



La **GREAM** formation

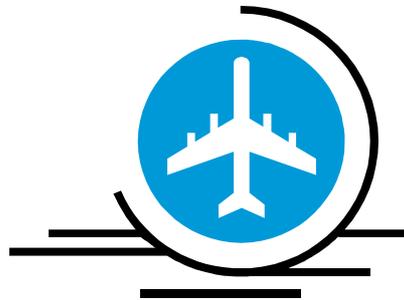
Groupe de recherche en économie appliquée et théorique

N° 004

"Réfléchir à changer"

Avril 2015

**Problèmes économétriques
et modélisation**



Massa Coulibaly

Table des matières

Introduction	1
1. Multicolinéarité	1
1.1. Comment savoir qu'il y a muticolinéarité ?.....	1
1.2. Comment y remédier ?.....	2
2. Dummy Variables.....	3
3. Hétéroscédasticité.....	5
3.1. <i>Comment déceler l'hétéroscédasticité ?</i>	5
3.1.1. Test de Bartlett.....	5
3.2.1. Test de Goldfeld-Qandt	6
3.2. <i>Comment y remédier ?</i>	7
4. Autocorrélation des erreurs.....	8
4.1. <i>Comment détecter l'autocorrélation ?</i>	9
4.2. Comment y remédier ?.....	10
4.2.1. Première procédure (OLS) des quasi - différences	12
4.2.2. Procédure itérative Cockrane-Orcutt	13
4.2.3. Procédure de Durbin	14
4.2.4. Procédure de Stone	14
4.2.5. Procédure (de balayage) Hildreth-Lu.....	14
4.3. La prévision dans les modèles à erreurs autocorrélées	15
5. Variables retardées	16
5.1. Pourquoi les modèles à variables retardées ?.....	16
5.1.1. Variables exogènes retardées.....	16
5.1.2. Variables endogènes retardées.....	17
5.1.3. Modèles autorégressifs à erreurs liées	17
5.2. Comment y remédier ?.....	19
5.2.1. Introduction du retard de Koyck	19
5.2.2. Retard (polynomial) d'Almon	19
6. Equations simultanées	20
6.1. Modèle.....	21
6.2. Identification.....	23
6.3. Méthodes d'estimation.....	24
6.3.1. Régression indirecte.....	24
6.3.2. Double moindres carrés (Theil)	25
6.3.3. Triple moindres carrés (Zellner)	25
6.3.4. FIML (Rubin)	26
6.4. Simulation.....	27
7. Introduction à la Modélisation économique	29
7.1. Modèles macro-économétriques	30
7.2. Les modèles VAR.....	31
7.3. Les modèles calculables d'équilibre général	32
7.4. Les modèles de cycles réels	34

Introduction

Les moindres carrés ordinaires (OLS) reposent sur un ensemble d'hypothèses. Le modèle de régression lui-même est une simplification de la réalité qui ne peut prendre en compte toute la richesse de cette réalité.

Dans la pratique de la modélisation (même seulement linéaire), il arrive que toutes les bonnes conditions (hypothèses, spécification de modèles, unicité de relation fonctionnelle) ne soient pas toujours remplies. Par rapport au cas "idéal" antérieur, lequel cas garantit, en général, aux estimateurs des propriétés optimales, il y a problème chaque fois qu'une ou plusieurs de ces conditions ne sont pas remplies. Nous en retiendrons ici les principaux cas de figures.

1. Multicolinéarité

Il y a multicolinéarité lorsque certaines ou toutes les variables explicatives sont colinéaires ou dépendantes entre elles. Il faut alors croire qu'il y a mauvaise spécification du modèle.

1.1. Comment savoir qu'il y a multicolinéarité ?

Dans les modèles à 3 variables, le coefficient de corrélation entre les deux variables explicatives (R_{23}) permet de tester la présence de colinéarité. Il en est ainsi même dans les modèles à plus de 2 variables explicatives lorsqu'on considère ces variables deux à deux pour ensuite, sur la base du test, éliminer, lorsque la corrélation est significative, celles qui sont les moins corrélées avec la variable dépendante.

Toujours dans ces modèles à plus de 2 variables explicatives, si on veut tester la multicolinéarité, on peut se servir du coefficient de détermination multiple R_i^2 ($i = 2, 3, \dots, k$) entre X_i et les $(k-1)$ variables de la matrice X, le test est donné par :

$$F_i = \frac{R_i^2}{1 - R_i^2} \frac{n - (k - 1)}{(k - 1) - 1}$$

De façon générale, dans le modèle linéaire général, il arrivera que le coefficient de détermination R^2 soit élevé alors que les régresseurs ne sont pas

individuellement significatifs. Justement parce que, avec la multicollinéarité, les estimateurs ont de larges erreurs types et donc des t (Student) faibles, e. g. :

$$Y = 20 + 0.6X_2 + 0.3X_3 \quad R_{1,23}^2 = 0.99$$

SE	(22)	(6)	(6)
t	(0.9)	(0.1)	(0.05)

1.2. Comment y remédier ?

1. Rejeter certaines variables
2. Acquérir des données complémentaires de manière à réduire la variance, ce qui n'est pas toujours possible vu le peu de disponibilité des statistiques
3. Respécifier la relation fonctionnelle sur la base d'informations supplémentaires e.g. en imposant certaines restrictions ou en ajoutant des variables muettes.

Ex.26 Soit le modèle: $Q = A L^\alpha K^\beta e^u$

Sous l'hypothèse des rendements d'échelle constants
i.e. $\alpha + \beta = 1 \Rightarrow \alpha = 1 - \beta$ d'où :

$$Q = AL^{1-\beta} K^\beta e^u \quad \Rightarrow \quad Y = \frac{Q}{L} = A \left(\frac{K}{L} \right)^\beta e^u = Ak^\beta e^u$$

où Y désigne la productivité moyenne du travail
k le ratio capital / main d'œuvre
 β l'élasticité - capital de la production.

Dans le modèle $Y = Ak^\beta e^u$, il ne peut plus y avoir le problème de colinéarité entre le capital et le travail, problème autrement fréquent.

2. Dummy Variables

Jusqu'ici, nous avons supposé que les variables contenues dans la matrice X sont des grandeurs cardinales données par la théorie économique. Nous pouvons à présent y inclure des variables qualitatives.

Pourquoi de telles variables ?

Pour représenter, soit :

des effets temporels (saison, événement naturel, etc.)

des effets spatiaux (région, milieu géographique, etc.)

des effets de variables qualitatives (sexe, statut social, etc.)

des effets de variables institutionnelles (réformes économiques, accords, etc.)

etc.

Ex.27 Soit le modèle de la consommation (C_t) en fonction du revenu (y_t) selon que l'on soit en période de paix (D_1) ou de guerre (D_2) :

$$(1a) \quad C_t = \alpha_1 + \beta Y_t + u_t$$

$$(1b) \quad C_t = \alpha_{21} + \beta Y_t + u_t \quad \text{avec } \alpha_2 > \alpha_1$$

$$(1) \quad C_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \beta Y_t + u_t$$

Où

$$D_1 = \begin{cases} 1 & \text{si paix} \\ 0 & \text{sin on} \end{cases} \quad D_2 = \begin{cases} 1 & \text{si guerre} \\ 0 & \text{sin on} \end{cases}$$

Si le programme de régression de (1) produit un terme constant (comme c'est automatiquement le cas dans la plupart des programmes informatiques), la procédure d'estimation ne marchera pas du fait de $(X'X)$ qui est ici une matrice singulière. C'est là le piège des dummy variables. Pour y remédier, on prend une seule dummy (dv) e.g.:

$$D = D_1 = \begin{cases} 1 & \text{si paix} \\ 0 & \text{si guerre} \end{cases}$$

Le modèle (1) devient : $C_t = \gamma_1 + \gamma_2 D_t + \beta Y_t + u_t$ (2)

Au regard des modèles (1) et (2) on a

	Modèle (1)	Modèle (2)	d'où
$E(C_t/D_t = 0)$	$= \alpha_2 + \beta Y_t$	$= \gamma_1 + \beta Y_t$	$\alpha_2 = \gamma_1$
$E(C_t/D_t = 1)$	$= \alpha_1 + \beta Y_t$	$= (\gamma_1 + \gamma_2) + \beta Y_t$	$\alpha_1 = \gamma_1 + \gamma_2$

Le test de $\gamma_1 (= \alpha_2)$ est de savoir si l'effet "guerre" est significativement différent de zéro.

Le test de $\gamma_2 (= \alpha_1 - \gamma_1 = \alpha_1 - \alpha_2)$ est de savoir s'il existe une différence d'effet significative de consommation selon qu'il s'agit d'une période de paix ou de guerre.

