Cluster-wise spatiotemporal panel models for global and local spillovers identification

Caterina Morelli
University of Valle d'Aosta

Paolo Maranzano
University of Milan-Bicocca

Raffaele Mattera
University of Campania "Luigi Vanvitelli"

Philipp Otto
University of Glasgow

Abstract

The increasing interconnectedness of the global economy has profoundly altered the way economic, environmental, and social phenomena evolve across space and time. Countries, regions, firms, and households are no longer isolated decision-making units; instead, they are embedded in dense networks of geographical proximity, trade relations, supply chains, migration flows, technological diffusion, and institutional linkages. Shocks arising in one location often propagate rapidly to others, generating complex patterns of dependence that challenge traditional econometric frameworks built on cross-sectional independence. At the same time, the strength, direction, and persistence of these spillovers depend on the structural characteristics of the interacting units. Differences in economic development, human capital, technological readiness, governance quality, infrastructure, and sectoral composition may shape both how strongly a unit influences its neighbors and how responsive it is to external stimuli. These considerations suggest that spatial dependence and structural heterogeneity are not separate phenomena, but rather two intertwined dimensions of the same underlying process.

The spatial econometrics literature has long emphasized the importance of modeling spatial dependence explicitly. Building on the insight that nearby units tend to be more strongly related than distant ones, formalized in Tobler's First Law of Geography (Tobler, 1970), spatial autoregressive and spatial Durbin models have become standard tools for capturing endogenous and exogenous spillovers across space (LeSage and Pace, 2009; Elhorst, 2010). In parallel, a large literature has highlighted the presence of spatial heterogeneity, meaning that structural relationships may vary systematically across locations (Goodchild, 2004; Zhu and Turner, 2022). Ignoring this dimension may lead to biased or inconsistent parameter estimates when the assumption of homogeneous effects is violated. More recently, efforts have been made to bridge spatial and factor-based approaches to cross-sectional dependence, recognizing that local spatial interactions and broader common components may coexist (Koopman et al., 2021; Chen et al., 2022; Su et al., 2023). Nevertheless, many empirical applications still rely on homogeneous spatial models

that impose identical coefficients across units, potentially obscuring meaningful differences in behavior and diffusion mechanisms.

This paper develops a novel heterogeneous spatiotemporal panel data model that integrates dynamic persistence, spatial spillovers, and structured group-level heterogeneity within a unified framework. The key idea is to allow the parameters governing both marginal effects and spatial transmission mechanisms to vary across latent clusters of units. Rather than assuming that all spatial units respond identically to shocks and covariates, the model partitions the cross-sectional dimension into a finite number of groups, each characterized by its own set of structural parameters. Importantly, spatial dependence is allowed to differ not only within clusters but also across clusters, thereby accommodating heterogeneous and potentially asymmetric inter-group interactions.

Formally, each unit belongs to one of G latent clusters that remain fixed over time. Conditional on this partition, the model includes several layers of dependence. First, contemporaneous endogenous spillovers are captured through a group-to-group spatial autoregressive matrix, which allows the outcome of a unit in cluster g to depend on the contemporaneous outcomes of neighboring units in cluster f, with the strength of this effect varying across all possible pairs of clusters. Second, temporal dynamics are introduced through cluster-specific autoregressive parameters, allowing persistence to differ across groups. Third, lagged spatial spillovers are incorporated through an additional group-to-group matrix governing the impact of past neighboring outcomes on current outcomes. Fourth, the marginal effects of covariates are allowed to vary across clusters, capturing structural heterogeneity in behavioral responses. Finally, spatial spillovers of covariates are also modeled in a group-specific manner, so that the influence of neighboring covariates may differ depending on both the origin and the destination cluster.

Suppose we have a panel data, with N units for T time stamp, thus y_{it} is the dependent variable for unit i at time t, $\mathbf{x}_{it}$ is the vector of k explanatory variables for the same unit and time. Moreover, let $\mathcal{P} = (\mathcal{C}_1, \dots, \mathcal{C}_G)$ being the partition in G number of clusters. Each unit in our sample belongs to one and only one cluster for the entire time period. Also, suppose we have spatial information regarding each unit i with respect to the other units j, which can be interpreted as a measure of spatial proximity or boundaries, and is quantified and defined as w_{ij}. To ensure model stability, the spatial weight matrix W must satisfy several assumptions LeSage (2008); LeSage and Pace (2009); Elhorst (2010).

