P-ISSN: 2709-6254	Journal of Development and Social Sciences	April-June 2022,Vol. 3, No. 2
O-ISSN:2709-6262	http://dx.doi.org/10.47205/jdss.2022(3-II)87	[972-985]

JDSS Journal of Development and Social Sciences
www.jdss.org.pk

RESEARCH PAPER

The Economic Geography of Road infrastructure in Pakistan: Exploratory Spatial Data Analysis (ESDA)

[1]Sabila Khatoon [2]Dr. Hafsa Hina

1. PhD Scholar, Department of Economics and Econometrics, Pakistan Institute of Development Economics, Islamabad, Pakistan
2. Assistant Professor, PIDE School of Social Sciences, Pakistan Institute of Development Economics, Islamabad, Pakistan

PAPER INFO

Received: March 18, 2022
Accepted: June 25, 2022
Online: June 27, 2022
Keywords: Exploratory Spatial Data Analysis (ESDA), New Economic Geography (NEG), Regional Connectivity Road Infrastructure
***Corresponding Author:**
sabila_15@pide.edu.pk

ABSTRACT

This paper presents an exploratory Spatial Data Analysis of road infrastructure using data of the district profiles of road infrastructure developed by the Urban Unit of the National Transport Research Centre (NTRC) during FY 2019-20. The indicators of the spatial autocorrelation (global and local Moran's I) were estimated using the neighborhood definition of distance-based spatial weight matrix. Using inverse distance weights matrices at 3 and 5 bandwidths, results show significant positive spatial autocorrelation of road infrastructure at the national level of Pakistan, revealing that districts with a low density of road infrastructure are in the geographic vicinity of the districts of the less-dense road network. Likewise, the districts with more extensive road networks are neighboring the districts with large road networks. Furthermore, it is found that districts lying in the quadrant of the low-low road infrastructure are geographically located in the province of KPK, Sindh, and Baluchistan. While the districts lying in the high-high quadrants are located in the province of Punjab, Pakistan. In addition, this study found insignificant positive spatial autocorrelation of road infrastructure when analyzed at the provincial scale of Punjab, Sindh, KPK, and Baluchistan.

Introduction

Transport infrastructure is a precursor for regional connectivity and economic integration. The road infrastructure is the most prominent medium of in-land transportation services in Pakistan compared to other types of transport (rail, air, and seaport) infrastructures. The empirical literature has yet not answered the road infrastructure and its relationship with the geography of Pakistan; and this paper filling this research-gap and presenting exploratory research on the space-dependence of road infrastructure in Pakistan. This paper analyzes the space dependence and spatial inequities of road infrastructure in Pakistan using district-level data.

This paper employed the up-to-date dataset of road infrastructure that has been developed by the Urban Unit of the National Transport Research Centre (NTRC) using digital resources and on-ground assessments during FY 2019-20 (NTRC, 2020). This dataset is extremely helpful for the comparability across the provinces and districts as it used the same definitions of road networks across provinces. The mettled high type roads (HTR) data for the district of Punjab include the road length of the national highways, motorways, district roads (provincial highways, R&B Sector roads, farm to market roads, sugar-cess Roads, and District Council roads). Given this bifurcation based on the types, data is not available for the other provinces of Sindh, Khyber Pakhtunkhwa (KPK), and

Baluchistan. The provinces of Sindh, KPK, and Baluchistan used the definition of blacktop (mettled) road as the HTR, and the data for national highways and motorways was not available separately for each district.

This paper is organized into six sub-sections. The first is an introduction, the second is the literature review and research gap, the third is the estimation technique, the fourth is data, the fifth is the ESDA, and the last is the conclusion.

Literature Review

The roads are classified made not only based on the capacity (such as low and or high type road) and material (such as asphalt, concrete, Bituminous, Gravel, etc.) but also on the basic functioning provided in an economy (such as inter-city and intra-city roads). The impact of road infrastructure impacting on an economy depending on the capacity utilization of the road infrastructure. The planning for investing in the development and up-gradation of the road infrastructure (planning, analysis, and operations) focuses on the capacity of roads. However, the capacity of roads is dependent on the volumes of trade flows, width of roadway, magnitude of upgrade and its length (Arasan & Arkatkar, 2011).

The literature widely acknowledges the role of a better-functioned transport system that generates productivity growth and increases the nation's competitiveness. Quality transport infrastructure directly increases the competitiveness of nations, as it reduces spatial frictions. Spulber (2007) describes these spatial frictions as the four 'Ts.' These are transaction costs, Tariff and nontariff costs, transport costs, and time costs. The quality of transport infrastructure reduces these four Ts, leading to improved competitiveness (Lafourcade & Thisse, 2008). In addition, it indirectly landscapes the economic geography via indirect spatial spillovers by generating the locational effects via the connectivity, accessibility, and integration of geographically dispersed markets. These locational specific border effects are indirect externalities also known as the spatial spillovers (Boarnet, 1998; Holtz-Eakin & Schwartz, 1995, Hu & Liu, 2010, Lafourcade & Thisse, 2008, Cohen, 2010, Álvarez et al., 2016).

