^2$: l'écart type des erreurs possibles α_i est le même d'un individu à un autre
- ② $\mathbb{E}(\alpha_j * \varepsilon_{it}/X) = 0 \forall i \text{ and } j$: Les deux erreurs ne sont pas corrélées.
- ③ $\mathbb{E}(\alpha_i * \alpha_j/X) = 0 \forall i \text{ and } j$: Les erreurs spécifiques ne sont pas corrélées d'un individu à un autre.
- ④ $\mathbb{E}(\eta_{it}) = \mathbb{E}(\alpha_i + \varepsilon_{it}) = 0$:
- ⑤ $\mathbb{E}(\eta_{it}^2) = \mathbb{E}((\alpha_i + \varepsilon_{it})^2) = \sigma_\varepsilon^2 + \sigma_\alpha^2$: homoscédasticité
- ⑥ $\mathbb{E}(\eta_i, \eta_j) = \mathbb{E}((\alpha_i + \varepsilon_{it})(\alpha_j + \varepsilon_{jt})) = \sigma_\varepsilon^2$: Autocorrelation indépendante du temps
- ⑦ $\forall t \neq s \mathbb{E}(\eta_{it}, \eta_{is}) = \mathbb{E}((\alpha_i + \varepsilon_{it})(\alpha_j + \varepsilon_{is})) = \sigma_\alpha^2$: Autocorrelation indépendante du temps
- ⑧ $\forall i \neq j \mathbb{E}(\eta_{it}, \eta_{jt}) = \mathbb{E}((\alpha_i + \varepsilon_{it})(\alpha_j + \varepsilon_{jt})) = 0$: Corrélation temporelle nulle
- ⑨ $\mathbb{E}(X_{it} * \eta_{it}) = 0$ (Hypothèse d'orthogonalité)

Estimateur: L'Estimateur des Moindres Carrés Généralisés (EMCG)

- Matrice des variances covariances de l'erreur du MCE

$$\Omega = \mathbb{E}(\eta * \eta') = \begin{bmatrix} \Sigma & . & \dots & 0 \\ 0 & \Sigma & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & \Sigma \end{bmatrix}$$

$$\Sigma_{T*T} = \begin{bmatrix} \eta_{11}^2 & \dots & \eta_{11}\eta_{1T} \\ \eta_{11}\eta_{12} & \dots & . \\ . & \dots & . \\ \eta_{11}\eta_{1T} & \dots & \eta_{1T}^2 \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_\alpha^2 & \sigma_\alpha^2 & \dots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\varepsilon^2 + \sigma_\alpha^2 & \dots & \sigma_\alpha^2 \\ . & . & \dots & . \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \dots & \sigma_\varepsilon^2 + \sigma_\alpha^2 \end{bmatrix}$$

$$\Sigma_{T*T} = \sigma_\varepsilon^2 * I_T + \sigma_\alpha^2 * i'_T * i_T$$

$$\Sigma_{T*T} = \sigma_\varepsilon^2 * I_T + \sigma_\alpha^2 * J_T$$

$$\Omega = \sigma_\varepsilon^2 * [I_N \otimes I_T] + \sigma_\alpha^2 * [I_N \otimes J_T]$$

Estimateur: L'Estimateur des Moindres Carrés Généralisés (EMCG)

$$\Omega = \sigma_\varepsilon^2 * [I_N \otimes I_T + \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} * [I_N \otimes J_T]]$$

$$\Omega = \sigma_\varepsilon^2 * [I_N \otimes (I_T - \frac{J_T}{T}) + \frac{\sigma_\alpha^2 + T * \sigma_\varepsilon^2}{\sigma_\varepsilon^2} * I_N \otimes \frac{J_T}{T}]$$

$$\Omega = \sigma_\varepsilon^2 * [W + \frac{\sigma_\alpha^2 + T * \sigma_\varepsilon^2}{\sigma_\varepsilon^2} * B]$$

$$\Omega^{-1} = \frac{1}{\sigma_\varepsilon^2} (W + \frac{\sigma_\varepsilon^2}{\sigma_\alpha^2 + T * \sigma_\varepsilon^2} * B) = \frac{1}{\sigma_\varepsilon^2} (W + \psi * B)$$

$$\Omega = (H' * H)^{-1} \quad H' * H = \Omega^{-1}$$

Modèle transformé $HY = HX\beta + H(\varepsilon + \alpha)$

L'estimateur MCG est donné par:

$$\hat{\beta}_{MCG} = (X' H' H X) X' H' H Y = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

$$\hat{\beta}_{MCG} = (X' H' H X) X' H' H Y = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

Estimateur: L'Estimateur des Moindres Carrés Généralisés (EMCG)

$$\hat{\beta}_{MCG} = (X' H' H X) X' H' H Y = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

$$\hat{\beta}_{MCG} = (X' (W + \psi * B) X) X' (W + \psi * B) Y$$

$$\hat{\beta}_{MCG} = (X' (W + \psi * B) X)^{-1} X' W Y + \psi * (X' (W + \psi * B)^{-1} X) X' B Y$$

