ENSEA Abidjan Année Scolaire 2020-2021

Ingénieurs Statisticiens

Economistes, 3eme année

ECONOMETRIE AVANCEE

Professeur KEHO Yaya

Septembre 2020

Abidjan, Côte d'Ivoire

Plan du Cours

Chapitre 1 : Modèle à variables instrumentales

Chapitre 2 : Modèle à décalages temporels

Chapitre 3 : Modèle à équations simultanées

Chapitre 4 : Modèle vectoriel autorégressif (VAR)

Chapitre 5 : Cointégration et modèle à correction d'erreur

Chapitre 1 : Modèle à variables instrumentales

1. Rappel des hypothèses du modèle linéaire ordinaire

$$y_t = a_0 + a_1 x_{1t} + \ldots + a_k x_{kt} + u_t$$

Soit

$$\mathbf{Y = Xa + u}$$

Hypothèses

- H1: $E(u_t)=0$
- H2: $V(u_t)=\sigma^2$
- H3: $E(u_t, u_{t'})=0 \quad t \neq t'$
- H4: $E(x_t u_t)=0$ →**Exogénéïté des X ou indépendance avec le terme d'erreur**
- H5: rang(X)=K<n
- H6: $\lim_{n\to\infty}\left(\frac{x'x}{n}\right) = \lim_{n\to\infty} S_X = V_X$ matrice finie et non singulière

Propriété de l'estimateur MCO de *a*

- ***Sans biais***

- ***Efficace*** : Utilise au mieux l'information disponible dans les données

- ***Convergent:*** Se rapproche de *a* quand la taille des données augmente

Que se passe-t-il si $E(x_t u_t) \neq 0$**?**

- Les variables explicatives ne sont pas exogènes

- L'estimateur des MCO est biaisé

- L'estimateur des MCO n'est pas convergent

Cas où cela peut arriver

- Variables explicatives mesurées avec erreurs
- Omission de variables explicatives Z corrélées avec celles prises en compte X
- Modèle autorégressif avec un terme d'erreur auto-corrélée
- Modèle à équations simultanées

→ Comment estimer le modèle de façon convergente?

2. Méthodes des variables instrumentales

- On veut régler un problème (la non-exogénéïté des explicatives X)

- On va se servir de variables Z dite instruments…

- …mais les instruments Z doivent vérifier:
  - C1: $E(Z'_t u_t)=0$ (Z exogènes)
  - C2: rang(Z)=H (pas de multicolinéarité)
  - C3: Z fortement correlée avec X
  - C4: $H \geq K$

Conditions d’identification de *a*

$$E(Z'_t u_t) = 0 \rightarrow E\left(Z'_t(y_t - X'_t a)\right) = 0$$

$$E(Z'_t y_t) = E(Z'_t X_t) a$$

On a un système de H équations à K inconnus

C4: $H \geq K$ assure l’identification des coefficients a

Estimateur à variables instrumentales

- On régresse par MCO les X sur les Z

$\rightarrow \hat{X}_t$

- On estime le modèle initial par MCO en remplaçant les X_t par $\hat{X}_t$:

$$y_t = \hat{X}_{t_t}{}' a + u_t$$

- L'estimateur de a est l'estimateur à variables instrumentales:

$$\hat{a}_{vi} = (X' P_Z X\,)^{-1} X' P_Z y$$

Cas particuliers

- Lorsque H=K, on a:

$$\hat{a}_{vi} = (Z'X)^{-1}Z'y$$

- Lorsque les X sont exogènes pour être leurs propres instruments (Z=X), on a:

$$\hat{a}_{vi} = (Z'X)^{-1}Z'y = (X'X)^{-1}X'y = \hat{a}$$

3. Propriétés de l'estimateur

- $\hat{a}_{vi}$ est en général biaisé

- $\hat{a}_{vi}$ est convergent

4.Test d'exogénéïté des variables explicatives

- $H_0: E\left(X'_t u_t\right) = 0$

Sous H_0 :

- $\hat{a}_{vi}$ et $\hat{a}_{mco}$ sont convergents vers *a*
- $\hat{a}_{vi} - \hat{a}_{mco} \rightarrow 0$

Le test d'Hausman est basé sur la différence entre les deux estimateurs; cette différence ne doit pas être significative sous H_0.

