Économétrie des Données de Panel

Ch 2. Modèles Linéaire Dynamiques

Pr. Philippe Polomé, Université Lumière Lyon 2

M1 APE Analyse des Politiques Économiques
M1 RISE Gouvernance des Risques Environnementaux

2016 – 2017

Introduction

- Le Ch. précédent a présenté des variantes du MRL avec
  - des EF ou des EA et
  - des régressuers **strictement exogènes**

$$E[\varepsilon_{it}|\alpha_i, x_{i1}, ..., x_{iT}] = 0,\ t = 1, ..., T$$

- Ce Ch. : Modèles linéaires en relaxant cette hyp.
  - Régresseurs **endogènes** $E[\varepsilon_{it}|x_{ijt}] \neq 0$ pour au moins un j

Plan

- Tous les estimateurs de ce Ch. sont des applications de la Méthode Généralisée des Moments (GMM)
- Cas général linéaire
  - Rappel en coupe transversale
  - Disponibilité des instruments en panel
- 2 applications
  - Hausman-Taylor
    - Essayer de récupérer des régresseurs invariants dans le temps
  - Arrellano-Bond
    - p.e. **variable dépendante retardée**
    - endogène en panel puisque autocorrélation

Outline

Le principe d'analogie

- Les estimateurs GMM sont basés sur le principe d'analogie
  - On suppose une ou pls conditions sur des moments de la population
  - On trouve des valeurs des paramètres t.q. ces conditions se réalisent dans l'échantillon
- Exemple classique de MM : estimation de la moyenne de la population (espérance)
  - lorsque y est iid d'espérance μ
- Dans la population $E[y-\mu]=0$ par définition
  - Replacer $E[\cdot]$ pour la population par $N^{-1}\sum_{i=1}^{N}(\cdot)$ pour l'échantillon définit le moment empirique correspondant :

$$\frac{1}{N}\sum_{i=1}^{N}(y_i-\mu)=0$$

- Résoudre pour μ définit l'estimateur MM $\hat{\mu}_{MM}=N^{-1}\sum_{i=1}^{N}y_i=\bar{y}$
  - L'estimateur MM de l'espérance est la moyenne de l'échantillon

Régression linéaire en coupe transversale

- MRL $y = x^{'}\beta + u$
  - x & β sont des vecteurs $K \times 1$
- Supposons $E\left[u|x\right] = 0$
  - Par la loi des espérances itérées (law of iterated expectations)
    - K conditions de moment inconditionnel $E\left[xu\right] = 0$
  - Donc, quand l'erreur a zéro moyenne conditionnalle / est "exogène" / orthogonale

$$E\left[x\left(y - x^{'}\beta\right)\right] = 0$$

- Estimateur MM de β = solution de ces mêmes conditions dans l'échantillon

$$\frac{1}{N}\sum_{i=1}^{N} x_i\left(y_i - x_i^{'}\beta\right) = 0$$

  - Ça donne $\hat{\beta}_{MM} = \left(\sum_i x_i^{'}x_i\right)^{-1}\sum_i x_i^{'}y_i$

Estimateur Variable Instrumentale (VI) linéaire

- Imaginons qu'il y ait endogénéité : $E[u|x] \neq 0$
- Si on a des instruments z t.q. $E[u|z] = 0$ **et**
  - Que ces instruments sont **bons**
    - fortement corrélés avec les régresseurs
  - Que $dim(z) = dim(x)$: exactement un instrument par régresseur
    - modèle dit "exactement identifié"
- Alors $\hat{\beta}_{MM} = \left(\sum_i z_i' x_i\right)^{-1} \sum_i z_i' y_i$ est consistant
  - alors que $\hat{\beta}_{OLS} = \left(\sum_i x_i' x_i\right)^{-1} \sum_i x_i' y_i$ est inconsistant
  - $\hat{\beta}_{MM}$ est l'estimateur Variable Instrumentale VI
    - une application du principe MM