Posons: $F_{Within} = (X' (W + \psi * B) X)^{-1} X' W X$

et $F_{Between} = \psi * (X' (W + \psi * B)^{-1} X) X' B X$

$$F_{Within} + F_{Between} = 1$$

$$F_{Within} * \hat{\beta}_{Within} = (X' (W + \psi * B) X)^{-1} X' W X (X' W X)^{-1} (X' W Y) = \\ (X' (W + \psi * B) X)^{-1} X' W Y$$

$$F_{Between} * \hat{\beta}_{Between} = \psi (X' (W + \psi * B) X)^{-1} X' B X (X' B X)^{-1} (X' B Y)$$

$$F_{Between} * \hat{\beta}_{Between} = \psi (X' (W + \psi * B) X)^{-1} X' B Y$$

Estimateur: L'Estimateur des Moindres Carrés Généralisés (EMCG)

$$F_{Between} * \hat{\beta}_{Between} = \psi(X'(W + \psi * B)X)^{-1} X' BX (X' B X)^{-1} (X' B Y)$$

$$F_{Between} * \hat{\beta}_{Between} = \psi(X'(W + \psi * B)X)^{-1} X' B Y$$

$$F_{Between} * \hat{\beta}_{Between} = (1 - F_{Within}) * \hat{\beta}_{Between}$$

$$\hat{\beta}_{MCG} = F_{Within} * \hat{\beta}_{Within} + (1 - F_{Within}) * \hat{\beta}_{Between}$$

- Si $\psi = 0$ (lorsque le nombre T des observations dans le temps tend vers l'infini), alors $F_{between}=0$ et l'estimateur MCG converge donc vers l'estimateur Within.
- Si $\psi = 1$ (ce qui se produit lorsque la variance des effets spécifiques tend vers 0, et donc l'estimateur des MCG converge vers l'estimateur MCO du modèle en pooling (aucune spécificité individuelle).
- $V(\hat{\beta}_{MCG}) = [X'\Omega^{-1}X]^{-1} = \sigma_\varepsilon^2 * [X'(W + \psi * B)X]^{-1}$

Estimateur: L'Estimateur des Moindres Carrés Généralisés (EMCG)

- L'estimateur $\hat{\beta}_{MCG}$ requiert nécessairement que soit connue la matrice des variances covariances soit connue (i.e σ_ε et σ_α soient connus).
- Ces paramètres sont inconnus : il faut les estimer.
- Plusieurs procédures existent pour les estimer.
- Hypothèse à tester avec le test de Hausman: la composante aléatoire spécifique α_i n'est pas corrélée avec les variables explicatives du modèle.

Propriétés des estimateurs

$\mathbb{E}(\alpha_i x_{i1}, \dots, x_{iK}) = 0$	$\mathbb{E}(\alpha_i x_{i1}, \dots, x_{iK}) \neq 0$		
LSDV	GLS	LSDV	GLS
T fixed, $n \rightarrow \infty$	Consistent	—	Consistent Not Consistent
$T \rightarrow \infty$ and $n \rightarrow \infty$	Consistent	BLUE	Consistent Consistent

Figure: Propriétés des estimateurs

Test de Hausman (1978):

- $H_0: \mathbb{E}(\alpha_i/X) = 0$ Vs $H_1: \mathbb{E}(\alpha_i/X) \neq 0$
- $Y_i = \sum_{j=1}^N \alpha_j * d_j + \beta * X_i + \varepsilon_i$
- Sous H_0 : $\hat{\beta}_{MCG}$ et $\hat{\beta}_{LSDV}$ sont consistents
- Sous H_0 : $\hat{\beta}_{MCG}$ est BLUE et atteint la borne de Cramer-Rao
- Sous H_1 : $\hat{\beta}_{MCG}$ n'est pas consistent
- $\mathbb{H} = (\hat{\beta}_{LSDV} - \hat{\beta}_{MCG})[V(\hat{\beta}_{LSDV}) - V(\hat{\beta}_{MCG})]^{-1}(\hat{\beta}_{LSDV} - \hat{\beta}_{MCG})$
- Sous H_0 $\mathbb{E}(\alpha_i/X) = 0$, on a $\mathbb{H} \rightarrow \chi^2(K)$

Test d'existence d'effets individuels

- Test d'existence d'effets individuels

$$\textbf{Modèle 1: } Y_{it} = \alpha_0 + \beta * X_{it} + \epsilon_{it}$$

$$\textbf{Modèle 2: } Y_{it} = \alpha_i + \beta * X_{it} + \epsilon_{it}$$

$$H_0 : \alpha_i = \alpha_0 \quad \forall i \in \{1, \dots, n\}$$

VS

$$H_1 : \exists i \in \{1, \dots, n\} \quad \alpha_i \neq \alpha_0$$

- $F = \frac{(SCR_1 - SCR_2) / (n - 1)}{SCR_1 / (n(T - 1) - K)}$
- SCR_1 représente la somme des carrés des résidus du modèle 1
- SCR_2 représente la somme des carrés des résidus du modèle 2

Tests d'homoscedasticité

- $H_0 : \sigma_i^2 = \sigma^2$
- En présence d'heteroscedasticité: l'estimateur ne peut plus être de variance minimale
- **Test de Breush-Pagan**
 - Step 1: Calculer les résidus estimés au carré.
 - Step 2: Régresser les résidus estimés au carré en fonction des variables explicatives du modèle.
 - Step 3: Calculer la statistique de test de Breush-Pagan ($BP = nT * R^2$) qui suit une loi de $\chi^2(K - 1)$ avec K le nombre de coefficients à estimer.
- **Test de White:** Une généralisation du test de Breush-Pagan.
 - La régression du step 2 inclut les carrés et les produits des variables explicatives.
 - Et la statistique de test devient: $BP = nT * R^2 \xrightarrow{} \chi^2(K - 1)$

