Journal of Modelling in Management

Spatial correlation across multiple economic sectors through the North American Industry Classification System: a multivariate Moran's I approach

Received 24 November 2025
Revised 1 April 2026
17 May 2026
Accepted 25 June 2026

Luigi Capoani
Department of Economics, Ca' Foscari University of Venice, Venezia, Italy, and

Andrea Scarso
European Youth Think-Tank, Strasbourg, France

Abstract

Purpose – This paper investigates spatial interdependencies among economic sectors in the USA using Multivariate Moran's I applied to county-level employment data classified according to the North American Industry Classification System (NAICS). The study aims to identify patterns of sectoral co-location, spatial clustering and regional economic interdependence, providing an exploratory framework for understanding the spatial organization of the US economy.

Design/methodology/approach – The study uses a spatial econometric framework based on queen contiguity matrices and Multivariate Moran's I. County-level employment data across 14 major NAICS sectors are analysed to detect spatial autocorrelation and intersectoral relationships. The analysis is complemented by a network-based interpretation in which sectors are represented as interconnected nodes linked through significant spatial correlations.

Findings – The results reveal distinct spatial economic clusters, including industrial agglomerations, consumer-oriented urban economies and financial and managerial hubs. The analysis also identifies negative spatial relationships between metropolitan knowledge-intensive sectors and territorially dispersed population-serving activities. Furthermore, some sectors occupy structurally central or bridging positions within the spatial network, suggesting differentiated roles in regional integration and specialization across the US economy.

Originality/value – This paper contributes to the literature by integrating Multivariate Moran's I with a network-based interpretation of sectoral spatial relationships. Unlike traditional spatial autocorrelation studies focused on single sectors or variables, the analysis examines interdependencies across multiple economic sectors simultaneously. The study provides an exploratory perspective on how sectoral co-location, divergence and structural centrality shape regional economic organization in the USA. By combining multivariate spatial econometrics with network analysis, the paper offers a novel framework for interpreting spatial economic structures and intersectoral connectivity.

Keywords Economic geography, Econophysics, Spatial autocorrelation, NAICS, Moran's I model

Paper type Research paper

JEL classification – A12, B1, B27, C31, F10, F20, R30

Journal of Modelling in Management
Emerald Publishing Limited
1746-5664
DOI 10.1108/JM2-11-2025-0653

1. Introduction

The spatial co-distribution and interdependence of economic sectors represent key dimensions of both regional and national development, as they shape productivity dynamics, knowledge spillovers and structural transformation processes. In this context, the present study applies a Multivariate Moran's I framework to NAICS industry clusters to identify and quantify patterns of cross-sectoral spatial dependence. The findings contribute to the existing literature by offering a novel empirical perspective on spatial economic interdependencies and policy-relevant insights for the design of targeted, data-driven interventions.

This analysis is grounded in an empirically oriented theoretical framework that draws on Michael Porter's (1998) proposition that the geographical concentration of industries within the same sector or in related activities enhances competitiveness and fosters innovation [1]. Rather than deploying these frameworks as explanatory mechanisms, however, this study operationalizes them through an empirical lens grounded in the spatial econometrics literature on industry co-location, intersectoral agglomeration and cluster formation. A cornerstone reference in this regard is the work of Ellison, Glaeser and Kerr (2010), who demonstrate that input–output linkages, labor market pooling and knowledge spillovers are the principal drivers of intersectoral co-location in US manufacturing. Their finding that pairwise coagglomeration indices systematically reflect production–chain proximity provides a direct empirical anchor for the cross-sectoral spatial correlations uncovered in the present analysis. Complementing this, Rosenthal and Strange (2001) show that labor market pooling is the most robust determinant of agglomeration across spatial scales in US manufacturing industries, further grounding the spatial clustering patterns observed here. More recently, Tian, Gottlieb and Goetz (2020) extend this co-location framework to multi-sector interactions measured across county borders using input–output information, providing a methodological precedent closely aligned with the present approach. For the specific case of consumer-facing service sectors, the distributional logic documented by Matti (2024), showing that health-care establishments follow residential population density across US counties, explains the spatial divergence observed between demand-driven services and metropolitan-agglomerated knowledge sectors.

To complement this theoretical perspective, Figure 1 provides an empirical illustration of contemporary industrial clustering in the USA, based on state science and technology institute (SSTI) analysis of Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) combined with U.S. Census Bureau geographic boundary data. Figure 1 operationalizes the notion of industrial spatial concentration, presenting a contemporary visualization of the spatial distribution of employment-based industrial concentrations across the USA, with clusters further differentiated by NAICS sectors.

Building on these insights, the study emphasizes spatial correlation as a key mechanism through which such advantages materialize. Spatial correlation reflects the interdependence of geographically distributed economic activities and the presence of spillover effects, namely, the diffusion of knowledge and innovation across regions and sectors, often mediated by knowledge networks and innovation systems (Audretsch and Belitski, 2022).

Spillovers illustrate how knowledge flows depend on both geographic proximity and the structure of intersectoral linkages, enabling firms to access external knowledge and stimulate their long-term performance (Bawa *et al.*, 2024). In this sense, spatial correlation fosters synergies arising from coordinated sectoral interactions, where collective dynamics generate outcomes that exceed those achievable by individual actors. These synergies are further reinforced by intersectoral learning processes, particularly through collaborations between industries and universities, which facilitate the transfer of codified knowledge and support innovation (Nie *et al.*, 2022; Liu *et al.*, 2019).

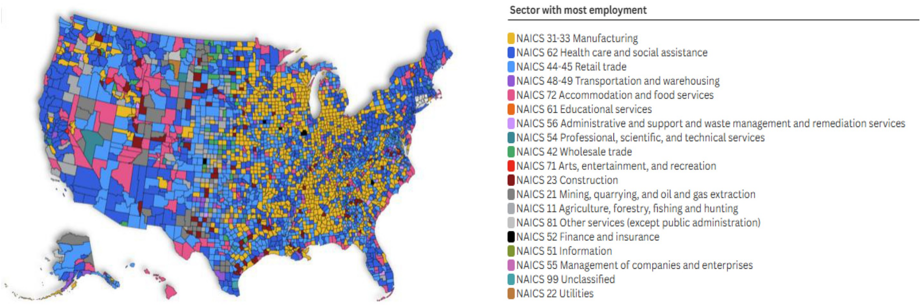

Figure 1. SSTI analysis of Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) Data, U.S. Census Bureau 2021 boundaries

As noted above, Ellison, Glaeser and Kerr (2010) identify input–output dependencies as the single most important determinant of coagglomeration among U.S. manufacturing industries. Importantly, the present study extends this framework to all major NAICS sectors (including services, finance, health care and retail) rather than focusing exclusively on manufacturing, thereby offering a broader picture of the US economic landscape.

Among the sectors most likely to exhibit this pattern are those occupying adjacent positions along production chains – particularly Construction and Wholesale Trade, whose activity is structurally linked to the location of manufacturing activity and building supply networks.

Additional empirical research explores the relationship between geographical proximity and spatial correlation dynamics with the productive structure of the economy and economic complexity. In particular, the degree of labor division and the intensity of sectoral interdependencies play a crucial role in determining the capacity of an economy to generate sustained development (Ketu and Ningaye, 2024; Stojkoski *et al.*, 2023). At the county level specifically, Pede *et al.* (2021) demonstrate that technological proximity and human capital are significant determinants of sectoral employment growth in US counties, using two-digit NAICS classifications. This finding corroborates the spatial clustering patterns uncovered in the present analysis.

Spatial correlation across economic sectors in the USA has been widely documented in the empirical literature, as it offers valuable insights into regional development dynamics, innovation patterns and industrial interconnections. A frequently cited example is Silicon Valley, often associated with strong innovation and economic growth driven by the spatial concentration of interconnected industries (Saxenian, 1996). Located in the San Francisco Bay Area, this region represents a major hub for technology companies, startups, universities and research centers. Its dense spatial configuration facilitates knowledge exchange, collaboration across companies and access to specialized talent, further supported by the presence of leading academic institutions such as Stanford University.

However, the economic benefits of spatial concentration depend not only on geographic proximity but also on the structure of relationships among industries. In particular, the concept of related variety suggests that regions hosting technologically related sectors tend to experience stronger growth, fostering knowledge spillovers (Boschma *et al.*, 2012). At the same time, effective knowledge exchange depends on cognitive proximity, as interacting actors must share sufficiently similar knowledge bases to enable meaningful interactions

(Balland *et al.*, 2015). These combined dynamics have been central to the emergence of Silicon Valley as a leading global center of technological innovation (Lécuyer, 2006).

Similar patterns can be observed in aerospace clusters in Huntsville, Alabama and Colorado Springs, Colorado (Garner Economics, 2017); automotive, steel and machinery production in the Midwest "Rust Belt" region (Warren, 2020); and financial services in New York's Wall Street district (Anagnostou, 2015). Taken together, these cases collectively illustrate how the clustering of related industries creates opportunities for collaboration, knowledge spillovers and strategic partnerships, ultimately reinforcing regional competitiveness. In this context, industrial diversification can further enhance economic stability, as fluctuations in one sector may be offset by the performance of others within the same area (Frigant and Lung, 2002).