Conditions de Moments supplémentaires

- des moments/instruments **additionels** peuvent améliorer **l'efficiencen**
  - mais demande une adaptation de MM
- Considérons que $dim\,(z) > dim\,(x)$
  - plus d'instruments que de régresseurs
- Quels instruments prend-t-on ? Toute sélection est **arbitraire**
  - soit z_1 et z_2 deux sous-ensembles de z t.q. $dim\,(z_1) = dim\,(z_2) = dim\,(x)$
  - Alors, $\hat{\beta}_{MM1} = \left(Z_1'X\right)^{-1} Z_1'Y \neq \left(Z_2'X\right)^{-1} Z_2'Y = \hat{\beta}_{MM2}$

Conditions de Moments supplémentaires

- Si on écrit les conditions de moment

$$E\left[Z\left(y - X'\beta\right)\right]$$

  - Donc : plus de conditions que de paramètres à estimer
- L'estimateur GMM **choisit** $\hat{\beta}$ de sorte à ce que le vecteur de conditions de moments dans l'échantillon

$$\frac{1}{N}\sum_i z_i\left(y_i - x_i'\beta\right)$$

soit aussi petit que possible en termes quadratiques
- C'est-à-dire $\hat{\beta}_{GMM}$ **minimise** :

$$Q_N(\beta) = \left[\frac{1}{N}\sum_i z_i\left(y_i - x_i'\beta\right)\right]' \mathbf{W}_N \left[\frac{1}{N}\sum_i z_i\left(y_i - x_i'\beta\right)\right]$$

où $\mathbf{W}_N$ est une matrice de poids dépendant de l'application

Conditions de Moments supplémentaires

- Comment choisir $\mathbf{W}_N$?
  - Soit $dim(z) = r$; $\mathbf{W}_N$ est $r \times r$, sdp et ne dépend pas de β
    - Essentiellement, W_N est un choix de pondération des instruments
    - Pour retrouver k instruments pondérés
  - Tout choix de $\mathbf{W}_N$ définit un estimateur consistant
    - mais avec différentes variances (quand $r > k$)
  - GMM spécifie le choix **optimal** de la matrice de poids $\mathbf{W}_N$
    - selon chaque cas particulier (autocorrélation, hétérocédasticité)
    - t.q. $\hat{\beta}_{GMM}$ a la plus **petite variance asymptotique**
    - 3 cas en panel

Outline

Hypothèses Panel

- Soit le modèle linéaire en panel

$$y_{it} = x_{it}\beta + u_{it} \tag{1}$$

x_{it} peut contenir des régresseurs invariants dans le temps et un intercept

- Pour le modèle de cette section, simplification :
  - **pas d'effet individuel** α_i
  - x_{it} comprend **seulement** des variables de la périodes courantes
    - Pas de retard
  - On peut voir cette simplification comme si les données étaient transformées
    - comme dans le Ch. 1 avec les lestimateurs˜
- En **gras** on empile les T observations pour le $i^{ème}$ agent

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{u}_i \tag{2}$$

EF et EA avec endogénéité

- Temporairement on remet les α_i
  - le modèle (1) $y_{it} = x_{it}\beta + u_{it}$ devient
$$y_{it} = \alpha_i + \mathbf{x}_{it}^{'}\beta + \epsilon_{it} \tag{3}$$
- Certains régresseurs dans x_{it} sont supposés **endogène**, donc $E\left[\mathbf{x}_{it}\left(\alpha_i + \epsilon_{it}\right)\right] \neq 0$
  - On appelle **EA** si $\exists$ instruments Z_i t.q. $E\left[\mathbf{Z}_i^{'}\left(\alpha_i + \epsilon_{it}\right)\right] = 0$
    - Alors on applique GMM selon les formules présentées + loin
  - On appelle **EF** s'il est **seulement** possible de trouver des instruments t.q. $E\left[\mathbf{Z}_i^{'}\epsilon_{it}\right] = 0$, mais $E\left[\mathbf{Z}_i^{'}\alpha_i\right] \neq 0$
    - Dans ce cas, il faut éliminer les EF α_i par différentiation comme dans le Ch. 1
    - et seuls les coefficients des régresseurs variables dans le temps sont identifiés
    - C'est la même discussion qu'au Ch. 1

Conditions de moment GMM en panel

- On revient au modèle sans les α_i
  - Donc on est soit en EA, soit en EF après élimination des α_i
- On suppose une matrice $T \times r$ d'instruments Z_i
  - $r \geq K$ est le nombre d'instruments / de conditions de moment t.q.

