NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ

Bulletin of Geography. Socio-economic Series, No. 66 (2024): 47-55
http://doi.org/10.12775/bgss-2024-0032

BULLETIN OF GEOGRAPHY. SOCIO–ECONOMIC SERIES

journal homepages:
https://apcz.umk.pl/BGSS/index
https://www.bulletinofgeography.umk.pl/

Exploring spatial distribution of retail stores in South Korea

Changro Lee

Kangwon National University, Department of Real Estate, Chuncheon, Gangwon-do, Republic of Korea,
e-mail: spatialstat@naver.com

How to cite:
Lee, C. (2024). Exploring spatial distribution of retail stores in South Korea. *Bulletin of Geography. Socio-economic Series*, 66(66): 47-55. DOI: http://doi.org/10.12775/bgss-2024-0032

Abstract. In location decisions, retailers generally choose to cluster stores to capture more customers. This study examines the spatial distribution of retail stores in South Korea, and provides empirical guidelines for practitioners in urban planning and property management. Three areas representing urban, semi-urban, and rural regions were chosen for the study and analyzed using the L-function and variogram analysis, which are standard evaluation tools for geographical clustering. Stores such as private academies were found to be located close to each other at the shortest distance, followed by real-estate offices and sports equipment stores. The L-function analysis showed that stores were clustered within a distance of up to 2 km in Seoul, 6 km in Gyeonggi province, and 10 km in Gangwon province. In addition, the identified scales correspond approximately to the spatial dependence ranges of land prices. These results are expected to be utilized to design commercial zones in urban planning and delineate cluster boundaries.

Article details:
Received: 04 January 2024
Revised: 04 September 2024
Accepted: 22 November 2024

Key words:
retail stores, geography, planning & development, clustering, spatial scale, L-function, variogram

Contents:

1. Introduction

Stores similar to each other tend to be located together in a close space. The geographical concentration of retailing activity and spatial scale of its clustering have been of great interest to government officials and business operators. If clusters on a relatively small scale are identified, the county or district governments may take an interest in those clusters and promote them by providing local-scale incentives. By contrast, agglomerations on larger scales may be encouraged by the central government or international institutions. Private business operators also determine where to open a new store by examining the geographical density and pattern of similar stores. Retailers locate their stores close to competitors to maximize agglomeration benefits or distantly from competitors to minimize competition.

This study explores the spatial patterns in the locations of retail stores (which include service providers and retailers of physical goods) in South Korea to provide insights to stakeholders in urban planning and property management. We use a toolbox for evaluating spatial concentration: Besag's L-function and variogram analysis. The extent of retail store clustering is measured at all relevant spatial scales using the L-function. Thereafter, the spatial dependence of land prices in the study area is investigated using variogram analysis. Finally, we examine the association between the clustering scales of stores and ranges of spatial dependence of land prices.

The locations of retail stores have primarily been studied from a business perspective, focusing on factors such as supply chain proximity for cost reduction and developing key performance indicators for different store locations, like sales per square meter or conversion rates. In contrast, this study examines the locational patterns in stores from a geographical perspective. Specifically, we investigate how stores are distributed across urban, semi-urban and rural landscapes, seeking patterns in their density. This study contributes to the existing literature in two significant ways. First, while prior geographical studies have mainly focused on the spatial patterns of economic activity in urbanized areas, this study expands the scope of geographical analysis to encompass semi-urban and rural regions. Second, this study attempts to connect the clustering scales of stores to the ranges of spatial dependence estimated from a land price analysis. Despite the critical role of land prices (and associated rents) in determining retail store locations, the association between the clustering scales and the land price ranges remains largely unexplored in the literature. By addressing these aspects, this study aims to fill a research gap in the understanding of retail spatial patterns from a comprehensive geographical perspective.

The remainder of this paper is organized as follows. In Section 2, we explain the concentration of retail stores and their spatial distribution. Section 3 describes the study areas and methods of spatial pattern analysis for evaluating store clustering. In Section 4, the application of the proposed methods to the study areas is presented along with a measurement of the extent of store clustering, and its implications. Section 5 provides a summary of the study and suggests directions for future research.