These examples also reveal an important structural distinction between sectors whose locational logic is driven by agglomeration in metropolitan cores – such as information technology – and sectors whose distribution follows broader residential population patterns, including health care and retail trade. The former tend to concentrate in a limited number of high-density urban environments, while the latter disperse more uniformly across the territory in response to demand-side rather than supply-side factors. This divergence in locational logic suggests that the two groups may exhibit structurally opposed spatial distributions at the county level, potentially generating negative cross-sectoral spatial correlations.

To empirically operationalize these concepts, the North American Industry Classification System (NAICS) provides a structured analytical framework. Developed jointly by the USA, Canada and Mexico (Kort, 1997), it enables a consistent categorization of economic activities and supports the empirical analysis of aggregation patterns, specialization processes and regional spillovers. The system is based on a hierarchical structure, moving from broad industry categories to increasingly detailed classifications. Each industry is assigned a unique six-digit code, with the first two digits indicating the highest-level industry category (Murphy, 1998).

Owing to its clarity and operational design, NAICS has also served as a benchmark for comparable classification systems worldwide, including the Statistical Classification of Economic Activities in the European Union (NACE) system (Kort, 1997). Moreover, NAICS-based data have proven particularly valuable in tracking structural economic transformations and identifying emerging or declining sectors at the local level. Its detailed classification scheme enables the precise identification of industry clusters, making it a key tool for informing targeted economic policies. For example, Kelton *et al.* (2008) used NAICS to construct national industry cluster templates, highlighting its practical relevance in applied regional analysis.

Industry clusters play a critical role not only in shaping economic structures and trade patterns but also in influencing the spatial organization of cities. However, when the scope of the analysis extends to multiple countries, the complexity of the underlying models increases considerably. In such cases, more sophisticated analytical tools are required to capture the nuances of international economic interactions. One such tool is the concept of multilateral resistance by Anderson and van Wincoop (2004), which examines how economic barriers between a country and the rest of the world influence bilateral trade flows. This approach expands the analytical scope of international trade, moving beyond the basic comparison between two countries. While rooted in international trade theory, it complements spatial autocorrelation metrics by capturing global interaction structures that are not directly observable through spatial dependence measures such as Moran's I. To better interpret the spatial autocorrelation patterns captured by Moran's I, it is useful to consider the role of agglomeration processes in spatial economic systems. The intensity of these effects typically declines with distance from their source; however, this attenuation can be mitigated through investments in transport infrastructure, which facilitate the formation and expansion of agglomeration economies. Transport costs, factor mobility and increasing returns at the corporate level can all shape the spatial organization of economic activity. Early foundations of spatial economic theory, including von Thünen (1826) and Weber (1929), already highlighted the importance of location decisions and cost minimization in shaping industrial distribution, the latter specifically introducing the role of agglomeration. These dynamics generate economies of scale, whereby co-located firms benefit from shared resources, specialized labor, knowledge spillovers and reduced costs, ultimately fostering regional economic development and urban growth (Johnson, 2008). As a result, clusters tend to extend along the entire production chain to include both production channels and customers, while also expanding horizontally to encompass manufacturers of complementary products and companies linked by common skills, technologies or inputs.

Building on these insights, contemporary empirical studies explore spatial dynamics in economic performance. For instance, Jeleskovic and Loeber (2023) show that industrial clusters significantly influence regional Gross Domestic Product (GDP) growth by analyzing a series of panel data from German Nomenclature of territorial units for statistics (NUTS 3) regions, highlighting both positive and negative effects depending on sectoral composition and regional characteristics. Similarly, Martinho (2011) provides evidence from the Portuguese economy by using the Verdoorn relation as an analytical framework. He demonstrates that geographical spillovers play a key role in shaping productivity across sectors, particularly in the service industry.

Therefore, spatial interdependencies between economic sectors prove crucial in shaping both regional and national economic development. Geographic proximity among industries facilitates the diffusion of knowledge, enhances productivity and stimulates innovation. More broadly, industrial clusters illustrate how the co-location of interconnected sectors fosters both competition and cooperation, reinforcing productivity gains and technological advancement.

This perspective is further strengthened by Stojkoski, Koch and Hidalgo (2023), who explore the interaction between three fundamental fields – trade, technology and research – as core dimensions of inclusive and sustainable growth, framed within the objectives of Inclusive Green Growth. While their Economic Complexity approach traditionally focuses on relatedness within an abstract product space, this study reframes these concepts in a geographical perspective, emphasizing how these dimensions intersect across space.

Within this framework, Multivariate Moran's I provides a suitable tool for capturing multidimensional spatial interactions. By extending traditional univariate measures, it allows for the analysis of cross-correlations between multiple variables across regions, thereby uncovering complex patterns of sectoral co-location. Recent developments in spatial econometrics have further refined this methodological approach. In particular, Yamada (2024) and Zhang *et al.* (2023) extended the application of Multivariate Moran's I to capture higher-order spatial dependencies that remain undetected in univariate settings. At the same time, it is important to acknowledge that the method is sensitive to the specification of spatial weights and may present interpretability challenges when multiple variables interact, thus requiring careful application within a broader methodological framework.

Recent contributions further emphasize the growing relevance of multivariate spatial approaches for economic policy. In particular, Pathmanathan *et al.* (2024) highlight how Multivariate Moran's I can provide a comprehensive framework for analyzing complex spatial structures and generating policy-relevant insights. More broadly, the adoption of quantitative spatial methods is crucial for understanding sectoral interdependencies, informing regional development strategies (Le Gallo and Ertur, 2019) and supporting targeted interventions aimed at enhancing competitiveness and innovation (Varga, 2006).

For instance, empirical research uses Standard Moran's I for spatial autocorrelation analysis to show heterogeneity of spatial patterns across regions, with some areas exhibiting strong sectoral clustering while others display weak or near-random spatial structures. Therefore, spatial dependence is interpreted as an indicator of the presence or absence of localized externalities and agglomeration processes. In this sense, regions characterized by significant positive spatial autocorrelation may reflect the existence of reinforcing sectoral linkages and agglomeration economies, whereas regions with weak spatial dependence may indicate a lack of structured intersectoral interactions, potentially requiring targeted policy interventions to foster coordination and the development of localized externalities (Sánchez Gamboa and Taddei Bringas, 2014).

Further empirical applications of Moran's I, alongside exploratory spatial data analysis, have explored the spatial agglomeration patterns of manufacturing and service sectors, using employment data and industry-level classifications to identify heterogeneous sectoral clustering dynamics across regions (Guillain and Le Gallo, 2010). These contributions demonstrate that industrial activities do not distribute randomly across space, but exhibit patterns of spatial concentration and dispersion that vary across sectors, reflecting differences in production structures, input linkages and local economic conditions. But more importantly, this line of research also highlights the limitations of relying solely on the Standard Moran's I, thus informing subsequent literature in complementing global statistics with local spatial indicators (Guillain and Le Gallo, 2010). However, these approaches remain primarily univariate and focus on the analysis of spatial concentration within individual sectors, without explicitly modeling intersectoral spatial dependencies, which require more recent multivariate extensions of spatial autocorrelation.

Complementing spatial econometrics, a rapidly expanding body of literature has adopted network-based approaches to investigate complex economic and financial systems, notably in the contributions of Xu and Zhang (2023, 2024a, 2024b). These studies analyze price co-movements across spatially distributed markets using correlation-based hierarchical and synchronization methods to uncover patterns of interdependence, heterogeneity and temporal evolution. In this setting, markets are represented as nodes and pairwise correlations as links, allowing the system to be analyzed as a network of interconnected economic units. Their findings highlight how economic variables, such as commodity prices or real estate indices, form interconnected systems characterized by clustered structures (i.e. groups of markets with similar dynamics) and varying degrees of synchronization across regions.

In parallel, Jackson (2008) conceptualizes economic systems as networks of strategic interactions rather than collections of isolated and independent agents. In this perspective, economic outcomes are not solely determined by individual characteristics or geographic proximity, but emerge from patterns of interdependence embedded in complex relational structures in which agents exchange and propagate information, knowledge and influence. In such systems, the structural importance of a node is not necessarily proportional to the intensity of its direct connections. Sectors that serve as bridges between otherwise weakly linked clusters – measurable through betweenness centrality – may generate systemic connectivity without exhibiting strong pairwise spatial correlations with their neighbors. Within this perspective, correlation-based methods have emerged as a natural tool to model such interdependencies. In particular, correlation matrices can be interpreted as weighted networks, where the strength of pairwise relationships defines the intensity of connections between nodes. This representation facilitates the identification of strongly interconnected market groups and the analysis of the hierarchical organization of economic systems, especially in the study of price co-movements across spatially distributed markets (Xu and Zhang, 2021). Building on this approach, spatial econometric measures can also be interpreted through a network lens. Multivariate Moran's I not only quantifies spatial cross-correlation but also implicitly induces a structure of interdependencies that can be viewed as a spatially embedded network. By generating a matrix of spatial relationships across variables and regions, the method creates a natural bridge between spatial econometrics and network analysis, enabling a more comprehensive understanding of how economic activities co-locate and interact in space. Within this context, the present study aims to contribute to the literature by applying Multivariate Moran's I to assess spatial cross-correlations between sectoral variables in the USA. Formally, the statistic extends the univariate Moran's I by evaluating the cross-correlation across multiple variables simultaneously, capturing the degree to which spatial patterns in one sector are associated with those in others. In this study, the resulting multivariate correlation matrix is used as the basis for capturing intersectoral spatial dependencies. This approach enables the identification of spatial patterns linking different economic sectors and provides an empirical foundation for further analysis of intersectoral dependencies and regional economic structure. Ultimately, the findings suggest that fostering strategic sectoral combinations and strengthening intersectoral cooperation can enhance productivity and overall economic performance.