$$E\left[\mathbf{Z}_i'\mathbf{u}_i\right] = 0 \tag{4}$$

- L'estimateur GMM basé sur ces conditions **minimise** une forme quadratique : $\hat{\beta}_{PGMM} =$

$$\left[\left(\sum_i \mathbf{X}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_i \mathbf{Z}_i'\mathbf{X}_i\right)\right]^{-1}\left(\sum_i \mathbf{X}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_i \mathbf{Z}_i'\mathbf{y}_i\right)$$

- Cet estimateur est consistant pour autant que les conditions de moment (4) tiennent
  - 3 cas

Cas 1. Panel GMM juste identifié

- Dans ce cas $r = K$, donc $dim(z) = dim(x)$
  - alors $\hat{\beta}_{PGMM}$ **se simplifie** en l'estimateur **VI** quel que soit $\mathbf{W}_N$

$$\hat{\beta}_{VI} = \left[\left(\sum_i \mathbf{X}_i' \mathbf{Z}_i\right)\right]^{-1} \left(\sum_i \mathbf{Z}_i' \mathbf{y}_i\right)$$

  - On voit bien que c'est la version panel de $\hat{\beta}_{VI} = \left(X'Z\right)^{-1} Z'Y$
- S'il y a des régresseurs exogènes
  - Ils sont leurs propres instruments
- Imaginons que $r > K$: plus d'instruments que de régresseurs
  - Il faut utiliser la formule avec la matirce de poids W_N
  - Il y a 2 cas (diapo suivante)

Cas 2 & 3. estimateurs PGMM suridentifiés optimaux

- Hyp. **pas d'hétéroscédasticité & pas d'autocorrélation**
  - $\hat{\beta}_{2SLS} = \left[\mathbf{X}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{y}$
    - C'est le cas équivalent à MC en 2 étapes (MC2E / 2SLS)
    - Stata ivreg ou menu stat→Panel→Endogenous covariates→IV regression
    - Gretl menu Model→Panel→Panel IV models
- Pas de telle hyp. (**robust**)
  - $\hat{\beta}_{2SGMM} = \left[\mathbf{X}'\mathbf{Z}\hat{\mathbf{S}}^{-1}\mathbf{Z}'\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{Z}\hat{\mathbf{S}}^{-1}\mathbf{Z}'\mathbf{y}$
    - $\hat{\mathbf{S}} = \frac{1}{N}\sum_i \mathbf{Z}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{Z}_i$ est un estimateur robuste de type White
    - $\hat{S}$ est consistant pour la matrice $r \times r$ $\mathbf{S} = plim\frac{1}{N}\sum_i \mathbf{Z}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{Z}_i$
  - C'est l'estimateur GMM en 2 étapes (2-Step GMM)
    - Premier pas est un estimateur consistant de β comme $\hat{\beta}_{2SLS}$
    - Ensuite on utilise les résidus $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\beta}_{2SLS}$ pour calculer $\hat{\mathbf{S}}$
    - R dans le package plm

Panel GMM suridentifié

- Dans beaucoup d'applications $\mathbf{Z}_i$ est composé de valeurs **retardées** des régresseurs
  - endogènes &/ou exogènes
- Imaginons qu'on dispose de r instruments
  - On peut souvent supposer que le premier retard de chaque régresseur est non-corrélé avec l'erreur courante
    - donc x_{it-1} sont disponible comme instruments **additionels** pour x_{it}
    - appelé **exogénéité faible** / **instruments prédéterminés**
    - On peut souvent poursuivre ainsi avec 2 retards, 3 retards...
    - On perd chaque fois une période d'observation, l'efficience baisse...
    - mais on augmente le nombre d'instruments, l'efficience augmente
  - Le modèle est alors très facilement **suridentifié**
    - Cela fait que même si on n'a pas d'endogénéité, panel GMM sera un estimateur plus efficient que MC

Inférence Panel-robuste

- $\hat{\beta}_{PGMM}$ est asymptotiquement normal
  - avec une matrice de var-cov asymptotique compliquée
- Un estimateur consistant de cette matrice existe
  - conditionnellement à un choix de $\mathbf{W}_N$
  - et on peut supposer l'indépendance entre i
- Un estimateur robuste de type White existe
  - Il permet de calculer des et robustes à l'Het. et l'Autoc.
  - Mais ça n'est pas généralement implémenté dans les logiciels
    - Pas dans Stata en général, sauf pour des cas particuliers (+ loin)
    - On ne verra que ces cas particuliers
- Alternativement, Bootstrap est faisable
- C'est comme dans le Ch. 1