Maintenant si on veut tester si l'effet de la paix est différent de zéro, l'hypothèse est : $H_0 : \alpha_1 = \gamma_1 + \gamma_2 = 0$

d'où :

$$t = \frac{\hat{\gamma}_1 + \hat{\gamma}_2}{SE(\hat{\gamma}_1 + \hat{\gamma}_2)} = \frac{\hat{\gamma}_1 + \hat{\gamma}_2}{\sqrt{V(\hat{\gamma}_1) + V(\hat{\gamma}_2) + 2COV(\hat{\gamma}_1, \hat{\gamma}_2)}}$$

où les éléments de $SE(\cdot)$ sont fournis par la matrice de variances - covariances des estimateurs.

Enfin, si on veut tester la différence éventuelle de pentes (consommations marginales) entre deux périodes, on peut adopter la spécification :

$$(3) \quad \begin{aligned} C_t &= \gamma_1 + \gamma_2 D_t + \gamma_3 (D_t Y_t) + \beta Y_t + u_t \\ &= \gamma_1 + \gamma_2 D_t + \gamma_3 X_t + \beta Y_t + u_t \end{aligned}$$

Ainsi si : $D = 0 \Rightarrow C_t = \gamma_1 + \beta Y_t + u_t$

$D = 1 \Rightarrow C_t = (\gamma_1 + \gamma_2) + (\gamma_3 + \beta) Y_t + u_t$

γ_3 reflète la différence de pente, d'où le test : $H_0 : \gamma_3 = 0$

Notes

- 1) L'approche des d.v. permet de désaisonnaliser les données économiques
- 2) Pour apprécier l'impact d'une d.v., il faut au préalable s'assurer qu'il n'y en a pas d'autres e.g. si on veut mesurer l'effet de l'ajustement structurel sur la croissance économique, il ne faut pas que cet effet renferme en même temps des effets de coup d'Etat, de guerre ou de toute autre perturbation qui ne soit composante de l'ajustement structurel.
- 3) Il peut arriver que dans un modèle, les d. v. jouent un rôle majeur dans la détermination de la variabilité expliquée; ce qui signifie que les variables économiques jouent un rôle mineur dans la série ou dans chacune de ses sous - périodes. La conséquence de ceci est que le pouvoir de prédiction du modèle peut ne pas être corrélé au R^2 de l'échantillon et donc nous amener à conclure incorrectement que l'explication économique des variables est satisfaite.
- 4) Le modèle peut contenir 2 ou plusieurs ensembles de d.v. tels que la consommation fonction des saisons, des catégories socioprofessionnelles, du sexe, etc.

3. Hétéroscédasticité

La perturbation peut croître avec l'accroissement de la variable dépendante i.e. $V(u_t)$ variable et non plus constante.

3.1. Comment déceler l'hétéroscédasticité ?

3.1.1. Test de Bartlett

- i). Répartir les n observations en G groupes de N_g observations chacun ($g = 1, 2, 3, \dots, G$)
- ii). Calculer la statistique

$$S = \frac{n \log\left(\sum N_g S_g^2\right) - \sum N_g \log S_g^2}{1 + \frac{1}{3(G-1)}\left(\sum \frac{1}{N_g} - \frac{1}{n}\right)}$$

$$\text{où } S_g^2 = \frac{1}{N_g} \sum (Y_g - \bar{Y}_g)^2$$

$$S \mapsto \chi^2(G-1)$$

H_0 : $V(u_t)$ cste i.e. homoscélasticité.

Si $S > \chi^2_{\text{tabulé}}$ $\Rightarrow \neg H_0$ i.e. il y a hétérosceasticité.

3.2.1. Test de Goldfeld-Quandt

Soit le modèle $Y_t = \beta_1 + \beta_2 X_t + u_t$

i). Ordonner les observations selon les valeurs de X et scinder en 2 groupes (petites valeurs et grandes valeurs) en omettant les d observations du milieu.

ii). Effectuer la régression pour chacun des 2 groupes.

$$\text{Chaque régression aura } \frac{n-d}{2} - 2 = \frac{n-d-4}{2} \text{ ddl}$$

où d est le nombre d'observations omises.

iii). Calculer

$$\frac{RSS_2}{RSS_1} \mapsto F\left(\frac{n-d-4}{2}, \frac{n-d-4}{2}\right)$$

H_0 : homosceasticité

Si $F > F_{\text{tabulé}}$ $\Rightarrow \neg H_0$ i.e. il y a hétérosceasticité.

3.2. Comment y remédier ?

Essentiellement par la méthode des moindres carrés généralisés (GLS).

Soit le modèle : $Y = X\beta + u$

Si $\hat{\beta}$ est l'estimateur de β :

P₁). $\hat{\beta}$ est linéaire en Y

$$\hat{\beta} = M Y = M(X\beta + u) = M X \beta + M u \quad \text{où } M = (m_{ij})_{k,n}$$

P₂). $\hat{\beta}$ est sans biais

$E(\hat{\beta}) = \beta \Leftrightarrow M X \beta + M E(u) = \beta \Rightarrow M X = I_k$ car $E(u) = 0$ par hypothèse.

Il s'ensuit : $\hat{\beta} = \beta + M u$

P₃). $\hat{\beta}$ est efficace

$$\begin{aligned} V(\hat{\beta}) &= E((\hat{\beta} - \beta)(\hat{\beta} - \beta)') \\ &= E(M u (M u)') \\ &= M E(u u') M' = M \Omega_u M' \quad \text{où } \Omega_u = (\sigma_{ij})_{n,n} \end{aligned}$$

Pour minimiser les variances des composantes du vecteur $\hat{\beta}$, il suffit de minimiser $\text{tr}(M \Omega_u M')$ sous la contrainte $M X = I_k$

Soit le lagrangien : $\mathcal{L} = \text{tr}(M \Omega_u M') - \text{tr}(\Lambda' (M X - I_k))$

où $\Lambda = (\lambda_{ij})_{k,k}$ est la matrice des multiplicateurs de Lagrange.

Selon les propriétés de l'opérateur "trace" (linéaire et commutatif) par rapport à la multiplication matricielle :

$$\begin{aligned} \text{tr}(M \Omega_u M') &= \text{tr}(M' M \Omega_u) && \text{- selon le format des matrices} \\ \text{tr}(\Lambda' M X) &= \text{tr}((\Lambda' M X)') = \text{tr}(X' M' \Lambda) \end{aligned}$$

$$\text{tr}(X' M' \Lambda) = \text{tr}(M' \Lambda X') \quad - \text{ selon le format.}$$

$$\text{D'où} \quad \mathfrak{f} = \text{tr}(M' M \Omega_u) - \text{tr}(2(M' \Lambda X')) + \text{tr}(2\Lambda')$$

$$\begin{aligned} \mathfrak{f}_M' &= \text{tr}(d M' \cdot M \Omega_u) + \text{tr}(M' \cdot dM \cdot \Omega_u) - \text{tr}(2d M' \cdot \Lambda X') \\ &= \text{tr}(d M' \cdot M \Omega_u) + \text{tr}(d M' \cdot M \Omega_u) - \text{tr}(2d M' \cdot \Lambda X') \\ &= \text{tr}(d M' (2 M \Omega_u - 2\Lambda X')) = 0 \end{aligned}$$

$\Rightarrow M \Omega_u = \Lambda X'$ qu'on postmultiplie par $\Omega_u^{-1} X$ pour avoir MX dont on sait qu'il est égal à I_k i.e. $MX = I_k$

$$\text{Soit } M \Omega_u (\Omega_u^{-1} X) = \Lambda X' (\Omega_u^{-1} X) \Leftrightarrow MX = \Lambda (X' \Omega_u^{-1} X) = I_k$$

Il s'ensuit alors que :

$$M \Omega_u = \Lambda X' = (X' \Omega_u^{-1} X)^{-1} X' \quad \text{d'où } M = (X' \Omega_u^{-1} X)^{-1} X' \Omega_u^{-1}$$

Soit finalement :

$$\hat{\beta} = MY = (X' \Omega_u^{-1} X)^{-1} (X' \Omega_u^{-1} Y)$$

l'estimateur BLUE des GLS qui comparé à celui des OLS indique qu'on a ici remplacé X' par $X' \Omega_u^{-1}$

On en déduit la matrice des variances - covariances de $\hat{\beta}$:

$$\Omega_{\hat{\beta}} = \sigma_u^2 (X' \Omega_u^{-1} X)^{-1}$$

L'estimation nécessite la connaissance de la matrice Ω_u . Nous y reviendrons avec l'autocorrélation des erreurs.

4. Autocorrélation des erreurs

L'autocorrélation des erreurs peut survenir de:

- la mauvaise spécification du modèle
- l'interpolation des observations statistiques
- l'omission de variables explicatives. Si une variable autocorrélée est

exclue de l'ensemble des variables explicatives du modèle, son influence sera reflétée dans la variable aléatoire u dont les valeurs seront autocorrélées, on parlera de quasi - autocorrélation.

