Does Economic Geography Matter for Pakistan?

A Spatial Exploratory Analysis of Income and Education Inequalities

Sofia Ahmed

Pakistan Institute of Development Economics

Paper submitted for the 27th PIDE-PSDE Annual Conference

Abstract

Generally, econometric studies on socio-economic inequalities consider regions as independent entities, ignoring the likely possibility of spatial interaction between them. This interaction may cause spatial dependency or clustering, which is referred to as spatial autocorrelation. This paper analyzes for the first time, the spatial clustering of income, income inequality, education, human development, and growth by employing spatial exploratory data analysis (ESDA) techniques to data on 98 Pakistani districts. By detecting outliers and clusters, ESDA allows policy makers to focus on the geography of socio-economic regional characteristics. Global and local measures of spatial autocorrelation have been computed using the Moran's *I* and the Geary's *C* index to obtain estimates of the spatial autocorrelation of spatial disparities across districts. The overall finding is that the distribution of district wise income inequality, income, education attainment, growth, and development levels, exhibits a significant tendency for socio-economic inequalities and human development levels to cluster in Pakistan (i.e. the presence of spatial autocorrelation is confirmed)[1].

Key words: Pakistan, spatial effects, spatial exploratory analysis, spatial disparities, income inequality, education inequality, spatial autocorrelation

---

[1] **Acknowledgements:** I would like to thank Dr Jannette Walde (University of Innsbruck Austria), Dr Maria Sassi (University of Pavia, Italy), Dr Alejandro Canadas (Mount St Marys University, USA),Dr. Giuseppe Arbia (University G. D'Annunzio of Chieti, Spatial Econometric Association), and Dr. Richard Pomfret (University of Adelaide) for their comments on an earlier version of this paper. I would also like to thank Khydija Wakeel and Muhammad Qadeer at the Planners Resource Centre Pakistan, for providing me with the shape files. Finally, I would like to acknowledge, the data management staff at the Pakistan Institute of Development Economics (PIDE) Islamabad, Federal Bureau of Statistics, Islamabad, and Dr. Amir Jahan Khan and Dr. Haroon Jamal (Sustainable Policy Development Centre, SPDC Karachi) for their generous data support.

Note: This is a preliminary version of this paper and comments are welcome. The author can be contacted at: sofia.ahmed@pide.org.pk

1. Introduction

From the industrial revolution to the emergence of the so-called knowledge economy, history has shown that economic development has taken place unevenly across regions. A region's economy is a complex mix of varying types of geographical locations comprising different kinds of economic structures, infrastructure, and human capital. In this context recent literature in regional sciences has highlighted how crucial it is to analyse socio-economic phenomena in the light of spatial concepts such as geography, neighbourhood, density, and distance (Krugman, 1991; Krugman and Venebles, 1995; Quah, 1996; Baldwin et al, 2003; van Oort, 2004; Kanbur and Venebles, 2005; World Development Report, 2009). Keeping these recent developments in view, this paper identifies, measures, and models the temporal relationship between space, economic inequalities, human development, and growth for the case of Pakistan[2]. Specifically, by using data at district level from 1998 and 2005, it utilizes spatial exploratory techniques to determine the effect of distance and contiguity among 98 of Pakistan's administrative districts on their human capital characteristics and inequalities[3]. This way it provides some of the first spatially explicit results for clustering of socio-economic characteristics across Pakistani districts[4].

Most of the existing research on Pakistan's socio-economy is based on a provincial level, and it neglects the role of social interactions the districts within the provinces[5]. This paper in particular investigates whether spatial clustering of income and average education levels can explain their distribution across Pakistani districts. District level research has become even more important as Pakistan has taken a major step towards fiscal decentralization with the enactment of the 18th Constitutional Amendment. Moreover the 7th National Finance Commission Award has allowed the transfer of more funds from the federation to the provinces which now have more authority over the provision of health, educational and physical infrastructure facilities. This fundamental shift towards the division

[2] Economic inequalities refer to education, earnings income inequalities in particular.

[3] Examples of studies similar to this paper include: Rey and Montouri (1999) on convergence across USA, Balisacan and Fuwa (2004) for income inequality in Philipines, Dall'erba (2004) analyses productivity convergence across Spanish regions over time, Dominicis, Arbia and de Groot (2005) analyses spatial distribution of economic activities in Italy, Pose and Tselios (2007) investigates education and income inequalities in the European Union, and Celebioglu and Dall'erba (2009) analyses spatial disparities in growth and development in Turkey.

[4] The only other exception includes Burki *et al* (2010) that has explicitly considered spatial dependencies in its analysis. However it has analysed 56 districts.

[5] Exceptions include Jamal and Khan (2003a, 2003b), Jamal and Khan (2008a, 2008b), Naqvi (2007),Arif *et al* (2010), Siddique (2008) and a few others. Except for Jamal and Khan (2003a, 2003b), Jamal and Khan (2007a, 2007b), most of them only study selected districts/villages from the same province e.g. Naqvi (2007) only analyses the districts/villages of Punjab.

of power between the centre and the provinces bears significant implications for the country's long term policy planning, management and implementation. As education and other public and social services become the sole domain of the provinces, there is a need for increased research at the district level.

Furthermore, Pakistan is also characterised with spatial disparities between its key socio-economic characteristics such as education, health, physical infrastructure, etc (Burki *et al*, 2010). While some districts have state of the art physical and human capital infrastructure, others have made little or no progress at all. This phenomenon is in line with the findings of the World Bank's World Development Report (2009) that has demonstrated how and why the clustering or concentration of people and production usually takes place in particular favourable areas (coasts, cities, etc) during the growth process in any country. For the case of Pakistan, the most developed districts are located in Northern and Central Punjab. It has been noted that Pakistani districts with a population density of more than 600 persons per square km are characterized by industrial clusters, superior education and health infrastructure and better sanitation facilities that serve as attractive pull factors, e.g., Karachi, Lahore, Peshawar, Charsadda, Gujranwala, Faisalabad, Sialkot, Mardan, Islamabad, Multan, Swabi, Gujrat and Rawalpindi (Khan, 2003). On the other hand, districts with lowest population densities (or those having below 30 persons per square km) are characterized by prevalence of various push factors such as; absence of job opportunities due to lower education and health facilities, poor agricultural endowments, barren or mountainous topography, and lack of limited presence of industrial units (Khan, 2003). Moreover, the fact that the highly (and medium) concentrated districts (except for Swat and Muzzaffargarh) are mostly clustered around metropolitan cities of Karachi and Lahore (Burki *et al*, 2010) demonstrates that a district's human and economic development is being shared by its neighboring districts, confirming that economic geography matters for Pakistan.

In the light of the above mentioned issues, this study empirically investigates the spatial clustering of economic inequalities, growth and development across Pakistani districts by utilizing ESDA techniques. The paper is organized as follows: Section 2 describes the data; Sections 3 and 4 provide a detailed overview of the methodology utilized; Section 5 presents the empirical results; finally Section 6 discusses the policy and methodological implications of the empirical results and concludes.

2. Data

For district wise average earnings income and education levels, this paper utilizes micro data from the Pakistan Social and Living Standards Measurement survey (PSLM) 2004-05. It is the only socio-economic micro data that is representative at the provincial and at the district level. Moreover, the sample size of the district level data is also substantially larger than the provincial level data contained in micro data surveys such as Household Income and Expenditure Survey (HIES) of Pakistan and the Labour Force Survey (LFS) of Pakistan. This has enabled researchers to draw socioeconomic information which is representative at lower administrative levels as well. The survey for 2004-05 provides district level welfare indicators for a sample size of about 76,500 households. It provides data on districts in all four provinces of Pakistan namely; Punjab, Sindh, Khyber Pakhtoonkhwa (KP), and Balochistan. The federally administered tribal areas (FATA region) along the Afghan border in the north-west and Azad Kashmir are not included in the data.

To analyse the spatial differences in district wise primary, secondary, and bachelor's education levels over time, this chapter has utilized the district level data from the 1998 Population Census of Pakistan. Since the data from PSLM (2004-05) is statistically comparable with the Pakistan Census Data (1998) the two data sets together provide a decent gap of 7 years to analyse the temporal changes in income and development characteristics across Pakistan.