The theory of new economic geography (NEG) acknowledges infrastructural spillovers generates as an outcome of the transport infrastructure, and that increases accessibility and inter-firm linkages which results in changes in productivity, within the regions and across spatial clusters (Venables, 2010). The recent advancements in spatial econometrics and techniques such as the exploratory spatial data analysis (ESDA) techniques helped economists incorporate natural geography into the econometric analysis (Anselin & Getis, 2010). More importantly, the traditional non-spatial econometric models lead to wrong inferences in the case of spatial dependence (Lesage, 1999). Therefore, the recent literature on transport infrastructure uses spatial econometric estimation techniques. However, the spatial lag or spatial error model has been used without often incorporating without testing the spatial dependence, leading to the wrong estimation of the spillover effect. Therefore, analysis of spatial dependence (or ESDA) is of essential to correctly define the neighborhood definition (Chen & Haynes, 2015). So far, to my best of knowledge, ESDA of road infrastructure in Pakistan has not yet been conducted, and this research will be beneficial for the econometric model construction for transport infrastructure in Pakistan. This technique is relatively new and provides high-end and useful insights with graphical maps to analyze the spatial dependence.

Estimation Techniques

This study analyzes the spatial dependence of road infrastructure in Pakistan. 'The spatial dependence is a collection of sample data means that observations at location 'i' depend on other observations at locations j when i ≠ j' (Lesage, 1999, p.11). Spatial autocorrelation is synonymous with spatial dependence in spatial econometrics. It is the degree of independent values observed in the geographic locations of the neighborhood. The Moran's I statistics is to assess the spatial autocorrelation. Moran's I statistic is based on the Pearson product-moment correlation coefficient, and geography is included via the spatial weight [W] matrix. It finds the correlation between two variables, the correlation of one variable with itself vis-à-vis a spatial weight matrix (Getis, 2010). Moran's-I focus on each observation as a difference from the mean of all observations.

The two leading indicators computed for this study. The first indicator is the Global Moran's I, known as the test for global spatial autocorrelation. The second indicator, the Local Moran's I or Local indicators of spatial association (LISA) cluster map, is to analyze details of the clustering effect.

Global Moran's I (Global Spatial Autocorrelation)

The Moran's I is computed by (eq. 1).

$$I = \frac{n}{S_0}\frac{\sum_i \sum_j W_{ij}.Z_i.Z_j}{\sum_i {Z_i}^2} \qquad \text{(eq. 1)}$$

Where I represent the Moran's statistic, $\boldsymbol{W_{ij}}$ are elements of the spatial weight matrix, $\boldsymbol{S_0} = \sum_{\boldsymbol{i}} \sum_{\boldsymbol{j}} \boldsymbol{W_{ij}}$ is the sum of all the weights, and n as the number of observations. It is a cross-product statistic between a variable and with its spatial lag, and the variable expressed in deviations from its mean. For an observation at location *i*, expressed as $Z_i = X_i - \bar{X}$ where $\bar{X}$ is the mean of a variable X . $Z_j = X_j - \bar{X}$ where $\bar{X}$ is the mean of a variable X .

Inference for Moran's I is based on a null hypothesis of spatial randomness. The statistic distribution under the null can be derived using an assumption of normality (independent normal random variates) (i.e., each value is equally likely to occur at any location).

$$p = \frac{R+1}{M+1}$$

R is the number of times the computed Moran's I from the spatial random data sets, and M equals the number of permutations, typically taken as 99, 999, etc., to yield a pseudo-p-value.

This software constructs the Moran Scatter Plot, and this tool of exploratory analysis is developed by Anselin (1996). When row-standardized weights ar used than, the sum of all the weights ($\boldsymbol{S_0} = \sum_{\boldsymbol{i}} \sum_{\boldsymbol{j}} \boldsymbol{W_{ij}}$) equals the number of observations ($\boldsymbol{n}$).

As a result, the expression for Moran's I simplify to:

$$I = \frac{\sum_i \sum_j W_{ij}\,.Z_i.Z_j}{\sum_i {Z_i}^2} = \frac{\sum_i (Z_i \times \sum_j W_{ij} Z_j)}{\sum_i {Z_i}^2}$$

Moran's scatter plot consists of a plot with the spatially lagged variable on the y-axis and the original variable on the x-axis. The slope of the linear fit to the scatter Plot equals Moran's I. The Plot is centered on the mean (of zero). All points to the right of the mean have $Z_i > 0$, and all points to the left have $Z_i < 0$. Visualization in the Moran scatter Plot is the classification of the *nature* of spatial autocorrelation into four categories. It can classify the values for the spatial lag above and below the mean as *high* and *low*. The scatter Plot is then decomposed into four quadrants. The upper-right and lower-left quadrants correspond with *positive* spatial autocorrelation (similar values at neighboring locations). This is referred as *high-high* and *low-low* spatial autocorrelation. The lower-right and upper-left quadrants correspond to *negative* spatial autocorrelation (dissimilar values at neighboring locations). This is referred as *high-low* and or *low-high* spatial autocorrelation.

LISA Cluster Map (Local Moran's I)

The LISA Cluster maps are developed based on the pioneered work of Anselin (1995). The global spatial autocorrelation indicators are designed to reject the null hypothesis of spatial randomness in favor of an alternative of *clustering*. However, such *clustering* is a characteristic of the complete spatial pattern and does *not* indicate the *location* of the clustering. Therefore, LISA is seen as having two essential characteristics First, it provides a statistic for each location with an assessment of significance. Second, it establishes a proportional relationship between the sum of the local statistics and a corresponding global statistic.

$$\boldsymbol{I_i} =. \frac{\sum_j \boldsymbol{W_{ij}} \boldsymbol{Z_i} \boldsymbol{Z_j}}{\sum_i \boldsymbol{Z_i}^2}$$

Here, the denominator is the same $\sum_i Z_i^2$ for every location. Therefore, we can consider it as a constant 'c'

$$\boldsymbol{I_i} = \boldsymbol{c}.\boldsymbol{Z_i} \sum_j \boldsymbol{W_{ij}} \boldsymbol{Z_j}$$

The preferred approach for hypothesis testing is a *conditional* permutation method, which is similar to the permutation approach considered in the Moran scatter Plot, except that the value of each $\boldsymbol{Z_i}$ held fixed at its location $\boldsymbol{i}$. The remaining $\boldsymbol{n-1}$ Z-values are then randomly permuted to yield a *reference distribution* for the local statistic (one for each location).