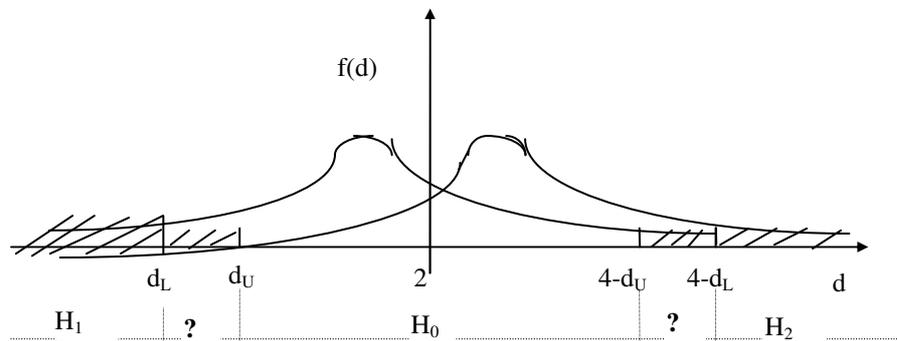
4.1. Comment détecter l'autocorrélation ?

✚ Durbin et Watson (1953) proposent la statistique

$$d = \frac{\sum (e_t - e_{t-1})^2}{\sum_1 e_t^2} \quad \text{où} \quad e_t = Y_t - \hat{Y}_t$$

Ils ont déterminé la densité $f(d)$ et montré que quelle que soit la suite de variables explicatives, les courbes de $f(d)$ oscillent entre 2 limites $f(d_L)$ et $f(d_U)$ symétriques par rapport à un axe d'abscisse 2.

Le test compare le d calculé avec d_L et d_U et leurs transformés $(4-d_L)$ et $(4-d_U)$ pour rechercher les possibilités d'autocorrélation (positive ou négative)



d_L et d_U sont 2 valeurs tabulées, fonction du nombre d'observations (n) et du nombre de variables explicatives (k').

Soit $u_t = \rho u_{t-1} + \varepsilon_t$ la relation qui lie les erreurs du modèle.

Pour de grands échantillons, on s'attend à ce que :

$$\sum_2 e_t^2 = \sum_2 e_{t-1}^2 = \sum_1 e_t^2$$

d'où

$$d = \frac{\sum_1 (e_t - e_{t-1})^2}{\sum_1 e_t^2} = \frac{\sum_1 e_t^2 + \sum_1 e_{t-1}^2 - 2 \sum_1 e_t e_{t-1}}{\sum_1 e_t^2}$$

$$= 1 + 1 - 2 \frac{\sum_1 e_t e_{t-1}}{\sqrt{\sum_1 e_t^2 \sum_1 e_{t-1}^2}} = 2(1 - \rho)$$

Le test proprement dit est

H_0 : absence d'autocorrélation ($\rho = 0$) contre

H_1 : autocorrélation positive ($\rho > 0$)

H_2 : autocorrélation négative ($\rho < 0$).

1) Si $0 < d < d_L$ $\Rightarrow H_1$

2) Si $d > 4 - d_L$ $\Rightarrow H_2$

3) Si $d_U < d < 4 - d_U$ $\Rightarrow H_0$

4) Si $d_L < d < d_U$ ou

$4 - d_U < d < 4 - d_L$ le test n'est pas concluant, il y a doute avec présomption de liaison entre les erreurs.

Ainsi, il n'y a pas d'autocorrélation si d est proche de 2 car $\rho = 0 \rightarrow d = 2$.

Note L'utilisation du test de DW nécessite que :

1. le nombre d'observations (n) soit important
2. le modèle étudié comporte un terme constant.

✚ (PCGIVE) ou le test du multiplicateur de Lagrange (LM)

Soit le modèle $u_t = \alpha_1 u_{t-1} + \dots + \alpha_m u_{t-m}$

$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$

Le LM calculé par PCGIVE $\mapsto \chi^2(m)$

Si $LM > \chi^2_{\text{tabulé}}$ $\Rightarrow \neg H_0$ i.e. il y a autocorrélation.

4.2. Comment y remédier ?

Nous avons vu que $\hat{\beta} = (X' \Omega_u^{-1} X)^{-1} (X' \Omega_u^{-1} Y)$ il faut donc disposer d'une estimation de Ω_u^{-1} pour obtenir l'estimateur $\hat{\beta}$.

Supposons que les erreurs suivent un processus autorégressif de premier ordre (PAPO)

$$u_t = \rho u_{t-1} + \varepsilon_t \text{ avec } |\rho| < 1$$

$$E(\varepsilon_t) = 0 \text{ et } E(\varepsilon_t \varepsilon_t') = \sigma_\varepsilon^2 \quad \forall t$$

Pour obtenir la matrice Ω_u écrivons :

$$\begin{aligned}
u_t &= \rho u_{t-1} + \varepsilon_t \\
&= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \rho^2 u_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \\
&= \rho^2(\rho u_{t-3} + \varepsilon_{t-2}) + \rho \varepsilon_{t-1} + \varepsilon_t \\
&= \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots \\
u_t &= \sum_{k=0}^{\infty} \rho^k \varepsilon_{t-k}
\end{aligned}$$

D'où

$$E(u_t^2) = \sum \rho^{2k} E(\varepsilon_{t-k}^2) = \sigma_\varepsilon^2 \sum \rho^{2k} = \frac{\sigma_\varepsilon^2}{1-\rho^2} = \sigma_u^2$$

$$E(u_t u_{t-1}) = E\left[\sum ((\rho u_{t-1} + \varepsilon_t) u_{t-1})\right] = \rho E(u_{t-1}^2) + E(\varepsilon_t u_{t-1}) = \rho \sigma_u^2$$

$$E(u_t u_{t-2}) = E\left[\sum ((\varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 u_{t-2}) u_{t-2})\right] = \rho^2 \sigma_u^2$$

d'où

$$E(u_t u_{t-\theta}) = \rho^\theta \sigma_u^2 = \rho^\theta \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

Note Le PAPO $u_t = \rho u_{t-1} + \varepsilon_t$ avec $|\rho| < 1$ est un processus stationnaire et donc stable.

Par définition, un processus est stationnaire lorsque tous ses moments sont indépendants du temps. Et lorsque seuls les moments du r^{e} ordre (e.g. $r = 2$) sont indépendants du temps, on dit que le processus est stationnaire à l'ordre r .

$$u_t = \rho u_{t-1} + \varepsilon_t$$

$$E(u_t)_1 = \rho E(u_{t-1}) + E(\varepsilon_t) = \rho E(u_{t-1}) \text{ stationnaire si et seulement si } E(u_t) = E(u_{t-1})$$

$$\text{D'où } (1-\rho)E(u_t) = 0 \Rightarrow E(u_t) = 0 \text{ avec}$$

$$E(u_t^2) = \frac{\sigma_\varepsilon^2}{1-\rho^2} = \sigma_u^2 \text{ variance du processus stationnaire}$$

$$\text{Maintenant de } E(u_t u_{t-\theta}) = \rho^\theta \sigma_u^2 = \rho^\theta \frac{\sigma_\varepsilon^2}{1-\rho^2} \text{ il s'ensuit :}$$

$$\Omega_u = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}$$

L'inverse de cette matrice s'obtient par transformation élémentaire sur Ω_u et I_n placées côte à côte:

$$\text{ligne : } L_i = L_i - \rho L_{i-1}$$

$$\text{colonne : } C_j = C_j - \rho C_{j-1}$$

$$\Omega_u^{-1} = \left(\frac{\sigma_\varepsilon^2}{1-\rho^2} \right)^{-1} \text{trsf}(\cdot) = \frac{1}{\sigma_\varepsilon^2} \begin{pmatrix} 1 & -\rho & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & \dots & 0 & 0 \\ 0 & -\rho & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -\rho & 1 \end{pmatrix}$$

Il existe des méthodes pratiques d'estimation de ρ et donc de détermination de Ω_u^{-1} pour enfin obtenir l'estimateur $\hat{\beta}$ de Aitken des GLS. Nous en retenons cinq :

4.2.1. Première procédure (OLS) des quasi - différences

Soit le modèle (1) : $Y = X\beta + u$

i). Par les OLS, on fait l'ajustement de (1) sans tenir compte de l'autocorrélation des erreurs.

