

Chapitre 3

Le modèle autorégressif spatial (SAR)

Corrigé

Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger
Exercices et corrigés · Économétrie spatiale (Master)

1 Exercices du chapitre

Correction — Exercice 1

- (a) Wy elle-même est bien exogène, mais Wy est une **fonction de y** , et y dépend en retour de ε à travers l'équation du modèle elle-même : Wy hérite donc mécaniquement de la dépendance de y à ε , même si W est fixée sans aucun lien avec les données.
- (b) Une omission de variable laisse les coefficients **non biaisés** lorsque la variable omise n'est pas corrélée aux régresseurs inclus, et seule l'inférence est affectée. Ici, le problème touche directement l'estimation du **coefficient ρ** lui-même : une correction des écarts-types ne changerait rien au biais du point estimé.
- (c) Parce que Wy est corrélé avec ε par construction (question a), la condition d'exogénéité des régresseurs requise par les MCO est violée pour Wy : l'estimateur MCO de ρ est donc **biaisé et non convergent**, et ce biais ne s'atténue pas quand $n \rightarrow \infty$.
- (d) Le **maximum de vraisemblance**, valide car il intègre explicitement le terme jacobien $|I - \rho W|$ issu du changement de variable vers la forme réduite, plutôt que de traiter Wy comme un régresseur ordinaire ; et les **variables instrumentales / GMM** (Kelejian-Prucha), valides car WX (et W^2X) sont corrélés avec Wy mais non corrélés avec ε sous les hypothèses usuelles, permettant une estimation par doubles moindres carrés adaptée au cas spatial.

Correction — Exercice 2

$$(a) I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \rho W = 0,3 \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0,1 & 0,1 & 0,1 \\ 0,1 & 0 & 0,1 & 0,1 \\ 0,1 & 0,1 & 0 & 0,1 \\ 0,1 & 0,1 & 0,1 & 0 \end{pmatrix}, \text{ et } \rho^2 W^2 \text{ obtenu en}$$

élevant W au carré puis en multipliant par 0,09.

(b) Effet sur y_A au premier tour : $(\rho W)_{AC} = 0,3 \times 1/3 = \mathbf{0,10}$.

(c) $(W^2)_{CC} = \sum_k w_{Ck}w_{kC} = 3 \times (1/3)^2 = 1/3$, donc $(\rho^2 W^2)_{CC} = 0,09 \times 1/3 \approx \mathbf{0,03}$ (effet de retour sur C elle-même). $(W^2)_{AC} = \sum_k w_{Ak}w_{kC} = w_{AB}w_{BC} + w_{AD}w_{DC} = 2 \times (1/3)^2 = 2/9$, donc $(\rho^2 W^2)_{AC} = 0,09 \times 2/9 \approx \mathbf{0,02}$.

(d) Effet total approché sur y_C : $1 + 0 + 0,03 = \mathbf{1,03}$. Effet total approché sur y_A : $0 + 0,10 + 0,02 = \mathbf{0,12}$. Le résultat exact étant 0,1299, l'approximation à l'ordre 2 (0,12) en est déjà très proche (écart $\approx 0,01$) : les termes d'ordre supérieur ($\rho^3 W^3$, etc.) contribuent de moins en moins, conformément à la décroissance géométrique du multiplicateur.

Correction — Exercice 3

(a) Dans les deux cas, la condition porte sur le **module** d'un paramètre autorégressif (ϕ pour le temps, ρ pour l'espace) devant rester strictement inférieur à 1 pour que le processus reste stable : au-delà, les chocs ne s'amortissent plus mais s'amplifient indéfiniment à mesure qu'ils se propagent, que ce soit dans le temps (AR(1)) ou de proche en proche dans l'espace (SAR).

(b) $\hat{\rho} = 1,15 > 1$ viole directement la condition de stabilité $|\rho| < 1$: le modèle SAR estimé est **explosif** et n'admet pas d'équilibre stable — un choc local s'amplifierait indéfiniment à chaque tour de propagation plutôt que de s'éteindre, ce qui rend le modèle économiquement incohérent, indépendamment de toute question de significativité statistique.

(c) $|-0,6| = 0,6 < 1$: ce résultat est **compatible** avec la condition de stabilité. Économiquement, un ρ négatif signifierait qu'une valeur élevée de y dans une région tend à s'accompagner de valeurs **plus faibles** chez ses voisines — un mécanisme de compétition ou de substitution entre unités voisines (par exemple, des entreprises voisines qui se disputent une même clientèle limitée), à l'opposé d'un mécanisme de renforcement mutuel.