Test d'Hausman

$$H = T(\hat{a}_{vi} - \hat{a}_{mco})'\hat{\Sigma}^{-1}(\hat{a}_{vi} - \hat{a}_{mco}) \rightarrow \chi 2(K)$$

Les logiciels donnent la *p-value*:

- *Si p-value*$<\alpha$: on rejette H_0
- *Si p-value*$> \alpha$: on accepte H_0

Chapitre 2 : Modèles à décalages temporels

Exemple:

$$C_t = a_0 + a_1 Y_t + a_2 C_{t-1} + a_3 Y_{t-1} + e_t$$

Pourquoi des variables retardées?

- ✓ Effets d'habitude ou de mémoire
- ✓ Délais de réaction
- ✓ Contraintes institutionnelles

1. Typologie des modèles à décalages temporels

Modèles Autorégressifs

$C_t = a_0 + a_1 C_{t-1} + e_t$ AR(1)

$C_t = a_0 + a_1 C_{t-1} + \ldots + a_k C_{t-p} + e_t$ AR(p)

Modèles à retards échelonnés

$C_t = a_0 + b_0 Y_t + b_1 Y_{t-1} + \ldots + b_p Y_{t-p} + e_t$

Modèles Autorégressifs à retards échelonnés ARDL (p, q)

$C_t = a_0 + a_1 C_{t-1} + \ldots + a_p C_{t-p} + b_0 Y_t + b_1 Y_{t-1} + \ldots + b_q Y_{t-q} + e_t$

2. Détermination du nombre de retards

- Test de significativité du dernier retard

- Critères d'information
  - Akaike (AIC)
  - Schwarz (SC)
  - Hannan-Quinn

➔ **Travaux Pratiques**

Chapitre 3: Modèles à équations simultanées

1. Exemple de spécification

$Log(C_t)=a_0+a_1Log(Y_t)+a_2Log(C_{t-1})+e_{1t}$

$Log(I_t)=b_0+b_1Log(Y_t)+b_2Log(I_{t-1})+b_3Log(G_t)+e_{2t}$

$Y_t=C_t+I_t+G_t$

C_t : Consommation privée;

Y_t : PIB;

I_t : Investissement privé

G_t : Dépenses publiques;

2. Statut des variables et problème d'endogénéïté

- Variables endogènes

- C_{t}, I_{t}, Y_{t}

- Variables exogènes et prédéterminées

- C_{t-1}, I_{t-1}, G_{t}

Y_t est une variable explication de C_t mais est endogène dans l'équation 3.

3. Conditions d'identification

Nombre de variables exclues de l'équation doit être supérieur ou égal au nombre d'équations moins 1.

4. Méthodes d'estimation

- Doubles moindres carrés

→ *Variables instrumentales*

- Maximum de vraisemblance

- Estimateur SUR

- Triples moindres carrés

→ *Variables instrumentales*

→ *Moindres carrés généralisés*

Chapitre 4: Modèles VAR

1. Série stationnaire?

- $E(X_t)=m$
- $V(X_t)=\varrho^2$
- $E(X_{t+h}-m)(X_t-m)=\gamma(h)$

X stationnaire est intégrée d'ordre 0 ou I(0).

Si $\Delta X_t = X_t - X_{t-1}$ est stationnaire, on dit que X est intégrée d'ordre 1 ou I(1).

2.Tests de stationnarité

Tests classiques

- Augmented Dickey-Fuller (ADF)
- Phillip-Perron (PP)
- KPSS

Si rupture structurelle:

- Test de Perron (1989)
- Test de Zivot et Andrews (1992)
- Lumsdaine et Papell (1997)
- Bai et Perron (1998)

3. Spécification d'un modèle VAR

X et Y 2 variables stationnaires liées par les relations:

$$Y_t=a_0+a_1X_{t-1}+\ldots+a_pX_{t-p}+b_0+b_1Y_{t-1}+\ldots+b_pY_{t-p}+e_{1t}$$

$$X_t=c_0+c_1X_{t-1}+\ldots+c_pX_{t-p}+d_0+d_1Y_{t-1}+\ldots+d_pY_{t-p}+e_{2t}$$

$\rightarrow$ VAR(p) bivarié

4. Détermination du nombre de retards

➔ Critères d'information

- Akaike (AIC)
- Schwarz (SC)
- Hannan-Quinn

5. Tests de causalité de Granger

X ne cause pas Y $\Leftrightarrow a_1=a_2\ldots=a_p=0$

Y ne cause pas X $\Leftrightarrow d_1=d_2\ldots=d_p=0$

→ Test de Fisher ou de Wald

Lorsque les variables ne sont pas stationnaires, on fait appel aux méthodes de cointégration