Outline

Motivation

- Habituellement, en panel, l'endogénéité
  - vient de régresseurs corrélés avec les effets individuels α_i
  - amène à l'inconsistance des estimateurs EA
- L'estimateur within est consistant
  - mais alors les coefficients des régresseurs invariants dans le temps ne peuvent être estimés
  - alors qu'il y a beaucoup d'études dont ce serait précisément le but

Spécification

- Modèle Hausman & Taylor

$$y_{it} = \mathbf{x}_{1it}^{'}\beta_1 + \mathbf{x}_{2it}^{'}\beta_2 + \mathbf{w}_{1i}^{'}\gamma_1 + \mathbf{w}_{2i}^{'}\gamma_2 + \alpha_i + \epsilon_{it} \quad (5)$$

où $\mathbf{x}_{1it}$ & $\mathbf{w}_{1i}$ ne sont pas corrélés avec α_i mais $\mathbf{x}_{2it}$ & $\mathbf{w}_{2i}$ le sont ; w indique les régresseurs invariants dans le temps

- C'est le modèle panel classique
  - sauf qu'on précise quels régresseurs sont corrélés avec α_i et lesquels sont invariants
  - Tous les régresseurs sont **non-corrélés** avec ε_{it}
- La transformation **Within** $\ddot{z}_{it} = z_{it} - \bar{z}_i$ élimine la corrélation avec α_i

$$\ddot{y}_{it} = \ddot{\mathbf{x}}_{1it}^{'}\beta_1 + \ddot{\mathbf{x}}_{2it}^{'}\beta_2 + \ddot{\epsilon}_{it}$$

- Mais aussi les régresseurs invariants dans le temps

Estimateur Hausman–Taylor consistant

- $\mathbf{x}_{2i}$ corrélé avec α_i
  - mais pas avec la transformation within $\ddot{x}_{2it} = x_{2it} - \bar{x}_{2i}$
    - Puisque la corrélation avec α_i ne peut être qu'avec la partie de $x_{2it}\,\forall t$ invariante au temps
  - Donc $\ddot{x}_{2it}$ peut être utilisé comme **instrument** pour x_{2it} endogène
- On prend pareillement $\ddot{x}_{1it}$ comme instrument pour x_{1it}
  - plutôt que x_{1it} lui-même comme on ferait d'habitude
    - Donc, même x_{1it} exogène est instrumenté
    - car cela sépare x_{1it} de sa partie invariante au temps $\bar{x}_{1i}$
  - Celle-ci ($\bar{x}_{1i}$) est utilisée comme instrument pour w_{2i} endogène
    - pourrait être un instrument faible
- w_{1i} exogène est utilisé comme instrument pour lui-même

Estimateur Hausman–Taylor

$$\begin{array}{cccccc} y_{it} = & \mathbf{x}'_{1it}\beta_1 + & \color{red}{\mathbf{x}'_{2it}}\beta_2 + & \mathbf{w}'_{1i}\gamma_1 + & \color{red}{\mathbf{w}'_{2i}}\gamma_2 + & \color{red}{\alpha_i} + \epsilon_{it} \\ & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \color{blue}{\ddot{\mathbf{x}}_{1it}} & \ddot{\mathbf{x}}_{2it} & \mathbf{w}'_{1i} & \color{blue}{\bar{\mathbf{x}}_{1i}} & \end{array}$$

- **Identification** des coef. des régresseurs invariant dans le temps γ
  - si # régresseurs **exogènes variant dans le temps** $\geq$ # régresseurs **endogènes invariant dans le temps**
  - c'est-à-dire si # de $x_1 \geq$ # de w_2
  - C'est-à-dire si # instruments est au moins égal au # de régresseurs w_2
- **Inefficient** puisque cet estimateur ignore la structure de corrélation panel de $(\alpha_i + \epsilon_{it})$
- HT est implémenté dans Stata
  - Mais apparemment sans les et panel-robustes