We conducted ESDA using GeoDa software and constructed Moran's I graph, Moran scattered Plot, and LISA cluster maps and this software uses the data visualization approach.

Spatial Weight Matrix

The results of the Moran's I depend on W's definition.

$\boldsymbol{W} = d_{ij}{}^{-\gamma}$ where $\gamma \geq 1$

Here, d_{ij} representing distance between two districts *i* and *j* while γ assumed as 1. Given the critical role of road infrastructure in generating economic interactions, and therefore distance-based W is preferred for this study. Furthermore, we preferred the inverse distance weight matrix, which gives more weightage to nearby spatial interactions than distant locations. Therefore, we developed an inverse Euclidean distance (row-standardized) power **W** for this study. we are using two bandwidths, to reduce the researcher's subjectivity bias while constructing **W**.

Data

For this study, we are using high accurate district profile data of the road infrastructure developed by the NTRC (2020) by conducting a ground survey in 2019/20. NTRC developed the road classification based on the number of lanes, carriageway type, traffic flows, and road material. All roads are categorized into Motorways, National

Highways, Highways, Primary Roads, Secondary Roads, and Local Roads. According to this report, the total road network of 493,088 km in Pakistan, including all provinces, AJ&K, Gilgit Baltistan, and Islamabad (NTRC, 2020). The province of Punjab has the largest road network (276631 km) compared total road network in all other provinces of Sindh, KPK, and Baluchistan (193401 km). Details are provided in table 1.

We constructed the spatial weights using shapefiles of Pakistan, which were obtained from an online free resource provided by United Nations Office for coordinating Humanitarian Affairs (OCHA). We used layer (2) the district-level polygons.

Table 1
Types and length of the road network: provinces of Pakistan (km)

	Express way	High way	Local road	Metro road	Motor way	National highway	Primary road	Secondary road	Total
Punjab	96	10606	215103	46	1418	1736	1247	46379	**276631**
Sindh	20	2934	67592	13	436	1854	769	16013	**89631**
KPK	88	2333	40863	0	100	1814	112	12819	**58129**
Baluchistan	0	1481	34347	0	377	3517	72	5847	**45641**
ICT	20	83	5378	18	6	41	106	342	**5994**
Total	**224**	**17437**	**363283**	**77**	**2337**	**8962**	**2306**	**81400**	**476026**

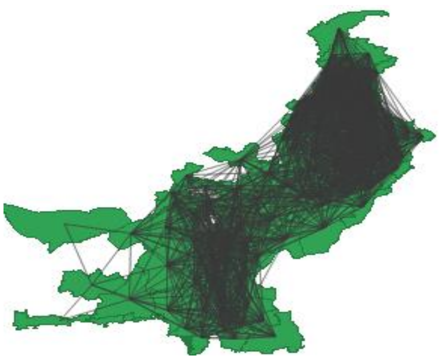
Source: NTRC (2020), 'Digitalization of Road Directory in the country, The Final Report.

Results and Discussion

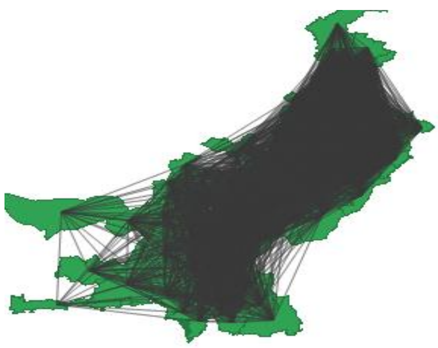
This section is comparting into five sub-sections. The first subsection provides an ESDA provide at districts level analysis at national level. While the following subsections provide the districts level ESDA for each province of Punjab, Sindh, KPK, and Baluchistan, respectively.

1. ESDA of Road Infrastructure: A Districts Level Analysis of Pakistan

An ESDA has been conducted for the 109 districts of Pakistan using **W** of 3 and 5 BW. The connectivity graph and maps are provided in Figure 1. The estimated two indicators for the detection of spatial dependence. First, the global Moran's I and second, is the LISA Cluster analysis, also known as 'local' spatial autocorrelation. We have estimated Moran's I and developed LISA cluster Map as given in Figure 2 & 3, respectively.

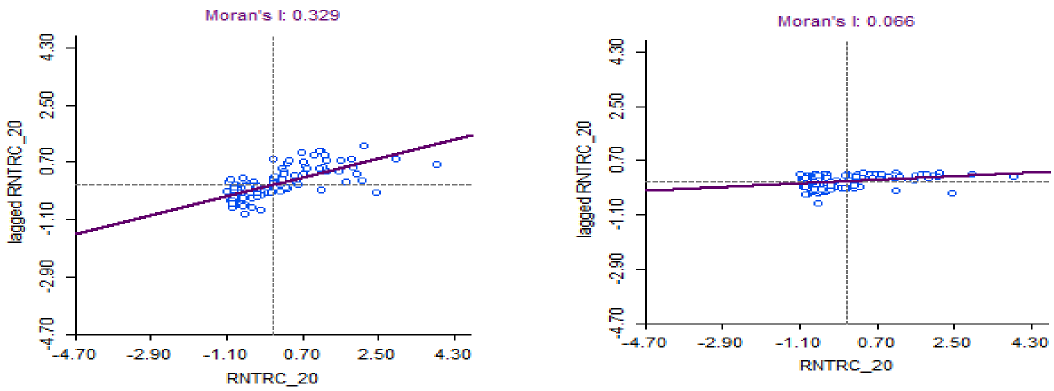


(**W**: 3 BW)