Finally, for investigating spatio-temporal differences in district wise income, GDP growth rate, and human development levels, this paper has taken its data from the National Human development Report (2003) and from Jamal and Khan (2007). Note that all income data from 2004-05 was deflated using the Pakistani Consumer Price Index (CPI) of 1998.

3. Methodology

Due to the abundance in data collected at a provincial or a rural/urban disaggregation, most socio-economic studies on Pakistan, are a province based analysis. Pakistani provinces however have extreme 'within' diversity in terms of their economic structures, development levels, cultures, language, natural resources and geography. Hence regional policy making requires analyzing socio–economic issues at an even smaller geographical disaggregation.

For this reason, the spatial unit of analysis chosen for this study is the 'districts' of Pakistan. In terms of geographical disaggregation Pakistan (excluding the Federally Administered Tribal Area (FATA) region and Azad Kashmir) has 4 levels consisting of 4 provinces (Punjab, Sindh, Khyber Pakhtoonkhwa (KP), and Balochistan), 107 districts, 377 sub-districts, and 45653 villages. A lower level unit of analysis is not being used because of two main reasons. Firstly, data on regional scales below the district level in Pakistan suffers from reliability issues. The second issue is more technical. In order to give information on 45,653 villages of Pakistan instead of 107 districts, the project would need a matrix of distance with $\frac{45{,}653 \times (45{,}653+1)}{2} = 1{,}042{,}121{,}031$ free elements to be evaluated, hence the utilization of district level data. Due to data constraints, this chapter analyzes 98 out of 107 districts in Pakistan (see Table A1).

3.1 Spatial economic analysis and spatial effects

A fundamental concept in geography is that proximate locations often share more similarities than locations far apart. This idea is commonly referred to as the 'Tobler's first law of geography' (Tobler, 1970). Classical statistical inference such as conventional regressions are inadequate for an in-depth spatial analysis since they fail to take into account spatial effects and problems of spatial data analysis such as spatial autocorrelation, identification of spatial clusters and outliers, edge effects, modifiable areal unit problem, and lack of spatial independence (Arbia, Benedetti, and Espa, 1996; Beck, Gleditsch, and Beardsley, 2006; Franzese and Hays, 2007)[6]. Moreover, as an uneven distribution of socio-economic economic characteristics is shaping the economic geography of most countries, spatial analysis also has increasing policy relevance (World Development Report—WDR, 2009). These reasons together necessitate the use of spatial exploratory and explanatory methods that can explicitly take spatial effects into account.

Spatial analysis investigates the presence (or absence) spatial effects which can be divided into two main kinds: spatial dependence and spatial heterogeneity. Spatial heterogeneity refers to the display of instability in the behaviour of the relationships under study. This implies that parameters and functional relationships vary across space and are not

[6] Modifiable Areal Unit Problem: When attributes of a spatially homogenous phenomenon (e.g. people) are aggregated into districts, the resulting values (e.g. totals, rates and ratios) are influenced by the choice of the district boundaries just as much as by the underlying spatial patterns of the phenomenon.

homogenous throughout data sets. Spatial dependence on the other hand, refers to the lack of independence between observations often present in cross sectional data sets. It can be considered as a functional relationship between what happens at one point in space and what happens in another. If the Euclidean sense of space is extended to include general space (consisting of policy space, inter-personal distance, social networks etc) it shows how spatial dependence is a phenomenon with a wide range of application in social sciences. Two factors can lead to it. First, measurement errors may exist for observations in contiguous spatial units. The second reason can be the use of inappropriate functional frameworks in the presence of different spatial processes (such as diffusion, exchange and transfer, interaction and dispersal) as a result of which what happens at one location is partly determined by what happens elsewhere in the system under analysis.

3.2 Quantifying spatial effects

Spatial dependence puts forward the need to determine which spatial units in a system are related, how spatial dependence occurs between them, and what kind of influence do they exercise on each other. Formally these questions are answered by using the concepts of neighbourhood expressed in terms of distance or contiguity.

Boundaries of spatial units can be used to determine contiguity or adjacency which can be of several orders (e.g. first order contiguity or more). Contiguity can be defined as linear contiguity (i.e. when regions which share a border with the region of interest are immediately on its left or right), rook contiguity (i.e. regions that share a common side with the region of interest), bishop contiguity (i.e. regions share a vertex with the region of interest), double rook contiguity (i.e. two regions to the north, south, east, west of the region of interest), and queen contiguity (i.e. when regions share a common side or a vertex with the region of interest) (LeSage, 1999). Other common conceptualizations of spatial relationships include inverse distance, travel time, fixed distance bands, and k-nearest neighbours.

The most popular way of representing a type of contiguity or adjacency is the use of the binary contiguity (Cliff and Ord, 1973; 1981) expressed in a spatial weight matrix (**W)**. In spatial econometrics **W** provides the composition of the spatial relationships among different points in space. The spatial weight matrix enables us to relate a variable at one point in space to the observations for that variable in other spatial units of the system. It is used as a variable while modelling spatial effects contained in the data. Generally it is based on using either distance or contiguity between spatial units. Consider below a spatial weight matrix for three units:

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{bmatrix}$$

where w_{12} or w_{ij} may be the inverse distance between two units i and j or it may be 0 and 1 if they share a border or a vertex. The **W** matrix displays the properties of a spatial system and can be used to gauge the prominence of a spatial unit within the system. The usual expectation is that values at adjacent locations will be similar.

3.3 The spatial weight matrix for Pakistan

The choice of the W matrix representation and its conceptualization has to be carefully based on theoretical reasoning and the historical factors underlying the concept or phenomenon under study.

This paper has employed two W matrices for Pakistan[7]. The first matrix is a simple binary contiguity W matrix (referred to as *BC matrix* from now onwards) based on the concept of Queen Contiguity i.e. if a district *i* shares a border *or* a vertex with another district *j*, they are considered as neighbours, and $w_{i,j}$ takes the value 1 and 0 otherwise. This matrix is also zero along its diagonal implying that a district cannot be a neighbour to itself. Hence it is a symmetric binary matrix with a dimension of 98x98 (98 being the total number of the districts being analyzed). This matrix precisely tells us the influence of geographically adjacent neighbours on each other. A simple binary contiguity matrix is a standard starting point and its influence is often compared with other types of W matrices.

The second W matrix developed for Pakistan is one based on inverse average road distance from a district *i* to the nearest district *j* which has a '*large city*' in it (referred to as *ID matrix* from now onwards). Out of the 98 districts being studied there are only 14 that come under the category of a district with a '*large size*' city as per the classification of the coding scheme for the PSLM survey. These include Islamabad as the federal capital city; Lahore, Faisalabad, Rawalpindi, Multan, Gujranwala, Sargodha, Sialkot, and Bahawalpur as districts

[7] Usually two or more weights matrices are utilized in spatial exploratory and econometric studies as a robustness measure. It is way of demonstrating whether strength of spatial effects are robust to changing definitions of neighbourhood.

with a ‘large size’ city in Punjab; Karachi, Hyderabad and Sukkur in Sindh; Peshawar in the North West Frontier Province and Quetta in Balochistan. This matrix is a symmetric non-binary matrix, again with a dimension of 98x98.

The reason for selecting road distance instead of train distance as is normally done in most studies on regional analysis is that in Pakistan, the road network is much better developed than the railway network . As a result, Pakistan’s transport system is primarily dependent on road transport which makes up 90 percent of national passenger traffic and 96 percent of freight movement every year (The Economic Survey of Pakistan, 2007-08). Inverse distance matrices have more explanatory power as partitions of geographic space especially when the phenomenon under study involves the exchange or transfer of information and knowledge (in our case income and education). It establishes a decay function that weighs the effect of events in geographically proximate units more heavily than those in geographically distant units. Since a country is not a plain piece of land, Euclidean distance calculations or distance as ‘the crow flies’ make little economic sense when we are trying to investigate the effect of distance from districts with a large city on regional human development characteristics. The effect of the density of country’s infrastructure network is an important influence for which reason road distances have been utilized. For this reason this paper has utilized the inverse of the average of the maximum and the minimum roads distance between a district and its nearest district with a ‘*large city*’.

Finally both the matrices are row-standardized, which is a recommended procedure whenever the distribution of the variables under consideration is potentially biased due to errors in sampling design or due to an imposed aggregation scheme.