On obtient un estimateur $\hat{\beta}^1$ de $\beta \Rightarrow \hat{Y}^1 = X\hat{\beta}^1$

ii). On estime ρ à l'aide de :

$$\hat{\rho} = \frac{\sum e_t e_{t-1}}{\sum e_{t-1}^2} \quad \text{où } e_t = Y_t - \hat{Y}^1$$

iii). On remplace ρ par $\hat{\rho}$ dans ce qui suit :

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$$

$$Y_{t-1} = \beta_1 X_{1t-1} + \beta_2 X_{2t-1} + \dots + \beta_k X_{kt-1} + u_{t-1}$$

$$\begin{array}{l} Y_t - \rho Y_{t-1} = \beta_1 (X_{1t} - \rho X_{1t-1}) + \dots + \beta_k (X_{kt} - \rho X_{kt-1}) \\ + u_t - \rho u_{t-1} \end{array}$$

$$\text{Soit : } \begin{cases} Y_2 - \rho Y_1 = \beta_1(X_{12} - \rho X_{11}) + \beta_2(X_{22} - \rho X_{21}) + \dots + \beta_k(X_{k2} - \rho X_{k1}) + \varepsilon_2 \\ \vdots \\ Y_n - \rho Y_{n-1} = \beta_1(X_{1n} - \rho X_{1n-1}) + \beta_2(X_{2n} - \rho X_{2n-1}) + \dots + \beta_k(X_{kn} - \rho X_{kn-1}) + \varepsilon_n \end{cases}$$

un système à (n-1) équations qui sous la forme vectorielle s'écrit :

$$MY = (MX)\beta + \varepsilon = (MX)\beta + Mu$$

$$\text{où } M = (m_{ij})_{(n-1)n} = \begin{pmatrix} -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix}$$

$$\Rightarrow \hat{\beta} = ((MX)'(MX))^{-1}(MX)'(MY) = (X'(M'M)X)^{-1}X'(M'M)Y$$

$$\text{avec } M'M = (m_{ij})_{n,n} = \begin{pmatrix} \rho^2 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix}$$

Note (M'M) ne diffère de Ω_u^{-1} que par ρ^2 au lieu de 1. Pour n suffisamment grand, les 2 estimateurs possèdent les mêmes propriétés.

4.2.2. Procédure itérative Cockrane-Orcutt

Soient les modèles :

$$(1): Y_t - \rho Y_{t-1} = \beta_1(X_{1t} - \rho X_{1t-1}) + \dots + \beta_k(X_{kt} - \rho X_{kt-1}) + \varepsilon_t$$

$$(2): Y_t - (\beta_1 X_{1t} + \dots + \beta_k X_{kt}) = \rho(Y_{t-1} - (\beta_1 X_{1t-1} + \dots + \beta_k X_{kt-1})) + \varepsilon_t$$

On régresse alternativement (1) et (2) en adoptant la procédure suivante :

i). Soit $\rho^{(0)}$ e.g. $\rho^{(0)} = 0$ une valeur a priori de ρ .

La régression sur (1) donne une première estimation des paramètres

$$\hat{\beta}_1^{(0)}, \hat{\beta}_2^{(0)}, \dots, \hat{\beta}_k^{(0)}$$

ii). On reporte dans (2) les $\hat{\beta}_j^{(0)}$ et on régresse pour déduire une nouvelle valeur $\rho(1)$ de ρ

iii). On reporte $\rho(1)$ dans (1) et on régresse à nouveau, pour trouver

$$\hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)}, \dots, \hat{\beta}_k^{(1)}$$

iv). Ainsi de suite.

On arrête la procédure lorsqu'il y a convergence des estimateurs $\hat{\beta}^{(i)}$ ou tout simplement lorsque ρ converge i.e. lorsque les valeurs trouvées après deux itérations successives ne diffèrent que d'un nombre donné aussi petit qu'on le veut e.g. 0.001 de différence seulement.

4.2.3. Procédure de Durbin

Soit

$$(1) : Y_t = \rho Y_{t-1} + \beta' X_t - \rho \beta' X_{t-1} + \varepsilon_t$$

$$\text{où } X_t = \begin{pmatrix} 1 \\ X_{2t} \\ \vdots \\ X_{kt} \end{pmatrix} \quad \text{et } \beta' = (\beta_1, \beta_2, \dots, \beta_k)$$

i). On estime ρ sur (1) sans tenir compte des restrictions sur les coefficients de Y_{t-1}, X_t, X_{t-1}

ii). On reporte la valeur estimée $\hat{\rho}$ de ρ dans

$$Y_t - \rho Y_{t-1} = \beta_1(X_{1t} - \rho X_{1t-1}) + \dots + \beta_k(X_{kt} - \rho X_{kt-1}) + \varepsilon_t$$

qu'on régresse pour obtenir l'estimateur $\hat{\beta}$ correspondant.

4.2.4. Procédure de Stone

Beaucoup moins précise que les autres, cette procédure consiste à régresser $Y_t - \rho Y_{t-1}$ sur $(X_{1t} - \rho X_{1t-1}), \dots, (X_{kt} - \rho X_{kt-1})$ pour deux valeurs extrêmes de ρ : $\rho = 0$ et $\rho = 1$

On obtient 2 estimateurs $\hat{\beta}^{(0)}$ et $\hat{\beta}^{(1)}$ qui encadrent l'estimateur recherché

$$\hat{\beta}^{(0)} < \hat{\beta} < \hat{\beta}^{(1)}$$

Pour s'appliquer, cette procédure exige que $\hat{\beta}^{(0)}$ et $\hat{\beta}^{(1)}$ soient proches.

4.2.5. Procédure (de balayage) Hildreth-Lu

Soit $\rho > 0$ tel que $\rho = \{0, 0.01, 0.02, \dots, 1\}$

i). On fait la régression pour chaque valeur de ρ

ii). On retient comme valeur de ρ celle qui minimise $RSS = \sum e_i^2$ et on en déduit les valeurs correspondantes $\hat{\beta}_j$

4.3. La prévision dans les modèles à erreurs autocorrélées

Soit le modèle $Y = X\beta + u$ avec $E(u) = 0$ et $E(uu') = \Omega_u$

Pour la période $(n+1)$, on donne le vecteur X_{n+1} des valeurs futures des variables explicatives et on se propose d'estimer Y_{n+1} noté $Y_0 = X_{n+1}\beta + u_{n+1}$

On suppose que

$$E(u_{n+1}) = 0, \quad E(u_{n+1}^2) = \sigma_{n+1}^2 \quad \text{et} \quad E(u_{n+1}u) = \begin{pmatrix} E(u_1 u_{n+1}) \\ E(u_2 u_{n+1}) \\ \vdots \\ E(u_n u_{n+1}) \end{pmatrix} = W$$

Evidemment $X_{n+1}\hat{\beta}$ n'est plus la meilleure prévision à cause de l'autocorrélation des erreurs qui oblige à prévoir aussi u_{n+1}

$$\hat{Y}_{n+1} = \hat{Y}_0 = X_{n+1}\hat{\beta} + \hat{u}_{n+1}$$

Définissons un prédicteur linéaire : $P = F'Y$

Si P doit être BLUP, F doit être choisi tel que la variance de prédiction soit minime:

$$\begin{cases} \min & \sigma_p^2 = E(P - Y_0)^2 \quad \text{sachant que} \\ & E(P) = Y_0 \end{cases}$$

$$\begin{aligned} Y_0 &= X_{n+1}\beta + u_{n+1} \\ P &= F'X\beta + F'u \end{aligned}$$

$$\begin{aligned} P - Y_0 &= (F'X - X_{n+1})\beta + F'u - u_{n+1} \\ E(P - Y_0) &= 0 \Rightarrow F'X - X_{n+1} = 0 \end{aligned}$$

d'où $P - Y_0 = F'u - u_{n+1}$ erreur de prévision

$$\begin{aligned} \sigma_p^2 &= E((P - Y_0)^2) \\ &= E((P - Y_0)(P - Y_0')) \\ &= E((F'u - u_{n+1})(F'u - u_{n+1}')) \end{aligned}$$

$$= E(F'uu'F - u_{n+1}u'F - F'uu_{n+1}' + u_{n+1}u_{n+1}') \\ \sigma_p^2 = F'\Omega_u F + \sigma_{u_{n+1}}^2 - 2F'W$$

$$\begin{cases} \min \sigma_p^2 \text{ sous la contrainte} \\ F'X - X_{n+1} = 0 \end{cases}$$

Soit $L = F'\Omega_u F + \sigma_{n+1}^2 - 2F'W - 2(F'X - X_{n+1})\Lambda$

$$\begin{cases} L_F = 0 \\ L_\Lambda = 0 \end{cases} \Leftrightarrow \begin{pmatrix} \Omega_u & X \\ X' & 0 \end{pmatrix} \begin{pmatrix} F \\ -\Lambda \end{pmatrix} = \begin{pmatrix} W \\ X_{n+1} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \hat{F} \\ -\hat{\Lambda} \end{pmatrix} = \begin{pmatrix} \Omega_u & X \\ X' & 0 \end{pmatrix}^{-1} \begin{pmatrix} W \\ X_{n+1} \end{pmatrix}$$

si les résidus suivent un PAPO, alors :

$$W = \rho \sigma_u^2 \begin{pmatrix} \rho^{n-1} \\ \rho^{n-2} \\ \vdots \\ 1 \end{pmatrix}$$

d'où

$$\hat{P} = X_{n+1} \hat{\beta} + \rho e^n$$

où e^n est le résidu de la n^{e} observation dans le modèle final.

5. Variables retardées

5.1. Pourquoi les modèles à variables retardées ?

On peut y trouver des justifications économiques. Les modèles autorégressifs sont des modèles qui incluent des variables décalées des exogènes et/ou des variables dépendantes dans l'ensemble des variables explicatives.