2. Literature review

2.1. Retail stores in a competitive environment

Retailing is a competitive low-margin business. However, most retail stores tend to be located near similar stores, leading to intense price competition in the market. This phenomenon has encouraged studies on why geographical clustering of stores is prevalent. According to the studies undertaken from this perspective, product assortment is the main driver that forces retailers to adopt a clustering strategy (Rooderkerk et al., 2013; Caro et al., 2014; Besbes & Sauré, 2016; Blanchet et al., 2016). The concentration of stores provides customers with more opportunities to find products they may like, thus attracting more customers. This effect is also referred to as the market size effect (Konishi, 2005). The more stores are located in a limited space, the more the market-size effect is likely to be. Another driver that retailers tend to group together is the presence of positive externalities: the more shops cluster, the greater the positive externalities such as increased foot traffic and a wider variety of nearby shops (Koster et al., 2019). However, price competition may simultaneously become fiercer as clustering intensifies (Netz & Taylor, 2002). Thus, identifying an optimal clustering level is often suggested as a path for future research in the literature.

The relationship between retail rent and distance to the city center has also been investigated extensively (You & Tseng, 2021; Orr & Stewart, 2022). For example, Ossokina et al. (2024) discovered a strong negative retail rent gradient from the city center to the outskirts, averaging -17% per 100 meters. This suggests that some peripheral

locations might become unviable for retail even at zero rent, resulting in their conversion to other types of property use.

Distance and time to travel to retail stores are other popular issues in the literature. Finding a relevant place near potential customers is considered to be the most significant driver of retail success. Customers always analyze travel time to a certain store and the money involved in the travel (Saini et al., 2010). Some customers find money less important than time because money is more fungible than time (Leclerc et al., 1995), whereas others think that money is more important because time is ambiguous to measure, and thus, they feel less accountable for how they spend their time (Okada, 2005). It can be inferred without much difficulty that a number of factors affect this trade-off between time and money allocation when a customer chooses a store. For example, customers view long distance and travel time less negatively when retailers guarantee the in-stock of products (Grewal et al., 2012).

At the micro level, a game theory model is often used to capture competition between stores. The authors in this line of research conducted and investigated a multiplayer location game in a closed-loop market and let each player choose an optimal point on the circle as its location such that its profit was maximized (Cabrera et al., 2009; Jang, 2015; Chen et al., 2020). In short, the market size effect of the number of, distance and travel time to, and optimal locations for stores have been prominent issues, and related studies have been undertaken mainly in the fields of economics and business management.

2.2. Spatial point pattern analysis of retail stores

Although previous studies on retail store locations are abundant, most have been carried out from the perspective of business management. Therefore, it is difficult to find studies focusing on the spatial point patterns in stores. Only a few studies have been carried out in the spatial analysis literature, the representative one being Duranton and Overman's (2008) study. They explored the point patterns of manufacturing industries in the UK and examined the difference in spatial distribution between manufacturers' types, such as affiliated and non-affiliated plants, or foreign-owned and domestic plants. After this study, two similar studies were undertaken: Jeong and Kim (2014) investigated spatial distribution of retail stores in South Korea and identified clustering patterns in food stores and dispersed patterns in sports equipment stores; Rodríguez Rangel et al. (2020) analyzed the intensity of tourist accommodations in Spain and identified a few highly clustered areas of accommodations.

Departing from spatial point pattern analysis, some studies have employed a network analysis approach. Wang et al. (2014) used a road network analysis tool and revealed that specialty stores preferred street centrality the most, followed by department stores and supermarkets. Rui et al. (2016) adopted the same network analysis approach and examined the competitive relationship between domestic and foreign brand chains in China. Recently, Tsoutsos and Photis (2020) collected image data from Google Street View and investigated the distribution of retail stores in cities of Greece. They found that population density and business characteristics significantly influenced the degree of clustering of retail stores.

This study was performed in the same manner as the point pattern analysis. However, this study focused on the scale of clustering in retail stores. That is, we attempted to identify different scales of clustering by region and industry type, which is clearly different from previous studies in the spatial analysis literature. Furthermore, we attempted to identify the association between the clustering scale of stores and land price distribution, which is expected to provide practical guidelines to stakeholders in the real-estate industry.

3. Data and methodology

3.1. Dataset and study area

Data on retail stores were collected from the Small Enterprise and Market Service (SEMS). SEMS is a division of the Ministry of Small Enterprises and Startups. SEMS uses the Standard Industrial Classification (SIC) codes to classify and support small business operators in South Korea. Information on three industries (three SICs) for 2021 was obtained from the SEMS: education and learning, real estate and sports (Note 1).

Three study areas were chosen for the analysis: Seoul, Gyeonggi province and Gangwon province. Seoul is the capital of South Korea, and is best described as a highly urbanized global city. Gyeonggi province is a region surrounding Seoul and is a mixture of urban and rural localities. Gangwon province is a region located east of Gyeonggi province that is best characterized by mountainous

and forested landscapes. These study areas were selected because they correspond to urban, semi-urban and rural territories, respectively, and thus are expected to show different distributions of stores depending on their local characteristics. Figure 1 illustrates these three study areas.