2 Exercice cumulatif (chapitre 3, chapitre 2, chapitre 1 et introduction)

Correction — Exercice 4

- (a) **Non.** Comme le rappelle l'introduction, l'indice de Moran détecte la présence d'une dépendance spatiale sans indiquer où elle loge : un I significatif est compatible aussi bien avec un modèle SAR qu'avec un modèle SEM.
- (b) Le test de Moran est un outil purement **statistique**; seul un raisonnement **économique** sur le mécanisme réel qui relie les régions (ici, la comparaison directe des prix entre voisines, une authentique rétroaction) permet de justifier la spécification SAR plutôt que SEM — exactement la distinction entre dépendance substantielle et corrélation nuisible posée dès l'introduction.
- (c) Avec $\hat{\rho} = 0,25$ et la matrice W à poids $1/3$: au premier tour, le choc se répercute sur chacune des trois autres régions à hauteur de $0,25 \times 1/3 \approx 0,083$; au deuxième tour, une fraction de cet effet revient affecter la région de départ et se propage encore aux autres, avec un poids environ trois fois plus faible qu'au tour précédent (car $\rho = 0,25 < 0,3$ de l'exercice 2, la décroissance est encore plus rapide) ; l'effet continue ainsi à s'atténuer géométriquement sans jamais s'annuler complètement, conformément à la condition de stabilité $|\rho| = 0,25 < 1$.
- (d) Parce que cette régression MCO initiale, comme le montre ce chapitre pour toute estimation SAR par MCO, souffre du même biais de simultanéité si la vraie dépendance loge dans y : les résidus utilisés pour le test de Moran, et plus encore les coefficients de cette régression, ne peuvent servir de point de départ fiable une fois que la chercheuse a elle-même choisi de croire à une dépendance substantielle logée dans y — il faut ré-estimer l'ensemble du modèle par ML ou IV/GMM plutôt que de partir des résultats MCO.

3 Applications en sciences économiques

Correction — Exercice 5

- (a) Le ML et l'IV/GMM reposent sur des hypothèses différentes (respectivement, une spécification distributionnelle complète pour le ML ; l'exogénéité des instruments WX , W^2X pour l'IV/GMM) : leur convergence vers des valeurs proches (0,45 contre 0,41) suggère que le résultat n'est pas un artefact des hypothèses propres à une seule méthode, renforçant la crédibilité du diagnostic de dépendance spatiale positive.
- (b) Le coefficient $\hat{\rho}$ seul ne mesure que l'intensité de la rétroaction **directe** entre voisins immédiats ; la question posée porte sur l'effet **total**, une fois pris en compte tous les tours de propagation successifs (voisins, voisins des voisins, etc.) jusqu'à un nouvel équilibre — exactement ce que somme le multiplicateur $(I - \rho W)^{-1}$.
- (c) Le modèle SAR estimé décrit une relation d'**équilibre** entre variables économiques réelles (caractéristiques du logement, du quartier, etc.) observées sur des données historiques : une hausse purement **affichée** et artificielle, sans changement des caractéristiques sous-jacentes qui déterminent réellement les prix, ne correspond à aucun des mécanismes économiques (comparaison réelle entre acheteurs, arbitrage de marché) que le coefficient $\hat{\rho}$ a capturés — le modèle ne prédit la propagation que de chocs de même nature que ceux observés dans les données ayant servi à l'estimer.

Correction — Exercice 6

- (a) $\hat{\rho} = 0,55$ signifie qu'une augmentation du taux d'adoption chez les coopératives voisines d'une coopérative donnée se répercute, via une véritable rétroaction bouclée, sur le taux d'adoption de cette coopérative elle-même — au-delà d'une simple ressemblance statistique, c'est un mécanisme de diffusion qui s'auto-entretient dans le système entier.
- (b) **Oui.** Puisque le modèle SAR formalise une authentique dépendance substantielle (un mécanisme de diffusion réel, comme le distingue l'introduction), le multiplicateur spatial $(I - \rho W)^{-1}$ garantit qu'un choc initial sur quelques coopératives pilotes se propage réellement, en cercles concentriques, au reste du système — contrairement au cas d'une simple corrélation nuisible où une telle politique ciblée serait vaine faute de mécanisme réel de propagation.

(c) Avec $n = 3\,500$ observations, le maximum de vraisemblance devient numériquement coûteux (la maximisation implique le calcul répété du déterminant jacobien $|I - \rho W|$ sur une matrice de très grande taille), alors que l'IV/GMM reste calculable à cette échelle au prix d'une perte d'efficacité acceptable — exactement le compromis rappelé dans ce chapitre entre les deux méthodes d'estimation.