To strengthen causal inference and account for spatial dependencies, this study also acknowledges the role of Spatial Durbin Models (SDM), as advocated by Elhorst (2014). Furthermore, the evolution of regional innovation systems is reflected in contemporary research by Jeleskovic and Loeber (2023) and Johnson (2008), demonstrating how industrial clusters drive regional GDP growth through systemic interdependencies and knowledge exchange. Finally, this research aligns with econophysics-inspired approaches to complexity, such as those explored by Ketu and Ningaye (2024), which emphasize how sectoral employment structures shape the resilience and complexity of modern economies.

Building on this theoretical and empirical foundation, three working hypotheses are advanced to guide the empirical analysis.

The first hypothesis concerns sectors embedded in production–chain relationships. As demonstrated by Ellison, Glaeser and Kerr (2010) and formalized through Leontief's (1936) Input–Output framework, industries connected through upstream–downstream linkages face strong incentives to co-locate to reduce transport costs, share specialized inputs and access common labor pools. This logic predicts positive spatial cross-correlation among sectors occupying adjacent positions along the production chain:

H1. Sectors embedded in production–chain relationships – particularly Construction and Wholesale Trade – exhibit significant positive spatial cross-correlation, consistent with co-location driven by input–output proximity and shared upstream–downstream linkages.

The second hypothesis concerns the structural divergence between sectors whose locational logic is driven by agglomeration in metropolitan cores and those whose distribution follows broader population patterns. As illustrated by the cases of Silicon Valley (Saxenian, 1996) and Wall Street (Anagnostou, 2015), knowledge-intensive activities tend to concentrate in a limited number of high-density urban environments, whereas demand-driven services such as health care and retail trade distribute more uniformly across the territory in response to residential population rather than production-side externalities. As noted above, this structural divergence points to negative cross-sectoral spatial correlations between the two groups at the county level.

H2. The information sector exhibits negative spatial cross-correlation with population-serving sectors, such as health care and retail trade, reflecting its concentration in a limited number of high-density metropolitan clusters structurally distinct from the more territorially dispersed distribution of demand-driven services.

The third hypothesis concerns the structural role of sectors within the spatial correlation network. Drawing on Jackson's (2008) network economics framework, in which agents derive influence not solely from the intensity of direct ties but from their intermediary position in the broader relational structure, it is possible that certain sectors generate systemic spatial connectivity without exhibiting strong pairwise co-location with their neighbors. Within a network constructed from the Multivariate Moran's I correlation matrix, such sectors would be identifiable through high betweenness centrality despite comparatively low direct correlation intensity.

H3. Sectors occupying bridging positions in the spatial correlation network exhibit high betweenness centrality despite comparatively low direct spatial correlation intensity with their neighbors.

The analysis now turns to the operational implementation of the proposed approach.

2. Methodological considerations

From an operational perspective, Multivariate Moran's I is computed by assessing the spatial cross-correlation between sectoral employment distributions, comparing values observed in each region with those of neighboring areas as defined by the spatial weights matrix. This formulation enables the identification of spatial clustering patterns and interdependencies across sectors, providing a quantitative measure of how economic activities co-locate in space.

A key strength of the index lies in its ability to jointly analyze multiple variables across space, allowing the detection of complex patterns of co-location and interdependencies that cannot be detected through univariate approaches. In regional economic analysis, this feature is particularly valuable for uncovering sectoral dynamics and spatial spillovers across interconnected industries. By comparing the distribution across neighboring regions, the index provides a synthetic measure of how economic structures evolve spatially (Yamada, 2024; Zhang *et al.*, 2023).

Despite its analytical power, the application of Multivariate Moran's I entails several methodological challenges that require careful consideration to ensure robust and interpretable results. First, outcomes are highly sensitive to the specification of the spatial weights matrix, which defines the structure of interregional interactions. Alternative specifications – such as contiguity-based or distance-based weights – may lead to differences in both the magnitude and direction of the estimated spatial relationships. This underscores the importance of grounding weight selection in both theoretical reasoning and empirical context.

Second, the multivariate nature of the index introduces a non-negligible computational burden. Analyzing multiple variables simultaneously requires large, high-quality data sets, making the approach more resource-demanding than univariate methods (LeSage and Pace, 2009).

A further limitation concerns its sensitivity to the spatial structure of the data, particularly the level of aggregation at which the analysis is conducted. The size and definition of spatial units can substantially influence results – a well-known issue referred to as the *Modifiable Areal Unit Problem*. Consequently, different choices of spatial boundaries or levels of disaggregation may lead to different conclusions, affecting both the identification and interpretation of spatial patterns. For instance, applying Local Moran's I to a single state such as California may fail to detect significant clusters, whereas expanding the analysis to all US counties may reveal patterns that were previously unobservable, as the Californian cluster may emerge among the broader array of clusters across the whole country. In such cases, the Local Moran Index – also known as the Local Spatial Autocorrelation analysis method – becomes a crucial statistical tool for detecting clusters and spatial anomalies (Anselin, 1995).

Moreover, the multivariate nature of the index introduces additional computational and interpretative complexity. Analyzing multiple variables simultaneously increases the dimensionality of the problem and may complicate the interpretation of results, especially when interactions between variables are non-linear or highly interdependent. Early contributions, among which Wartenberg (1985), addressed this issue by proposing methods to summarize multivariate spatial structures, conceptually related to dimension-reduction techniques such as Principal Component Analysis.

From an inferential perspective, Multivariate Moran's I should be regarded primarily as an exploratory tool. While it effectively identifies spatial clustering and intersectoral correlation, it does not provide direct evidence of causal relationships. To enhance interpretability and maintain analytical rigor, it should therefore be complemented with more structured econometric models, such as the Spatial Durbin Model (SDM), which explicitly accounts for spatial dependence and enables causal interpretation (Elhorst, 2014). In this study, the index is used as a preliminary analytical step, aimed at detecting spatial regularities that can inform subsequent modeling strategies.

The assessment of statistical significance represents another crucial aspect of the analysis. For Global Moran's I, significance can be evaluated using either analytical approximations or permutation-based approaches. The latter are generally preferred because of their flexibility and fewer distributional assumptions, as they approximate the null distribution through repeated random permutations of the data. In line with standard practice, this study adopts a permutation-based approach with 999 replications, allowing the computation of pseudo p-values under the null hypothesis of spatial randomness. While the analysis focuses primarily on the structure of multivariate spatial correlations, the interpretation emphasizes relative patterns rather than reporting detailed global Moran's I statistics (such as Z-scores). This choice is consistent with the exploratory nature of the study.

As previously mentioned, Local Moran's I (Anselin, 1995) provides a complementary perspective to the global measure by identifying spatial clusters and outliers at the level of individual regions. Unlike the global statistic, which summarizes overall spatial dependence (Zhang *et al.*, 2008), the local indicator enables the detection of geographically concentrated patterns, such as regional industrial clusters or spatial anomalies. For instance, while Global Moran's I captures national trends, local statistics can identify geographically concentrated clusters – such as manufacturing, steel and machinery industries in the Midwest "Rust Belt" – which are essential for understanding how local industrial composition shapes regional economic performance. Although both indices rely on the same principles, they serve distinct but complementary analytical purposes. The statistical significance of local indicators is typically assessed through Monte Carlo permutation tests (Hope, 1968), which compare observed values with simulated distributions generated under random spatial arrangements.

Overall, these considerations underline the importance of adopting a cautious and structured approach when applying Multivariate Moran's I. While the index provides a powerful tool for exploring spatial interdependencies and supporting network-based interpretations of economic systems, its results should be interpreted in conjunction with complementary methods and within a well-defined empirical framework.

3. Empirical analysis of employment in the US market

In the empirical section of this study, the analysis focuses on the US labor market using county-level data for the year 2022, providing a static snapshot of the spatial distribution of employment across sectors. The primary data set consists of data from the *Bureau of Economic Analysis (BEA)* and the *U.S. Census Bureau*. In particular, employment data from the BEA provide county-level job counts by industry, including both full-time and part-time employment. These estimates are constructed by the BEA using multiple underlying sources, primarily including data from the U.S. BLS and the Internal Revenue Service, and are adjusted to ensure consistency with BEA national and regional accounts (U.S. Bureau of Economic Analysis, 2026). While the BEA data set covers multiple geographic units, including counties and metropolitan areas, the present analysis is restricted to county-level observations to ensure spatial consistency. Additionally, geographic boundary data, provided in shapefile format by the U.S. Census Bureau (updated to 2024), are used to define county-level spatial units and ensure accurate spatial representation. Supplementary employment data sourced from the U.S. Census Bureau complement the analysis, providing annual average employee counts across NAICS sectors, ensuring comparability across regions of different sizes. Finally, the data set was selected based on the principles of relevance, reliability and representativeness, in accordance with the best practices used in statistical analysis.