Figure 2 shows the locations of the educational and learning stores in the study areas. Stores of other types are not shown due to space limitations.

3.2. L-function and variogram

This study examines the spatial distribution of stores based on the distance between stores. First, the average distance from each store to its nearest store is calculated, which provides information on the locational proximity of stores in the same industry. Second, Besag's L-function (Besag, 1977) is used to describe the clustering degree in stores for all distance bands. Third, we adopt a variogram to measure the degree of spatial dependence of land prices and identify the association between the scale of store clustering and spatial range of land prices.

The L-function is estimated as follows: first, the number of stores at different distance bands for each store is counted; second, the mean of the counts for each distance band is calculated; and third, the mean is divided by the overall store density. The overall store density is often denoted as λ=N/A, where N is the total number of stores and A is the area of the study region. The function calculated thus far is referred to as the K-function. In short, the K-function is the ratio of the mean of the counts for each distance band to the overall store density. K values greater than K_expected indicate a clustering of data points at a given distance, whereas K values less than K_expected indicate a dispersion. K_expected is a function generated under the complete spatial randomness (CSR) assumption (Note 2). The K-function equation is as follows:

$$\hat{K}(d) = \frac{1}{\hat{\lambda}^2 \cdot A} \sum_{i=1}^{N} \sum_{i \neq j} I_d(d_{ij}) \quad (1)$$

where $I_d(d_{ij})$ is an indicator function that takes the value of 1 if the distance between two points i and j is less than or equal to d, and 0 otherwise. d represents a specified distance or distance threshold: it is measured in meters in this study. N is the total number of data points (stores) in the study area. Both i and j represent individual data points, specifically each store in this study. The L-function is simply an arithmetic transformation of K and $K_{expected}$ so that the latter always lies on a diagonal line. The equation used is as follows:

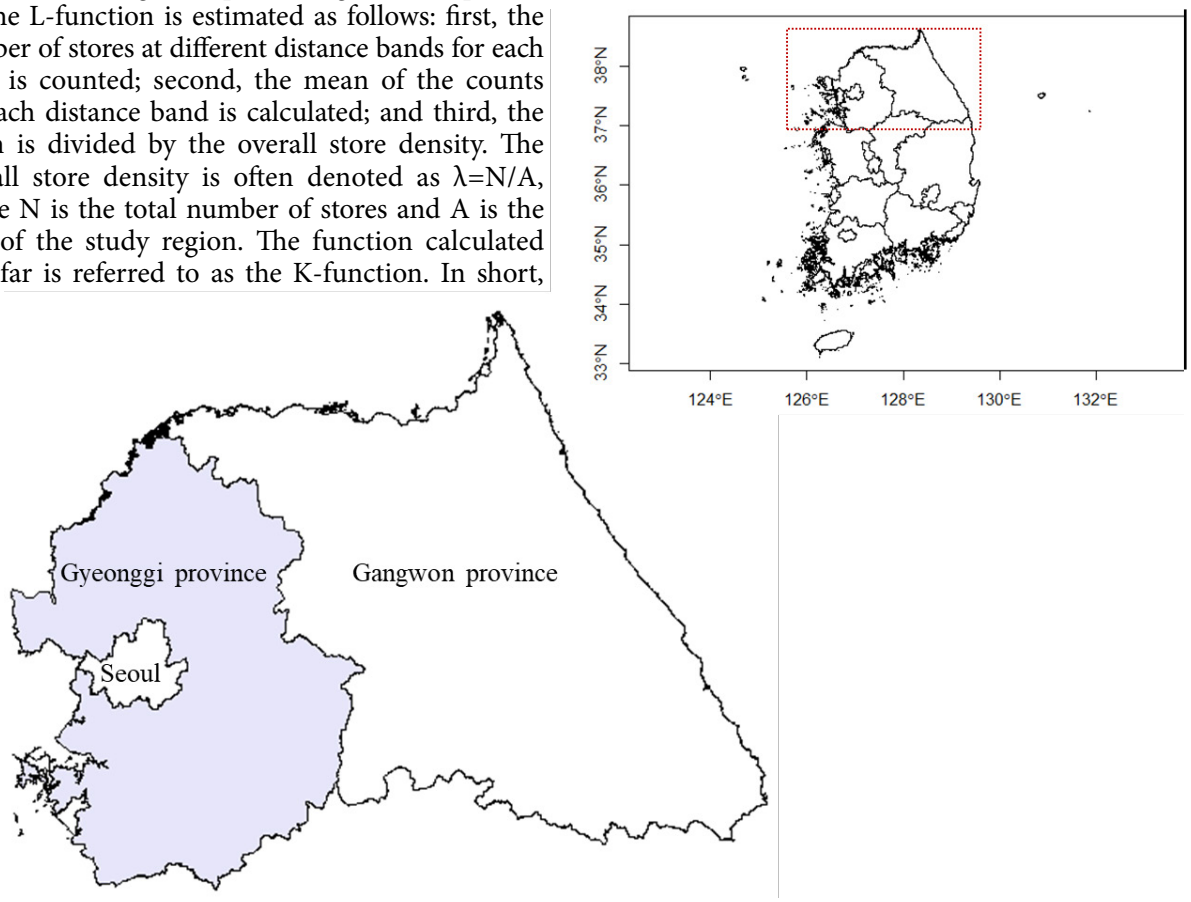

Fig. 1. Study areas
Source: author

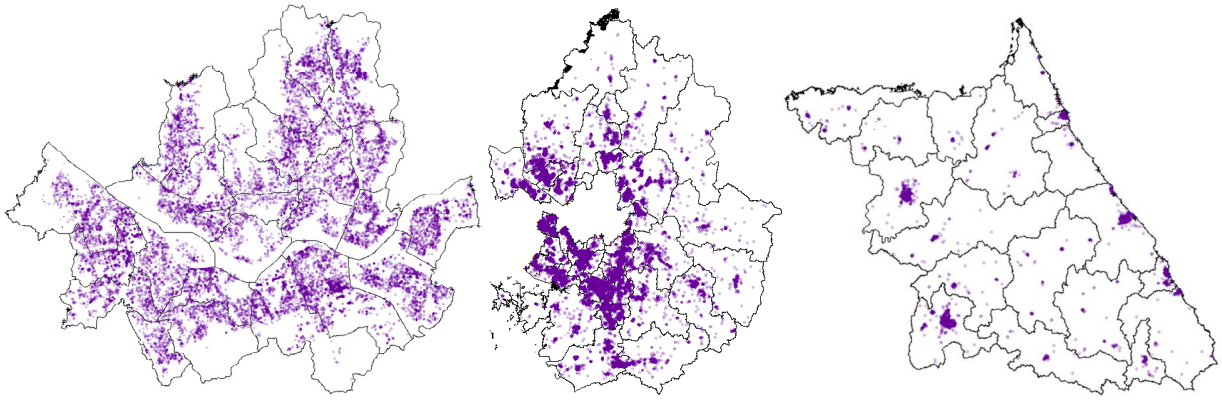

Fig. 2. Locations of educational and learning stores: Seoul (left), Gyeonggi province (middle), Gangwon province (right)
Source: author

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d \qquad (2)(\text{Note 3})$$

This transformation makes it easy to discern even a small difference between K and $K_{expected}$. From L values above a diagonal line, clustering is inferred at distance d, whereas L values below the diagonal line indicate a dispersion (inhibition) at distance d (Kosfeld et al., 2011).

A variogram is a function that depicts the degree of the spatial correlation of a spatial phenomenon. It shows differences in pairs of data points separated by lag distance. These lags are analyzed for their mean squared differences. The degree of dissimilarity between $Z(x)$ and $Z(x+h)$ is formulated using the variogram function $\gamma(h)$ as follows:

$$2\gamma(h) = Variance\,[Z(x+h) - Z(x)] \qquad (3)$$

where $Z(x)$ is a spatial process at location x and h indicates the lag distance. If a process, such as land price formation, has a strong spatial correlation, $\gamma(h)$ will increase, eventually reaching a saturation point. If $\gamma(h)$ displays a flat pattern, it indicates spatial independence of the process (Note 4).

This study investigates whether land prices in a study area have a spatial correlation. Because data from the SEMS do not contain land price information, the prices of benchmark lots for 2021 were obtained from the Ministry of Land, Infrastructure and Transport (MOLIT) (Note 5) and used to compute variogram functions. MOLIT publicly releases annual price data for these benchmark lots. These lots are sampled across the entire territory of South Korea, and their prices are surveyed and determined by real-estate experts, including property appraisers and brokerage agents. The number of benchmark lots collected from the web page is 4,737 in Seoul, 3,605 in Gyeonggi province and 1,081 in Gangwon province. Where land prices are found to have a spatial correlation, the association between this correlation and the scale of store clustering is further analyzed.