Exemple : Baltagi and Khanti-Akom (1990)

- 595 obs. d'individus sur 1976–1982
  - Du Panel Study of Income Dynamics (PSID)
    - use http ://www.stata-press.com/data/r11/psidextract
    - r11 pour moi car j'ai Stata 11

lwage	log-salaire, supposé fonction de :		y	Inst
ed	IT	années d'éducation	w_2	$\bar{\mathbf{x}}_1$
wks	VT	le temps que la personne a tavaillé pour la firme	x_2	$\ddot{\mathbf{x}}_2$
exp	VT	expérience de travail		
occ	VT	occupation - type de fonction je suppose	x_1	$\ddot{\mathbf{x}}_1$
	si la personne (0/1)			
smsa	VT	... vit dans une grande agglomération	x_1	$\ddot{\mathbf{x}}_1$
south	VT	... vit dans le sud		
ind	VT	... est dans l'industrie		
ms	VT	... est mariée	x_2	$\ddot{\mathbf{x}}_2$
union	VT	... est syndiquée		
fem	TI	... est un homme	w_1	w_1
blk	TI	... est African-American		

Endogénéité

- Les IT *fem*, *blk* sont exogènes : w_1
- Les VT *exp*, *exp*^{2}, *wks*, *ms*, et *union*
  - Peuvent tous être corrélés avec les effets individuels inobservés
    - $=$ sont endogènes
  - Ces variables présentent-elles suffisamment de variation within-panel pour être leurs propres instruments ?
    $\ddot{x}_{2it} = x_{2it} - \bar{x}_{2i}$
    - On regarde les variations within / between
    - xtsum exp exp2 wks ms union $\rightarrow$ assez faible, mais ok
- On suppose que les VT *occ*, *south*, *smsa*, *ind* sont toutes exogènes : x_1
  - $\bar{x}_1$ est utilisé comme instrument pour l'endogène IT *ed* : w_2
    - L'instrument pour x_1 is $\ddot{x}_1$
  - Corrélation suffisante pour identifier le coefficient de *ed* ?
    - correlate occ south smsa ind ed $\rightarrow$ risque instrument faible

Régression

- xthtaylor lwage occ south smsa ind exp exp2 wks ms union fem blk ed, endog(exp exp2 wks ms union ed)
- Décomposition de la varience en σ_μ et σ_ϵ : 0.9418 et 0.1518, respectivement
  - indiquant qu'une large fraction de la variance totale de l'erreur est attribuée à μ_i
  - Les stat z indique que plusieurs coefficients pourraient ne pas être significativement différent de zéro
  - Les coefficients des variables IT *fem* et *blk* ont des et relativement grands
  - L'et du coefficient de *ed* est relativement petit

Outline

Dynamique

- Les régresseurs comprennent **un retard** de la variable dépendante

$$y_{it} = \gamma y_{i,t-1} + \mathbf{x}_{it}^{'}\beta + \alpha_i + \epsilon_{it}, \; i = 1, \ldots, N, \; t = 1, \ldots, T \quad (6)$$

- On suppose $|\gamma| < 1$
  - Dans les applications, cela peut être testé en utilisant des tests de racines unitaires panel
  - $\neg R$ racine unitaire, alors y_{it} est une marche aléatoire (random walk)
    - L'inférence n'est pas valide
  - Pas dans ce cours

Corrélation entre y_{it} & $y_{i,t-1}$

- On a à présent une corrélation sérielle dans y_{it}
  - directement via $y_{i,t-1}$
  - en plus d'indirectement via la persistence donnée par α_i
  - Ces 2 causes amènent à **différentes interprétations** de la corrélation dans le temps
- Du modèle (6) avec $\beta = 0$
  - $y_{it} = \gamma y_{i,t-1} + \alpha_i + \epsilon_{it}$, on a

$$\begin{aligned} Cor\left[y_{it}, y_{i,t-1}\right] &= Cor\left[\gamma y_{i,t-1} + \alpha_i + \epsilon_{it}, y_{i,t-1}\right] \\ &= \gamma Cor\left[y_{i,t-1}, y_{i,t-1}\right] + Cor\left[\alpha_i, y_{i,t-1}\right] \\ &= \gamma + \frac{1-\gamma}{1 + (1-\gamma)\,\sigma_\epsilon^2 / (1+\gamma)\,\sigma_\alpha^2} \end{aligned}$$