4. Exploratory Spatial Data Analysis

Exploratory spatial analysis aims to look for “associations instead of trying to develop explanations” (Haining, 2003: 358)**.** This chapter applies exploratory spatial data analysis (ESDA) techniques to district wise data on income, education, growth and development levels in order to detect the presence of spatial dependence. ESDA describes and visualizes spatial distributions, “identifies spatial outliers, detects agglomerations and local spatial autocorrelations, and highlights the types of spatial heterogeneities” (van Oort 2004, 107; Haining, 1990; Bailey and Gatrell, 1995; Anselin, 1988; Le Gallo and Ertur, 2003).The particular ESDA techniques employed in this study include the computation of Moran's *I* and Geary's *C* spatial autocorrelation statistics. They demonstrate the spatial association of data collected from points in space and measures similarities and dissimilarities in observations across space in the whole system (Anselin, 1995). However due to the presence of uneven spatial clustering, the Local Indicators of Spatial Association which measure the contribution of individual spatial units to the global Moran's I statistic have also been utilized (*Ibid*). The results are illustrated using Moran scatter plots that have been generated to demonstrate the spatial distribution of district wage and education levels across Pakistan.

4.1 Measures of spatial autocorrelation:

i) Global spatial autocorrelation

Spatial autocorrelation occurs when the spatial distribution of the variable of interest exhibits a systematic pattern (Cliff and Ord, 1981). Positive (negative) spatial autocorrelation occurs when a geographical area tends to be surrounded by neighbours with similar (dissimilar) values of the variable of interest. As previously mentioned, this paper utilizes two measures Moran's *I* and Geary's *C* statistics to detect the global spatial autocorrelation present in the data[8]. The Moran's *I* is the most widely used measure for detecting and explaining spatial clustering not only because of its interpretative simplicity but also because it can be decomposed into a local statistic along with providing graphical evidence of the presence of absence of spatial clustering.

It is defined as:

$$I = \frac{n}{S_0} \cdot \frac{\sum_i^n \sum_j^n w_{i,j}\,(y_i - \bar{y})(y_j - \bar{y})}{\sum_i^n (y_i - \bar{y})^2} \tag{1}$$

where y_i is the observation of variable in location i , $\bar{y}$ is the mean of the observations across all locations, n is the total number of geographical units or locations, $w_{i,j}$ is one of the elements of the weights matrix and it indicates the spatial relationship between location *i* and location j.

[8] Another well known measure of spatial autocorrelation is Getis and Ord's *G* statistic, see Anselin (1995a, p.22-23).

S_0 is a scaling factor which is equal to the sum of all the elements of the W matrix :

$$S_0 = \sum_i^n \sum_j^n w_{i,j} \tag{2}$$

S_0 is equal to n for row standardized weights matrices (which is the preferred way to implement the Moran's I statistic), since each row then adds up to 1. The first term in equation (1) then becomes equal to 1 and the Moran's I simplifies to a ratio of spatial cross products to variance.

Under the null hypothesis of no spatial autocorrelation, the theoretical mean of Moran's I is given by:

$$E(I) = -1/(n-1) \tag{3}$$

The expected value is thus negative and will tend to zero as the sample size increases as it is only a function of n (the sample size). Moran's I ranges from -1 (perfect spatial dispersion) to +1 (perfect spatial correlation) while a 0 value indicates a random spatial pattern. If the Moran's I is larger than its expected value, then the distribution of y will display positive spatial autocorrelation i.e. the value of y at each location i tends to be similar to values of y at spatially contiguous locations. However, if I is smaller than its expected value, then the distribution of y will be characterized by negative spatial autocorrelation, implying that the value of y at each location i tends to be different from the value of y at spatially contiguous locations. Inference is based on z-values computed as:

$$z_I = \frac{I - E(I)}{sd(I)} \tag{4}$$

i.e. the expected value of I is subtracted from I and divided by its standard deviation. The theoretical variance of Moran's I depends on the assumptions made about the data and the nature of spatial autocorrelation. This paper presents the results under the randomization assumption i.e. each value observed could have equally occurred at all locations[9]. Under this assumption z_I asymptotically follows a normal distribution, so that its significance can be evaluated using a standard normal table (Anselin 1992a). A positive (negative) and

[9] The other two assumptions include the assumption of normal distribution of the variables in question (normality assumption) or a randomization approach using a reference distribution for I that is generated empirically (permutation assumption). For details and formulas of the randomization assumption, see Sokal *et al*. 1998).

significant z- value for Moran's I accompanied by a low (high) p-value indicates positive (negative) spatial autocorrelation[10].

The second measure of spatial autocorrelation that has been utilized is the Geary's C which is defined as:

$$C = \frac{(N-1)\sum_i \sum_j w_{i,j}(X_i - X_j)^2}{2W\sum_i (X_i - \bar{X})^2} \tag{5}$$

where N is the number of spatial units (districts in our case); X is the variable of interest; $w_{i,j}$ represents the spatial weights matrix, where W is the sum of all $w_{i,j}$. The value of Geary's C lies between 0 and 2. Under the null hypothesis of no global spatial autocorrelation, the expected value of C is equal to 1. If C is larger (smaller) than 1, it indicates positive (negative) spatial autocorrelation. Geary's C is more sensitive to local spatial autocorrelation than Moran's I. Inference is based on z-values, computed by subtracting 1 from C and dividing the result by the standard deviation of C:

$$z_c = \frac{c-1}{sd(c)} \tag{6}$$

The standard deviation of C is computed under the assumption of total randomness, implying that z_c is asymptotically distributed as a standard normal variate (Anselin, 1992a; Pissati, 2001).

Finally, the results of the Moran's *I* and Geary's *C* are dependent on the specification of the weights matrix. Although interpretations change depending on whether the matrix was based on the use of physical distance or economic distance, a "pattern of decreasing spatial autocorrelation with increasing orders of contiguity (distance decay) is commonly witnessed in most spatial autoregressive processes regardless of the matrix specification" (van Oort, 2004: 109).

ii) Local spatial autocorrelation

Since the Moran's *I* and Geary's *C* are global statistics based on simultaneous measurements from many locations, they only provide broad spatial association measurements, ignore the location specific details, and do not identify which local spatial clusters (or *hot spots*)

---

[10] Negative spatial autocorrelation reflects lack of clustering, more than even the case of a random pattern. The checkerboard pattern is an example of perfect negative spatial autocorrelation.

contribute the most to the global statistic. As a remedy, local statistics commonly referred to as 'Local Indicators of Spatial Association (LISA)'are used along with graphic visualization techniques of the spatial clustering such as a Moran's Scatterplot (Fotheringham et al, 2000; Haining, 2003).

The Moran scatterplot is derived from the global Moran *I* statistic. Recall that the Moran's I formula when we use a row standardized matrix can be written as:

$$I=\frac{\sum_i^n (y_i-\bar{y})\ (\sum_i^n w_{i,j}\ (y_j-\overline{y}))}{\sum_i^n (y_i-\bar{y})^2} \tag{7}$$

This is similar to the formula for a coefficient of the linear regression b, with the exception of $(\sum_i^n w_{i,j}\ \ (y_j - \overline{y}))$, which is the so-called spatial lag of the location i.

Therefore *I* is formally equivalent to the regression coefficient in a regression of a location's spatial lag (Wz) on the location itself. This interpretation is used by the Moran's scatterplot, enabling us to visualize the Moran's I in a scatterplot of Wz versus z, where $z = y_i - \bar{y})/(y_i)$.Moran's *I* is then the slope of the regression line contained in the scatterplot. A lack of fit in this scatterplot indicates local spatial associations (local pockets/non-stationarity). This scatterplot is centered on 0 and is divided in four quadrants that represent different types of spatial associations.

5. Empirical Results

5.1 *Spatial autocorrelation estimates for district-wise income inequality levels*

Our first empirical estimation involves calculating measures of spatial dependence for district income inequality (measured as Gini coefficient of average district earnings income) in the year 2004-05. Table 1 provides the results of Moran's *I* statistic and Geary's *C* statistic for district income inequality levels using the two weight matrices. In both the cases, the null hypothesis of no spatial dependence of income inequality between districts is rejected at the significance level of 1% as the measures demonstrate a weakly positive spatial autocorrelation amongst district inequality levels (0.21 under BC matrix specification and 0.25 under ID matrix specification). The results for Geary's C statistic have been reported in Table A2a in the Appendix. This implies that income inequality in one district is not strongly spatially associated with income inequality in its neighbouring districts in the case of Pakistan.