4. Results and implications

4.1. Results

Table 1 shows the average distances between the nearest neighbor stores. The distances tend to increase as industry categories progress from education and learning to real estate to sports. Educational and learning stores are located close to each other at the smallest scale, which indicates that stores such as private academies and study rooms capture the market size effect (product assortment) most effectively by clustering in a limited space. Students usually require educational services for several subjects, from language to mathematics, and the clustering of academies providing educational services for different subjects can attract more students to an academy-clustered area. By contrast, sports equipment stores are located close to each other at the largest scale, implying that these stores tend to avoid price competition as much as possible.

Table 2 and Figure 3 show the spatial scales identified using the L-functions. In Figure 3, the L-functions in all the study areas lie above the diagonal lines, indicating a significant clustering tendency of retail stores. The strength of clustering was evaluated by the gap between the L-function

Table 1. Distances between the nearest neighbor stores

Industry category	Seoul	Gyeonggi province	Gangwon province
Education and learning	20 m (26 m*)	Less than 1 m (46 m)	18 m (161 m)
	(n = 22,765**)	(n = 40,451)	(n = 5,049)
Real estate	32 m (29 m)	35 m (60 m)	69 m (245 m)
	(n = 14,254)	(n = 24,096)	(n = 2,086)
Sports	79 m (54 m)	83 m (115 m)	154 m (374 m)
	(n = 4,636)	(n = 6,728)	(n = 1,060)

* Standard errors; ** the number of stores

Table 2. Spatial scales of stores and the range of land prices

Analysis method	Industry category	Seoul	Gyeonggi province	Gangwon province
L-function	Education and learning	2.0 km	6.8 km	10.0 km
	Real estate	2.0 km	5.9 km	10.0 km
	Sports	2.0 km	5.7 km	10.0 km
Variogram	Range of land prices	2.1 km	5.2 km	8.2 km

Source: own elaboration

and diagonal line; Gangwon province shows the strongest clustering tendency for all distance bands. Gangwon province is a rural region, and highly urbanized cities are rare and located sparsely across the region, as shown in Figure 2. Therefore, the strong clustering in Gangwon province could be attributed to the sparse distribution of the urbanized cities where the majority of stores are located.

In Figure 3, Seoul displays a growing spatial clustering in all three industry categories, within a radius of ~2 km. In Gyeonggi province, the L-functions reveal significant clustering at the medial spatial scale – that is, between 5.7 and 6.8 km, as also shown in Table 2. In Gangwon province, the L-functions increase sharply up to a distance of 8 km, and then plummet. In Gangwon province, the L-functions meet diagonal lines at ~10 km. Beyond this threshold, stores tend to be more dispersed than expected under the CSR assumption. In short, a clear clustering of stores is observed within a distance of up to 2 km in Seoul, ~6 km in Gyeonggi province, and 10 km in Gangwon province.

The bottom row of Table 2 shows ranges identified from the variogram analysis: 2.1 km, 5.2 km, and 8.2 km for Seoul, Gyeonggi province, and Gangwon province, respectively. Figure 4 shows the corresponding variogram graphs. The ranges in the variograms show values similar to those identified in the L-function analysis, as shown in Table 2. Thus, the radii of the circles around the stores are closely related to the ranges of spatial dependence of land prices. Stores appear to form clusters within a distance constrained by the spatial dependence of land prices. If the land price level is drastically different between two localities (i.e., no spatial dependence of prices between two areas), it is difficult to consider the two areas to belong to the same submarket, and, therefore, clustering is unlikely to occur in these areas.

4.2. Implications

The excessive education fever of the South Korean people is well known, making South Korea among the most highly educated countries in the world: in 2020, 64% of 25- to 34-year-old men had tertiary qualifications and 76% of their female peers (OECD, 2021). Accordingly, a large number of educational and learning stores, such as private academies, are enjoying a booming market, and such stores were found to form clusters at the smallest scale to maximize their market size effect. By contrast, sports equipment stores are clustered at the largest scale to alleviate price competition. These findings can provide an enhanced understanding of store management to future business owners,