Our model can be defined as follows:

$$\begin{aligned} y_{it} = &\sum_{j=1}^{N}\sum_{f=1}^{G} \rho_{gf} w_{ij} y_{jt} I(j \in \mathcal{C}_f) + \tau_g y_{i,t-1} + \sum_{j=1}^{N}\sum_{f=1}^{G} \eta_{gf} w_{ij} y_{j,t-1} I(j \in \mathcal{C}_f) \\ &+ \ \mathbf{x}'_{it}\boldsymbol{\beta}_g + \sum_{j=1}^{N}\sum_{f=1}^{G} w_{ij}\mathbf{x}'_{jt}\boldsymbol{\delta}_{gf} I(j \in \mathcal{C}_f) + \epsilon_{it}, \quad i \in \mathcal{C}_g \end{aligned} \tag{1}$$

The model includes heterogeneous spatial and temporal effects across clusters. The parameter ρ_{gf} captures contemporaneous spatial spillovers from cluster f to cluster g, distinguishing between within-cluster and cross-cluster dependence. The parameter τ_g measures group-specific temporal persistence. Lagged spatial spillovers are captured by η_{gf}, allowing for both within- and cross-cluster dynamic interactions. The vector $\boldsymbol{\beta}g$ represents cluster-specific marginal effects of covariates, while $\boldsymbol{\delta}gf$ captures spatial spillovers of covariates across and within clusters. The error term ϵ_{it} is assumed to be normally distributed with constant variance.

This structure encompasses a rich set of possible configurations. Within-cluster spillovers may be absent, homogeneous, or heterogeneous across groups. Cross-cluster spillovers may be zero, symmetric, or asymmetric, allowing for directional influence patterns. For example, economically advanced regions may exert stronger influence on less developed regions than the reverse, or regions with similar industrial structures may interact more intensively within clusters than across clusters. By explicitly modeling these possibilities, the framework captures complex diffusion patterns that standard homogeneous models cannot represent.

We provide the derivation and interpretation of direct and indirect effects. In spatial models, regression coefficients do not coincide with marginal effects because spatial autoregressive feedback propagates shocks through the network. Here, this propagation is further enriched by cluster-specific dynamics and group-to-group spillovers. The authors derive the full matrix of partial derivatives linking changes in covariates to changes in outcomes, accounting for spatial and temporal feedback. The resulting block structure distinguishes intra-group effects from inter-group transmission, allowing identification of average, cluster-specific, and cross-cluster effects.

Estimating the proposed model requires addressing two main challenges: the presence of multiple spatial dependence parameters and the unknown cluster membership. To tackle these issues, we develop a two-step estimation procedure. In the first step, conditional on a given partition, we estimate the structural parameters using a quasi maximum likelihood approach tailored to the heterogeneous spatial dynamic panel setting. The method builds on established QML techniques for dynamic spatial panels (Yu et al., 2008; Lee and Yu, 2010a,b), extending them to accommodate multiple group-to-group spatial operators. By decomposing the model into non-spatial and spatial blocks and concentrating the likelihood with respect to the spatial parameters, the procedure remains computationally feasible even in moderately large panels.

In the second step, we update cluster assignments given the current parameter estimates. Unlike conventional cluster-wise regression approaches that treat observations as conditionally independent (Bonhomme and Manresa, 2015), our updating rule accounts explicitly for spatial dependence. The likelihood contribution of each unit incorporates not only its own fit but also the influence of neighboring units, reflecting the spatially interconnected nature of the data. This spatially aware clustering mechanism reduces the risk of assigning units to clusters that are inconsistent with the observed dependence structure. The two steps are iterated until convergence, defined in terms of stabilization of parameter estimates or improvements in the log-likelihood function.

To evaluate the performance of the proposed estimation strategy, we conduct an extensive Monte Carlo simulation experiment. The simulation design considers alternative configurations of within- and cross-cluster spillovers, including scenarios with no cross-cluster interaction, symmetric cross-cluster dependence, and asymmetric directional spillovers. Because fully general spatial dynamic Durbin models may face identification challenges when all dependence parameters are included simultaneously (Anselin et al., 2008; Elhorst, 2010), we also analyze a series of restricted model variants that reflect common empirical specifications. The results show that the estimator performs well in recovering the heterogeneous spatial and temporal parameters across a wide range of settings. While estimation variability increases with model complexity, the method successfully captures the relative magnitude and direction of within- and cross-cluster interactions and provides accurate estimates of the associated direct and indirect effects.