As discussed in the theoretical framework, spatial autocorrelation measures are key tools for assessing the extent to which similar values are geographically clustered. In the context of employment distribution, a positive Moran's I value indicates that counties with similar levels of employment are geographically proximate, suggesting the presence of spatial clustering patterns. Conversely, negative values of Moran's I imply spatial dispersion, where areas with high employment levels are surrounded by areas with lower levels and vice versa. In line with this approach, previous studies (e.g., Khan and Siddique, 2021) have applied spatial autocorrelation techniques to analyze income inequalities in the USA, finding significant spatial dependence with patterns that evolved over time.

In parallel, the analysis of sectoral structure is restricted to two-digit NAICS codes, which represent the highest level of aggregation in the classification system and ensure sufficient geographic coverage across US counties. Sector selection follows a data-driven filtering procedure to ensure robustness and spatial representativeness. Specifically, variables corresponding to two-digit NAICS codes have been extracted by identifying columns matching the two-digit NAICS structure, while NAICS 99 (Nonclassifiable Establishments) was excluded, as it does not represent a substantive economic sector. Only the "Total" enterprise-size category (Enterprise Size = "1: Total") was retained, ensuring that employment figures reflect total sectoral activity independent of firm size composition. To further enhance data reliability, a data-quality filter was applied, leading to the exclusion of sectors with valid observations in fewer than 70% of US counties from the analysis. The final sample resulted in 14 sectors with sufficient geographic coverage for robust spatial analysis. Missing observations in the remaining data set were addressed through a geographic k-nearest-neighbor imputation procedure ($k = 10$, haversine distance), using the median value of neighboring counties. This ensures that isolated suppressed or missing entries in the original BEA data do not lead to the exclusion of entire spatial units, while preserving local spatial structure in the data.

In this study, to ensure the robustness of the analysis, areas that do not share contiguous borders with other regions of the USA were excluded. In particular, observations from Hawaii, AK, the U.S. Virgin Islands, the Commonwealth of the Northern Mariana Islands, Guam, American Samoa and Puerto Rico (both for state and county analysis) were omitted, despite their occasional presence in the original data sets. This choice is motivated by both practical and methodological considerations. From a methodological standpoint, spatial units without neighboring entities would be assigned zero weights, as the calculation of Moran's I relies on the existence of physical continuity. In this study, contiguity is defined using a binary Queen criterion, whereby two regions are considered neighbors if they share either a boundary or a vertex with another administrative unit of the same type. Formally, the spatial weights matrix is defined as $W = [w_{ij}]$, where $w_{ij} = 1$ if regions i and j are contiguous, and $w_{ij} = 0$ otherwise. The matrix is then row-standardized to ensure comparability across regions with different numbers of neighbors. All computations were implemented in Python using the *libpysal* library (PySAL Development Team) for spatial weights construction, a custom NumPy-based implementation for the Multivariate Moran's I computation (Wartenberg, 1985; Yamada, 2024), and the *NetworkX* library for network construction and centrality measures. This specification is widely used in spatial econometrics and is consistent with the geographical structure of US counties. The choice of a contiguity-based specification allows for a direct interpretation of spatial interactions and ensures consistency with established practices in spatial econometrics, particularly in light of the known sensitivity of Moran's I to the choice of the spatial weights matrix (Yamada, 2024).

While the present analysis focuses on a contiguity-based approach, alternative specifications – such as inverse distance matrices or k-nearest neighbors – represent valid extensions and may be explored in future research to assess the robustness of the results.

By incorporating multiple variables, Multivariate Moran's I provides a nuanced understanding of how different economic sectors co-locate and interact, shaping income patterns across geographic regions. Therefore, it is particularly relevant for examining the interrelationship between employment distribution and the economic sectors classified under NAICS.

Overall, combining spatial autocorrelation measures – both univariate and multivariate Moran's I – with NAICS classifications provides a comprehensive approach to examining the spatial dynamics of income distribution in the USA. The methodological framework is designed to ensure transparency and replicability, with clearly defined data sources, spatial structures and network construction procedures. This approach allows for the identification of regional clusters and the exploration of sectoral interdependencies, which are assessed through a correlation matrix highlighting key NAICS sector pairs and further illustrated using network representations. Together, these tools inform the design of targeted economic policies and interventions, with particular attention to sector pairings and the industrial clusters they form. The following section presents the results derived from this analytical framework, focusing on insights drawn from the correlation matrix and network representation to highlight the most significant industrial clusters.

4. Insights obtained from the Moran's I analysis

The application of the Multivariate Moran's I methodology to the spatial distribution of economic sectors across US counties, based on the NAICS, reveals a wide range of meaningful correlation patterns that highlight both clustering tendencies and geographic specialization. Overall, the analysis uncovers a complex structure of intersectoral dependencies, manifesting as either positive spatial associations (indicating co-location industries with a similar economic strength) or negative correlations, reflecting structural and locational divergence. As Multivariate Moran's I captures spatial co-variation rather than causal relationships, the following interpretations should be understood as indicative of potential spatial patterns and structural associations, rather than as evidence of causal spillovers or direct economic interactions.

To construct the network representation, the matrix of spatial correlations derived from the Multivariate Moran's I is interpreted as a weighted adjacency matrix. Under this approach, spatial correlation is treated not only as a statistical measure but also as a structural relation, where Moran-based dependencies translate into network linkages. Each node corresponds to a NAICS sector, while edges capture the strength of spatial correlation between sectors. To enhance interpretability, only statistically significant correlations above a given threshold are retained. This makes it possible to identify, from the resulting network structure, groups of strongly interconnected sectors. The threshold is selected to balance network density and interpretability, thereby avoiding both overly dense and excessively sparse structures. While alternative thresholds may affect the number of links, the main structural patterns remain qualitatively stable.

A noteworthy insight concerns the presence of positive spatial correlations among several complementary economic sectors. Given the sensitivity of Moran-type statistics to the underlying spatial weight matrix and data structure, correlation magnitudes are interpreted in relative rather than absolute terms. All values should therefore be understood within the overall distribution of pairwise correlations observed across the 14 sectors included in the analysis. Among the most notable positive associations, Arts, Entertainment and Recreation (NAICS 71) displays a robust tendency to co-locate with Information (NAICS 51) and Real Estate and Rental and Leasing (NAICS 53), suggesting shared urban and metropolitan spatial footprints. This pattern is consistent with the evidence from Nakajima *et al.* (2012), who find that consumer services such as accommodation and arts-related activities exhibit stronger clustering than manufacturing sectors at close distances in US urban areas. Construction (NAICS 23) shows strong positive spatial associations with other services (NAICS 81) and wholesale trade (NAICS 42), a pattern consistent with the co-location of building activity and its supporting trade and service functions. This finding aligns directly with Ellison, Glaeser and Kerr (2010), who identify input–output linkages as the single most important driver of pairwise co-agglomeration among US industries: Construction and Wholesale Trade occupy adjacent positions in the building materials supply chain, generating structural incentives for co-location that are reflected in their positive spatial cross-correlation. Similarly, health care and social assistance (NAICS 62) displays a marked positive association with retail trade (NAICS 44–45), reflecting their shared presence in suburban and mixed-use environments. This spatial co-distribution is consistent with Matti (2024), who demonstrates that health care service locations in US counties are primarily driven by residential population density, which is the same demand-side locational logic that shapes retail trade distribution. The broader pattern of consumer-service clustering across US urban areas is further documented by the spatial analysis of Wang and Wen (2021), who find that health care, retail trade and accommodation and food services systematically co-locate within the same employment clusters at the sub-metropolitan level, precisely mirroring the positive associations uncovered here. While these patterns are consistent with the presence of complementarities and potential spillovers, they do not constitute direct evidence of such mechanisms, which would require formal spatial econometric modeling, such as Spatial Autoregressive (SAR) or SDM.

The analysis also reveals a set of negative spatial correlation patterns, indicating spatial decoupling between certain sectors. The most pronounced negative relationship emerges between arts, entertainment and recreation (NAICS 71) and health care and social assistance (NAICS 62), with a coefficient of approximately −0.07. This suggests that counties with a high concentration of arts and recreational activities tend to exhibit lower levels of health-care employment and vice versa. This divergence likely reflects different locational logics: arts-related activities tend to concentrate in dense urban cores and tourism-oriented areas, whereas health-care services follow broader population distributions, including suburban and peri-urban contexts. Information (NAICS 51) similarly displays negative spatial associations with both health care (NAICS 62) and retail trade (NAICS 44–45), consistent with its concentration in a limited number of high-density metropolitan clusters that differ structurally from the more spatially dispersed distribution of retail and health services. This structural divergence finds direct empirical support in the literature on the geography of US service sectors. Gutiérrez Posada *et al.* (2015), analyzing US county-level employment data, document that knowledge-intensive services (including information and professional services) agglomerate strongly in metropolitan cores from the 1980s onward, while population-serving activities such as health care and retail display a substantially more uniform spatial distribution. The negative spatial correlations identified in the present analysis are therefore not merely statistical artifacts, but reflect a documented structural divide in the locational logic of US service sectors. More generally, these results remain descriptive and do not constitute causal evidence of spatial specialization dynamics: establishing causal mechanisms would require the application of complementary spatial econometric approaches beyond the scope of this analysis.