(**W:** 5 BW)

Figure 1 Connectivity Graphs and Maps: Districts of Pakistan



I: 0.329; Pseudo P-value: 0.001 (3 BW) *I: 0.066; Pseudo P-value: 0.001 (5BW)*

Figure 2 Global Moran's I of Road infrastructure in Pakistan

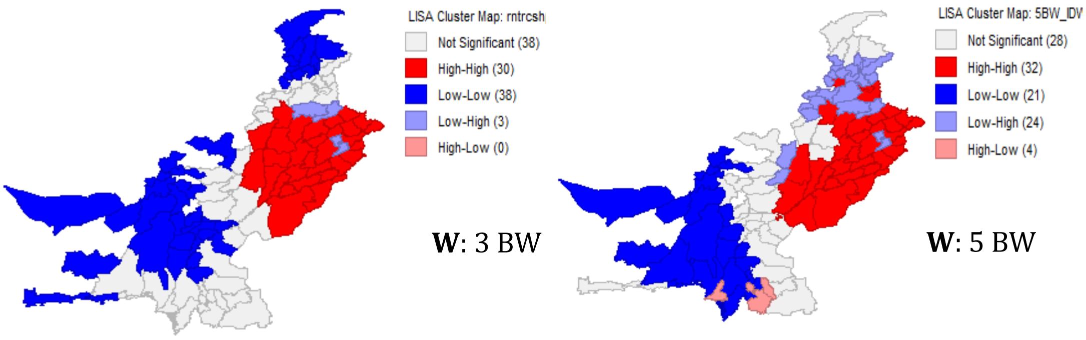


Figure 3 LISA Cluster Map (Local Moran's I) of Road infrastructure in Pakistan

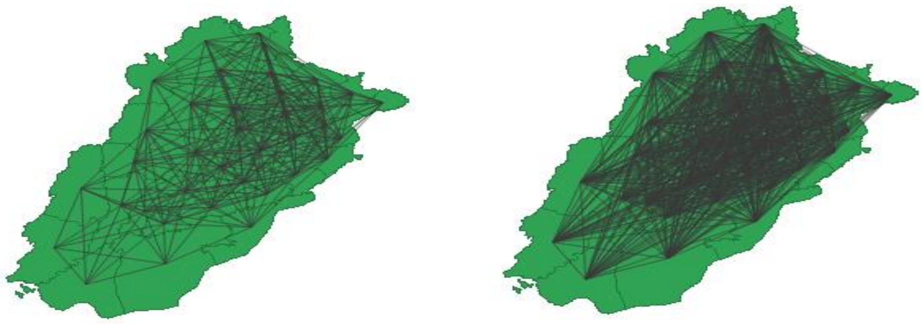
The results signify the existence of spatial dependence and positive spatial autocorrelation (highly significant) of road infrastructure in Pakistan. The Moran's I scatter Plot elucidates that the road infrastructure is not uniformly distributed across districts but lies in the quadrants of high-high, low-high, and high-low regions. It seems a strong clustering effect exists, and therefore, the Local Moran's I, the LISA, provides a detailed analysis. The LISA cluster map depicts an extreme polarization and clustering of road infrastructure in Pakistan. The LISA cluster using **W** of 3 BW depicts 30 districts spatially clustered as high-value and 38 districts clustering as low values of road infrastructure compared to average with other districts of Pakistan. On the other hand, using **W** the 5 BW, LISA cluster map highlights 21 districts with the least road infrastructure are geographically clubbed as low-value clusters. While 32 districts with more extensive road networks are geographically clubbed as high-value clusters. These polarized clusters are statistically significant at 5% significance using 999 random computations. The low-value clusters lie in Baluchistan, Sindh, and KPK, while the high-value cluster lie in the province of Punjab except the Nowshera, a district of KPK. More details are in table 1.

Table 1
Highly significant Spatial clusters of Road infrastructure in Pakistan

***W**: 3 BW*	
Low-Low	**High-High**
KPK Buner, Charsada, Chitral, Kohistan, Lower Dir, Malakand PA, Mardan, Peshawar, Pishin, Shangla, Swat, Tor ghar, Upper Dir. **Sindh** Dadu, Jacobabad, Larkana, Naushero Feroze, Qambar Shahdatkot, SB Abad, Shikarpur, Sukkur, **Balochistan** Chagai, Gwader, Jaffarabad, Jhal Magsi, Kacchi, Kalat, Kashmore, Kharanm, Khuzdar, Mastung, Nasirabad, Panjgur, Quetta, Sibi, Sohbatpur, Zhob and Ziarat.	**Punjab** Bahawalnagar, Bahwalpur, Bhakkar, Chiniot, DG Khan, DI Khan, Faisalabad, Gujranwala, Gujrat, Hafizabad, Jhang, Kasur, Khanewal, Khusab, Lahore, Layyah, Lodhran, Mandi Bahauddin, Mianwali, Multan, Muzaffargarh, Narowal, Okara, Pakpattan, Sahiwal, Sarghoda, Sheikhupura, Sialkot, Toba tek Singh, Vehari,
***W**: 5 BW*	
Low-Low	**High-High**
Balochistan Awaran, Chagai, Jhal Magsi, Kalat, Kharan, Khuzdar, Lasbela, Mastung, Panjgur, Pishin, Qambar Shahdadkot, Quetta, Ziarat, **Sindh** Dado, Jamshoro, Matiari, Naushehro Feroze, SB Abad, Tando Allah Yar, Tando Muhammad Khan, Thatta.	**Punjab** Bahawalnagr, Bhalwalpur, Chiniot, Dera Ghazi Khan, Faisalabad, Gujranwala, Gujrat, Hafizabad, Islamabad, Jhang, Kasur, Khanewal, Khushab, Lahore, Lodhran, Mandi Bahauddin, Mianwali, Multan, Muzaffargarh, Narowal, Okara, Pakpattan, Rahm yar Khan, Rajanpur, Rawalpindi, Sahiwal, Sargodha, Sheikhpura, Sialkot, Toba Tek Sings, Vehari **KPK** Nowshera