The proposed framework is particularly relevant for environmental and regional applications. In the context of greenhouse gas emissions, for instance, regions differ markedly in their economic structure, energy mix, technological capacity, and policy environment. These differences affect both the determinants of emissions and the way emission shocks propagate across space. A homogeneous spatial model may mask important asymmetries and lead to misleading policy implications. By identifying latent clusters of regions with similar structural characteristics and by modeling heterogeneous spillovers across these clusters, the present approach provides a more realistic representation of emission dynamics and spatial externalities.

More broadly, the model contributes to the growing literature on spatiotemporal econometrics by offering a flexible yet interpretable framework that integrates local spatial interactions, dynamic persistence, and structured heterogeneity. It complements recent attempts to jointly account for spatial and factor dependence (Koopman et al., 2021; Chen et al., 2022; Su et al., 2023) by introducing a cluster-based perspective that emphasizes group-level transmission mechanisms. The combination of heterogeneous parameters, block-structured spillovers, and spatially informed clustering provides a powerful tool for analyzing complex systems characterized by both interdependence and diversity.

The framework is well suited to applications in environmental economics and regional science, where understanding heterogeneous diffusion patterns and asymmetric spillovers is essential for effective policy design.

REFERENCES

Luc Anselin, Julie Le Gallo, and Hubert Jayet. *Spatial Panel Econometrics*, pages 625–660. 01 2008. ISBN 978-3-540-75889-1. doi: 10.1007/978-3-540-75892-1_19.

Stéphane Bonhomme and Elena Manresa. Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3):1147 – 1184, 2015. doi: 10.3982/ECTA11319. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84930701347&doi=10.3982%2fECTA11319&partnerID=40&md5=85d9497a93302077edaf408c61409777.

Jia Chen, Yongcheol Shin, and Chaowen Zheng. Estimation and inference in heterogeneous spatial panels with a multifactor error structure. *Journal of Econometrics*, 229(1):55 – 79, 2022. doi: 10.1016/j.jeconom.2021.05.003. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85107415238&doi=10.1016%2fj.jeconom.2021.05.003&partnerID=40&md5=70ac5d23e604efd74bca505d7697cca3.

J. Paul Elhorst. Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, 5 (1):9 – 28, 2010. doi: 10.1080/17421770903541772. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-77951588750&doi=10.1080%2f17421770903541772&partnerID=40&md5=989ce1cd6549a430702052c8cc8b517b.

Michael F. Goodchild. The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*, 94(2):300 – 303, 2004. doi: 10.1111/j.1467-8306.2004.09402008.x. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-28444486932&doi=10.1111%2fj.1467-8306.2004.09402008.x&partnerID=40&md5=d605533b3f672a5caa262fde7991da5f.

Siem Jan Koopman, Julia Schaumburg, and Quint Wiersma. Joint modelling and estimation of global and local cross-sectional dependence in large panels, 2021.

Lung-Fei Lee and Jihai Yu. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2):165–185, 2010a. doi: doi.org/10.1016/j.jeconom.2009.08.001.

Lung-Fei Lee and Jihai Yu. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory*, 26(2):564 – 597, 2010b. doi: 10.1017/S0266466609100099. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-77951095647&doi=10.1017%2fS0266466609100099&partnerID=40&md5=df080f13e8236cce04927f6d6c248776.

James LeSage and R. Kelley Pace. *Introduction to spatial econometrics*. 2009. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045659213&partnerID=40&md5=f36c5e559bdadaa988c0541a6335d735.

James P. LeSage. An introduction to spatial econometrics. *Revue d'Economie Industrielle*, 123(3):19 – 44, 2008. doi: 10.4000/rei.3887. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84975123004&doi=10.4000%2frei.3887&partnerID=40&md5=99d9b43555794e6bf01be1a88a5dba7a.

Liangjun Su, Wuyi Wang, and Xingbai Xu. Identifying latent group structures in spatial dynamic panels. *Journal of Econometrics*, 235(2):1955 – 1980, 2023. doi: 10.1016/j.jeconom.2023.02.007. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85150284978&doi=10.1016%2fj.jeconom.2023.02.007&partnerID=40&md5=9b64906a16238a2c87132760a208cb52.

W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240, 1970.

Jihai Yu, Robert de Jong, and Lung fei Lee. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, 146(1):118–134, 2008. doi: doi.org/10.1016/j.jeconom.2008.08.002.

A-Xing Zhu and Matthew Turner. How is the third law of geography different? *Annals of GIS*, 28(1):57 – 67, 2022. doi: 10.1080/19475683.2022.2026467. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85122804468&doi=10.1080%2f19475683.2022.2026467&partnerID=40&md5=f501b5d5a40bde7806736b8d21c4f98e.