The interpretation of these results is further enriched by the network analysis derived from the Multivariate Moran's I correlation matrix, which provides a structural representation of intersectoral dependencies. This section interprets the results through a dual analytical perspective, combining spatial econometrics and network theory. While Multivariate Moran's I captures the intensity of spatial co-variation and interdependencies, the network representation highlights the structural roles of sectors within the broader economic system. By modeling sectors as nodes within a graph and encoding their spatial correlations as weighted edges, it becomes possible to visualize the emergence of coherent economic clusters.

Four key groupings were identified through the analysis:

(1) Urban institutional services: including health care, professional services and other service activities. These are predominantly located in urban settings and characterized by a strong institutional and population-serving presence.

(2) Consumer-driven urban economy: comprising retail and wholesale trade, as well as accommodation and food services, emphasizing the consumer-oriented dimension of urban economic activity.

(3) Industrial agglomerations: encompassing manufacturing and construction, where industrial production and related activities are spatially concentrated.

(4) Urban financial and managerial hubs: covering finance, real estate and knowledge-intensive business services. They are centered around urban economic cores and metropolitan business environments.

Beyond this descriptive classification, the network perspective provides additional insights into the functional roles of these groupings. In particular, the identified clusters should not be interpreted merely as co-located sectors, but as interconnected sub-systems characterized by different degrees of internal cohesion and external connectivity. From a network standpoint, some clusters – such as industrial agglomerations and urban institutional services – function as relatively cohesive modules, where strong internal linkages support localized spillovers and specialization dynamics. In contrast, others, particularly those related to consumer-driven activities, play a more integrative role by linking production, distribution and final demand across regions. This distinction is crucial for understanding the structure of regional economic systems: highly cohesive clusters may enhance local productivity but may also reinforce sectoral lock-in, whereas more connected clusters contribute to system-wide integration and resilience by facilitating cross-sector interactions. As a result, the economic relevance of a sector cannot be assessed solely in terms of its size or correlation intensity but must also consider its position within the broader network structure.

The results also carry important policy implications. On the one hand, observed co-localization patterns underline the importance of strategic complementarities and infrastructure synergies; on the other hand, the spatial disjunctions between certain sectors reflect the enduring logic of territorial specialization. The policy implications are manifold. In particular, integrating spatial econometric evidence with network-based centrality measures allows policymakers to identify strategically relevant sectors not only by size but also by their systemic importance within the spatial economic structure. Regions characterized by strong intersectoral clustering may benefit from integrated infrastructure investments and coordinated development strategies. A network-based interpretation further suggests that policy interventions should be tailored to the structural role of sectors within the economic system. Policies targeting highly cohesive clusters may focus on enhancing innovation and specialization, while interventions aimed at bridging or highly connected sectors should prioritize maintaining intersectoral linkages and preventing systemic fragmentation. This approach highlights the importance of combining policies with a network-aware perspective to strengthen regional economic resilience.

Conversely, areas marked by economic divergence may require tailored interventions that account for the specific locational constraints of their dominant sectors. Ultimately, the Multivariate Moran's I approach offers a robust analytical framework for capturing the spatial logic underpinning economic structures. By quantifying interdependencies across sectors and situating them within the broader territorial context, it offers valuable insights for spatial planning, regional development and economic policy formulation.

The heatmap in Figure 2 illustrates the spatial correlations among 14 major U.S. economic sectors at the two-digit NAICS level, computed using the Multivariate Moran's I under Queen contiguity. Each cell represents the bivariate spatial correlation between a pair of sectors, with warmer colors (red tones) indicating positive co-location tendencies and cooler colors (blue tones) indicating spatial divergence. Consistent with the exploratory nature of the analysis, the correlations are moderate in magnitude throughout, ranging from approximately -0.07 to $+0.08$. This range reflects the inherent constraints of a contiguity-based spatial structure applied at the county level across a large and economically diverse country. Beyond individual pairwise relationships, the heatmap reveals the presence of structured correlation blocks, where groups of sectors display similar spatial behavior. These blocks provide a more aggregated view of intersectoral dependencies and allow for a direct comparison with the cluster configuration identified in the network analysis. The most pronounced positive correlations – shown in deeper red – emerge between sectors that share urban spatial footprints or supply-chain proximity. Notably, Arts, Entertainment and Recreation (71) exhibits a strong co-location pattern with Information (51) and Real Estate and Rental and Leasing (53), suggesting a tendency for these activities to cluster in economically dynamic urban environments. Similarly, Construction (23) shows strong

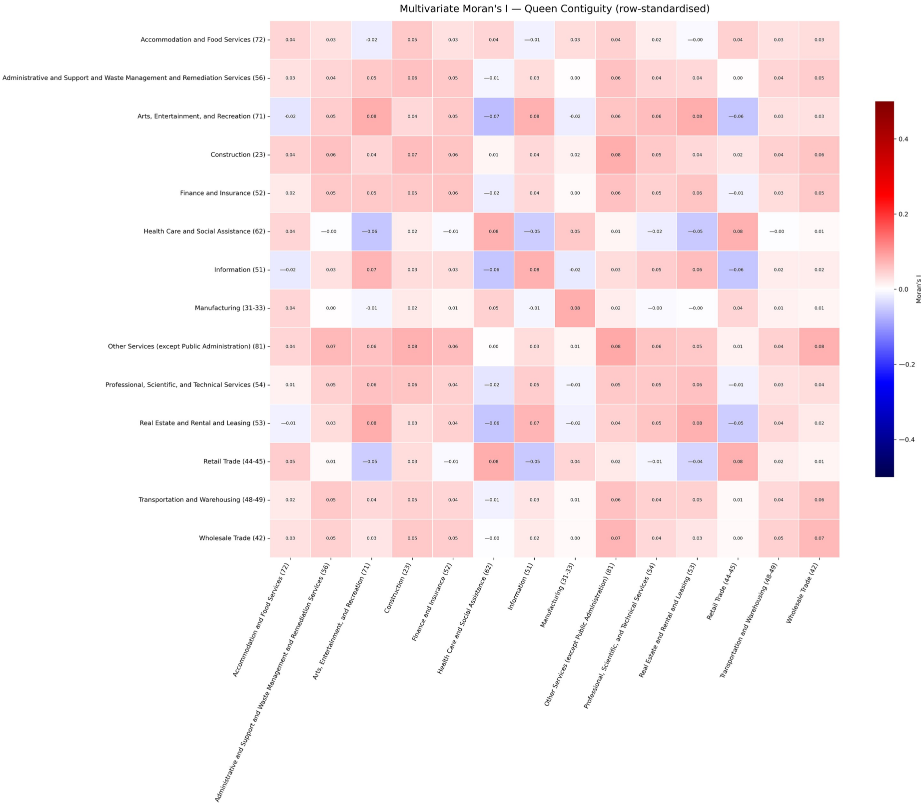

Figure 2. Multivariate spatial correlation matrix heatmap of US economic sectors
Source: Authors' editing

positive spatial association with Other Services (81) and Wholesale Trade (42), a pattern consistent with the co-location of building activity and the trade and service functions that support it. Health Care and Social Assistance (62), in turn, displays a marked positive association with Retail Trade (44–45), reflecting the shared suburban and mixed-use spatial contexts of these two sectors. When viewed as a whole, these positive associations form identifiable red blocks that largely correspond to the "urban financial and managerial hubs," "consumer-driven urban economy" and parts of the "urban institutional services" clusters described earlier. This consistency suggests that sectors grouped together in the network representation also exhibit similar spatial correlation patterns in the heatmap, reinforcing the robustness of the clustering structure. The most pronounced negative correlations – visible as blue cells – point to structural spatial divergence between sectors with contrasting locational requirements. The strongest negative relationship is observed between arts, entertainment and recreation (71) and health care and social assistance (62), with a coefficient of approximately −0.07, suggesting that counties with a high relative presence of arts and recreational activities tend to exhibit lower concentrations of health-care employment and vice versa. Information (51) also displays negative spatial associations with both health care (62) and retail trade (44–45), consistent with the concentration of information-sector activity in a limited number of high-density metropolitan clusters that structurally differ from the more spatially dispersed distribution of retail and health services. These negative relationships are not isolated but instead form broader blue regions within the matrix, separating groups of sectors with fundamentally different spatial logics. In particular, they reflect the divide between highly specialized, metropolitan-based activities and more population-serving or territorially distributed sectors, thereby capturing patterns of spatial specialization identified earlier in the analysis. Overall, the heatmap provides a comprehensive, matrix-based representation of spatial interdependencies across sectors, showing that these relationships are not isolated pairwise phenomena but part of a broader, structured pattern of associations. The presence of coherent blocks of positive and negative correlations suggests that sectors tend to organize into groups characterized by similar spatial behaviors, reflecting underlying economic structures rather than random variation. These patterns offer preliminary empirical support for the existence of distinct sectoral groupings within the US economic space, which are further explored in the subsequent analysis.

Figure 3 depicts the spatial autocorrelation structure of employment across US economic sectors, based on Multivariate Moran's I. Nodes represent sectors (classified according to NAICS codes), while edges capture statistically significant spatial correlations between sector pairs. Red edges denote positive associations, whereas blue edges indicate negative ones. Edges are retained only when the corresponding bivariate Moran's I is statistically significant at $p < 0.05$ (two-tailed, 999 permutations) and exceeds a minimum absolute threshold of $|I| \geq 0.05$, thereby retaining 74 of the 91 off-diagonal pairs. Edge thickness is proportional to the magnitude of the correlation, with thicker edges reflecting stronger associations.