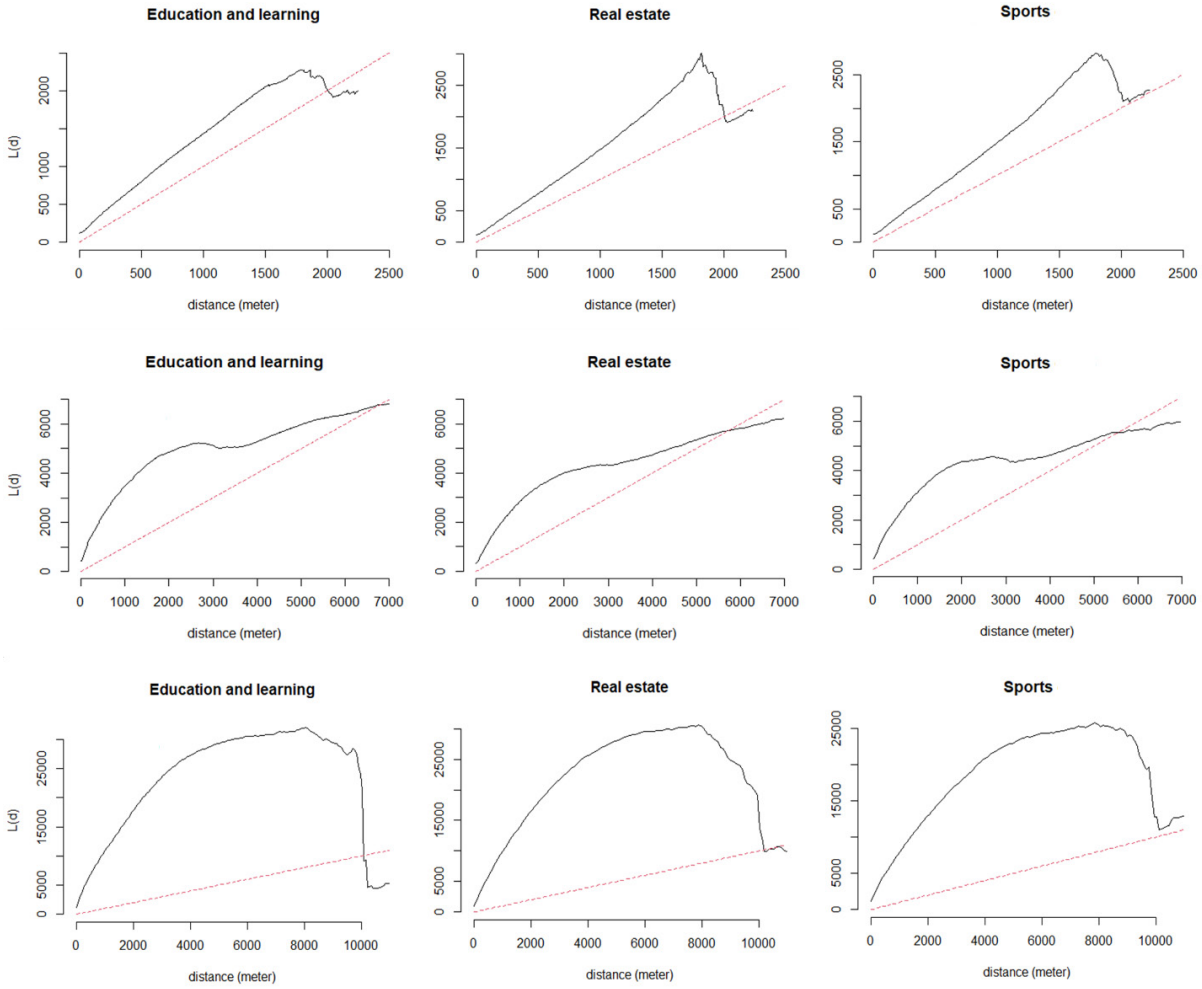

Fig. 3. L-functions by region and industry category: Seoul (upper), Gyeonggi province (middle), Gangwon province (bottom)
Source: own elaboration

