

Chapitre 8

Applications, régression géographiquement pondérée et limites

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) **Hétérogénéité** : c'est le coefficient lui-même (le rendement de l'éducation) qui varie dans l'espace, pas une dépendance entre observations — cas typique de la GWR.
- (b) **Dépendance** : un mécanisme relie directement les observations entre elles — cas SAR.
- (c) **Dépendance** (avec une composante temporelle) : une diffusion entre unités voisines — cas SAR ou SAR-panel dynamique (chapitre 7).
- (d) **Hétérogénéité** : le coefficient (pluviométrie → rendement) diffère selon des caractéristiques locales (type de sol), sans qu'aucune observation n'influence directement une autre — cas GWR.
- (e) **Oui** : les deux hypothèses (dépendance entre observations, coefficient variable dans l'espace) ne s'excluent pas mutuellement et peuvent être simultanément vraies pour un même phénomène — d'où l'existence d'extensions récentes combinant les deux logiques, mentionnées dans ce chapitre.

Correction — Exercice 2

- (a) **Effet d'échelle** : la corrélation change lorsqu'on modifie le niveau d'agrégation (communes puis provinces), à zonage comparable.
- (b) **Effet de zonage** : à échelle égale (même nombre approximatif d'unités), c'est le **découpage** lui-même, et non le niveau d'agrégation, qui change les résultats.
- (c) Parce qu'un problème d'endogénéité admet des solutions générales et reproductibles (variables instrumentales, panel, discontinuité), alors que le découpage administratif ou géographique est, par nature, un choix largement **arbitraire** au regard du phénomène économique étudié, sans critère objectif unique pour le fixer une fois pour toutes.
- (d) Tester la robustesse des conclusions à **plusieurs découpages plausibles**, et rester prudent sur toute conclusion qui ne survit qu'à un seul découpage particulier.

Correction — Exercice 3

- (a) **Non**. La corrélation observée au niveau régional (agrégé) ne se transpose pas nécessairement au niveau individuel — c'est précisément la définition de l'erreur écologique.
- (b) Par exemple, la région pourrait attirer à la fois des immigrants **et** des natifs particulièrement qualifiés (un pôle économique dynamique), sans qu'aucun lien causal individuel n'existe entre statut migratoire et emploi : le faible taux de chômage régional refléterait alors la composition socio-économique globale de la région, pas un avantage spécifique des immigrants sur les natifs.
- (c) Avec des données **individuelles**, on pourrait directement comparer le taux de chômage des immigrants à celui des natifs *au sein* de la même région, en contrôlant leurs caractéristiques propres — l'agrégation régionale, en additionnant des individus hétérogènes, masque justement cette possibilité de comparaison directe.
- (d) Il s'agit d'un **troisième type** de difficulté, distinct de la dépendance et de l'hétérogénéité spatiales : un problème d'**inférence** lié au niveau d'agrégation des données lui-même, plutôt qu'à une propriété statistique du modèle économétrique.

2 Exercice cumulatif (chapitre 8 et chapitre 1)

Correction — Exercice 4

- (a) Les deux mises en garde partagent la même structure logique : un choix méthodologique — la matrice W pour le chapitre 1, le découpage des unités pour le MAUP — est largement **arbitraire** au regard du phénomène économique étudié, et ce choix peut à lui seul déterminer la conclusion obtenue, sans qu'aucun critère objectif ne permette de trancher définitivement en faveur d'un choix unique.
- (b) 3 matrices $W \times 2$ découpages = 6 combinaisons distinctes à examiner en principe.

- (c) Une conclusion qui résiste à l'ensemble des 6 combinaisons est nettement plus **crédible** qu'une conclusion obtenue avec un seul choix arbitraire : elle n'est plus un artefact d'une spécification particulière, mais un résultat robuste à deux sources indépendantes d'arbitraire méthodologique.
- (d) **Seulement le risque associé au choix spécifique de W et de découpage testés** : tester 6 combinaisons parmi une infinité de choix possibles ne garantit jamais qu'**aucune** autre combinaison plausible ne changerait la conclusion — la discipline de robustesse réduit le risque sans jamais l'éliminer complètement, ce qui est cohérent avec l'affirmation du cours selon laquelle le MAUP n'a « pas de solution générale ».

3 Applications en sciences économiques

Correction — Exercice 5

- (a) Le SAR capture la **dépendance** entre biens voisins (un mécanisme de comparaison de prix) tandis que la GWR capture l'**hétérogénéité** du rendement d'une caractéristique (la surface habitable) selon le quartier — les deux hypothèses (dépendance entre observations, coefficient variable dans l'espace) peuvent être simultanément vraies, comme le rappelle l'exercice 1(e), d'où leur complémentarité plutôt que leur substituabilité.
- (b) Une bande étroite capture une variation locale fine mais avec très peu d'observations par régression locale, ce qui gonfle la **variance** des estimations locales (résultats peu fiables, bruités). Le cours recommande de calibrer la largeur de bande par **validation croisée**, en minimisant l'erreur de prédiction hors échantillon.
- (c) **Oui**. Les résultats de la GWR, comme tout résultat spatial, dépendent potentiellement du découpage retenu pour définir les quartiers ciblés par la politique de rénovation : un découpage différent des mêmes zones urbaines pourrait faire ressortir des priorités de rénovation sensiblement différentes, exactement le risque de l'effet de zonage du MAUP.

Correction — Exercice 6

- (a) Si les spillovers de croissance entre régions voisines sont réels (une région dynamique tire ses voisines vers le haut) mais ignorés, le test de convergence risquerait d'attribuer à tort à la seule dynamique de rattrapage individuel un phénomène qui provient en partie d'une interdépendance régionale non modélisée — exactement le type de confusion mise en garde dès l'introduction du cours entre dépendance substantielle et facteur omis.
- (b) Il s'agit du **MAUP**, plus précisément de son **effet de zonage** (nombre d'unités comparable, 12 contre 16, mais découpage différent).
- (c) L'étude devrait rapporter les résultats obtenus sous **les deux découpages** (et idéalement d'autres découpages plausibles), signaler explicitement si la conclusion qualitative sur la convergence résiste ou non à ce changement, et rester prudente dans ses conclusions si le résultat s'avère sensible au découpage retenu — la même discipline de robustesse que celle recommandée pour le choix de W au chapitre 1.