The network structure visually supports the cluster configuration discussed in the previous section. Groups of closely connected nodes correspond to the identified economic clusters, including service-oriented sectors such as other services (NAICS 81), arts,

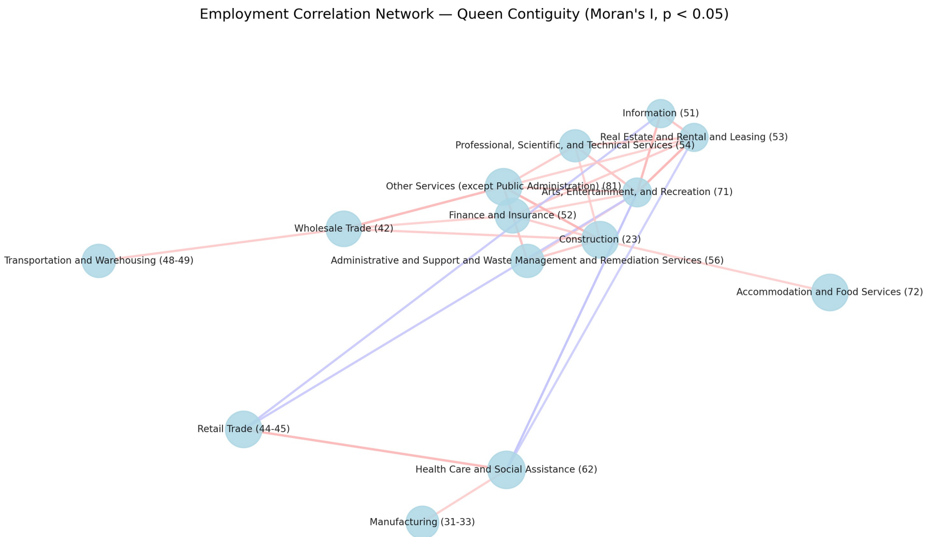

Figure 3. Employment spatial correlation network based on Moran's I with thresholding
Source(s): Authors' editing

entertainment and recreation (71), finance and insurance (52) and real estate and rental and leasing (53), alongside production-related groupings involving construction (23) and wholesale trade (42). These clusters appear as densely connected subgraphs, suggesting strong internal spatial interdependencies. The graph also highlights sectors that occupy relatively central positions within the network – most notably construction (23), wholesale trade, professional, scientific and technical services (54) and finance and insurance – which may reflect their structural role in linking production, distribution and service-oriented activities. However, this interpretation should be understood strictly in network-topological terms rather than as evidence of causal economic influence.

The analysis provides substantial insights into the structural roles played by different economic sectors within the inter-industry network, from which several important observations emerge. To further characterize the network, centrality measures are computed on a weighted, undirected graph, where edge weights correspond to the magnitude of spatial correlations. Normalization procedures are applied where appropriate – particularly for closeness centrality – to ensure comparability across sectors and to account for the size and topology of the network. Importantly, centrality should be interpreted as a structural property of the network rather than as a direct proxy for economic importance or sectoral performance.

First, sectors such as other services (81) and arts, entertainment and recreation (71) display consistently high centrality across multiple dimensions. Other services ranks first in both strength (weighted degree) and closeness, while arts, entertainment and recreation ranks highest in betweenness and eigenvector centrality, suggesting that it is both well embedded within influential clusters and structurally pivotal for information and resource flows across the network. Construction (23) and finance and insurance (52) also exhibit high scores across strength, eigenvector and betweenness centrality, suggesting a potentially foundational role in intersectoral linkages. These results should nonetheless be interpreted with caution, as centrality reflects network topology rather than causal economic influence; causal claims would require validation through spatial econometric models such as SAR, SEM or SDM.

Second, the analysis identifies sectors that may perform a bridging function within the network – most notably retail trade (44–45) and health care and social assistance (62). These sectors display exceptionally high betweenness centrality (ranking fifth and second, respectively) despite relatively low strength (14th and 13th) and negative eigenvector centrality scores. This indicates that their structural importance does not derive from the volume or intensity of their direct connections, but from their position as intermediaries linking otherwise disconnected parts of the network. From a spatial econometric perspective, this structural profile is consistent with the role of channels through which spatial co-variation patterns may propagate across regions and industries – although the direction and causality of such processes cannot be inferred from Moran's I alone. This interpretation aligns with the literature on knowledge networks and innovation systems, where intermediary nodes facilitate the diffusion of ideas, specialized skills and resources across otherwise disconnected industrial clusters. From a resilience perspective, however, the concentration of intersectoral connectivity in a limited number of bridging sectors may introduce structural vulnerabilities. Drawing on the concepts of path dependence and lock-in (David, 1985; Arthur, 1989), disruptions affecting high-betweenness sectors could plausibly propagate throughout the network, weakening intersectoral linkages and potentially leading to fragmentation. This highlights the importance of promoting diversification and redundancy in intersectoral connections as a means of mitigating the risks associated with economic lock-in.

Finally, several sectors – including administrative and support and waste management services (56), transportation and warehousing (48–49), manufacturing (31–33) and accommodation and food services (72) – display a dual profile: they exhibit relatively strong local connectivity, as reflected in their comparatively high closeness centrality combined with zero betweenness centrality. This suggests that they are well integrated within local industrial clusters but do not function as structural connectors between them. This distinction is analytically relevant for policy design. Sectors with high betweenness centrality represent both strategic opportunities and sources of systemic vulnerability: targeted investments in these sectors may yield system-wide benefits, while disruptions affecting them risk propagating across the network. Conversely, sectors with strong local integration but limited bridging capacity may contribute to cluster cohesion without significantly affecting cross-cluster connectivity.

From a methodological perspective, the interpretation of these centrality measures is inherently dependent on the normalization scheme used. This is particularly relevant for closeness centrality: without normalization, sectors in larger or more isolated components may appear artificially less central because of structural properties of the graph rather than genuine differences in connectivity. Normalization ensures more meaningful comparisons across sectors and helps identify those that are truly efficient in accessing the broader network. Ultimately, this analysis reveals the multifaceted nature of centrality in economic networks, offering a nuanced understanding of industrial interdependencies that can inform targeted interventions – whether for economic stimulus, risk mitigation or strategic investment – while remaining grounded in the descriptive and exploratory nature of the multivariate spatial approach used.

5. Conclusions

The concept of spatial autocorrelation is fundamental to understanding the interconnections between economic sectors within geographically defined regions. This study shows how spatial correlation is associated with patterns of economic organization linked to growth, innovation and resilience, without implying direct causal relationships. Empirical evidence from industrial clusters across the USA – such as Silicon Valley for technology, the aerospace industry in Huntsville and Colorado Springs, the Rust Belt for manufacturing and the financial sector in New York – highlights the crucial role of geographic concentration in driving productivity, competitiveness and overall economic development (Saxenian, 1996; Garner Economics, 2017; Warren, 2020; Anagnostou, 2015). The empirical analysis broadly supports the three hypotheses advanced in the introduction: positive spatial cross-correlation is confirmed among production–chain linked sectors, the information sector displays negative associations with population-serving activities consistent with divergent locational logics, and sectors occupying bridging network positions exhibit high betweenness centrality despite low direct correlation intensity – a pattern that carries the theoretical implication discussed below. From a methodological perspective, the application of spatial econometric measures, particularly Multivariate Moran's I, provides a robust framework for assessing spatial interdependencies among economic sectors. As demonstrated by Yamada (2024), this approach captures multidimensional spatial clustering and its implications for economic performance. While effective in identifying broad spatial patterns, it also presents limitations, including computational complexity and difficulties in isolating sector-specific relationships (Nguyen and Vu, 2019; LeSage and Pace, 2009). To address these issues and strengthen causal interpretation, complementary models such as the SDM (Elhorst, 2014) can be used.

Beyond the USA, existing research emphasizes the broader relevance of economic complexity and sectoral interdependencies for inclusive and sustainable growth (Ketu and Ningaye, 2024; Stojkoski *et al.*, 2023). Evidence from Germany and Portugal similarly confirms that industrial clustering plays a key role in shaping regional GDP dynamics (Jeleskovic and Loeber, 2023; Martinho, 2011), further reinforcing the general applicability of spatial correlation in economic analysis.

Integrating spatial autocorrelation measures with industry classification systems such as the NAICS further enhances the understanding of sectoral dynamics. This combined framework allows for a structured assessment of regional economic variations and supports the identification of development patterns (Kort, 1997), providing a solid analytical foundation for policy design.

From a policy perspective, the network structure identified in this study offers a basis for designing targeted and differentiated interventions grounded in observed sectoral interdependencies. Rather than relying on generic development strategies to be uniformly applied across space, policy design should be informed by the specific patterns of cross-sectoral spatial dependency revealed by the analysis.