- La 2º égalité suppose $Cor\left[\epsilon_{it}, y_{i,t-1}\right] = 0$
- La 3º égalité s'obtient dans le cas particulier des EA
  - avec $\epsilon_{it} \sim iid\left[0, \sigma_\epsilon^2\right]$ & $\alpha_{it} \sim iid\left[0, \sigma_\alpha^2\right]$

2 raisons possible de corrélation entre y_{it} & $y_{i,t-1}$

1. **Véritable dépendance à l'état (True state dependence)**
    - Quand la corrélation dans le temps est due au mécanisme causal que $y_{i,t-1}$ détermine y_{it}
    - Cette dépendance est relativement grande si
        - l'effet individuel est relativement petit $\alpha_i \simeq 0$
        - ou lorsque σ^2_α est petit par rapport à σ^2_ϵ car alors $Cor\left[y_{it}, y_{i,t-1}\right] \simeq \gamma$
2. Corrélation **spurieuse** entre y_{it} & $y_{i,t-1}$, sans mécanisme causal,
    - due à de **l'hétérogénéité inobservée** α_i
        - donc $\gamma = 0$
        - mais $\hat{\gamma}_{OLS} \neq 0$ car $Cor\left[y_{it}, y_{i,t-1}\right] = \sigma^2_\alpha / \left(\sigma^2_\alpha + \sigma^2_\epsilon\right)$ comme dans le Ch. 1
    - due des question de **séries temporelles** : y_{it} is I(1)
        - Pas dans ce cours

Véritable dépendance à l'état vs. hétérogénéité inobservée

- Les 2 cas permettent une corrélation arbitrairement proche de 1 (100%)
  - parce que soit $\gamma \to 1$ ou $\sigma_\alpha^2/\sigma_\epsilon^2 \to 0$
  - Mais ces 2 explications ont des implications politiques radicalement différentes
    - On prend l'exemple des revenus y_{it}
- Explication "Véritable dépendance à l'état"
  - Les revenus y_{it} sont toujours au-dessus de la moyenne (ou en-dessous)
    - même après avoir contrôlé pour les régresseurs x_{it}
    - car les revenus futurs sont déterminés par les revenus passés
- Expication hétérogénéité inobservée
  - γ est en réalité petit mais des régresseurs important ont été **omis** de x_{it},
    - ce qui amène à un α_i élevé
    - qui fait que $\hat{\gamma}_{LS}$ semble élevé (facteurs confondants)

Véritable dépendance à l'état vs. hétérogénéité inobservée

- C'est-à-dire, les gens sont-ils pauvres (ou riches) parce que
  - Ils ont été pauvres (ou riches) ?
    - Dans ce cas, il faut traiter la pauvreté en transférant de l'argent
  - Ou bien ont-ils des caractéristiques individuelles qui font qu'ils sont pauvres ?
    - Dans ce cas, la pauvreté pourrait être traitée par exemple en améliorant l'éducation

Inconsistance des estimateurs du Ch. 1

- **Tous** les estimateurs du Ch.1 sont **inconsistants** lorsqu'on inclue un retard de la variable dependante
- p.e. **MCO** de y_{it} sur $y_{i,t-1}$ et $\mathbf{x}_{it}$
  - Erreur $(\alpha_i + \epsilon_{it})$, corrélée avec $y_{i,t-1}$ par α_i
- Estimateur **Within** : $y_{it} - \bar{y}_i$ sur $(y_{i,t-1} - \bar{y}_i)$ et $(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)$ avec erreur $(\epsilon_{it} - \bar{\epsilon}_i)$
  - $y_{i,t-1}$ correlée avec $\epsilon_{i,t-1}$ et donc avec $\bar{\epsilon}_i$
- Inconsistance aussi pour l'estimateur **EA** du Ch 1
  - puisqu'il s'agit d'une combinaison linéaire de within et between

Modèle différences premières

- Le modèle dyn. (6) en diff. premières, $t = 2, \ldots, T$:

$$y_{it} - y_{i,t-1} = \gamma \left(y_{i,t-1} - y_{i,t-2}\right) + \left(\mathbf{x}_{it} - \mathbf{x}_{i,t-1}\right)' \beta + \left(\epsilon_{it} - \epsilon_{i,t-1}\right)$$