Test d'absence d'autocorrelation

- $Y_{it} = \alpha_i + \beta * X_{it} + \varepsilon_{it}$
- $\hat{\varepsilon}_{it} \rightarrow AR(1) : \hat{\varepsilon}_{it} = \rho * \hat{\varepsilon}_{it-1} + \eta_{it}$
- $H_0 : \rho = 0$ vs $H_1 : (\rho) < 0$
- **Test de Durbin-Watson**
 - $DW = \frac{\sum_{i=1}^N \sum_{t=1}^T (\hat{\varepsilon}_{it} - \hat{\varepsilon}_{it-1})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{\varepsilon}_{it}^2}$
- **Barghava, Franzini et Narendranathan (1982) (BFN)**
 - $BFN = \frac{\sum_{i=1}^N \sum_{t=1}^T (\hat{\varepsilon}_{it} - \hat{\varepsilon}_{it-1})^2}{\sum_{i=1}^N \sum_{t=1}^T (\hat{\varepsilon}_{it} - \bar{\varepsilon}_i)^2}$
- **Baltagi and Wu**
 - $BW = \frac{NT^2}{T-1} * [\frac{\sum_{i=1}^N \sum_{t=1}^T \hat{\varepsilon}_{it} \hat{\varepsilon}_{it-1}}{\sum_{i=1}^N \sum_{t=1}^T \hat{\varepsilon}_{it}^2}]^2 \rightarrow \chi^2(1)$

Modèles dynamiques en Panel

Overview

- ① Introduction
- ② Biais de nickel
- ③ Estimateurs en panel dynamique
 - Approche par les variables instrumentales
 - Approche GMM

Introduction

- Modèle dynamique: Les valeurs passées des variables influencent leurs valeurs courantes.
- $Y_{it} = \gamma * Y_{it-1} + \beta * X_{it} + \alpha_i + \epsilon_{it}$
- Que se passe-t-il lorsque nous procedons comme le cas statique?
 - La variable dépendante décalée apparaît comme variable explicative, l'exogénéité stricte des régresseurs n'est plus valable
 - Les propriétés des estimateurs MLE et MCG dépendent de T et n.
 - Les valeurs initiales d'un processus dynamique posent un autre problème. Il s'avère qu'avec une formulation à effets aléatoires, l'interprétation d'un modèle dépend de l'hypothèse d'observation initiale.

Introduction

Avantages des modèles dynamiques en panel

- ① **Prise en compte des effets d'inertie** : Utile pour des phénomènes comme la croissance économique, où les conditions présentes dépendent fortement des performances passées.
- ② **Prise en compte de l'endogénéité**: grâce aux techniques d'instrumentation.
- ③ **Meilleure prise en compte de la théorie**: Les panels fournissent une meilleure identification et un pouvoir explicatif plus élevé.

Limites

- ① **Complexité des estimations** : Les modèles comme le GMM nécessitent des hypothèses fortes (orthogonalité des instruments).
- ② **Instruments faibles et sensibilité des résultats aux instruments utilisés** : Trop d'instruments peuvent affaiblir la validité statistique.

Introduction

Applications courantes

- ① Analyse de la croissance économique**
- ② Études sur l'investissement et l'emploi**
- ③ Évaluation des politiques publiques**
- ④ Modélisation des comportements financiers ou des dynamiques commerciales**

Biais de Nickel

- L'estimateur LSDV est consistent dans le cas d'un modèle à effets fixes et des modèles à effets aléatoires.
- L'estimateur LSDV est n'est pas consistent en modèle dynamique (CQFD ici).
- Le biais de l'estimateur LSDV en panel dynamique est connu sous le nom de biais de Nickel (1981).
 - Avec $Y_{it} = \gamma * Y_{it-1} + \alpha_i + \epsilon_{it}$: On a
 - $\text{plim}_{n \rightarrow +\infty} (\hat{\gamma}_{LSDV} - \gamma) = -\frac{1+\gamma}{T-1} * (1 - \frac{1}{T} \frac{1-\gamma^T}{1-\gamma}) * (1 - \frac{2\gamma}{(1-\gamma)(T-1)} (\frac{1-\gamma^T}{1-\gamma}))^{-1}$
 - $\text{cov}(\bar{y}_{i..}, \bar{\varepsilon}_i) \approx \frac{T-1}{T^2} \sigma_\varepsilon^2$