Table 1: Global Autocorrelation results for Income Inequality—Moran's I (2005)

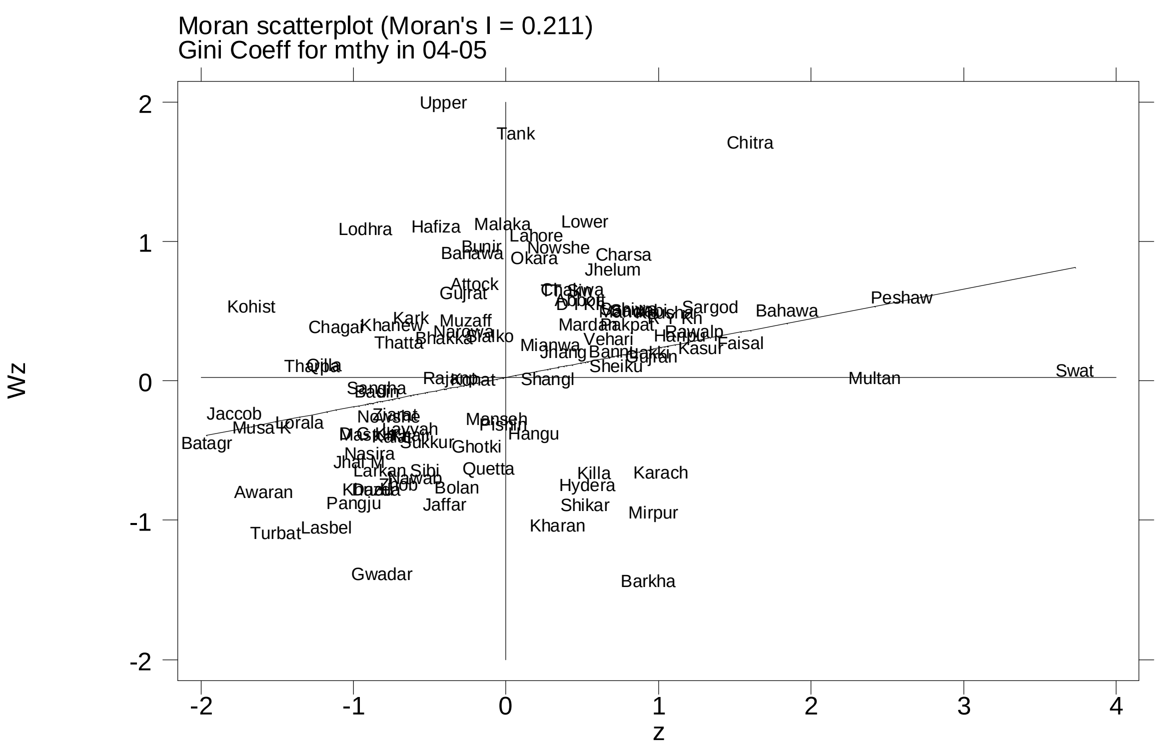
Weight Matrix	I	II
i ≠ *j*	$w_{i,j} = 0 \text{ or } 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
i = j	$w_{i,i} = 0$	
Moran's *I*	0.211	0.257
E(I)	-0.010	-0.010
Sd(I)	0.074	0.103
Z	2.985	2.601
p-value	0.003	0.009

5.2 Local spatial association between district-wise income inequality levels

The Moran scatterplot provides a more disaggregated view of the nature of the global autocorrelation. It not only provides us information on the presence of clusters in the data but also on the outliers contained in it (see Figure 1). This scatterplot is divided into four quadrants, each of which represents a different type of spatial association. The upper right quadrant (High-High zone) represents spatial clustering of a district with a high level of the variable under study (income inequality in our case) around neighbours that also have high values of income inequality as demonstrated by the high values of both, the Z-score and the Wz (the spatial lag). The upper left quadrant (Low z – High Wz zone) represents spatial clustering of a district with a low level of income inequality with neighbouring districts that have a high income inequality levels. The lower left quadrant (Low z – Low Wz zone) represents spatial clustering of a district with a low income level around neighbours that also have low incomes. The lower right quadrant (High z – Low Wz zone) represents spatial clustering of a high income inequality district with neighbours that have low income inequality levels.

Figure 1 illustrates the results obtained in Col I of Table 1 via a Moran scatterplot for Gini coefficient of district per capita incomes using the binary contiguity weights matrix. It shows a positive global Moran's I (z-score = 2.98), which is represented by the slope of the black line. Due to the weakly positive spatial autocorrelation, we are unable to detect any substantial clusters of high (or low) inequality districts in particular for the year 2005. Similarly, Figure A8 (see Appendix) also shows a Moran scatterplot for Gini coefficient of district per capita incomes, however it has utilized an inverse distance weights matrix instead. The overall spatial autocorrelation is although statistically significant, it still remains weak.

Figure 1. Spatial Autocorrelation of District Income Inequality using the BC matrix



5.3 Spatial association between district-wise education levels

The role of human capital in generating growth is important since the distribution of income is mainly driven by the distribution of human capital within a country (Golmm and Ravikuman, 1992; Saint-Paul and Verdier, 1993; Galor and Tsiddon, 1997). Hence the operation of human capital externalities and knowledge spillovers plays an important role in generating regional dependencies and disparities. It has been demonstrated that regions located in an economic periphery experience lower returns to skill attainment and hence have reduced incentives for human capital investments and agglomerations. However spatial externalities do not spread without limits (Darlauf and Quah, 1999) as a result of which closely related economies or regions tend to have similar kinds of human capital externalities and technology levels as compared to the more distant ones (see Quah, 1996; Mion, 2004). This section investigates the spatial disparities in education levels across Pakistan, the extent to which neighbouring districts share similar levels of education, and examines whether district human development level inequalities are spatially associated.

In order to do so, this paper uses the average district wise education attainment level (which is measured as the average number of schooling years completed in a district) as a proxy for human capital. It is expected that neighbours of districts with high education attainment should also have high educational awareness and hence similar if not equal attainment levels. Again the Moran's *I* global and local indices along with a Moran scatterplot and Geary's C statistic have been utilized.

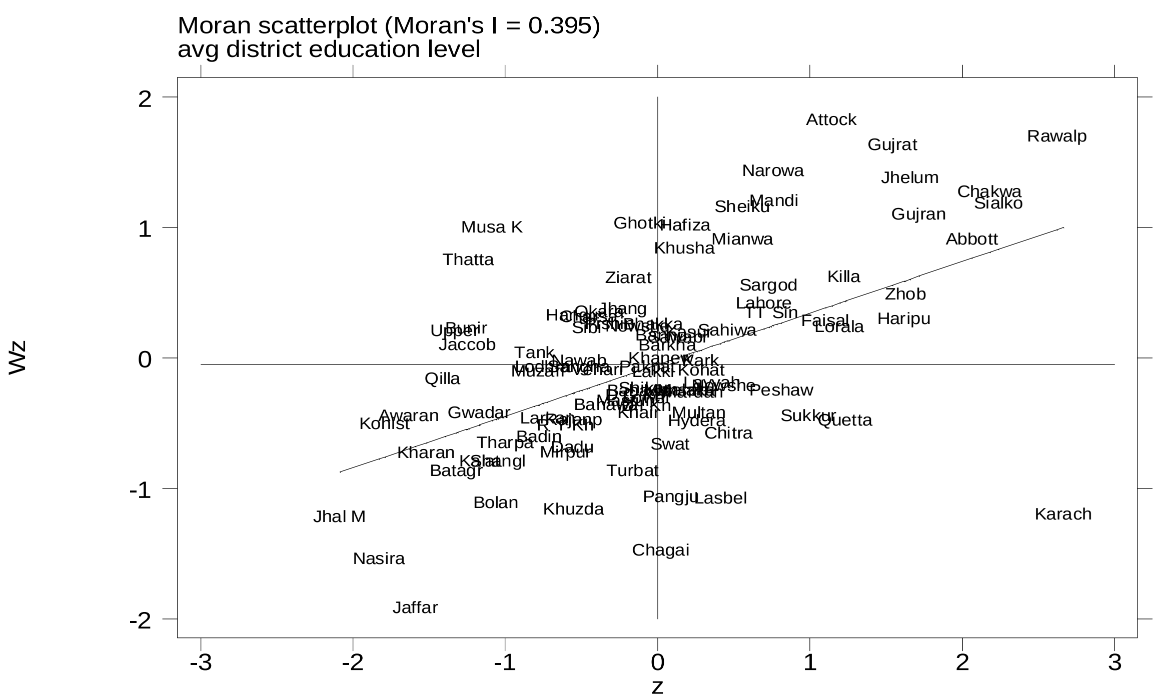
Our results indicate that there exists a greater possibility of knowledge spillovers between districts that share a border, as compared to when they do not (see Table 2). The global Moran's I for average district education level (measured as the average education attainment of a district's citizens) is positive and statistically significant when neighbourhood is defined in terms of contiguity, however it is negative and statistically insignificant when neighbourhood is defined in terms of proximity. These results imply that for a Pakistani district, sharing a border with a district whose individuals have a high (low) education level, '*may'* result in rising (lowering) its own education levels.