- MCO sur ce modèle est inconsistant parce que $y_{i,t-1}$ corrélé avec $\epsilon_{i,t-1}$
    - donc le régresseur $(y_{i,t-1} - y_{i,t-2})$ corrélé avec l'erreur $(\epsilon_{it} - \epsilon_{i,t-1})$
- Donc l'estimateur D1 du modèle dyn, est aussi inconsistant
- Par contre, on peut utiliser VI
    - avec $y_{i,t-2}$ comme instrument pour $(y_{i,t-1} - y_{i,t-2})$
    - $y_{i,t-2}$ instrument valide puisque non-corrélé avec $(\epsilon_{it} - \epsilon_{i,t-1})$
        - Ça pourrait encore dépendre de la corrélation sérielle des erreurs ϵ_{it}
    - $y_{i,t-2}$ est un "bon" instrument puisque corrélé à $(y_{i,t-1} - y_{i,t-2})$

Estimation plus efficiente du modèle en différences premières

- L'estimateur VI précédent est **juste identifié**
  - il demande qu'au moins **3** périods of data soient disponibles pour chaque individu
- Une estimation **plus efficiente** est possible
  - En utilisant des retards **supplémentaires** de la variable dépendante comme instruments
  - L'estimateur devient alors **sur-identifié**
    - estimation par 2SLS ou 2SGMM

Estimateur Arellano–Bond

- L'estimateur panel GMM qui fait ça est appelé

Arellano–Bond $\hat{\beta}_{AB} =$

$$\left[\left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i'\tilde{\mathbf{X}}_i\right)\right]^{-1}\left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i'\tilde{\mathbf{y}}_i\right)$$

- avec
  - $\tilde{\mathbf{X}}_i$ est une matrice $(T-2)\times(K+1)$ avec $t^{ème}$ ligne $\left(\Delta y_{i,t-1}, \Delta \mathbf{x}_{it}^{'}\right)$, $T = 3, \ldots, T$
  - $\tilde{\mathbf{y}}_i$ est un vecteur $(T-2)\times 1$ avec $t^{ème}$ ligne Δy_{it}
    - On est bien dans le modèle différences premières
  - $\mathbf{Z}$ est défini à la prochaine diapo

Estimateur Arellano–Bond

- $\mathbf{Z}_i$ est une matrice $(T-2) \times r$ d'instruments :

$$\mathbf{Z}_i = \begin{bmatrix} \mathbf{z}'_{i3} & 0 & \cdots & 0 \\ 0 & \mathbf{z}'_{i4} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{z}'_{iT} \end{bmatrix}$$

avec souvent $\mathbf{z}'_{it} = \left[y_{i,t-2}, y_{i,t-3}, \ldots, y_{i1}, \Delta \mathbf{x}'_{it} \right]$

- Donc on rajoute un instrument à chaque période

Estimateur Arellano–Bond

- Des retards de $\mathbf{x}_{it}$ ou de $\Delta\mathbf{x}_{it}$ peuvent de plus être utilisé comme instruments
  - et pour T suffisamment grand, on peut limiter le nombre de retards de y_{it} qui sont utilisés comme instrument
    - p.e. pas plus que $y_{i,t-4}$
- 2SLS et 2SGMM correspondent à différentes matrices de poids W_N
  - selon le traitement de l'hétéroscédasticité et de l'autocorrélation
    - Voir section précédente

Exemple

- Données Arellano-Bond disponible par webuse abdata
  - Year = t, n = log of employment, w = log of real wage, k = log of gross capital, ys = log of industry output, unit = firm index (id)
- Structure panel comprise par Stata
  - Vérification par Menu : Panel data : Declare dataset to...
    - Panel ID = id ; time=year
- Estimateur Arellano-Bond
  - Menu : panel data : Dynamic... :
  - xtabond n w k ys, lags(1) vce(robust) artests(2)
  - Différence importante avec "robust vce"
  - C'est GMM "2SLS" pour panel
    - assume pas autocorrélation & homoscé.
    - comme vu + haut
    - Pour 2SGMM utilisr xtdpd (linear dynamic PD)
    - On verra ça en M2