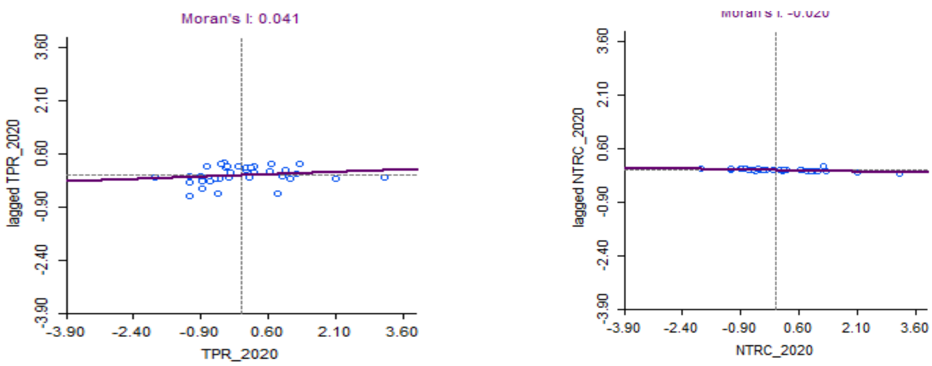
2. ESDA of Road Infrastructure: A Districts Level Analysis of Punjab, Pakistan

The researcher analyzed the spatial dependence of road infrastructure within Punjab using road length data of 36 districts and Islamabad Capital Territory (ICT). The data of ICT is also included given the geographical adjacency. For analyzing the spatial dependency **W** of 2 BW and 5 BW constructed and connectivity graphs are provided in Figure 4. Moran's I and LISA cluster maps are provided in Figure 5 & 6, respectively.



W:2 BW) **W**:5 BW

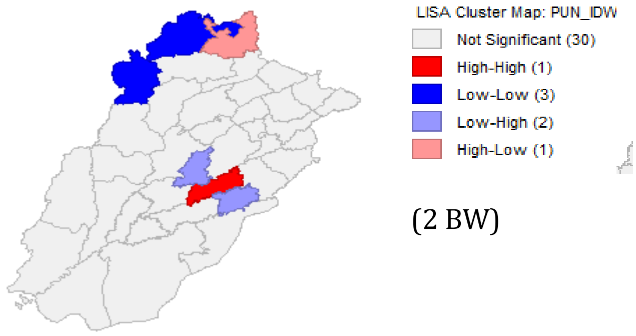
Figure 4 Connectivity Graphs: ICT and District of Punjab, Pakistan



I: 0.041, (pseudo p-values ranges 0.065 to 0.077), **W**: 2 BW

I: -0.020, (pseudo p-values ranges 0.040 to 0.070), **W**: 5 BW

Figure 5 Global Moran's I of Road Infrastructure of Punjab and ICT



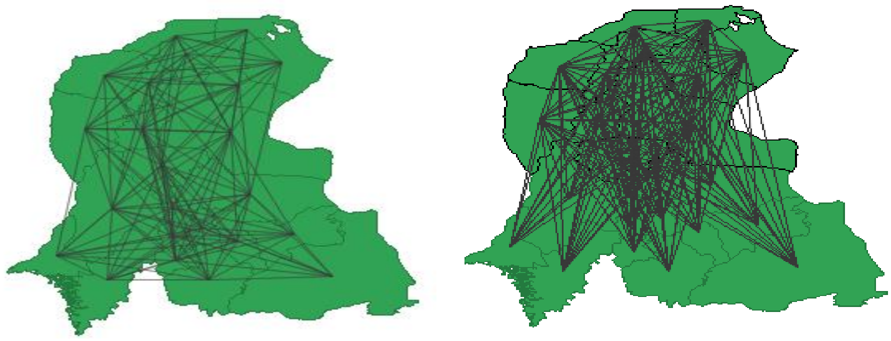
(2 BW)

Figure 3 LISA Cluster Map (Local Moran's I) of Punjab and ICT

Using W with 2 BW and 5 BW, estimated Moran's I are 0.041 and -0.020, are not statistically significant at 5 percent (using randomization option of 999) that depicts weak spatial dependence of road infrastructure within Punjab. The LISA Cluster map analysis depicts districts in four colors, and red-colored districts depict the high-high clusters. It means the districts with high-road infrastructure are neighboring with high road infrastructure. At the same time, blue colored district depicts the statistically significant clusters of low values. It means districts neighboring with a low-road network on average. The LISA depicts one significant cluster as high values for the district of Sahiwal and two low-value significant clusters of Miawali and Attock (see Figure 6).