The positive pattern for spatial autocorrelation for average district education levels demonstrated by the BC matrix shows more clusters with low education levels (in the case of Balochistan) and high education levels (in the case of Punjab) as compared to outliers. Districts in northern Punjab emerge in the High-High quadrant and confirm our assumption about high human capital districts being located close to each other (Figures 2 and A5). Similar empirical findings have also been put forward in a recent study on agglomeration patterns of industries across Pakistani districts in a study by Burki and Khan (2010).

Table 2: Global Autocorrelation results for Education Attainment—Moran's I (2005)

Weight Matrix	I	II
$i \neq j$	$w_{i,j} = 0\ or\ 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
$i = j$	$w_{i,i} = 0$	
Moran's *I*	0.395	-0.003
E(I)	-0.010	-0.01
Sd(I)	0.075	0.103
Z	5.440	0.072
p-value	0.000	0.943

Figure 2. Spatial Autocorrelation of District Education Levels using the BC matrix



The neighbouring districts of Karachi and Thatta emerge as the most significant outliers when we analyze the local Moran's I values using the BC and the ID matrices. While Karachi falls into the High-Low zone, Thatta falls in the Low-High zone. However, the fact that being a neighbour with Karachi (a district with one of the highest average education levels in

Pakistan) does not translate in Thatta having improved human capital characteristics is not very surprising. Regional science and regional economics literature has demonstrated that the economic influence and knowledge spillover effects of coastal cities (such as Karachi) are quite different from the pattern of spillovers generated by landlocked regions (Glaeser *et al*, 1992; Henderson, 2003). The overall spatial pattern of autocorrelation is quite diffused when we use the ID matrix for analysis (see Figure A5). However under both the neighbourhood structures Rawalpindi, Abbottabad, Chakwal and Jhelum emerge as a statistically significant cluster of districts with high average education attainment levels.

5.4 The dynamics of spatial association between district-wise income inequality and education levels

This section analyses the temporal change in the spatial distribution of district wise real per capita GDP growth rate, district wise per capita incomes, and district human development levels between 1998 and 2005. It also examines the spatial association between district wise primary, secondary, and bachelors education levels in 1998.

Figures A3a, A3b, A3c, and A3d in the Appendix each demonstrates a Moran scatterplot which provides a disaggregated picture of the nature of spatial autocorrelation for district per capita income in 1998 and 2005, using the BC and ID matrix respectively. The spatial lag (Wz) in this situation is a weighted average of the incomes of a district's neighbouring districts. The scatter plots in both the years (using both the matrices) demonstrate that the overall pattern of spatial dependence between district income levels has remained positive and statistically significant. However, the overall value of the global Moran's I statistic has reduced from being 0.81 to 0.38 between 1998 and 2005 when the results are reported using the BC matrix. Similarly, the value of global Moran's I statistic has reduced from being 0.91 to 0.51 between 1998 and 2005 under the results produced using the ID matrix.

Furthermore a spatial analysis of the growth rate between 1998 and 2005, also indicates a positive and a statistically significant spatial autocorrelation pattern when neighbourhood is defined in terms of contiguity but a statistically insignificant pattern when neighbourhood is defined in terms of proximity as measured by the ID matrix (see Table 3). This implies that districts with a high (low) real GDP growth rate may be spatially associated with their contiguous neighbouring districts which also have high (low) real GDP growth rates.

Table 3. Spatial Autocorrelation of per capita GDP Growth Rate between 1998—2005

GDP Growth Rate (1998-2005)		
	BC matrix	*ID matrix*
Moran's I	0.430	0.140
E(I)	-0.010	-0.010
Sd(I)	0.071	0.099
Z	6.204	1.524
P-value	0.000	0.128

Source: Author's own calculations

Moreover, since our macro-data from 1998 provides district wise statistics on individual education attainment levels (measured as the percentage of individuals having completed an education level), it has allowed us to analyse whether education levels in neighbouring districts are spatially associated or how the distance from large neighbouring cities (or provincial capitals) affects the incentives to obtain education in a district. Table 4 demonstrates that whether neighbourhood is measured in terms of geographic proximity (using ID matrix) or in terms of geographic contiguity (using BC matrix), there exists a positive and highly significant spatial autocorrelation for levels of education below high-school (i.e primary, matric i.e. grade 10, and inter i.e. grade 12). However, for higher levels (Bachelors and above), geographic contiguity to a district with a high percentage of graduates could be more influential than the distance from the provincial capital or the nearest large city.

Finally, although spatial association between district development levels (as measured by the Human Development Index (HDI) calculated by the UNDP in NHDR, 2003) has reduced between 1998 and 2005 from 0.40 to 0.311, it still remains positive and significant (see Table 5). These results for Pakistani districts again confirm the findings of the new economic geography literature that a region's development levels, depend on the development levels prevailing in its neighbouring regions.

Table 4. Spatial Autocorrelation for Education Levels (1998)

Primary Education				*Matric*				*Higher Education—Bachelors*		
	BC	ID			BC	ID			BC	ID
Moran's I	0.494	0.559		Moran's I	0.391	0.247		Moran's I	0.327	-0.014
E(I)	-0.010	-0.010		E(I)	-0.010	-0.010		E(I)	-0.010	-0.010
Sd(I)	0.075	0.103		Sd(I)	0.074	0.102		Sd(I)	0.074	0.102
Z	6.745	5.501		Z	5.443	2.523		Z	4.582	-0.038
P-value	0.000	0.000		P-value	0.000	0.012		P-value	0.000	0.969
Geary's C	0.497	0.983		Geary's C	0.610	0.703		Geary's C	0.610	1.643
E(c)	1.000	1.000		E(c)	1.000	1.000		E(c)	1.000	1.000
Sd(c)	0.079	0.244		Sd(c)	0.085	0.379		Sd(c)	0.086	0.392
Z	-6.401	-0.069		Z	-4.573	-0.783		Z	-4.538	4.193
P-value	0.000	0.945		P-value	0.000	0.434		P-value	0.000	0.000
Source: Author's own calculations. BC: Binary Contiguity Matrix, ID: Inverse Distance Matrix										

Table 5. HDI Spatial Autocorrelation using the Binary Contiguity Matrix

District Human Development Index (HDI)		
	1998	*2005*
Moran's I	0.405	0.311
Standard deviation (I)	0.075	0.074
Z-value	5.573	4.341
P-value	0.000	0.000

Source: Author's calculations using data from NHDR (2003).

6. Conclusions

This paper has performed an exploratory analysis of socio-economic disparities across Pakistan for the first time and has provided useful insights for the conduct of economic regional policy in Pakistan. It has investigated the spatial distribution of income inequality, income, education, growth and development levels for 98 districts between 1998 and 2005. The overall finding that emerges from this chapter is that the distribution of district wise income inequality, income, education attainment, growth, and development levels, exhibits a significant tendency to cluster in space (i.e. the presence of spatial autocorrelation is confirmed), thereby highlighting the importance of understanding economic geography in the context of Pakistan.