Biais de Nickel

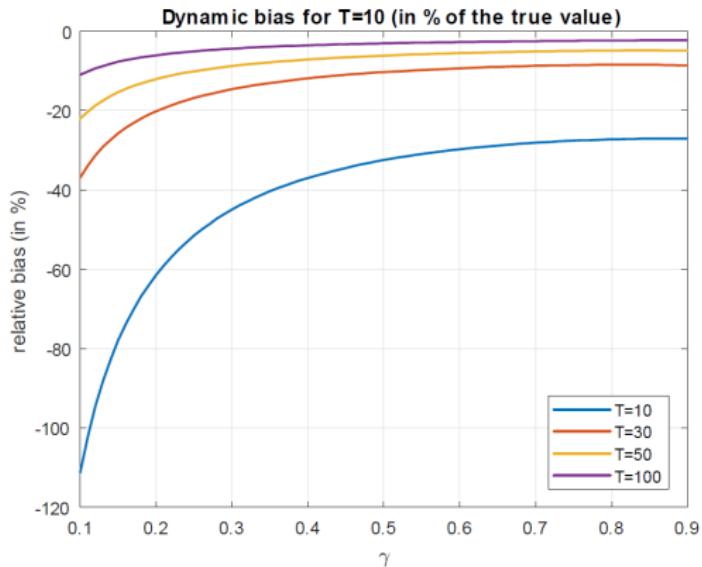


Figure: Biais de Nickel

Exercice: Reproduire ce graphique avec
 $\gamma = 0.5, \sigma_\varepsilon = 1, \varepsilon_{it} \rightarrow \mathbb{N}(0, 1), \alpha_i^* \rightarrow \mathbb{U}[-1, 1]$

Rappel

- ① $Y = X\beta + \varepsilon$
- ② Endogénéité: $\mathbb{E}(\varepsilon|X) \neq 0_{N \times 1}$
- ③ $p\lim \frac{1}{N} X' \varepsilon = \mathbb{E}(X_j \varepsilon_j) = \gamma \neq 0_{N \times 1}$
- ④ $\mathbb{E}(\hat{\beta}_{OLS}) \neq \beta$
- ⑤ $\hat{\beta}_{OLS} = \beta + Q^{-1}\gamma$ avec $Q = p\lim \frac{1}{N} X' X \neq \beta$

Instruments

- ① Soit $Z = (Z_1, Z_2, \dots, Z_H)$, H variables instrumentales ou instruments.
Ces instruments sont bons lorsqu'ils respectent ces deux propriétés
- ② Exogénéité: $\mathbb{E}(\varepsilon|Z) = 0_{N \times 1}$
- ③ Pertinence: Ils sont corrélés avec les variables indépendantes.
 $\mathbb{E}(X_{jk} * Z_{jh}) \neq 0$

Condition d'orthogonalité: $\mathbb{E}(Z_j(Y_j - X'_j\beta)) = 0_H$

- ① Si $H = K$, l'estimateur par variables instrumentales est donné par:

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$$
- ② Si $H > K$ (Sur-Identification), alors nous ne pouvons pas estimer par $\hat{\beta}_{IV}$. Alternatives: 2SLS or GMM
- ③ Si $H < K$ (Sous-Identification)

Instruments: 2SLS

- ① Projeter X sur Z : $\hat{X} = Z(Z'Z)^{-1}Z'X$
- ② Projeter Y sur \hat{X}
- ③ L'estimateur 2SLS est donné par $\hat{\beta}_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'Y$ vspace8
- ④ $\hat{\beta}_{2SLS} = (XZ(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y$

Approche IV de Anderson et HSIAO (1982)

- ① $Y_{it} = \gamma * Y_{it-1} + \beta' X_{it} + \rho'_i \omega_i + \alpha_i + \epsilon_{it}$
- ② γ le paramètre autoregressif
- ③ β un vecteur de K_1 paramètres de variables explicatives dépendant du temps
- ④ ρ un vecteur de K_2 paramètres de variables explicatives indépendant du temps
- ⑤ σ_ε^2 et σ_α^2 sont les variances des termes d'erreur.

Approche IV de Anderson et HSIAO (1982): Procédure

- ① Etape 1: Transformation en différence première
- ② Etape 2: Choix des instruments
- ③ Etape 3: Estimation de ρ , γ et β
- ④ Etape 4: Estimation des variances σ_ϵ^2 et σ_α^2

Approche IV de Anderson et HSIAO (1982): Procédure

① Etape 1: Transformation en différence première

$$\textcircled{1} \quad Y_{it} - Y_{it-1} = \gamma * (Y_{it-1} - Y_{it-2}) + \beta'(X_{it} - X_{it-1}) + \varepsilon_{it} - \varepsilon_{it-1}$$

② Etape 2: Choix des instruments

① Conditions de Validité des instruments

① Exogénéité: $\mathbb{E}(Z_{jt}(\varepsilon_{it} - \varepsilon_{it-1})) = 0$

② Pertinence $\mathbb{E}(Z_{jt}(Y_{it-1} - Y_{it-2})) \neq 0$

② Instruments proposés par Anderson and HSIAO (1982)

① $Z_{it} = Y_{it-2}$

② $Z_{it} = Y_{it-2} - Y_{it-3}$

Approche IV de Anderson et HSIAO (1982): Procédure

① Etape 2: Choix des instruments

- ① Premier instrument: $Z_i = (Y_{it-2}, (X_{it} - X_{it-1})')'$
- ② Deuxième instrument: $Z_i = ((Y_{it-2} - Y_{it-3}), (X_{it} - X_{it-1})')'$

② Etape 3: Estimation de ρ , γ et β

- ①
$$\begin{pmatrix} \hat{\gamma}_{IV} \\ \hat{\beta}_{IV} \end{pmatrix} = (Z'X)^{-1}Z'Y$$
- ② $\hat{\rho} = (\sum_{i=1}^n \omega_i \omega_i')^{-1} (\sum_{i=1}^n \omega_i h_i)$ Avec $h_i = \bar{Y}_i - \hat{\gamma}_{IV} \bar{Y}_{i,-} - \hat{\beta}_{IV} \bar{X}_i$
- ③ Ces estimateurs sont consistants lorsque n or T tend vers infini.

