

# Chapitre 5

Modèles généraux et choix de spécification

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) Sous  $\theta = 0$  (aucun régresseur décalé spatialement), le SDM se réduit exactement au SAR.
- (b) Sous la restriction non linéaire  $\theta = -\rho\beta$  (la restriction du facteur commun), le SDM devient équivalent au SEM.
- (c)  $\theta = -\hat{\rho}\hat{\beta} = -0,35 \times 2,5 = -0,875$ .
- (d)  $\hat{\theta} = -0,92$  est très **proche** de la valeur attendue  $-0,875$  sous la restriction du facteur commun (écart  $\approx 0,045$ ). Cela suggère, de façon informelle, que les données sont **compatibles** avec un simple modèle SEM plutôt qu'avec la généralité complète du SDM — une conclusion qu'un test formel du rapport de vraisemblance permettrait de confirmer rigoureusement.

### Correction — Exercice 2

- (a)  $LM_\rho = 2,10 < 3,841$  et  $LM_\lambda = 1,50 < 3,841$  : aucun n'est significatif, les **MCO suffisent**.
- (b)  $LM_\rho = 8,30 > 3,841$  est significatif,  $LM_\lambda = 1,20 < 3,841$  ne l'est pas : retenir le modèle **SAR**.
- (c) Les deux tests simples sont significatifs ( $6,70$  et  $7,90 > 3,841$ ), il faut donc examiner les versions robustes :  $LM_\rho^* = 1,10 < 3,841$  (non significatif),  $LM_\lambda^* = 5,40 > 3,841$  (significatif) : retenir le modèle **SEM**.
- (d) Si les deux versions robustes  $LM_\rho^*$  et  $LM_\lambda^*$  restaient significatives simultanément, la procédure recommanderait le modèle combiné **SAC/SARAR**, qui superpose un lag spatial dans  $y$  et une erreur autorégressive spatiale.

### Correction — Exercice 3

- (a) La démarche ascendante d'Anselin part des MCO et n'introduit de la complexité spatiale que si les tests LM la justifient ; la démarche descendante d'Elhorst part du modèle le plus général (SDM) et le restreint seulement si les données le permettent.
- (b) La démarche ascendante risque de rester bloquée sur un modèle trop restreint (par exemple SAR ou SEM simple) faute d'avoir jamais testé l'alternative plus large (SDM ou SAC) qui aurait pu être la vraie spécification.
- (c) La démarche descendante est plus coûteuse en calcul et en degrés de liberté, puisqu'elle estime d'emblée le modèle le plus riche (avec tous les régresseurs décalés spatialement  $WX$ ) même quand une spécification plus simple aurait suffi.
- (d) Parce qu'aucune des deux démarches n'est unanimement supérieure : elles reposent sur des logiques différentes et un accord entre leurs conclusions renforce nettement la confiance dans la spécification finalement retenue, tandis qu'un désaccord signale un résultat fragile qui mérite un examen plus approfondi.

## 2 Exercice cumulatif (chapitre 5, chapitre 4, chapitre 3, chapitre 2 et chapitre 1)

### Correction — Exercice 4

- (a)  $LM_\rho = 5,2$  et  $LM_\lambda = 6,8$  sont tous deux significatifs ( $> 3,841$ ) : il faut examiner les versions robustes.  $LM_\rho^* = 0,9$  n'est pas significatif,  $LM_\lambda^* = 4,1$  l'est : la procédure recommande le modèle **SEM**.
- (b) **Oui** : le résultat statistique (SEM recommandé) est parfaitement cohérent avec le raisonnement économique du chapitre 4 — une variable omise (qualité du sol) spatialement corrélée, sans mécanisme direct de propagation entre parcelles, est précisément la signature qu'un modèle SEM est censé capturer.
- (c) L'indice de Moran détecte la présence d'une dépendance spatiale globale sans distinguer sa source ; les tests  $LM_\rho$  et  $LM_\lambda$ , en testant spécifiquement  $H_0 : \rho = 0$  contre une alternative SAR d'une part,

$H_0 : \lambda = 0$  contre une alternative SEM d'autre part, apportent exactement la distinction qui manquait — ils permettent de trancher *statistiquement* entre les deux mécanismes, là où Moran ne pouvait offrir qu'un raisonnement économique a priori.

(d) Avec  $LM_\rho^* = 4,5$  et  $LM_\lambda^* = 3,9$  tous deux significatifs, la procédure recommanderait le modèle combiné **SAC/SARAR**. Ce modèle est plus coûteux à estimer car il superpose un lag spatial dans  $y$  et une erreur autorégressive spatiale, ce qui fragilise l'identification et rend les résultats plus sensibles au choix exact des matrices de pondération  $W_1$  et  $W_2$  — la généralité s'achète toujours en puissance statistique et en robustesse, comme le rappelle ce chapitre.

### 3 Applications en sciences économiques

#### Correction — Exercice 5

(a)  $LM_\rho = 12,4$  et  $LM_\lambda = 9,8$  significatifs, donc examen des versions robustes :  $LM_\rho^* = 6,1$  significatif,  $LM_\lambda^* = 0,7$  non significatif : la procédure recommande le modèle **SAR**.

(b) **Oui** : un véritable effet de comparaison entre biens voisins (les acheteurs comparent les prix affichés dans le voisinage avant de faire une offre) est exactement le type de mécanisme de rétroaction directe dans  $y$  que le modèle SAR est censé capturer, cohérent avec le résultat statistique.

(c) **Oui, plausiblement**. La démarche d'Elhorst, en estimant un SDM complet puis en testant la restriction du facteur commun ( $\theta = -\rho\beta$ ), devrait ici rejeter cette restriction (puisque le SAR, non le SEM, est la spécification pertinente) et converger vers une conclusion compatible avec le SAR — les deux démarches, bien que méthodologiquement différentes, sont conçues pour converger sur les mêmes données lorsque le signal est suffisamment net.

#### Correction — Exercice 6

(a) Les deux tests simples et les deux versions robustes étant tous significatifs, la procédure recommande le modèle combiné **SAC/SARAR**.

(b) Ce résultat est cohérent avec la coexistence des deux mécanismes soupçonnés : la diffusion atmosphérique réelle entre zones voisines correspond au terme  $\rho Wy$  (dépendance substantielle logée dans  $y$ ), tandis que le régime de vents dominants régional, non observé, correspond au terme  $\lambda Wu$  (corrélation nuisible logée dans l'erreur) — exactement la structure que le SAC/SARAR est conçu pour capturer simultanément.

(c) Le chapitre 5 met en garde sur le « prix de cette généralité » : le modèle SAC/SARAR est plus difficile à estimer (identification plus fragile, calcul plus lourd) et ses résultats deviennent particulièrement sensibles à la spécification exacte des matrices  $W_1$  (pour le lag) et  $W_2$  (pour l'erreur) — une mise en garde d'autant plus pertinente ici que la chercheuse pourrait légitimement choisir des critères de voisinage différents pour la diffusion atmosphérique (probablement liée aux vents dominants, donc directionnelle) et pour le facteur climatique régional (probablement plus diffus).