Specifically the following main findings emerge from this chapter. First, the province of Punjab contains the largest cluster of high per capita income districts in both 1998 and 2005. Second, district wise income inequality levels demonstrate weak spatial association. Moreover district education levels reveal high spatial association, and districts with a high (low) real GDP growth rate have been spatially associated with contiguous neighbouring districts which also have high (low) real GDP growth rates between 1998 and 2005.Third, there exists positive spatial dependence for education levels below bachelors (i.e. primary, matric i.e. grade 10, and inter i.e. grade 12). However, for higher levels (Bachelors and above), geographic contiguity to a district with a high percentage of graduates, is more influential than the distance from the provincial capital or the nearest large city. This result is corroborated by the findings from Burki and Khan (2010) which confirms that districts located away from urban centers are also the ones with lowest education levels in Pakistan. Our empirical analysis also reveals that except for Lahore, none of the other 3 provincial capitals of Pakistan (Karachi, Peshawar, Quetta) have high knowledge spillovers. While this finding is not surprising for Karachi, since coastal cities have different spillover mechanisms as compared to landlocked cities, it indicates that infrastructure and cluster development can facilitate increased knowledge spillovers at least from the centers of economic activity in Pakistan if not from all large city districts. Finally, spatial association of district wise Human Development Indicators confirms that a district's development levels may depend on the development levels prevailing in its neighbouring districts in Pakistan.

The methodological implication of the above mentioned results is that studies which utilize Ordinary Least Squares to investigate intra- Pakistan socio-economic issues could possibly be producing inaccurate statistical inferences. By assuming spatial-independence, they may produce estimates that are biased and over estimated, since our results show that observations for socio-economic district characteristics do tend to cluster in Pakistan. The main policy implication that emerges from our results is that growth and development policies need to focus on infrastructure and cluster development that can cater to large segments of the population. This is particularly because the spatial pattern of income inequality, district incomes, education levels, and development levels shows how development in Pakistan is concentrated in Punjab (in particular Northern Punjab especially in terms of human development indicators).

The presence of possible spatial spillovers as demonstrated in this paper also implies that cluster development can play an extremely important role in generating knowledge externalities, domestic commerce, and employment creation by bringing work and knowledge to people instead of them travelling to it. Pakistan already has many pseudo-clusters that have developed over time. Examples include the IT cluster ‘Karachi’, textile and leather cluster ‘Faisalabad’, automotive manufacturing cluster ‘Port Qasim’, furniture cluster ‘Gujranwala’, light engineering cluster ‘Gujrat’, sports and surgical cluster ‘Sialkot’, heavy industries cluster ‘Wah’ and even light weapons manufacturing cluster ‘Landikotal’. An emphasis on regional and industrial regeneration policies can play a crucial role in reducing spatial disparities and enhancing the regional advantages of these districts (Planning Commission, 2011). Finally, this paper has highlighted the importance of additional research on Pakistan that takes into account spatial effects. Since it has only considered spatial changes in socio-economic phenomena in 8 years between 1998 and 2005, an immediate possibility could be to extend this spatio-temporal analysis may include extending it over a longer period of time. Another possibility may involve a spatial econometric analysis of the effect of a district’s inequality, income and education levels on its growth. While the presence of spatial clustering of income and education in Pakistan (as demonstrated in this paper) could support the use of a spatial lag model to capture the spillover of inequality between districts, missing data on district incomes or omitted variables could also necessitate the use of a spatial error model (which reflects spatial autocorrelation in measurement errors) in analyzing the effect of inequality on district income levels.

APPENDIX

Table A1. List of Districts

	PUNJAB		SINDH	67	Chitral
				68	Malakand Agency
1	Rawalpindi	35	Hyderabad	69	Shangla
2	Jhelum	36	Dadu	70	Bannu
3	Chakwal	37	Badin	71	Lakki Marwat
4	Attock	38	Thatta	72	D I Khan
5	Gujranwala	39	Mirpur Khas	73	Tank
6	Mandi Bahauddin	40	Sanghar	74	Bunir
7	Hafizabad	41	Tharparkar		
8	Gujrat	42	Sukkur		BALOCHISTAN
9	Sialkot	43	Ghotki	75	Quetta
10	Narowal	44	Khair pur	76	Sibi
11	Lahore	45	Nawab shah	77	Nasirabad
12	Kasur	46	Larkana	78	Kalat
13	SheikuhuPura	47	Jaccobabad	79	Pishin
14	Okara	48	Shikarpur	80	Qilla Abd
15	Faisalabad	49	Nowshero Feroz	81	Bolan
16	Jhang	50	Karachi	82	Pangjur
17	TT Singh			83	Barkhan
18	Sargodha		KP	84	Chagai
19	Khushab	51	Peshawar	85	Jaffarabad
20	Mianwali	52	Charsadda	86	Jhal Magsi
21	Bhakkar	53	Nowshera	87	Mastung
22	Multan	54	Kohat	88	Awaran
23	Khanewal	55	Kark	89	Gwadar
24	Lodhran	56	Hangu	90	Turbat
25	Vehari	57	Mardan	91	Kharan
26	Sahiwal	58	Sawabi	92	Ziarat
27	Pakpattan	59	Abbottabad	93	Khuzdar
28	Bahawalpur	60	Haripur	94	Killa Saif
29	Bahawalnagar	61	Mansehara	95	Lasbella
30	R Y Khan	62	Batagram	96	Loralai
31	D G Khan	63	Kohistan	97	Musa Khel
32	Muzaffar grah	64	Swat	98	Zhob
33	Layyah	65	Lower Dir		
34	Rajanpur	66	Upper Dir		

Table A2a. Global autocorrelation results for income inequality—Geary's C (2005)

Weight Matrix	I	II
i ≠ j	$w_{i,j} = 0\ or\ 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
i = j	$w_{i,i} = 0$	
Geary's C	0.824	1.458
E(C)	1.000	1.000
Sd(C)	0.082	0.324
Z	-2.138	1.413
p-value	0.033	0.158

Source: Author's Calculations

Table A2b. Global autocorrelation results for district per capita income— BC Matrix

Weight Matrix	1998	2005
i ≠ j	$w_{i,j} = 0\ or\ 1$	$w_{i,j} = 0\ or\ 1$
i = j	$w_{i,i} = 0$	
Moran's *I*	0.818	0.380
E(I)	-0.010	-0.010
Sd(I)	0.103	0.101
Z	8.048	3.856
p-value	0.000	0.000

Source: Author's Calculations

Figure A3a. Moran Scatterplot real per capita district income, 1998 (BC matrix)

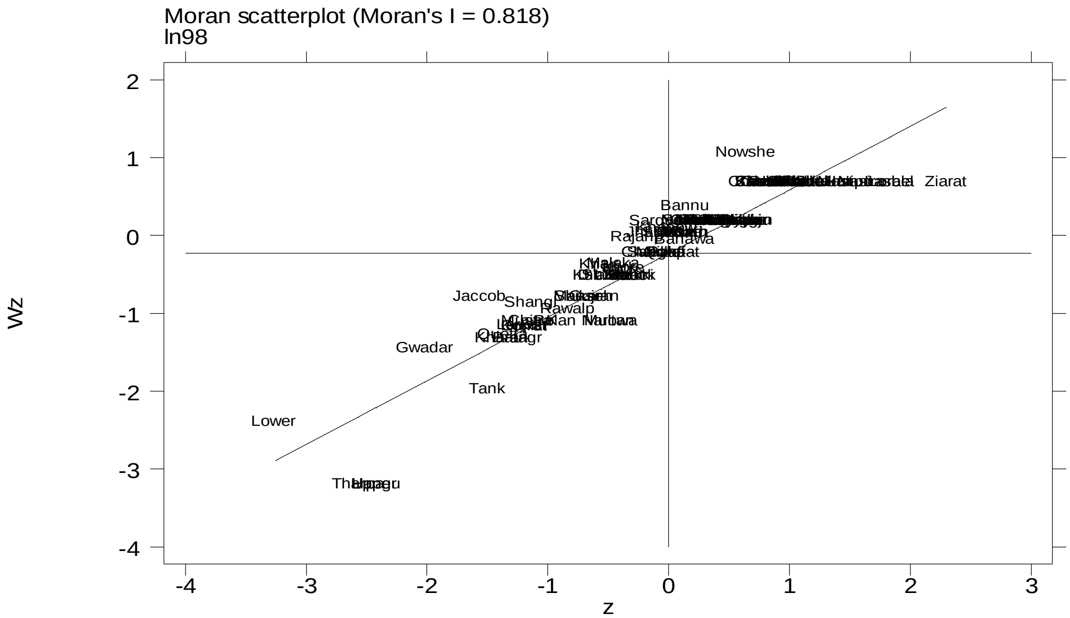


Figure A3b. Moran scatterplot for real per capita district income, 2005 (BC matrix)