Approche IV de Anderson et HSIAO (1982): Procédure

① Etape 4: Estimation des variances σ_ε^2 et σ_α^2

$$\textcircled{1} \quad \hat{\sigma}_\varepsilon^2 = \frac{1}{n(T-1)} \sum_{t=1}^T \sum_{i=1}^n \hat{\varepsilon}_{it}^2$$

$$\textcircled{2} \quad \hat{\sigma}_\alpha^2 = \frac{1}{n} \sum_{i=1}^n \left(\bar{Y}_i - \hat{\gamma}_{IV} \bar{Y}_{i,-1} - \hat{\beta}'_{IV} \bar{X}_i - \rho' \omega_i \right) - \frac{1}{T} \hat{\sigma}_\varepsilon^2$$

② Propriétés des estimateurs des variances σ_ε^2 et σ_α^2

- ① L'estimateur $\hat{\sigma}_\varepsilon^2$ est consistant lorsque n or T tend vers infini.
- ② L'estimateur $\hat{\sigma}_\alpha^2$ est consistant lorsque n tend vers infini.

Approche GMM

Overview

- ① L'approche GMM repose sur les conditions d'orthogonalité entre les variables retardées et le terme d'erreur
- ② Soit en différences premières ou 'en niveau.'
- ③ Si le modèle dynamique est présenté en différences premières, les instruments sont en niveau (Y_{it-2-j}),
- ④ Et lorsqu'il est présenté en niveau, les instruments sont en différences premières (ΔY_{it-2-j})

Avantages et limites

- ① Réduit les problèmes de faiblesse des instruments observé
- ② Plus adapté lorsque les variables explicatives sont persistantes dans le temps.
- ③ Plus complexe et très sensible à la prolifération des instruments

Approche GMM en différences premières d'Arellano et Bond (1991)

① Etape 1: Transformation en différence première

$$\textcircled{1} \quad Y_{it} - Y_{it-1} = \gamma * (Y_{it-1} - Y_{it-2}) + \beta'(X_{it} - X_{it-1}) + \varepsilon_{it} - \varepsilon_{it-1}$$

② Etape 2: Choix des instruments

① Conditions de Validité des instruments pour $j \geq 0$

- ① Exogeneité: $\mathbb{E}(Y_{it-2-j}(\varepsilon_{it} - \varepsilon_{it-1})) = 0$
- ② Pertinence $\mathbb{E}(Y_{it-2-j}(Y_{it-1} - Y_{it-2})) \neq 0$

② GMM idea

- ① $\mathbb{E}(Y_{it-2-j}(\varepsilon_{it} - \varepsilon_{it-1})) = 0$ pour $j=1\dots m$
- ② Cette relation ci-dessus peut être utilisée comme une condition des moments pour estimer β , γ , σ_ε^2 et σ_α^2
- ③ $\mathbb{E}(Y_{it-2-j}(\Delta Y_{it} - \gamma_0 \Delta Y_{it-1} - \beta'_0 \Delta X_{it})) = 0$ avec $\Delta = (1 - L)$ est un opérateur retard

Approche GMM en différences premières d'Arellano et Bond (1991)

- ① $\mathbb{E}(q_{it}(\varepsilon_{it} - \varepsilon_{it-1})) = 0$
- ② $q_{it} = (Y_{i0}, \dots, Y_{it-2}, x_i')$
- ③ $\Delta = (1 - L)$ est un opérateur retard
- ④ $\mathbb{E}(Q_i(\Delta \varepsilon_i)) = 0$

$$⑤ Q_i = \begin{bmatrix} q_{i2} & \cdot & \dots & 0 \\ 0 & q_{i3} & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & q_{iT} \end{bmatrix}$$

- ⑥ $T(T-1)(K_1+1/2)$ conditions des moments

Approche GMM en différences premières d'Arellano et Bond (1991)

- ① L'estimateur GMM cherche à trouver $\hat{\theta}_{GMM} = (\hat{\beta}_{GMM}, \hat{\gamma}_{GMM})$ qui annule $\frac{1}{N} \sum_{i=1}^N (Q_i(\Delta \varepsilon_i))$

$$g(\theta) = g(\beta, \rho) = \frac{1}{N} \sum_{i=1}^N (Q_i(\Delta \varepsilon_i))$$

- ② La méthode des moments généralisés (GMM) minimise un objectif de la forme

$$m(\theta) = \|\mathbf{g}(\theta)'\|_{\mathbb{W}}$$

- ③ Avec $g(\theta)$ le vecteur des moments et \mathbb{W} (matrice de pondération) une matrice définie positive

$$m(\theta) = g(\theta)' \mathbb{W} g(\theta) = g(\beta, \gamma)' \mathbb{W} g(\beta, \gamma).$$

- ④ L'objectif est de résoudre

$$\hat{\theta}_{GMM} = (\hat{\beta}_{GMM}, \hat{\gamma}_{GMM}) = \arg \min m(\theta).$$

Approche GMM System de Blundel et Bond (1998)