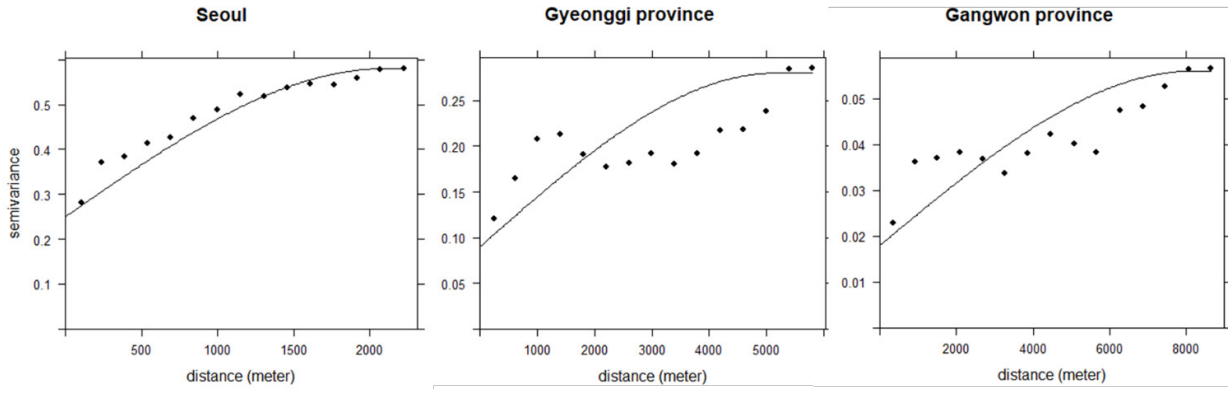

Fig. 4. Empirical (dot) and theoretical (curve) variograms of land prices
Source: own elaboration

branch managers of franchise business and property managers.

The clustering scales were found to vary with the study area and industry category. As shown in Figure 3, the spatial scales of stores appeared to be more heavily influenced by the region than by the industry category. In addition, the results show that the clustering scales of stores identified by the L-function generally correspond to the ranges revealed by the variograms based on land prices. This indicates that the clustering of commercial properties tends to form at a distance band similar to the range of the spatial dependence of land prices. These findings can be utilized by local government officers and property developers when designing commercial zones in urban planning and delineating cluster boundaries.

5. Conclusion

Three study areas were chosen for the analysis of spatial store distribution: Seoul, Gyeonggi province and Gangwon province. In all the areas, educational and learning stores were found to be located close to each other at the shortest distance, followed by real-estate shops and sports equipment stores. According to the L-function analysis, clustering scales of stores were found to be 2 km, ~6 km and 10 km in Seoul, Gyeonggi province and Gangwon province, respectively. In addition, these scales generally corresponded to the spatial dependence ranges of land prices. These results are expected to provide useful guidelines for practitioners in urban planning and property management. For example, when property developers select locations for new educational facilities or learning centers, they should prioritize areas with existing clusters of similar business to leverage the clustering effect. In contrast, when choosing locations for real-estate shops and sports equipment stores, it is important to balance the benefits of being near similar businesses with the opportunity to tap into other market segments.

Only a few industry categories were analyzed in this study: education and learning, real estate, and sports. The study area was also limited to three regions. This was due to the availability of the data. Thus, the analysis in this study needs to be extended to other business types and areas to enhance the generalizability of the findings. An analysis of additional information, such as store size and sales volume, is also expected to provide a more accurate explanation of store distribution, if this information is available in future research.

Notes

1. The education and learning industry includes private academies, after-school tutoring services, nurseries, and study rooms; real-estate industry includes brokerages, housing marketing, property appraisal, consulting, construction, and facility management businesses; sports industry includes sporting goods shops, recreational goods shops, and related general merchandise stores.
2. In $\check{K}(d)$, the hat (^) indicates that this is the empirical K-function. While the original input variables are measured in meters, the output unit of the K-function (and L-function) does not correspond to meters. This is because its calculation involves summing up indicator functions and then normalizing them by the intensity of the point pattern and the study area's size. Only relative comparisons are meaningful when interpreting K-function (and L-function) output values. Refer to Dixon et al. (2012) for details of the CSR assumption.
3. By employing K(d) in equation (2), the transformed theoretical K-function is compared against the empirical K-function, denoted as K $\check{K}(d)$.
4. The factor of 2 in the variogram equation is a consequence of the statistical property of variance. Z(x+h) and Z(x) can be treated as two random variables. Thus, when the variance of their difference is calculated, this equation is derived: Variance [Z(x+h) - Z(x)] = Variance [Z(x+h)] + Variance [Z(x)] - 2 × Covariance [Z(x+h), Z(x)]. The "- 2 ×" in front of the covariance term is what leads to the 2γ(h) in the variogram equation. Refer to Isaaks and Srivastava (1989) and Gómez-Hernández (2005) for details on the variogram.
5. The dataset is available at: https://www.data.go.kr.

References

Besag, J. (1977). Contribution to the discussion on Dr Ripley's paper. *Journal of the Royal Statistical Society*, 39: 193–195.

Besbes, O. & Sauré, D. (2016). Product assortment and price competition under multinomial logit demand. *Production and Operations Management*, 25(1): 114–127.

Blanchet, J., Gallego, G. & Goyal, V. (2016). A markov chain approximation to choice modeling. *Operations Research*, 64(4): 886–905.