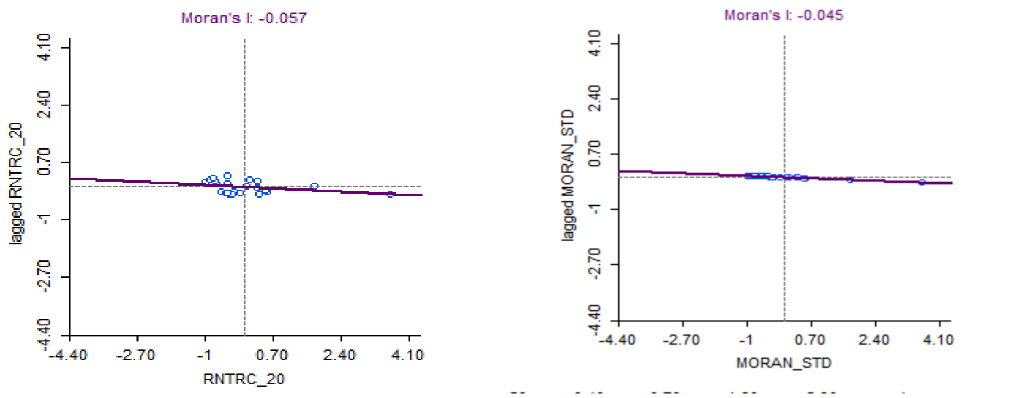
3. ESDA of Road Infrastructure: A districts level analysis of Sindh, Pakistan

The analysis of the spatial dependence of road infrastructure in 23 districts of Sindh province has been conducted using **W** of 2 and 5 BW. The connectivity graphs and maps are provided as Figure 7, and the results of Moran's I and LISA cluster map is given in Figure 8 & 9, respectively.



(**W**: 2 BW) (**W**: 5 BW)

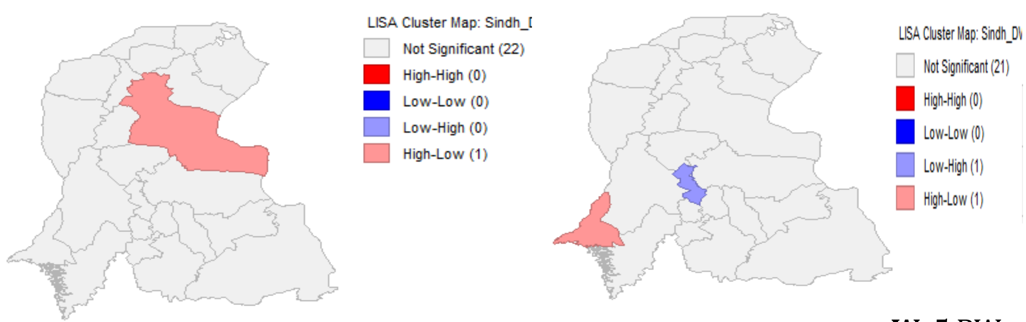
Figure 7 Connectivity Graphs: Districts of Sindh, Pakistan



I: -0.057, (pseudo p-values ranges 0.464 to 0.477), **W:** 2BW

I: -0.045 (pseudo p-values ranges 0.060 to 0.076), **W:** 5BW

Figure 8 Global Moran's I: Road Infrastructure in Districts of Sindh, Pakistan



W: 2 BW **W**: 5 BW

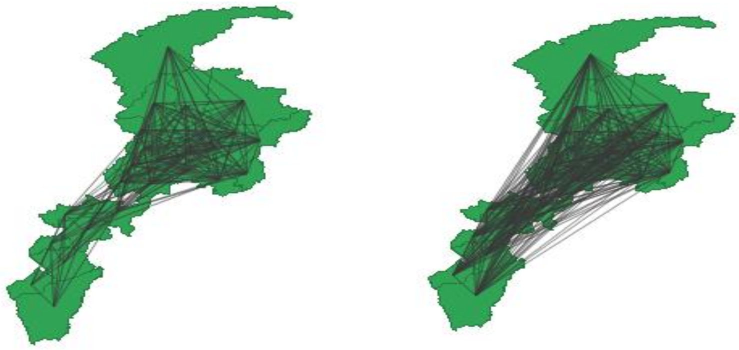
Figure 9 Local Moran's I: Road Infrastructure in Districts of Sindh

The results show an insignificant positive spatial autocorrelation of road infrastructure within the province of Sindh (as Moran's I is not statistically significant at

5% using random computations of 999). The district with high-road networks is neighboring with districts with low road network, and its vice versa and LISA cluster map depicted no significant clustering effects (Figure 9)

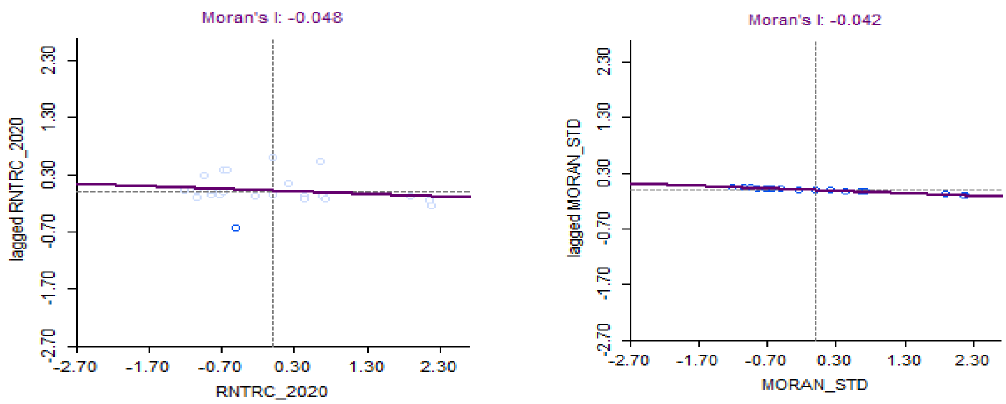
4. ESDA of Road Infrastructure: A district-level analysis of KPK, Pakistan

The spatial dependence test of the road infrastructure in 25 districts of KPK using W of 2 BW and 5 BW has been conducted. The connectivity maps are provided as figures 10. The estimated global Moran's I and LISA cluster map are given in Figure 11 & 12, respectively.