① Cette approche combine deux systèmes d'équations: les équations en différences premières et les équations en niveau

② Les instruments dans le modèle en différences premières

$$\Delta Y_{it} = \gamma * \Delta Y_{it-1} + \beta' \Delta X_{it} + \Delta \varepsilon_{it}$$
 sont exprimés en niveau

③ Les instruments dans le modèle en niveau ($Y_{it} = \gamma * Y_{it-1} + \beta' X_{it} + \varepsilon_{it}$) sont exprimés en différences premières

④

$$\begin{cases} \Delta Y_{it} &= \gamma * \Delta Y_{it-1} + \beta' \Delta X_{it} + \Delta \varepsilon_{it} \\ Y_{it} &= \gamma * Y_{it-1} + \beta' X_{it} + \varepsilon_{it} \end{cases}$$

⑤ Blundel et Bond (1998) ont montré que l'estimateur GMM System est plus performant que celui en différences premières, ce dernier donne des résultats biaisés dans des échantillons finis lorsque les instruments sont faibles.

Comment l'approche GMM est mise en pratique?

Two-Step GMM (GMM en deux étapes)

Principe

- ① **Step 1:** La matrice de pondération \mathbb{W} est initialement supposée égale à l'identité.
- ② **Step 2:** Mettre à jour la matrice de pondération \mathbb{W} en utilisant les résidus de la première étape pour calculer une estimation plus efficace de la matrice de variance-covariance

Avantages

- ① Plus efficace que la méthode à une étape

Limites

- ① Les résultats peuvent être biaisés si le nombre d'instruments est élevé par rapport à la taille de l'échantillon
- ② Les statistiques de test de Hansen peuvent être sensibles

Comment l'approche GMM est mise en pratique? Iterative GMM (GMM itératif)

Principe

- ➊ La matrice de pondération \mathbb{W} est mise à jour à chaque itération jusqu'à convergence.

Avantages

- ➋ Produit des estimations asymptotiquement plus efficaces

Limites

- ➌ Calculs intensifs, surtout pour les grands ensembles de données

Commandes stata

- ➍ **xtabon, xtabon2, xtdpdgmm, xtdpdqml**

L'approche LSDVC (Least Squares Dummy Variable Corrected) de Bruno (2005)

Principe

- ① Combine l'estimateur LSDV en panel dynamique, l'estimateur GMM d'Arellano et Bond (1991) et l'estimateur d'Anderson et HSIAO (1982).

Avantages

- ① Moins d'instruments requis
- ② Simple à appliquer

Limites

- ① Hypothèse d'exogénéité des variables explicatives
- ② Adapté qu'uniquement lorsque N très grand

Approche Bias-Corrected (Breitung, Kripfganz et Hayakawa, 2021)

Principe

- ① Concentre sur la correction du biais d'estimation des paramètres dans les modèles dynamiques de panel, en particulier pour les panels avec une petite dimension temporelle (**Comande stata: xtdpdbc**).
- ② Le BC est particulièrement efficace lorsque T est petit, contrairement aux estimateurs GMM qui peuvent être biaisés dans ce cas.

Avantages

- ① Meilleure correction des biais d'attrition et de la petite dimension temporelle, estimation robuste avec des erreurs corrigées

Limites

- ① Les corrections de biais ne sont pas nécessaires avec un T grand.

Tests de validité des instruments

- ① Test de Sargan (1958)
- ② Test de Hansen (une version robuste du test de Sargan) “

Tests de validité des instruments

① Test de Sargan (1958)

- Hypothèse nulle (H_0) : Les instruments sont valides (ils ne sont pas corrélés avec les résidus). VS Hypothèse alternative (H_1) : Certains instruments sont invalides (ils sont corrélés avec les résidus).
- Statistique du test: Chi-deux avec un degré de liberté égal au nombre d'instruments moins le nombre de paramètres estimés
- Limites: Sensibilité à l'hypothèse d'homoskédasticité

Tests de validité des instruments

- ① Test de Hansen (version robuste du test de Sargan)
 - Hypothèse nulle (H_0) : Les instruments sont valides VS Hypothèse alternative (H_1) : Certains instruments sont invalides
 - Statistique du test: Chi-deux avec un degré de liberté égal au nombre d'instruments moins le nombre de paramètres estimés
 - Avantages: Plus fiable que le test de Sargan en présence d'hétéroscléasticité ou d'autocorrélation

- ① Modèles Non-linéaires en Panel
- ② Modèles à Corrélation Spatiale
- ③ Modèles de Données de Panel Non-stationnaires
- ④ Modèle de gravité en panel
- ⑤ Modèle de frontière stochastique en données de panel

Modèles logit, probit, Tobit (choix binaire)

① Logit

$$P(y_{it} = 1 \mid X_{it}, \mu_i) = \frac{\exp(X'_{it}\beta + \mu_i)}{1 + \exp(X'_{it}\beta + \mu_i)}$$

② Probit

$$P(y_{it} = 1 \mid X_{it}, \mu_i) = \Phi(X'_{it}\beta + \mu_i)$$

③ Tobit

$$y_{it}^* = X'_{it}\beta + \mu_i + \epsilon_{it}, \quad y_{it} = \max(0, y_{it}^*)$$

④ Poisson

$$P(y_{it} \mid X_{it}, \mu_i) = \frac{\exp(-\lambda_{it})\lambda_{it}^{y_{it}}}{y_{it}!}, \quad \lambda_{it} = \exp(X'_{it}\beta + \mu_i)$$

Modèles de comptage (Poisson et Binomial negatif)

① Poisson

$$P(y_{it} | X_{it}, \mu_i) = \frac{\exp(-\lambda_{it}) \lambda_{it}^{y_{it}}}{y_{it}!}, \quad \lambda_{it} = \exp(X'_{it}\beta + \mu_i)$$

② Binomial Negatif

Une généralisation du modèle de Poisson pour tenir compte de la sur-dispersion.