5.1.1. Variables exogènes retardées

On suppose que l'effet d'une variable est distribué sur plusieurs périodes de temps :

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_S X_{t-S} + u_t$$

Les OLS donnent des BLUE si le modèle est correctement spécifié. Dans la pratique, plusieurs difficultés surgissent :

- i). rien (ni même la théorie économique) ne permet de préciser la longueur du décalage (retard). On la détermine en examinant la signification des coefficients de plusieurs valeurs échelonnées de X.

Mais cela soulève plusieurs difficultés statistiques, entre autres, les multiples valeurs de X sont fortement corrélées, ce qui conduit à de très imprécises estimations des coefficients de décalage et à des difficultés à effectuer des inférences sur ces coefficients

- ii). plus le décalage est élevé, plus faible sera le nombre d'observations. Ainsi si l'échantillon est petit, nous ne pourrons estimer les paramètres faute de ddl adéquats pour permettre les tests de signification habituels
- iii). le problème de multicollinéarité, compte tenu de la forte corrélation qui peut exister entre valeurs successives de la même variable. Cette forte colinéarité fera que les estimateurs seront imprécis et leurs erreurs types grands d'où mauvaise spécification du modèle.

Pour contourner ces difficultés, plusieurs méthodes ont été suggérées visant toutes à réduire le nombre de variables décalées. Cela se fait en imposant des restrictions sur les β et en construisant de nouvelles variables (W_j) à partir d'une combinaison de variables décalées :

$$W_j = \sum_{i=0}^s i^j X_{t-i} \quad j = 1, 2, \dots, r$$

5.1.2. Variables endogènes retardées

Soit le processus autorégressif :

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_S Y_{t-S} + u_t$$

ou plus généralement le modèle :

$$Y_t = \beta_1 Y_{t-1} + \dots + \beta_S Y_{t-S} + b_1 X_{1t} + b_2 X_{2t} + \dots + b_k X_{kt} + u_t$$

En appliquant les méthodes habituelles d'estimation, les estimateurs obtenus n'ont pas les mêmes propriétés statistiques que dans le cas du modèle linéaire général.

5.1.3. Modèles autorégressifs à erreurs liées

Considérons 2 modèles :

$$(1) \begin{cases} Y_t = \beta Y_{t-1} + u_t \\ u_t = \rho u_{t-1} + \varepsilon_t \end{cases}$$

$$Y_t - \rho Y_{t-1} = \beta(Y_{t-1} - \rho Y_{t-2}) + \varepsilon_t$$

Soit
$$Y_t = (\beta + \rho)Y_{t-1} - \beta\rho Y_{t-2} + \varepsilon_t \quad (1')$$

Les OLS sur (1) donnent :
$$\hat{\beta} = \frac{\sum Y_t Y_{t-1}}{\sum Y_{t-1}^2}$$

Or de (1') on a :

$$\sum Y_t Y_{t-1} = (\beta + \rho) \sum Y_{t-1}^2 - \beta\rho \sum Y_{t-1} Y_{t-2} + \sum \varepsilon_t Y_{t-1}$$

$$\Rightarrow \hat{\beta} = \frac{\sum Y_t Y_{t-1}}{\sum Y_{t-1}^2} = \beta + \rho - \beta\rho \frac{\sum Y_{t-1} Y_{t-2}}{\sum Y_{t-1}^2} + \frac{\sum \varepsilon_t Y_{t-1}}{\sum Y_{t-1}^2}$$

$$\lim_{t \rightarrow \infty} \hat{\beta} = \beta + \rho - \beta\rho \lim_{t \rightarrow \infty} \hat{\beta}$$

d'où
$$\lim_{t \rightarrow \infty} \hat{\beta} = \frac{\beta + \rho}{1 + \beta\rho}$$

lorsque ρ varie entre -1 et 1, il en est de même pour $\lim \hat{\beta}$ donc $\hat{\beta}$ comporte un biais.

$$(2) \quad \begin{cases} Y_t = \beta Y_{t-1} + bX_t + \alpha + u_t \\ u_t = \rho u_{t-1} + \varepsilon_t \end{cases}$$

$$Y_t - \rho Y_{t-1} = \beta(Y_{t-1} - \rho Y_{t-2}) + b(X_t - \rho X_{t-1}) + \alpha(1-\rho) + \varepsilon_t$$

$$Y_t = (\beta + \rho)Y_{t-1} - \beta\rho Y_{t-2} + bX_t - \rho bX_{t-1} + \alpha(1-\rho) + \varepsilon_t$$

Soit

$$Y_t = a_1 Y_{t-1} - a_2 Y_{t-2} + bX_t + b_1 X_{t-1} + d + \varepsilon_t$$

sans s'occuper des contraintes qui lient les a_j et b_j on peut estimer, par les OLS, ρ :

$$\hat{\rho} = -\frac{\hat{b}_1}{\hat{b}}$$

En reportant $\hat{\rho}$ dans (2), on obtient les estimateurs $\hat{\beta}$, \hat{b} , $\hat{\alpha}$.

5.2. Comment y remédier ?

5.2.1. Introduction du retard de Koyck

On suppose $\beta_j = \gamma^j \beta_0$ $j = 1, 2, \dots, k$ et $0 < \gamma < 1$

Le modèle $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_k X_{t-k} + u_t$ devient :

$$Y_t = \alpha + \beta_0 X_t + (\gamma \beta_0) X_{t-1} + \dots + (\gamma^k \beta_0) X_{t-k} + u_t$$

$$\rightarrow \gamma Y_{t-1} = \gamma \alpha + \gamma \beta_0 X_{t-1} + \gamma^2 \beta_0 X_{t-2} + \dots + \gamma^k \beta_0 X_{t-k-1} + \gamma u_{t-1}$$

$$Y_t - \gamma Y_{t-1} = \alpha(1-\gamma) + \beta_0 X_t + \dots + u_t - \gamma u_{t-1}$$

On suppose que k est grand et que donc : $\gamma^k \beta_0 X_{t-k} = \gamma^k \beta_0 X_{t-k-1}$

d'où $Y_t = \alpha(1-\gamma) + \gamma Y_{t-1} + \beta_0 X_t + v_t$ lequel modèle présente malheureusement l'autocorrélation des erreurs, $\text{cov}(v_t, v_{t-1}) = -\gamma \sigma_u^2 \neq 0$ du coup les estimateurs seront biaisés et non consistants.

5.2.2. Retard (polynomial) d'Almon

Soit $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_s X_{t-s} + u_t$

Au lieu d'estimer directement les β_j par les OLS, on les estime indirectement.

Soit $\beta_j = f(j)$ une fonction polynomiale en j de la forme :

$$f(j) = a_0 + a_1 j + \dots + a_j^r$$

Les étapes d'estimations des β_j sont les suivantes :

i). spécifier le degré r (supposé faible, $r = 2, \dots$) du polynôme et le nombre S de décalage

ii). donner à j les valeurs successives entières: $0, 1, 2, \dots, S$

Soit donc le système (S)
$$\begin{cases} \beta_0 = f(0) = a_0 \\ \beta_1 = f(1) = a_0 + a_1 + a_2 + \dots + a_r \\ \beta_2 = f(2) = a_0 + 2a_1 + 2^2 a_2 + \dots + 2^r a_r \\ \vdots \\ \beta_s = f(S) = a_0 + S a_1 + S^2 a_2 + \dots + S^r a_r \end{cases}$$

Soit
$$\beta_j = \sum_{i=0}^r j^i a_i \quad j = 0, 1, 2, \dots, S$$

iii). estimer les a_i par les OLS sur le modèle transformé :

$$Y_t = a_0 W_0 + a_1 W_1 + a_2 W_2 + \dots + a_r W_r + u_t$$

où les W_i sont des combinaisons linéaires des X tels que

$$\begin{aligned} W_0 &= X_t + X_{t-1} + X_{t-2} + \dots + X_{t-S} \\ W_1 &= X_{t-1} + 2 X_{t-2} + \dots + S X_{t-S} \\ W_2 &= X_{t-1} + 2^2 X_{t-2} + \dots + S^2 X_{t-S} \\ &\vdots \\ W_r &= X_{t-1} + 2^r X_{t-2} + \dots + S^r X_{t-S} \end{aligned}$$

$$\text{Soit } W_i = \sum_{j=0}^S j^i X_{t-j} \quad i = 0, 1, 2, \dots, r$$

iv). substituer les LSE des a_i obtenus dans le système (S) pour trouver les estimateurs

$$\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_S$$

6. Equations simultanées

Jusqu'ici nous n'avons considéré que des modèles à une seule équation dans laquelle plusieurs variables prédéterminées expliquent une variable endogène. En économie, il est plus réaliste de construire des modèles comportant plusieurs équations en même temps.