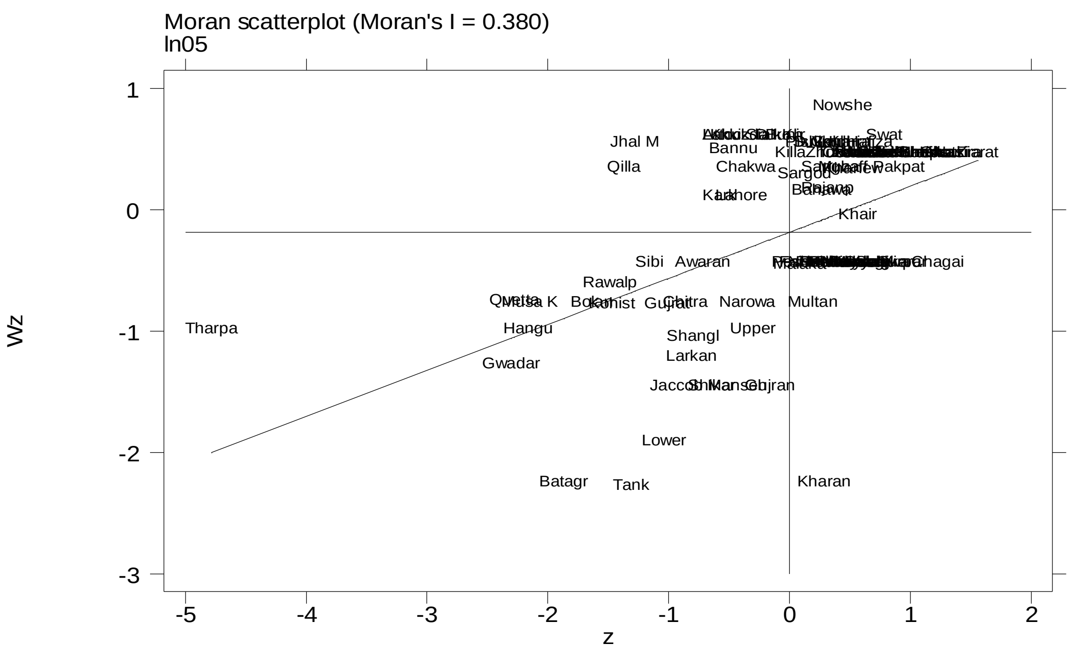


Figure A3c. Moran scatterplot district real per capita income, 2005 (ID matrix)

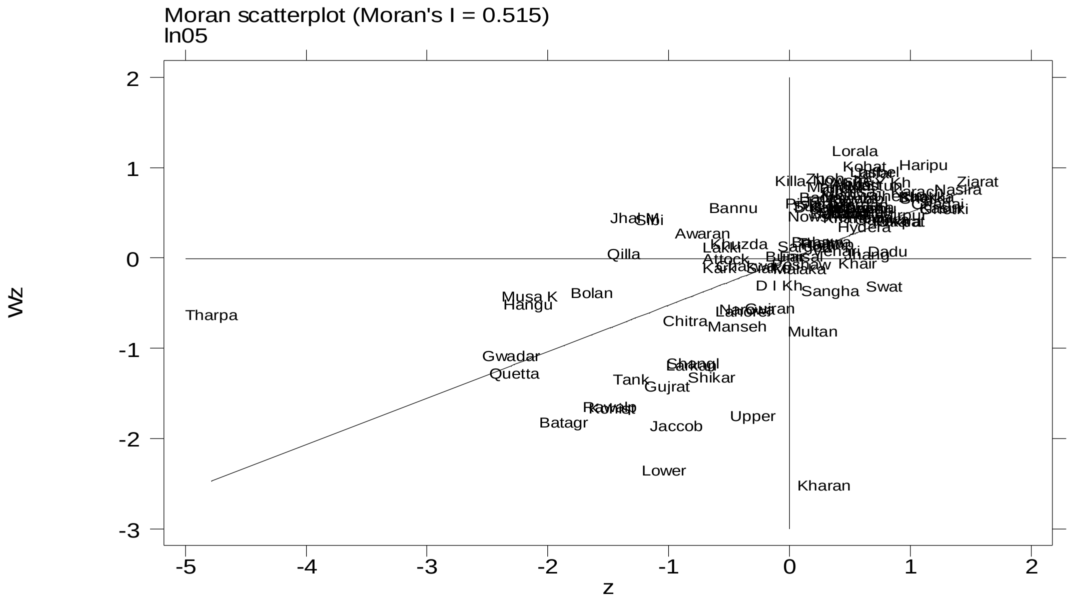


Figure A3d. Moran scatterplot district real per capita income, 1993 (ID matrix)

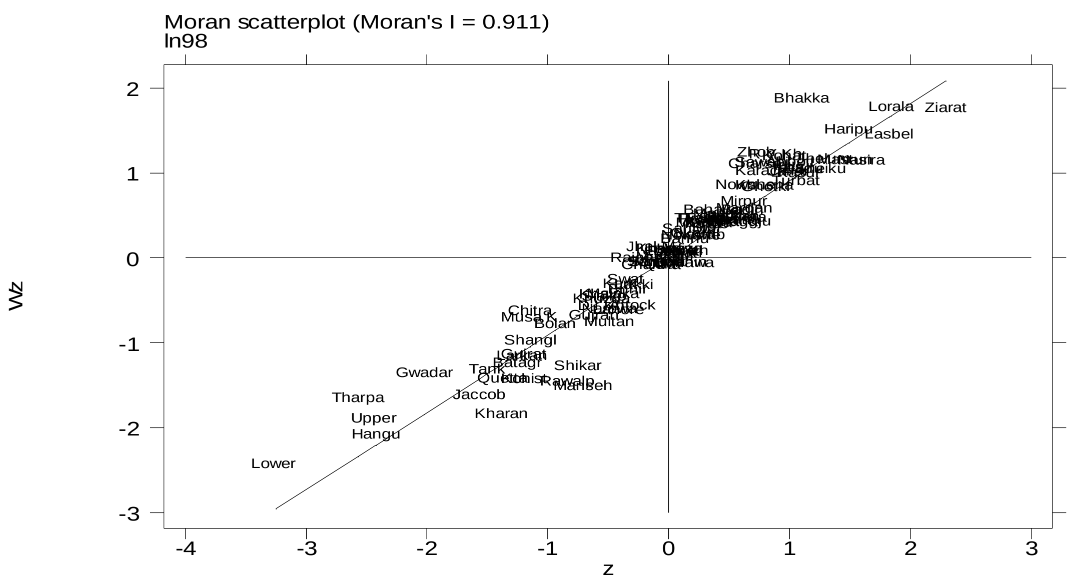


Figure A4. Moran scatterplot for average district education level using the ID matrix

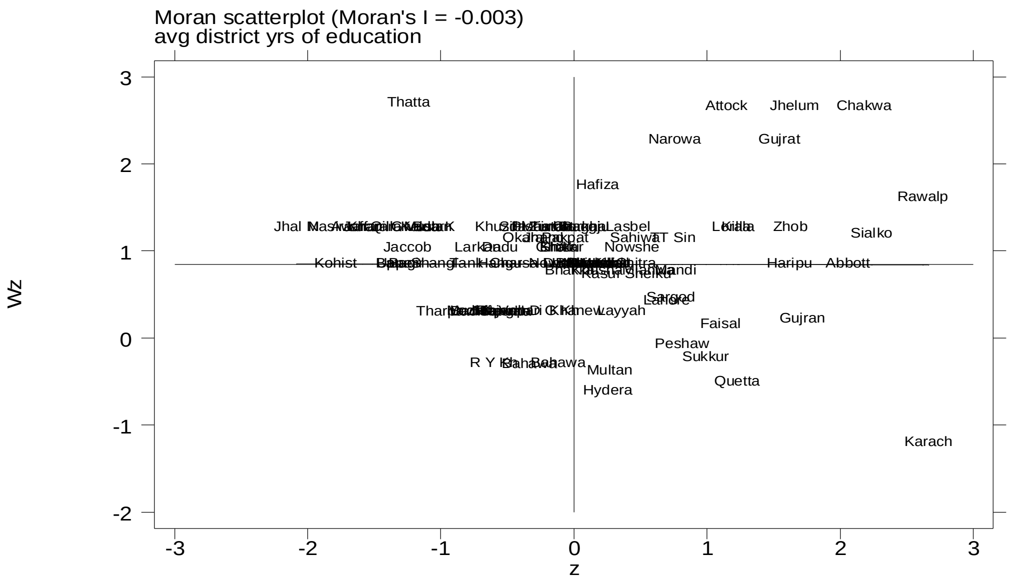


Table A5. Global autocorrelation results for education attainment—Geary's C (2005)