Modèles non-linéaires à seuil en panel, Hansen (1999)

① Exemple d'un modèle à deux régimes

$$y_{it} = \begin{cases} X'_{it}\beta_1 + \mu_i + \epsilon_{it}, & \text{si } q_{it} \leq \gamma, \\ X'_{it}\beta_2 + \mu_i + \epsilon_{it}, & \text{si } q_{it} > \gamma. \end{cases} \quad (1)$$

② Estimation du seuil

$$\hat{\gamma} = \arg \min_{\gamma} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - X'_{it}\beta(\gamma) - \mu_i)^2.$$

③ Estimation des paramètres

$$\hat{\beta}_1 = \arg \min_{\beta_1} \sum_{q_{it} \leq \hat{\gamma}} (y_{it} - X'_{it}\beta_1 - \mu_i)^2,$$

$$\hat{\beta}_2 = \arg \min_{\beta_2} \sum_{q_{it} > \hat{\gamma}} (y_{it} - X'_{it}\beta_2 - \mu_i)^2.$$

Principe, formulation et exemples

- ① **Principe** Modelisation des données où les observations ne sont pas uniquement influencées par des facteurs individuels et temporels, mais aussi par des effets géographiques ou spatiaux
- ② **Formulation (fixed or random effects spatial models)**

$$Y_{it} = \tau Y_{it-1} + \psi WY_{it-1} + \rho WY_{it} + \beta X_{it} + \alpha_i + \gamma_t + \nu_{it} \quad (2)$$

- τ : coefficient associé à Y_{it-1} , variable dépendante retardée,
- ψ : coefficient associé à WY_{it-1} , dépendance spatiale retardée,
- W : matrice de contiguïté spatiale,
- ρ : coefficient associé à WY_{it} , dépendance spatiale contemporaine,
- α_i : effet individuel spécifique à l'unité i ,
- γ_t : effet temporel spécifique à la période t ,
- ν_{it} : terme d'erreur idiosyncratique,

- ③ **Exemples:** Analyse de la transmission des prix, Analyse des flux commerciaux, analyse de la productivité agricole

Hypothèses des Modèles Spatiaux

- **Stationnarité** : la relation spatiale est constante dans l'espace.
- **Exogénéité des régresseurs** : les variables explicatives ne sont pas corrélées avec le terme d'erreur.
- **Matrice de pondération spatiale (W)** :
 - Doit être définie a priori (distance, voisinage, etc.).
 - Symétrique ou non, selon le contexte.
- Les erreurs suivent une distribution normale.

Approches d'Estimation

- 1 **Méthode des Moindres Carrés (OLS)** : Non adaptée en présence de corrélation spatiale.
- 2 **Maximum de Vraisemblance (ML) :**
 - Appropriée pour les modèles SAR et SEM.
 - Estime directement les paramètres spatiaux (ρ, ψ).
- 3 **Méthode des Moments Généralisés (GMM) :**
 - Plus flexible.
 - Utile pour corriger des spécifications erronées.
- 4 **Bayésienne** : Exploite des distributions a priori pour estimer les paramètres.
- 5 Commandes stata: **xtspatial, xsmle**

Principe, Formulation et exemple

1 **Principe** Les modèles ARDL permettent d'inclure des retards de la variable dépendante ainsi que des variables explicatives dans l'estimation, ce qui permet d'analyser les effets à court et à long terme. En plus, ils permettent d'analyser les variables stationnaires et non stationnaires avec un ordre d'intégration égal à 1.

2 Formulation (ARDL en panel)

$$Y_{it} = \alpha_i + \gamma_t + \sum_{p=1}^q \theta_p Y_{it-p} + \sum_{q=1}^r \phi_q X_{it-q} + \epsilon_{it} \quad (3)$$

- θ_p et ϕ_q : coefficients des termes autoregressifs et distribués des variables dépendantes et explicatives,
- ϵ_{it} : terme d'erreur idiosyncratique.

3 Exemples: Analyse de l'intégration des marchés, Analyse de la croissance

Hypothèses des Modèles ARDL

- 1 Stationnarité des résidus (après différenciation si nécessaire).
- 2 Relation de cointégration entre les variables (long terme).
- 3 Absence de corrélation des résidus.

Approches d'Estimation

- 1 Modèle PMG (Pooled Mean Group) : Autorise des coefficients hétérogènes à court terme mais homogènes à long terme.
- 2 Modèle MG (Mean Group) : Autorise des coefficients hétérogènes dans toutes les dimensions.
- 3 Modèle DFE (Dynamic Fixed Effects) : Suppose des coefficients fixes à travers les entités.
- 4 Commande stata: **xtpmg**

Principe, formulation et exemple

1 **Principe** Le modèle de frontière stochastique permet de mesurer l'**efficacité technique** et d'identifier les facteurs qui influencent l'efficacité technique et d'isoler la partie aleatoire.