W: 2 BW **W**: 5 BW

Figure 10 Connectivity Graphs and Maps: Districts of KPK, Pakistan



I: -0.048, (Pseudo p-values ranges 0.448 to 0.494), **W:** 2 BW

I: -0.042, (Pseudo p-values ranges 0.114 to 0.136), **W:** 5 BW

Figure 11 Global Moran's I: Road Infrastructure in Districts of KPK and ICT

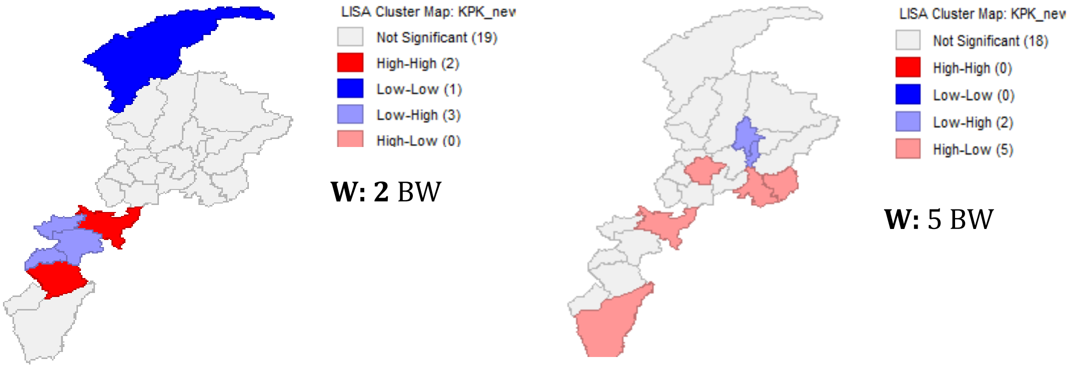
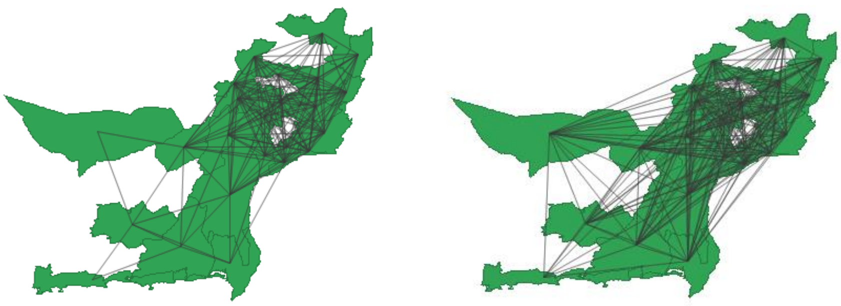


Figure 12 Local Moran's I: Road Infrastructure in Districts of KPK and ICT

The estimated Moran's-I is -0.048 and -0.042, and Moran's graphs lie in the upper left (low-High) and lower right (High-low) quadrants depicting a weak clustering effect. The randomized permutation also signals an insignificant Moran's I, so we can accept the null hypothesis and conclude that the spatial randomness of road infrastructure in the districts of KPK. The LISA cluster map at 5 BW shows two significant spatial outliers districts but doesn't find any significant spatial clusters. While using 2 BW, one significant spatial cluster of low-value and two significant spatial clusters of high value exist. The district of Chitral lies in the low-value cluster (on average), and the districts of Lakki-Mawat and Kohat are districts of high-values spatial cluster road infrastructure compared to their neighborhood (on average).

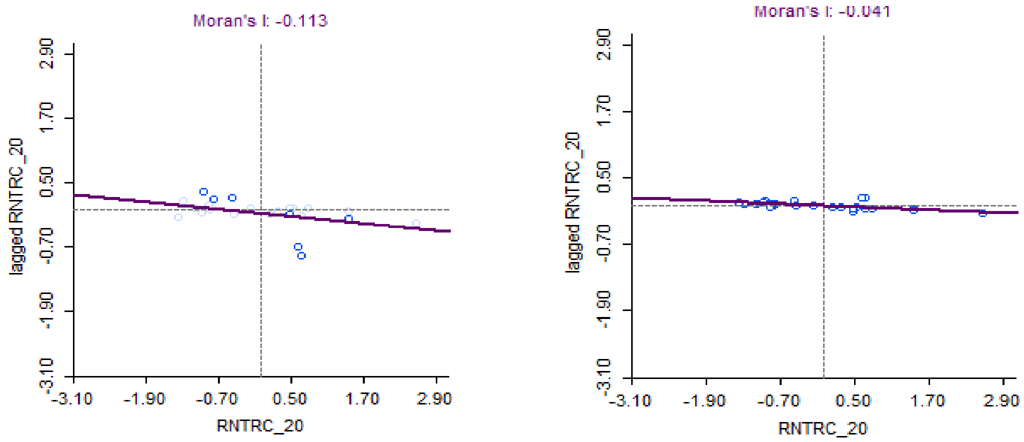
5. ESDA of road infrastructure: A districts level analysis of Baluchistan, Pakistan

An ESDA of road infrastructure for the existing road infrastructure in 24 districts of Baluchistan has been conducted using W of 3 BW, and 5 BW. The connectivity graphs are provided in Figure 13. The global and local Moran's I result are provided in Figure 14 and 15, respectively.