Ex.27 Soit le modèle keynésien simple qui met en relation consommation (C), revenu (Y) et investissement (I) :

$$(S) \begin{cases} (1) C = \alpha + \beta Y_t + u_t & \text{equation de comportement} \\ (2) Y_t = C_t + I_t & \text{equation comptable (ou d'identité)} \end{cases}$$

On ne peut pas estimer les paramètres de (1) sans tenir compte de (2).

L'estimation se fait à partir des deux équations qui forment le système d'équations simultanées. Sous cette forme, le modèle est dit être sous forme structurelle ; ses équations n'expriment alors que les hypothèses de la théorie économique.

En combinant (1) et (2) on obtient :

$$\begin{cases} C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + \frac{1}{1-\beta} u_t \\ Y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + \frac{1}{1-\beta} u_t \end{cases}$$

Avec ces deux équations, le modèle est dit écrit sous forme réduite :

$$(R) \quad \begin{cases} C_t = a_1 + b_1 I_t + \varepsilon_t \\ Y_t = a_2 + b_2 I_t + \varepsilon_t \end{cases}$$

Le modèle (R) comporte :

- 2 variables endogènes : Y, C
- 1 variable exogène : I

Il permet l'application des OLS sur chaque équation prise séparément. Les estimateurs qu'on en obtient sont généralement BLUE. Malheureusement ils n'ont pas toujours de signification économique.

Ex.28 Soit un 2^e modèle avec C (consommation des ménages), G (dépenses publiques), r (taux d'intérêt réel),

$$(S) \quad \begin{cases} C_t = \beta_0 + \beta_1 G_t + \beta_2 r_t + \varepsilon_t \\ I_t = \alpha_0 + \alpha_1 G_t + \alpha_2 r_t + \varepsilon_t \\ Y_t = C_t + I_t + G_t \end{cases}$$

On en déduit

$$(R) \quad \begin{cases} C_t = b_0 + b_1 G_t + b_2 r_t + \varepsilon_t \\ I_t = a_0 + a_1 G_t + a_2 r_t + \varepsilon_t \\ Y_t = \gamma_0 + \gamma_1 G_t + \gamma_2 r_t + \varepsilon_t \end{cases}$$

modèle à: 3 variables endogènes: C, I, Y
2 variables exogènes : G, r.

6.1. Modèle

Un modèle à équations simultanées sera dit complet s'il y a autant de variables endogènes qu'il y a d'équations structurelles et s'il peut être résolu. Sa solution est appelée forme réduite qui décrit les variables endogènes en terme de variables exogènes et du terme aléatoire.

Supposons un modèle à :

- ✓ L équations structurelles i.e. L variables conjointement dépendantes
- ✓ k variables prédéterminées i.e. variables exogènes plus variables endogènes retardées
- ✓ N_j variables explicatives par équation structurelle j.

Il y aura donc en tout $(k + L)$ valeurs pour chacune des n observations.

Nous avons :

- ✓ la matrice des variables prédéterminées :

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} \end{pmatrix} = (X_{ij})_{nk}$$

- ✓ la matrice des variables endogènes :

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1L} \\ Y_{21} & Y_{22} & \cdots & Y_{2L} \\ \vdots & \vdots & & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nL} \end{pmatrix} = (Y_{ij})_{nL}$$

- ✓ la variable aléatoire :

$$u = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1L} \\ u_{21} & u_{22} & \cdots & u_{2L} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nL} \end{pmatrix} = (u_{ij})_{nL}$$

Soient les paramètres :

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1L} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2L} \\ \vdots & \vdots & & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{kL} \end{pmatrix} = (\beta_{ij})_{kL}$$

$$\gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1L} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2L} \\ \vdots & \vdots & & \vdots \\ \gamma_{L1} & \gamma_{L2} & \cdots & \gamma_{LL} \end{pmatrix} = (\gamma_{ij})_{LL}$$

γ est une matrice carrée puisqu'il y a autant d'équations qu'il y a de variables conjointement dépendantes

Soit finalement le modèle :

$$Y\gamma + X\beta = u \quad (S)$$

si γ est une matrice non singulière, alors (S) donne :

$$Y = u\gamma^{-1} - X\beta\gamma^{-1} \quad (\text{R})$$

6.2. Identification

Le problème est d'estimer de façon unique les paramètres du modèle structurel à partir des estimateurs du modèle réduit, les premiers ayant une signification économique précise.

Soit l'équation j du système d'équations (S) :

$$Y_j = Z_j\delta_j + \varepsilon_j \quad \text{à } N_j \text{ variables explicatives}$$

On ne peut appliquer les OLS du fait que Z_j et ε_j sont corrélés

$$Y_j = Z_j\delta_j + \varepsilon_j \quad \text{qu'on prémultiplie par } X'$$

$$X' Y_j = (X'Z_j) \delta_j + X'\varepsilon_j$$

(on montre que $(X'Z_j)$ et $(X'\varepsilon_j)$ sont stochastiquement indépendantes)

➔ $\hat{\delta} = d_j = (X'Z_j)^{-1}X' Y_j$ si $X'Z_j$ matrice carrée non singulière (tout le problème étant de savoir si cette condition est satisfaite)

🚩 Condition d'identification

$$(X'Z_j) = (C_{ij})_{kN_j}$$

- i). Une équation structurelle est identifiée si $k = N_j$
- ii). Une équation structurelle est suridentifiée si $k > N_j$
- iii). Une équation structurelle n'est pas identifiée si $k < N_j$

Dans les deux premiers cas i.e. $k \geq N_j$, on peut recourir à la méthode des doubles moindres carrés et le logiciel TSP est bien adapté à de telles estimations.

Par contre, dans le dernier cas, il faut recourir à d'autres spécifications puisqu'alors, on ne peut estimer les paramètres.

Note Dans la pratique, il arrive que bien de coefficients β_{ij} et γ_{ij} soient nuls, toutes les variables n'apparaissant pas dans toutes les équations. Ces coefficients nuls sont appelés restrictions a priori. Et si dans un modèle, il

existe $LL = L^2$ restrictions a priori (condition d'ordre), le problème de l'identification revient à résoudre un système à kL équations, kL inconnues. Le modèle est alors dit "**juste identifiable**".

Ex.29 Soit le modèle :

$$\begin{cases} D_t = a_0 + a_1 p_t + a_2 R_t + \varepsilon_t \\ S_t = b_0 + b_1 p_t + b_2 p_{t-1} + v_t \\ D_t = S_t \end{cases}$$

Soit $k = 4$ variables prédéterminées : p , $p(-1)$, R , E

- L'équation D_t a 3 variables explicatives: $k > N_j \Rightarrow$ équation suridentifiée
- L'équation S_t a 4 variables explicatives: $k = N_j \Rightarrow$ équation identifiée.

6.3. Méthodes d'estimation

Elles sont fonction de l'identification des modèles d'équations simultanées. On en distingue :

- la régression indirecte
- les doubles moindres carrés (2LS)
- les triples moindres carrés (3LS)
- le maximum de vraisemblance à information limitée (FIML).

6.3.1. Régression indirecte

Nous avons vu que le modèle d'équations simultanées pouvait s'écrire:

$$Y = u\gamma^1 - X\beta\gamma^1 \quad \text{qu'on peut encore écrire :}$$

$$Y = XA + \varepsilon$$

Lorsque chaque équation est identifiable, on estime les coefficients de A par les moindres carrés pour en déduire les estimateurs de γ et β

Lorsque des équations sont suridentifiées, la méthode n'est plus applicable. On recourt aux 2LS.

6.3.2. Double moindres carrés (Theil)

La méthode s'applique aux équations suridentifiées ainsi qu'aux équations juste identifiées pour lesquelles les résultats sont les mêmes que ceux obtenus par la régression indirecte. L'application se fait en deux étapes

- i). on régresse les variables endogènes sur les variables prédéterminées du modèle. On obtient alors les valeurs ajustées des endogènes.
- ii). on remplace les variables endogènes par leurs valeurs ajustées (qui ne sont plus stochastiques) et on fait la régression.

6.3.3. Triple moindres carrés (Zellner)

Soient les modèles :

$$(1) : Y\gamma + X\beta = u$$

$$(2) : Y = XA + \varepsilon$$

Les 3LS supposent que toutes les équations de la forme structurelle (1) soient suridentifiées ou au moins identifiables. L'application se fait en trois étapes

- i). on régresse chaque variable endogène du modèle sur toutes les exogènes i.e. on estime la forme réduite (2) pour obtenir une estimation préliminaire des endogènes $Y_1^*, Y_2^*, \dots, Y_L^*$
- ii). on réécrit les équations de la forme structurelle (1) en remplaçant les Y_j prédéterminés par Y_j^* (mais pas des endogènes Y_j des équations structurelles) :

$$Y^*\gamma + X\beta = u \quad (1^*)$$

On applique à (1^{*}) les OLS pour obtenir une première estimation des coefficients de γ et β . Ici s'arrêtent les 2LS.