The empirical evidence presented here allows for four distinct, actionable lines of policy intervention, differentiated by the spatial and structural profile of the targeted region. First, for regions characterized by strong positive spatial correlation between construction (NAICS 23) and wholesale trade (NAICS 42) – indicative of active production-chain co-agglomeration – targeted policy interventions should focus on deepening input–output integration within these clusters through supply chain coordination programs, shared logistics infrastructure and dedicated industrial zones that co-locate contractors, materials suppliers and business service providers. The empirical finding that this pair exhibits one of the highest bivariate Moran's I values in the analysis implies that the productive spillovers between these two sectors are spatially concentrated and potentially self-reinforcing: policies that reduce internal transaction costs within the cluster (such as dedicated freight corridors, interoperable procurement platforms, or joint workforce training programs across NAICS 23 and NAICS 42) are likely to generate returns that exceed those achievable through sectoral support in isolation. This approach reflects the evidence from Neumark and Simpson (2014) that place-based industrial policies generate the greatest returns when they reinforce existing agglomeration tendencies rather than attempting to generate clusters from scratch in greenfield locations. Second, the strong positive spatial association between retail trade (NAICS 44–45) and health care and social assistance (NAICS 62) in suburban and mixed-use environments points to a specific opportunity for integrated service planning. Policymakers in counties characterized by this co-location pattern should consider mixed-use zoning frameworks that allow health and retail facilities to share parking infrastructure, accessibility services and digital scheduling systems, reducing duplication of investments and enhancing service accessibility for residents. Crucially, the spatial distribution of both sectors follows residential demand patterns (Matti, 2024), meaning that access inequalities are systematically reproduced in underserved communities. Data-driven planning tools that combine Moran's I spatial clustering indicators with health service area and retail trade area boundaries can help identify specific counties where this co-location is absent, signaling unmet demand and justifying targeted infrastructure investment or regulatory incentives to attract both retail and health care providers simultaneously. Third, the strong negative spatial correlation between Information (NAICS 51) and both health care and retail trade reveals a structural divide between knowledge-intensive metropolitan clusters and population-serving territorial economies. Regions on the "wrong side" of this divide (counties with high health care or retail employment but low information-sector presence) face the structural constraints documented by Gutiérrez Posada *et al.* (2015): persistent spatial exclusion from the agglomeration dynamics that underpin high-productivity growth. For these regions, policies focused on digital diffusion and technology adoption are more likely to succeed than attempts to directly attract information sector firms, which are subject to strong lock-in effects in existing metropolitan clusters. Specifically, public investments in broadband infrastructure, digital skills training programs within community colleges and innovation vouchers for small firms in health care and retail to adopt digital coordination tools represent actionable interventions that can generate productivity gains within the existing sectoral structure of these counties, rather than requiring a structural transformation that empirical evidence suggests is difficult to engineer. Fourth, the diversification and regional growth literature (Incoronato, 2024; Neumark and Simpson, 2014) emphasizes that manufacturing-based clusters risk “boom-and-bust” trajectories if they fail to develop complementary sectors over time. For US counties exhibiting high positive spatial correlation between Manufacturing (NAICS 31–33) and adjacent production sectors, this finding implies that industrial policy should not be limited to supporting the dominant sector but should actively encourage related diversification – for instance, by supporting the emergence of professional and technical services (NAICS 54) and finance and insurance (NAICS 52) activities that increase the network centrality of the cluster and reduce its vulnerability to sector-specific shocks. The network centrality analysis presented in this paper provides a quantitative basis for identifying which sectors, if introduced or strengthened within a given regional cluster, would generate the greatest improvement in overall network connectivity and resilience.

Network centrality analysis further refines these policy implications by identifying distinct functional roles of sectors within the economic system. Core sectors, characterized by high strength and eigenvector centrality – particularly other services (NAICS 81), construction (NAICS 23) and finance and insurance (NAICS 52) – represent key anchors of regional production systems. Policy support for these sectors should go beyond generic infrastructure investments and instead focus on sector-specific bottlenecks: for construction, this means streamlining permitting processes and investing in skilled trades training; for finance and insurance, it means improving regulatory environments and digital financial infrastructure that extend credit access to underserved industries within the cluster. In contrast, sectors with high betweenness centrality despite low direct connection intensity – specifically retail trade (NAICS 44–45) and health care and social assistance (NAICS 62) – serve as structural bridges connecting otherwise disconnected parts of the economic network. These sectors represent critical leverage points where targeted investments generate system-wide multiplier effects. For retail trade, investments in omnichannel logistics infrastructure and last-mile connectivity directly strengthen its bridging function between production and final demand. For health care, investments in interoperable health information systems and community health worker programs enhance the sector’s capacity to connect populations across diverse economic geographies, thereby reinforcing the intersectoral connectivity it provides. However, the concentration of intersectoral connectivity in a limited number of bridging sectors also implies systemic vulnerability: as David (1985) and Arthur (1989) emphasize, path-dependent lock-in can rapidly propagate through network connections, meaning that disruptions to high-betweenness sectors risk generating cascading fragmentation. Policies aimed at increasing redundancy – for instance, by developing secondary bridging sectors such as transportation and warehousing (NAICS 48–49) or administrative services (NAICS 56) as alternative connectors – are therefore essential to mitigate structural fragility and support long-run regional economic resilience.

Taken together, these findings highlight the importance of adopting a network-based and spatially grounded perspective for understanding regional economic dynamics. By combining spatial econometrics with network analysis, this study provides a comprehensive framework for identifying both opportunities and vulnerabilities in complex economic systems, offering valuable guidance for future research and policy design.

Funding
The authors were solely responsible for the design, execution and analysis of the study. No external body has supported this research.

Ethics statement
This article does not contain any author-led studies involving human participants.

Authors contributions
The authors were solely responsible for the conception and design of the study, the acquisition and analysis of the data and the drafting of the manuscript.

Data availability
Data and materials supporting the results of this study are available from the authors upon reasonable request.

Consent to participate
Not applicable as this study did not involve human participants.

Consent to publish
Not applicable as this study did not involve human participants.

Note

[1.] Further theoretical foundations include Leontief's (1936) Input–Output framework, which formalizes intersectoral linkages through production and exchange relationships, and Krugman's (1991) New Economic Geography approach, which emphasizes the role of spatial clustering in enabling industries to exploit economies of scale and reduce production costs.

References

Anagnostou, L. (2015), *The New York City Financial Services Cluster – Research Paper*, Darla Moore School of Business, University of SC.

Anderson, J.E. and van Wincoop, E. (2004), "Trade costs", *Journal of Economic Literature*, Vol. 42 No. 3, pp. 691-751, doi: 10.1257/0022051042177649.

Anselin, L. (1995), "Local indicators of spatial association—LISA", *Geographical Analysis*, Vol. 27 No. 2, pp. 93-115.

Arthur, W.B. (1989), "Competing technologies, increasing returns, and lock-in by historical events", *The Economic Journal*, Vol. 99 No. 394, pp. 116-131.

Audretsch, D.B. and Belitski, M. (2022), "The knowledge spillover of innovation", *Industrial and Corporate Change*, Vol. 31 No. 6, pp. 1329-1357.

Balland, P.A., Boschma, R. and Frenken, K. (2015), "Proximity and innovation: from statics to dynamics", *Regional Studies*, Vol. 49 No. 6, pp. 907-920.

Bawa, S., Benin, I.W. and Almudaihesh, A.S. (2024), "Innovation networks and knowledge diffusion across industries: an empirical study from an emerging economy", *Sustainability*, Vol. 16 No. 24, p. 11308.

Boschma, R., Minondo, A. and Navarro, M. (2012), "Related variety and regional growth in Spain", *Papers in Regional Science*, Vol. 91 No. 2, pp. 241-257.

David, P.A. (1985), "Clio and the economics of QWERTY", *The American Economic Review*, Vol. 75 No. 2, pp. 332-337.

Elhorst, J.P. (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, Springer, Vol. 479, p. 480.

Ellison, G., Glaeser, E.L. and Kerr, W.R. (2010), "What causes industry agglomeration? Evidence from coagglomeration patterns", *American Economic Review*, Vol. 100 No. 3, pp. 1195-1213, doi: 10.1257/aer.100.3.1195.

Frigant, V. and Lung, Y. (2002), "Geographical proximity and supplying relationships in modular production", *International Journal of Urban and Regional Research*, Vol. 26 No. 4, pp. 742-755.

Garner Economics (2017), *The Aerospace Industry Cluster in U.S. metros*, Garner Economics, LLC.

Guillain, R. and Le Gallo, J. (2010), "Agglomeration and dispersion of economic activities in and around Paris: an exploratory spatial data analysis", *Environment and Planning B: Planning and Design*, Vol. 37 No. 6, pp. 961-981.

Gutiérrez Posada, D., Rubiera Morrollón, F. and Vieira Rivas, D. (2015), "How services increased the economic gap between the rural and urban US", World Economic Forum, available at: https://www.weforum.org/stories/2015/07/how-services-increased-the-economic-gap-between-the-rural-and-urban-us/

Hope, A.C. (1968), "A simplified monte Carlo significance test procedure", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 30 No. 3, pp. 582-598.

Incoronato, L. (2024), "Place-based industrial policies and local agglomeration in the long run [job market paper]", available at: https://lincoronato.github.io/files/jmp_incoronato.pdf

Jackson, M.O. (2008), "Networks and economic behavior", *Annual Review of Economics*, Vol. 1 No. 1, pp. 489-513.