2 Formulation du Modèle

$$y_{it} = f(x_{it}; \beta) \cdot \exp(-u_{it}) \cdot \exp(v_{it}) \quad (4)$$

- y_{it} : Production observée de l'unité i au temps t .
- x_{it} : Intrants (facteurs de production) pour l'unité i au temps t .
- β : Paramètres du modèle à estimer.
- u_{it} : **Inefficacité technique**, suit une distribution semi-normale
- v_{it} : **Erreur aléatoire**, suit une distribution normale .
- Transformation logarithmique: $\ln(y_{it}) = \ln(f(x_{it}, \beta)) + v_{it} - u_{it}$.

Hypothèses du modèle

- 1 La frontière de production est stochastique.
- 2 Les inefficacités u_{it} suivent une distribution asymétrique.
- 3 Les erreurs v_{it} et u_{it} sont indépendantes.
- 4 Les données sont correctement spécifiées (modèle fonctionnel approprié, ex. Cobb-Douglas ou Translog).

Approches d'estimation

- 1 Maximum de vraisemblance (MLE)
- 2 Méthodes bayésiennes
- 3 Approche en deux étapes (déconseillée pour sa faible efficacité statistique)
- 4 Commande stata: **sfpanel**

Principe, formulation et exemple

1 **Principe** Le modèle de gravité est un outil clé en économie internationale pour analyser les flux commerciaux entre pays ou régions. Ce modèle est inspiré de la loi de gravitation de Newton : les flux commerciaux augmentent avec la taille économique et diminuent avec la distance.

2 Formulation du Modèle

$$\log(F_{ij,t}) = \beta_0 + \beta_1 \log(GDP_i) + \beta_2 \log(GDP_j) - \beta_3 \log(Distance_{ij}) + \gamma X_{ijt} + \epsilon_{ij,t} \quad (5)$$

- $F_{ij,t}$: Flux commerciaux entre i et j à l'année t .
- GDP_i, GDP_j : PIB des pays i et j .
- $Distance_{ij}$: Distance entre les deux pays.

3 **Exemple:** Mesure de l'effet d'un Accord Commercial sur le Commerce Bilatéral

Hypothèses du Modèle de Gravité

- 1 Les flux bilatéraux dépendent positivement des "masses économiques" des partenaires.
- 2 Les flux diminuent avec la distance entre les partenaires.
- 3 Les effets spécifiques des pays (fixes ou aléatoires) et des périodes peuvent influencer les flux.
- 4 Les chocs ou biais non observés sont aléatoires et indépendants des variables explicatives.

Approches d'Estimation

1 Effets fixes (FE) :

- Contrôle des caractéristiques inobservées spécifiques aux pays ou aux paires.
- Exemple : Effets fixes par pays exportateurs et importateurs.

2 Effets aléatoires (RE) :

- Suppose que les effets spécifiques ne sont pas corrélés avec les variables explicatives.

3 Poisson Pseudo Maximum Likelihood (PPML) :

- Adapté aux données avec beaucoup de zéros.
- Estimation robuste pour éviter les biais dans la log-linéarisation.
- Commandes stata: **xtreg, ppmlhdfe, xtpoisson**

Produit de kronecker

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}$$

Figure: Produit de kronecker

Estimateur LSDV

$$Q_{(T,T)} = I_T - \frac{1}{T}ee' = \begin{pmatrix} 1 - \frac{1}{T} & -\frac{1}{T} & \dots & -\frac{1}{T} & -\frac{1}{T} \\ -\frac{1}{T} & 1 - \frac{1}{T} & \dots & -\frac{1}{T} & -\frac{1}{T} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{1}{T} & -\frac{1}{T} & \dots & 1 - \frac{1}{T} & -\frac{1}{T} \\ -\frac{1}{T} & -\frac{1}{T} & \dots & -\frac{1}{T} & 1 - \frac{1}{T} \end{pmatrix}$$

Figure: Opérateur within

Estimateur LSDV ou Within

$$\hat{\beta}_{LSDV} = \left(\sum_{i=1}^n X_i' Q X_i \right)^{-1} \left(\sum_{i=1}^n X_i' Q y_i \right)$$

Figure: Estimateur LSDV

Projets

- Modèles à variable dépendante qualitative (Logit et Probit) en données de panel
- Modèle de frontière stochastique en données de panel
- Modèle ARDL en données de panel
- Modèle VAR en données de panel
- Modèle de gravité

References

- 1 Anderson, T. W., Hsiao, C. (1982). "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, 18(1), 47-82.
- 2 Arellano et Bond (1991) : Application à Difference GMM pour les données de panel.
- 3 Arellano et Bover (1995) : Introduction des moments supplémentaires pour System GMM.
- 4 Baltagi B. H. , (2005), "Economic analysis of panel data", Ed. J. Wiley
- 5 Blundell et Bond (1998) : Raffinement de System GMM avec Two-Step GMM.
- 6 Bruno, G. S. F. (2005). "Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals." *Stata Journal*, 5(4), 473-500.

References

Christophe Hurlin: <https://sites.google.com/view/christophe-hurlin/teaching-resources/panel-data-econometrics>