Cabrera, J.G., Smith, N.R., Kalashinkov, V. & Cobas-Flores, E. (2009). Mathematical model for retail stores location in competitive environments. *International Journal of Innovative Computing, Information and Control*, 5(9): 2511–2521.

Caro, F., Martínez-de-Albéniz, V. & Rusmevichientong, P. (2014). The assortment packing problem: Multiperiod assortment planning for short-lived products. *Management Science*, 60(11): 2701–2721.

Chen, X., Song, Q. & Zhao, Z. (2020). Two-Player Location Game in a Closed-Loop Market with Quantity Competition. Complexity, 2020(1): 4325454.

Dixon, P.M., El-Shaarawi, A.H. & Piegorsch, W.W. (2012). Ripley's K function. *Encyclopedia of Environmetrics*, 3: 1796–1803.

Duranton, G. & Overman, H.G. (2008). Exploring the detailed location patterns of UK manufacturing industries using microgeographic data. *Journal of Regional Science*, 48(1): 213–243.

Gómez-Hernández, J.J. (2005). *Geostatistics*. In: Hydrogeophysics. Springer, Dordrecht.

Grewal, D., Kopalle, P., Marmorstein, H. & Roggeveen, A.L. (2012). Does travel time to stores matter? The role of merchandise availability. *Journal of Retailing*, 88(3): 437–444.

Isaaks, E.H. & Srivastava, M.R. (1989). Applied geostatistics, 551.72 ISA.

Jang, M.S. (2015). Optimum location in spatial competition. Master's Thesis. University of Seoul.

Jeong, D.S. & Kim, H.B. (2014). An analysis of the distribution and factors influencing sales by business types. *GRI Review*, 16(2): 101–122.

Konishi, H. (2005). Concentration of competing retail stores. *Journal of Urban Economics*, 58(3): 488–512.

Kosfeld, R., Eckey, H.F. & Lauridsen, J. (2011). Spatial point pattern analysis and industry concentration. *The Annals of Regional Science*, 47(2): 311–328.

Koster, H.R., Pasidis, I. & van Ommeren, J. (2019). Shopping externalities and retail concentration: Evidence from Dutch shopping streets. *Journal of Urban Economics*, 114: 103194.

Leclerc, F., Schmitt, B.H. & Dube, L. (1995). Waiting time and decision making: Is time like money? *Journal of consumer research*, 22(1): 110–119.

Netz, J.S. & Taylor, B.A. (2002). Maximum or minimum differentiation? Location patterns of retail outlets. *Review of Economics and Statistics*, 84(1): 162–175.

OECD, (2021). Education GPS. Available at: http://gpseducation.oecd.org (Accessed: 25 August 2022).

Okada, E.M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of marketing research*, 42(1): 43–53.

Orr, A.M. & Stewart, J.L. (2022). Property use diversity and spatial accessibility within urban retailing centres: drivers of retail rents. *Journal of Property Research*, 39(4): 365–392.

Ossokina, I.V., Svitak, J. & Teulings, C.N. (2024). The urban economics of retail. *Regional Science and Urban Economics*, 104026.

Rodríguez Rangel, M.C., Sánchez Rivero, M. & Ramajo Hernández, J. (2020). A spatial analysis of intensity in tourism accommodation: An application for extremadura (Spain). *Economies*, 8(2): 28.

Rooderkerk, R.P., Van Heerde, H.J. & Bijmolt, T.H. (2013). Optimizing retail assortments. *Marketing Science*, 32(5): 699–715.

Rui, Y., Yang, Z., Qian, T., Khalid, S., Xia, N. & Wang, J. (2016). Network-constrained and category-based point pattern analysis for Suguo retail stores in Nanjing, China. *International Journal of Geographical Information Science*, 30(2): 186–199.

Saini, R., Rao, R.S. & Monga, A. (2010). Is that deal worth my time? The interactive effect of relative and referent thinking on willingness to seek a bargain. *Journal of Marketing*, 74(1): 34–48.

Tsoutsos, M.C. & Photis, Y.N. (2020). Spatial point pattern analysis of urban retail stores: the case of twelve large-and medium-sized Greek cities. *European Journal of Geography*, 11(4): 36–63.

Wang, F., Chen, C., Xiu, C. & Zhang, P. (2014). Location analysis of retail stores in Changchun, China: A street centrality perspective. *Cities*, 41: 54–63.

Yuo, T.S.T. & Tseng, T.A. (2021). The environmental product variety and retail rents on central urban shopping areas: A multi-stage spatial data mining method. *Environment and Planning B: Urban Analytics and City Science*, 48(8): 2167–2187.