- iii). on remplace les paramètres estimés par leurs valeurs dans (1) pour estimer les résidus

$$Y\hat{\gamma} + X\hat{\beta} = \hat{u} \quad \text{puis} \quad \hat{\varepsilon} = \hat{u}\hat{\gamma}^{-1}$$

ce qui permet d'estimer la matrice Ω_ε

$$\Omega_\varepsilon = \frac{1}{n} \begin{pmatrix} \sum \hat{\varepsilon}_1^2 & \sum \hat{\varepsilon}_1 \hat{\varepsilon}_2 & \dots & \sum \hat{\varepsilon}_1 \hat{\varepsilon}_L \\ \sum \hat{\varepsilon}_1 \hat{\varepsilon}_2 & \sum \hat{\varepsilon}_2^2 & \dots & \sum \hat{\varepsilon}_2 \hat{\varepsilon}_L \\ \vdots & \vdots & \ddots & \vdots \\ \sum \hat{\varepsilon}_1 \hat{\varepsilon}_L & \sum \hat{\varepsilon}_2 \hat{\varepsilon}_L & \dots & \sum \hat{\varepsilon}_L^2 \end{pmatrix}$$

On applique les GLS au modèle (2) : $Y = XA + \varepsilon$.

Soit $\hat{A} = (X' \Omega_\varepsilon^{-1} X)^{-1} X' \Omega_\varepsilon^{-1} Y$ l'estimateur des 3LS.

6.3.4. FIML (Rubin)

Le FILM s'applique aux équations suridentifiées ou à la rigueur identifiables.

i). Soit l'équation j suridentifiée du modèle (1) à :

- L variables endogènes
- h variables prédéterminées \Rightarrow (k-h) variables prédéterminées absentes dans cette équation.

$$\begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1L} \\ Y_{21} & Y_{22} & \cdots & Y_{2L} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{n1} & Y_{n1} & \cdots & Y_{nL} \end{pmatrix} \begin{pmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{lj} \end{pmatrix} + \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1h} \\ X_{21} & X_{22} & \cdots & X_{2h} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n1} & \cdots & X_{nh} \end{pmatrix} \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \\ \vdots \\ \beta_{hj} \end{pmatrix} = \begin{pmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{nj} \end{pmatrix}$$

$$\Leftrightarrow Y\gamma_j + X_h\beta_j = u_j \quad (1)$$

$$\Rightarrow Z = Y\gamma_j = u_j - X_h\beta_j = u_j + X_h(-\beta_j) = X_h\beta_h + u_j \quad (2)$$

On applique les OLS sur (2) pour obtenir :

$$\hat{\beta}_h = (X_h' X_h)^{-1} X_h' Z$$

$$\begin{aligned} \text{RSS} &= e_j' e_j = (Z - X_h \hat{\beta}_h)' (Z - X_h \hat{\beta}_h) = V_j \\ &= Z' Z - Z' X_h (X_h' X_h)^{-1} X_h' Z \\ &= \beta_j' Y' Y \beta_j - \beta_j' Y' X_h (X_h' X_h)^{-1} X_h' Y \beta_j \end{aligned}$$

$$V_j = \beta_j' (M_H) \beta_j$$

ii). Maintenant si l'on considère le modèle sans tenir compte des variables prédéterminées absentes, on écrira :

$$Z = X \beta_h + u_j \quad (2')$$

$$\rightarrow V = \text{RSS} = \beta_j' (M) \beta_j$$

iii). On fait le rapport des 2 variances résiduelles :

$$\rho = \frac{V_j}{V} = \frac{\beta_j'(M_h)\beta_j}{\beta_j'(M)\beta_j} \quad (3)$$

$\rho \geq 1$ puisque V tient compte de toutes les variables prédéterminées, la variance résiduelle est par conséquent plus faible.

✚ Le FIML consiste précisément à trouver un estimateur $\hat{\beta}_j$ de β_j tel que ρ soit le plus proche possible de l'unité. Ce qui revient à rapprocher V_j de V

$$\begin{aligned} & \min (V_j - \rho V) \\ & \frac{d(V_j - \rho V)}{d\beta_j} = 0 \quad \Leftrightarrow \quad 2\beta_j'(M_h - \rho M) = 0 \end{aligned}$$

On obtient ainsi système d'équations homogènes en ρ qui admet une solution non banale si et seulement si :

$$\det(M_h - \rho M) = 0 \quad (4)$$

On obtient un polynôme en ρ de degré L dont on cherche la racine ρ la plus proche de l'unité.

De $\beta_j'(M_h - \rho_0 M) = 0$ on en déduit $\hat{\beta}_j$ puis:

$$\begin{aligned} \hat{Z} &= -X_h \hat{\beta}_j \quad \text{puis l'estimateur } \hat{\beta}_h \text{ de } \beta_h : \\ \hat{\beta}_h &= (X_h' X_h)^{-1} X_h' \hat{Z} \quad (5) \end{aligned}$$

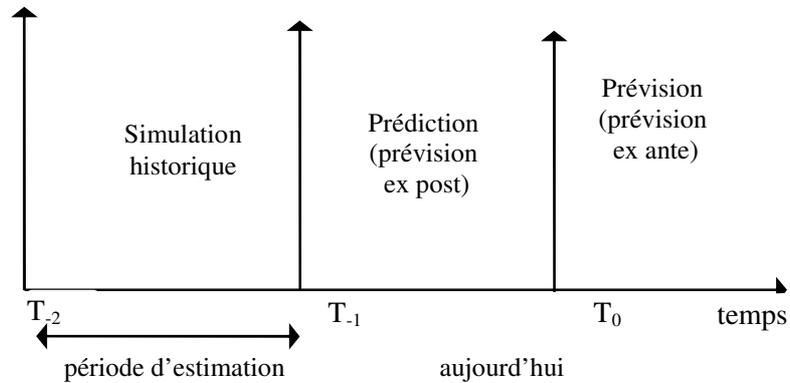
Note Le FIML est aussi dit méthode de minimum de rapport de variances - LVR (least variance ratio) puisque consistant à minimiser ρ .

6.4. Simulation

Les équations simultanées peuvent être utilisées à :

- tester et évaluer des modèles
- faire l'analyse historique des politiques économiques
- faire de la prévision/simulation.

Sur ce dernier point, il faut préciser que l'horizon temporel dépend de l'objectif poursuivi :



Pour la simulation, il faut :

- faire des hypothèses sur la trajectoire des variables exogènes
- initialiser, au-delà de la période d'estimation, les variables endogènes en corrigeant les décalages.

✚ Comment évaluer un modèle de simulation ?

Comme pour les équations individuelles, on peut recourir aux tests habituels, F, t, d, Seulement la vitalité des équations individuelles ne garantit pas celle du modèle (à équations simultanées) dans son ensemble.

Les tests suggérés dans la pratique sont :

- **RMSE** (Rooth Mean Square Error): Espérance quadratique des erreurs, qui doit être minime:

$$RMSE = \sqrt{\frac{1}{n} \sum (Y_t^s - Y_t^o)^2} \quad \text{avec}$$

Y_t^s valeur simulée de Y_t

Y_t^o valeur observée de Y_t

- **PRMSE** (Percentage Rooth Mean Square Error), la RMSE en pourcentage

$$PRMSE = \sqrt{\frac{1}{n} \sum \left(\frac{Y_t^s - Y_t^o}{Y_t^o} \right)^2}$$

- **U** (Coefficient d'inégalité de Theil)

$$U = \frac{\sqrt{\frac{1}{n} \sum (Y_t^s - Y_t^o)^2}}{\sqrt{\frac{1}{n} \sum (Y_t^s)^2 + \frac{1}{n} \sum (Y_t^o)^2}}$$

le modèle est dit prédire parfaitement si $U \rightarrow 0$
 et dit accuser des faiblesses notables si $U \rightarrow 1$

7. Introduction à la Modélisation économique

Spécifier un modèle c'est construire des relations entre variables

Les modèles sont des outils de prévision et d'analyse de politiques économiques.