Jeleskovic, V. and Loeber, S. (2023), "How industrial clusters influence the growth of the regional GDP: a spatial-approach", arXiv Preprint, arXiv:2401.10261.

Johnson, B. (2008), "Cities, systems of innovation and economic development", *Innovation*, Vol. 10 Nos 2-3, pp. 146-155.

Kelton, C.M., Pasquale, M.K. and Rebelein, R.P. (2008), "Using the North American industry classification system (NAICS) to identify national industry cluster templates for applied regional analysis", *Regional Studies*, Vol. 42 No. 3, pp. 305-321.

Ketu, I. and Ningaye, P. (2024), "Sectoral employment shares shape economic complexity: empirical evidence from African countries", *Global Journal of Emerging Market Economies*, Vol. 16 No. 2, pp. 168-187.

Khan, M.S. and Siddique, A.B. (2021), "Spatial analysis of regional and income inequality in the United States", *Economies*, Vol. 9 No. 4, p. 159.

Kort, J.R. (1997), "The North American industry classification system in BEA's economic accounts", *Federal Register*, Vol. 249, pp. 82-227.

Krugman, P. (1991), "Increasing returns and economic geography", *Journal of Political Economy*, Vol. 99 No. 3, pp. 483-499, doi: 10.1086/261763.

Le Gallo, J. and Ertur, C. (2019), "Heterogeneous reaction versus interaction in spatial econometric regional growth and convergence models", In *Handbook of Regional Growth and Development Theories*, Edward Elgar Publishing, pp. 451-466.

Lécuyer, C. (2006), *Making Silicon Valley: Innovation and the Growth of High Tech*, MIT Press, pp. 1930-1970.

Leontief, W.W. (1936), "Quantitative input and output relations in the economic systems of the United States", *The Review of Economics and Statistics*, Vol. 18 No. 3, pp. 105-125.

LeSage, J. and Pace, R.K. (2009), *Introduction to Spatial Econometrics*, Chapman and Hall/CRC.

Liu, Y., Li, L. and Zheng, F.T. (2019), "Regional synergy and economic growth: evidence from total effect and regional effect in China", *International Regional Science Review*, Vol. 42 Nos 5-6, pp. 431-458.

Martinho, V.J.P.D. (2011), "Spatial autocorrelation and Verdoorn law in the Portuguese nuts III", arXiv Preprint, arXiv:1110.5578.

Matti, C. (2024), "The location of urban healthcare services: evidence from phoenix yelp reviews", *Southern Economic Journal*, Vol. 90 No. 3, doi: 10.1002/soej.12550.

Murphy, J.B. (1998), "Introducing the North American industry classification system", *Monthly Labor Review*, Vol. 121 No. 7, pp. 43-47.

Nakajima, K., Saito, Y.U. and Uesugi, I. (2012), "Measuring economic localization: evidence from Japanese firm-level data", *Journal of the Japanese and International Economies*, Vol. 26 No. 2, pp. 201-220, doi: 10.1016/j.jjie.2012.02.002.

Neumark, D. and Simpson, H. (2014), Place-based policies (NBER Working Paper No. 20049). National Bureau of Economic Research, doi: 10.3386/w20049.

Nguyen, T.T. and Vu, T.D. (2019), "Identification of multivariate geochemical anomalies using spatial autocorrelation analysis and robust statistics", *Ore Geology Reviews*, Vol. 111, p. 102985.

Nie, L., Gong, H., Zhao, D., Lai, X. and Chang, M. (2022), "Heterogeneous knowledge spillover channels in universities and green technology innovation in local firms: stimulating quantity or quality?", *Frontiers in Psychology*, Vol. 13, p. 943655.

Pathmanathan, D., Dabo, I.M., Khoo, T.H., Ali-Hassan, A. and Dabo-Niang, S. (2024), "Spatial principal component analysis and Moran statistics for multivariate functional areal data", arXiv Preprint, arXiv:2408.08630.

Pede, V.O., Florax, R.J.G.M., de Groot, H.L.F. and Barboza, G. (2021), "Technological leadership and sectorial employment growth: a spatial econometric analysis for U.S. counties", *Economic Notes*, Vol. 50 No. 1, p. e12178, doi: 10.1111/ecno.12178.

Porter, M.E. (1998), "Clusters and the new economics of competition", *Harvard Business Review*, Vol. 76 No. 6, pp. 77-90.

Rosenthal, S.S. and Strange, W.C. (2001), "The determinants of agglomeration", *Journal of Urban Economics*, Vol. 50 No. 2, pp. 191-229, doi: 10.1006/juec.2001.2230.

Sánchez Gamboa, J.M. and Taddei Bringas, C. (2014), "Regions and spatial distribution of economic activities in Sonora", *Estudios Sociales (Hermosillo, Son)* Vol. 22 No. 43, pp. 187-215.

Saxenian, A. (1996), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128, with a New Preface by the Author*, Harvard University Press.

Stojkoski, V., Koch, P. and Hidalgo, C.A. (2023), "Multidimensional economic complexity and inclusive green growth", *Communications Earth and Environment*, Vol. 4 No. 1, p. 130.

Tian, Z., Gottlieb, P.D. and Goetz, S.J. (2020), "Measuring industry co-location across county borders", *Spatial Economic Analysis*, Vol. 15 No. 1, pp. 92-113. doi: 10.1080/17421772.2020.1673898.

U.S. Bureau of Economic Analysis (2026), "Employment by county, metro, and other areas", available at: www.bea.gov/data/employment/employment-county-metro-and-other-areas

Varga, A. (2006), "The spatial dimension of innovation and growth: Empirical research methodology and policy analysis", *European Planning Studies*, Vol. 14 No. 9, pp. 1171-1186.

von Thünen, J.H. (1826), *Der Isolierte Staat in Beziehung Auf Landwirtschaft Und Nationalökonomie*, Perthes, Hamburg.

Wang, B. and Wen, B. (2021), "The spatial distribution of businesses and neighborhoods: what industries match or mismatch what neighborhoods?", *Habitat International*, Vol. 117, p. Article 102440, doi: 10.1016/j.habitatint.2021.102440.

Warren, W.J. (2020), *Beyond the Rust Belt–The Neglected History of Rural Wild West Industrialization after World War II*, Cornell University Press and Books, pp. 72-102, doi: 10.7591/9781501751318-006.

Wartenberg, D. (1985), “Multivariate spatial correlation: a method for exploratory geographical analysis”, *Geographical Analysis*, Vol. 17 No. 4, pp. 263-283.

Weber, A. (1929), *Theory of the Location of Industries*, University of Chicago Press.

Xu, X. and Zhang, Y. (2021), “Network analysis of corn cash price comovements”, *Machine Learning with Applications*, Vol. 6, p. 100140, doi: 10.1016/j.mlwa.2021.100140.

Xu, X. and Zhang, Y. (2023), “Spatial-Temporal analysis of residential housing, office property, and retail property price index correlations: evidence from ten Chinese cities”, *International Journal of Real Estate Studies*, Vol. 17 No. 2, pp. 1-13. doi: 10.11113/intrest.v17n2.274.

Xu, X. and Zhang, Y. (2024a), “Network analysis of comovements among newly-built residential house price indices of seventy Chinese cities”, *International Journal of Housing Markets and Analysis*, Vol. 17 No. 3, pp. 726-749, doi: 10.1108/IJHMA-09-2022-0134.

Xu, X. and Zhang, Y. (2024b), “Network analysis of price comovements among corn futures and cash prices”, *Journal of Agricultural and Food Industrial Organization*, Vol. 22 No. 1, pp. 53-81, doi: 10.1515/jafio-2022-0009.

Yamada, H. (2024), “Moran’s I for multivariate spatial data”, *Mathematics*, Vol. 12 No. 17, p. 2746, doi: 10.3390/math12172746.

Zhang, C., Luo, L., Xu, W. and Ledwith, V. (2008), “Use of local Moran’s I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland”, *Science of The Total Environment*, Vol. 398 Nos 1-3, pp. 212-221.

Zhang, C., Lv, W., Zhang, P. and Song, J. (2023), “Multidimensional spatial and correlation analysis and its application based on improved Moran I”, *Earth Science Informatics*, Vol. 16 No. 4, pp. 3355-3368.

Further reading

Bakshi, B.R. (1998), “Multiscale PCA with application to multivariate statistical process monitoring”, *AIChE Journal*, Vol. 44 No. 7, pp. 1596-1610.

Ionela, G.-P. and Bele, I. (2017), *Developing a Growth Pole: theory and Reality*, doi: 10.18515/dBEM.M2017.n01.ch22.

Glaeser, E. L. (Ed.) (2010), *Agglomeration Economics*, University of Chicago Press.

Hidalgo, C.A. and Hausmann, R. (2009), “The building blocks of economic complexity”, *Proceedings of the National Academy of Sciences*, Vol. 106 No. 26, pp. 10570-10575.

Higgins, B. (2017), “François Perroux”, In B. Higgins and D.J. Savoie (Eds), *Regional Economic Development*, Routledge, pp. 31-47.

Wang, J. and Kashlak, A.B. (2021), “Local statistics for spatial panel models with application to the US electorate”, arXiv Preprint, arXiv:2110.10622.

Corresponding author
Luigi Capoani can be contacted at: luigi.capoani@unive.it