(W: 3 BW) (W: 5 BW)

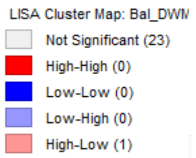
Figure 13 Connectivity Graphs and Maps: Districts of Baluchistan, Pakistan



Pseudo p-values ranges (0.110 to 0.137) DWM (**W**: 3BW)

Pseudo p-values ranges (0.346 to 0.434) (**W**: 5BW)

Figure 4 Global Moran's I: Road infrastructure in districts of Baluchistan



(**W:** 5BW)

Figure 13 LISA cluster map: Road infrastructure in districts of Baluchistan

The ESDA revealed an insignificant positive spatial autocorrelation of road infrastructure within the districts of Baluchistan (using 999 randomized computations). At the same time. LISA cluster map also elucidated insignificant spatial clusters in Baluchistan.

Conclusion

The ESDA has been conducted at the national and province scales using the district-level data of road infrastructure in Pakistan for survey-based data obtained during 2019-20. The main results are;

- On the national scale, road infrastructure is significantly spatially dependent within 109 districts of Pakistan. ESDA revealed a significant positive spatial autocorrelation reveals the road infrastructure is clustered (low-low and high-high). These low-value clusters are located in the Baluchistan and North-East districts of KPK, while the high-value cluster lies in the Punjab and North-west districts of KPK.

- At the regional scale of Punjab, the ESDA of the road infrastructure revealed an insignificant positive spatial autocorrelation. At the same time, the LISA cluster map depicts the districts of Sahiwal as the significant low-low cluster and the Mianwali and Attock as the significant high-high clusters.

- At the regional scale of Sindh, the ESDA of road infrastructure depicts an insignificant positive spatial autocorrelation. Still, data reveals significant negative spatial autocorrelation that doesn't depict any clustering effect of road infrastructure within districts of Sindh.

- At the regional scale of KPK, ESDA of road infrastructure signaled insignificant positive spatial autocorrelation at the regional scale. The LISA analysis depicts that the district Chitral has lower road infrastructure than the road infrastructure (on average) in its neighbors. In contrast, the districts of Lakki-Mawat and Kohat are located as high-value spatial clusters relatively.

- At the regional scale of Baluchistan, the ESDA of road infrastructure revealed an insignificant positive spatial autocorrelation and LISA clusters.

This study highlights the regional imbalances (spatial clusters) in terms of road infrastructure provisioning in Pakistan. On the national level analysis, this is inferred that the road infrastructure is not randomly distributed across the geographic space of Pakistan but spatially clustered. At the same time, at regional level, within each province, the distribution of road infrastructure is not clustered significantly.

Conclusively, this paper is providing an important insight for policy makers as it carefully highlighting the district level inequities of road infrastructure, which is spatially clustered as high-high and low-low districts in Pakistan. However, this study is only presenting an exploratory data analysis and a benchmark study, and it highly recommends incorporating the bordered or spatial spillover effects for the district level analysis in Pakistan.

References

Álvarez, I. C., Barbero, J., & Zofío, J. L. (2016). A spatial autoregressive panel model to analyze road network spillovers on production. *Transportation Research Part A: Policy and Practice*, *93*, 83–92. https://doi.org/10.1016/j.tra.2016.08.018

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin, L. (1996). 'The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association'. In M. Fischer, H. Scholten, and D. Unwin (eds.), *Spatial Analytical Perspectives on GIS*. London: Taylor and Francis.

Arasan, V. T., & Arkatkar, S. S. (2011). Derivation of capacity standards for intercity roads carrying heterogeneous traffic using computer simulation. *Procedia - Social and Behavioral Sciences*, *16*, 218–229. https://doi.org/10.1016/j.sbspro.2011.04.444

Boarnet, M. G. (1998). Spillovers and the locational effects of public infrastructure. *Journal of Regional Science*, *38*(3), 381–400. https://doi.org/10.1111/0022-4146.00099

Chen, Z., & Haynes, K. E. (2015). Spatial impact of transportation infrastructure: a spatial econometric CGE approach. In *Regional science matters* (pp. 163-186). Springer, Cham.

Cohen, J. P. (2010). The broader effects of transportation infrastructure: Spatial econometrics and productivity approaches. *Transportation Research Part E: Logistics and Transportation Review*, *46*(3), 317–326. https://doi.org/10.1016/j.tre.2009.11.003

Getis, A. (2010). Spatial Autocorrelation. In *Handbook of Applied Spatial Analysis*. https://doi.org/10.1007/978-3-642-03647-7

Haughwout, A. (2005). Infrastructure and Social Welfare in Metropolitian America. *Economic Policy Review*, *7*(3), 1–16.

Holtz-Eakin, D., & Schwartz, A. E. (1995). Spatial productivity spillovers from public infrastructure: Evidence from state highways. *International Tax and Public Finance*, *2*(3), 459–468. https://doi.org/10.1007/BF00872777

Hu, A., & Liu, S. (2010). Transportation, economic growth and spillover effects: The conclusion based on the spatial econometric model. *Frontiers of Economics in China*, *5*(2), 169–186. https://doi.org/10.1007/s11459-010-0009-0

Lafourcade, M., & Thisse, J.-F. (2008). New economic geography: A guide to transport analysis. *PSE Working Papers*, *33*(0), 0–35. http://halshs.archives-ouvertes.fr/halshs-00586878/fr/

LeSage, J. P. (1999). The theory and practice of spatial econometrics. *University of Toledo. Toledo, Ohio*, *28*(11).

NTRC, (2020). *Digitalization of Roads Directory in the Country* (Report: No. NTRC-333), The Urban Unit National Transport and Research Centre.

Venables, A. J. (2010). New economic geography. In *Economic growth* (pp. 207-214). Palgrave Macmillan, London.