Le premier modèle est proposé par Jan Tinbergen (1936) devant l'Association économique hollandaise où le prix Nobel (1969) cherchait à savoir s'il était possible de relancer l'économie hollandaise en plein crise sans détériorer la balance commerciale. Pour ce faire, il retraça au moyen de relations mathématiques, le fonctionnement de l'économie pour aider à analyser l'effet des politiques économiques et/ou à faire de la prévision

On distingue 4 types de modèles:

- les modèles macro-économétriques
- les modèles VAR (Vectoriel auto-régressif), du début des années 90 pour la prévision économique
- les modèles calculables d'équilibre général, l'étude du long terme et des mesures des politiques économiques structurelles e.g. introduction d'un nouvel impôt, libéralisation des échanges, etc.
- les modèles RBC (Real business cycles) ou modèles de cycles réels, pour l'analyse de l'effet des politiques conjoncturelles de stabilisation

Si les modèles macro-économétriques sont d'inspiration théorique keynésienne (la demande détermine le niveau de l'activité économique), les autres modèles sont d'inspiration néo-classique à la Lucas. A la suite de Mankiw et Romer, les MCEG et les MRBC sont des versions keynésiennes ou non walrasiennes

Le passage du modèle théorique au modèle appliqué se fait par étapes successives dont:

- la construction du modèle théorique
- le chiffrage qui se fait soit par estimation économétrique soit par calibrage
- l'utilisation du modèle

Pour l'étude des différents modèles, il faut passer en revue:

- l'historique
- ses fondements théoriques
- les méthodes empiriques de construction du modèle
- des exemples d'utilisation
- ses apports et limites

Les étapes de la prévision, qui suppose le modèle construit, estimé et validé:

- définir l'horizon de projection
- donner une prévision pour les variables exogènes sur toute la période de précision, sur la base d'un ensemble d'hypothèses cohérentes; pour des prévisions à long terme, le prévisionniste envisage plusieurs scénarios sur l'environnement international et la politique économique
- vérifier si chaque équation retrace bien le comportement observé ou s'il apparaît un biais (erreur systématique) sur les dernières observations disponibles, et alors introduire des variables d'écart pour supprimer de telle erreur
- fournir une prévision spontanée, qui ne sera pas publiée
- faire la prévision par itérations

Les critères de fiabilité de la prévision (définis par l'économiste hollandais Theil (1958)):

- Erreur absolue moyenne (MAE – Mean absolute error) en tant que moyenne des valeurs absolues des erreurs de prévision pour une variable donnée
- Erreur quadratique moyenne (MSE – Mean squared error), moyenne des carrés des erreurs de prévision
- Racine carrée de l'erreur quadratique moyenne (RMSE – Root mean squared error)
- Voir Massa pour les autres critères

7.1. Modèles macro-économétriques

3 innovations du début des années 40:

- la diffusion du schéma IS/LM de Hicks qui transpose la pensée keynésienne en modèle macro-économique quantitatif
- l'essor de l'économétrie permettant de quantifier les paramètres du schéma keynésien
- les calculateurs électroniques nécessaires à la résolution numérique des modèles

Ex. modèle Klein-Goldberger (1955): 17 équations dont 12 de comportement et 5 comptables

A la base, il y a:

Un schéma IS/LM, équilibre sur le marché des biens et services et sur le marché de la monnaie, pour des niveaux donnés de prix et de salaires

Une cohérence comptable des équations par l'équilibre ressources-emplois:

$$Y = C + I + G + \Delta S + X - M$$

Critiques faites aux modèles:

- Ne pas tenir compte des effets patrimoniaux sur le comportement des agents
- Accorder peu de poids au comportement d'offre des entreprises
- Tenir peu compte des anticipations des agents
- Utiliser des fondements théoriques obsolètes
- Lucas (1976) critique sévèrement que les modèles macro-économiques ne permettent pas de saisir la réaction d'une économie à un changement de politique économique e.g. dévaluation ou baisse de charges sociales employeurs, puisque les paramètres sont constants au lieu de dépendre de la politique économique

Dans le modèle Multimod du FMI tente de pallier ces insuffisance e.g. l'investissement dépend de Tobin's Q qui rapporte la somme actualisée des profits futurs attendus d'une unité de capital et le coût de cette unité de capital, qui doit être supérieur à UN pour que l'investissement soit rentable.

7.2. Les modèles VAR

Sims a remis en cause la succession des étapes de construction des modèles macro-économétriques:

- Formulation et estimation d'une équation structurelle pour chaque variable du modèle
- Empilement de ces équations qui donne la forme structurelle du modèle

- Transformation de la forme structurelle en une forme réduite en vue de la résolution numérique

Sims postule plutôt une forme réduite linéaire simple ou modèle VAR, qui comporte pour chaque variable endogène une équation explicative sur les réalisations passées de toutes les variables endogènes du modèle. Il ne s'agit dès lors que de définir le nombre de variables et le nombre de retards. Ce choix est limité par le nombre d'observations disponibles.

Si k variables et p retards, alors $np+1$ paramètres par équation, or il faut que le nombre d'observations soit beaucoup plus grand que le nombre de paramètres à estimer. C'est pourquoi les modèles VAR ne retiennent pas plus de 10 variables (composantes PIB, prix des biens et des facteurs et emploi) et 5 à 6 retards.

Puisqu'il y a que des variables endogènes, l'étape de prévision des exogènes est ignorée et puisqu'il n'y a pas non plus de variables d'écart, c'est la prévision spontanée qui est publiée.

7.3. Les modèles calculables d'équilibre général

On introduit dans les modèles la dynamique et les anticipations, ce sont les prix qui s'ajustent de façon à assurer l'équilibre simultané sur tous les marchés et chaque agent agit de façon à maximiser une fonction spécifique (utilité pour le consommateur, profit pour le producteur) sous une contrainte spécifique (budget pour le consommateur, technologie pour le producteur). Les conditions sous lesquelles l'équilibre existe portent sur les préférences des agents et sur les possibilités techniques de la production.

A la différence des modèles macro-économétriques dont les paramètres sont estimés par l'économétrie, l'estimation des MCEG se fait par calibrage i.e.:

- construire une base de données pour une année ou une décennie quelconque
- chercher à faire correspondre l'équilibre du marché avec cette base de données, et donc trouver tous les paramètres qui permettent cette correspondance
- utiliser le modèle estimé pour analyser les effets d'un changement dans la politique économique
- mesurer la sensibilité des résultats à la valeur des paramètres structurels choisis.

La MCS est un tableau conventionnel constitué, en ligne et en colonne, de différents comptes. Les comptes en ligne reçoivent les paiements des comptes

en colonne. La construction de la MCS est basée sur le principe de la comptabilité à double entrée, ce qui garantit un équilibre entre le total des ressources et le total des emplois.

Tableau 1. MCS 2010 du Mali en différents blocs

	Facteurs	Agents	Branches	Marché intérieur	Exportation	Investissement	Total
Facteurs			Bloc 1				
Agents	Bloc 2	Bloc 3		Bloc 4	Bloc 5		
Branches				Bloc 6	Bloc 7		
Marché intérieur		Bloc 8	Bloc 9			Bloc 10	
Exportation		Bloc 11					
Accumulation		Bloc 12					
Total							

- Bloc 1 Rémunération des facteurs par branche de production
- Bloc 2 Rémunération des agents par facteur
- Bloc 3 Transferts de revenus entre agents
- Bloc 4 Prélèvement sur la vente des biens et services sur le marché domestique de la part des agents (Etat et collectivités locales)
- Bloc 5 Prélèvement sur la vente des biens et services sur le marché d'exportation de la part des agents (Etat et collectivités locales)
- Bloc 6 Revenu des branches sur leurs ventes sur le marché domestique
- Bloc 7 Revenu des branches sur leurs ventes sur le marché d'exportation
- Bloc 8 Dépense des agents en achat de biens et services sur le marché domestique
- Bloc 9 Consommation intermédiaire des branches dans les différents produits composites domestiques (produits localement et importés)
- Bloc 10 Dépense d'investissement dans les différents produits composites domestiques (produits localement et importés)
- Bloc 11 Recettes d'exportation par produit payées par les agents (Reste du monde)
- Bloc 12 Epargne des agents

Exemples d'utilisation de MCEG:

La structure fiscale optimale, étant donné que l'impôt modifie le comportement des ménages et des entreprises et donc l'équilibre macro-économique.

Le choix d'un système de financement des retraites, où le maintien de l'âge et des prestations nécessite un relèvement des cotisations ce qui équivaut à un sacrifice générationnel car ceux qui subiront aujourd'hui les taux de cotisation élevés percevront une pension sans rapport avec l'effort à eux demandé

7.4. Les modèles de cycles réels

Une succession de chocs (une sécheresse, la rébellion chez soi ou chez le partenaire, un naufrage, etc.) engendre des cycles économiques

Les étapes:

- décomposer l'évolution économique en tendance et cycle (i.e. les fluctuations autour de la tendance)
- construire et calibrer un MEGC
- bombarder le modèle de chocs aléatoires
- vérifier que les fluctuations issues du modèle retracent celles des variables de l'économie réelle.

Références bibliographiques

1. R. Giraud, N. Chaix (1989): Econométrie, Ed. PUF
2. J. Johnson (1972): Econometric methods, Ed. Mcgraw-hill
3. Y. Kouadio, K. M. Zéhia : Atelier intensif d'économétrie (document de travail), Réseau de recherche sur les politiques industrielles en Afrique / CODESRIA, juillet 1994
4. A. Koutsoyiannis (1988): Theory of econometrics, Ed. Macmilan
5. G.S. Maddala (1977): Econometrics, Ed. Macgraw-hill
6. E. Malinvaud (1981): Méthodes statistiques de l'économétrie, Ed. Dunod
7. G. Vangrevilinghe (1973): Econométrie, Ed. Hermann