Chapitre 5: Cointégration et modèle à correction d'erreur

1. Variables cointégrées, c'est quoi?

X et Y sont dites cointégrées ssi:

- X et Y sont I(1)
- Il existe α et β tel que $e_t = Y_t\alpha + X_t\beta$ est I(0)

2. Tests de cointégration

2.1 Méthode de Engle et Granger (1987)

- On estime par MCO: $Y_t = X_t \beta + e_t$
- On teste la stationarité des résidus e_t
- Si e_t est I(0) alors X et Y sont dites cointégrées

2.2 Méthode de Johansen(1988)

VAR(k): $\boldsymbol{X_t = \Pi_1 X_{t-1} + \cdots + \Pi_k X_{t-k} + \varepsilon_t}$

qui peut s'écrire sous la forme VECM(k-1):

$$\boldsymbol{\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i X_{t-i} + \varepsilon_t}$$

→ Utilise la méthode du maximum de vraisemblance

→ Teste le rang de la matrice autorégressive Π

rang(Π)=nombre de relations de cointégration

Statistiques de test

- Statistique de la trace
- Statistique de la valeur propre maximale

→ Tests sensibles au nombre de retards *k*, au nombre de variables et à la présence de termes déterministes (constante, trend, dummy)

2.3 Bounds test de Pesaran et al. (2001)

Dans Engle-Granger et Johansen: variables intégrées du même ordre (I(1)

Johansen: même nombre de retards pour toutes les variables

L'approche de Pesaran *et al*. (2001):

- Variable dépendante I(1)
- Variables explicatives I(0) et I(1)
- Nombre de retards différents sur les variables
- Robuste au problème d'autocorrélation et d'endogénéïté
- Application sur échantillons de taille réduite

Stratégie du bounds test

- Estimer par MCO

$$\Delta Y_t = \gamma + \sum_{i=1}^{p} m_{1i} \Delta Y_{t-i} + \sum_{i=1}^{p} m_{2i} \Delta X_{t-i} + \rho_1 Y_{t-1} + \rho_2 X_{t-1} + \mu_t$$

- Tester la restriction

$$H_0: \ \rho_1 = \rho_2 = 0$$

→ **Statistique de Fisher**

- Les valeurs critiques sont tabulées par Pesaran *et al.* (2001)

I(0): borne inférieure

I(1): borne supérieure

- Si *F* > I(1): on rejette H0, il y a possibilité de cointégration

- Si F<I(0): on accepte H0, il n'y a pas cointégration

- Si H_0: $\rho_1 = \rho_2 = 0$ est rejetée, il convient de tester la significativité de ρ_1

Si $\rho_1 = 0$: on ne peut conclure à l'existence d'une relation de long terme entre les variables.

si $\rho_1 \neq 0$: il y a relation de cointégration entre les variables, qui est dégénérée sous $\rho_2 = 0$.

3. Modèle à correction d'erreur

Si X et Y sont cointégrées alors elles admettent une forme à correction d'erreur :

$$\Delta Y_t = a_0 + \lambda_1 e_{t-1} + a_1 \Delta X_{t-1} + \ldots + a_{p-1} \Delta X_{t-p+1} + b_1 \Delta Y_{t-1} + \ldots + b_{p-1} \Delta Y_{t-p+1} + u_{1t}$$

$$\Delta X_t = c_0 + \lambda_2 e_{t-1} + c_1 \Delta X_{t-1} + \ldots + c_{p-1} \Delta X_{t-p+1} + d_1 \Delta Y_{t-1} + \ldots + d_{p-1} \Delta Y_{t-p+1} + u_{2t}$$

avec :

$e_{t-1} = Y_{t-1} - X_{t-1}\beta$: écart à l'équilibre

λ_1 et λ_2 : forces de rappel vers la relation de long terme:

λ_1 ou $\lambda_2 < 0$

4. Tests de causalité de Granger

Causalité de long terme

X ne cause pas Y $\Leftrightarrow \lambda_1 = 0$

Y ne cause pas X $\Leftrightarrow \lambda_2 = 0$

Causalité de court terme

X ne cause pas Y $\Leftrightarrow a_1 = a_2 \ldots = a_{p-1} = 0$

Y ne cause pas X $\Leftrightarrow d_1 = d_2 \ldots = d_{p-1} = 0$

Merci de votre aimable attention

Rendez-vous pour les Travaux Pratiques