Weight Matrix	**I**	**II**
i ≠ *j*	$\boldsymbol{w_{i,j} = 0 \text{ or } 1}$	$\boldsymbol{w_{i,j} = \frac{1}{d_{i,j}}}$
i* = *j	$\boldsymbol{w_{i,i} = 0}$	
Geary's *C*	0.584	1.092
E(*C*)	1.000	1.000
Sd(*C*)	0.080	0.275
Z	-5.230	0.336
p-value	0.000	0.737

Source: Author's Calculations

Figure 6a. Moran's scatterplot for primary education using the BC matrix

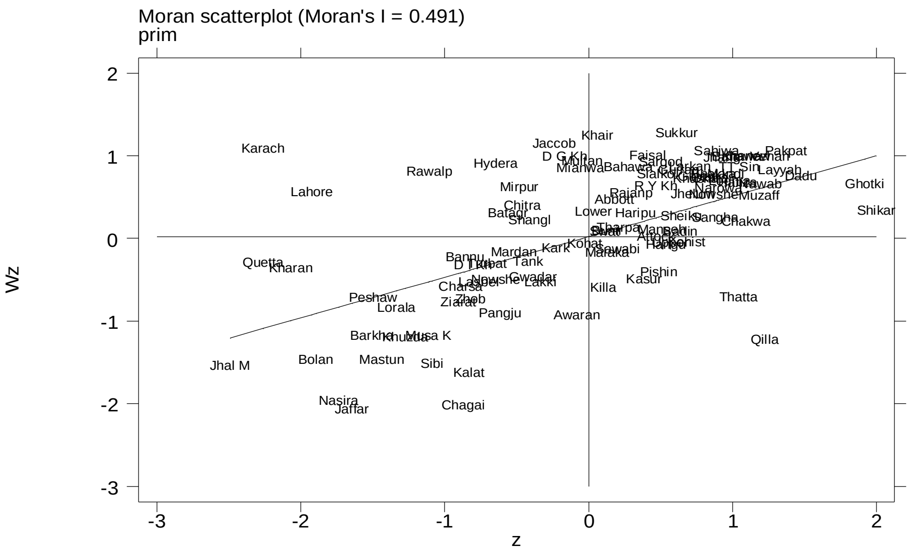


Figure 6b. Moran's scatterplot for primary education using the ID matrix

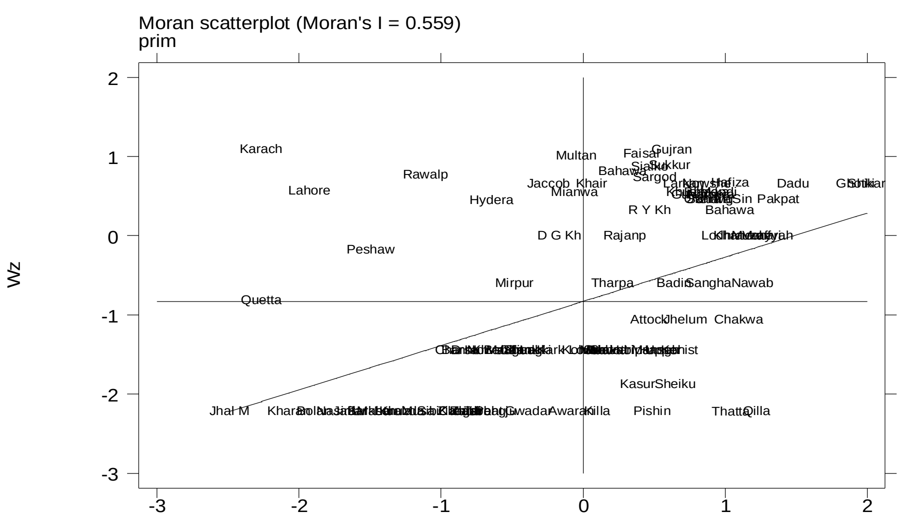


Figure A7a. Moran's scatterplot for higher education in 1998 using the BC matrix

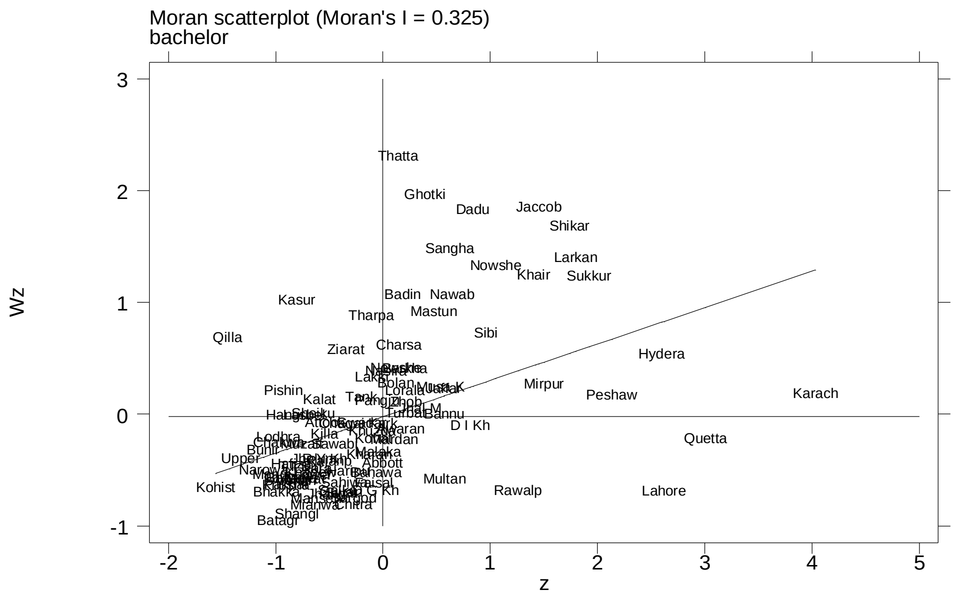


Figure A7b. Moran's scatterplot for higher education in 1998 using the ID matrix

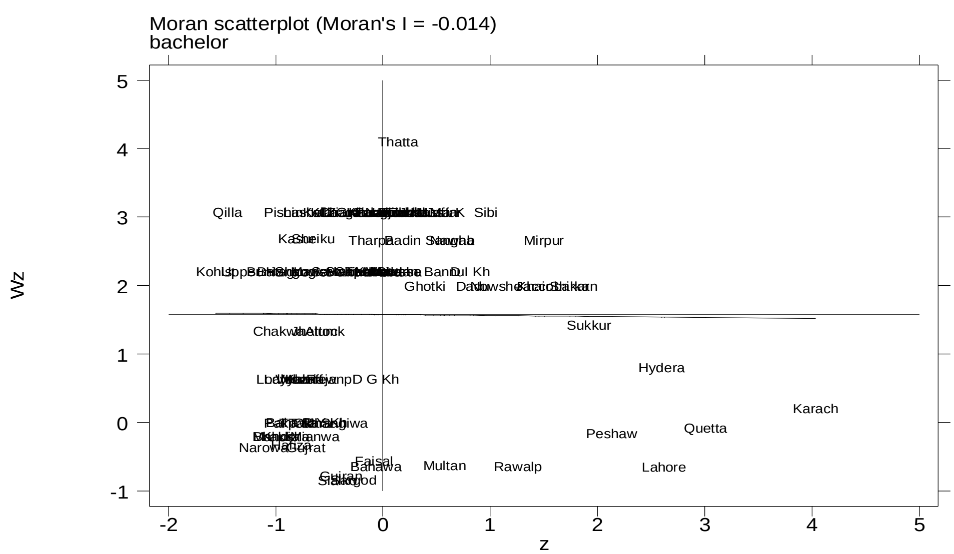
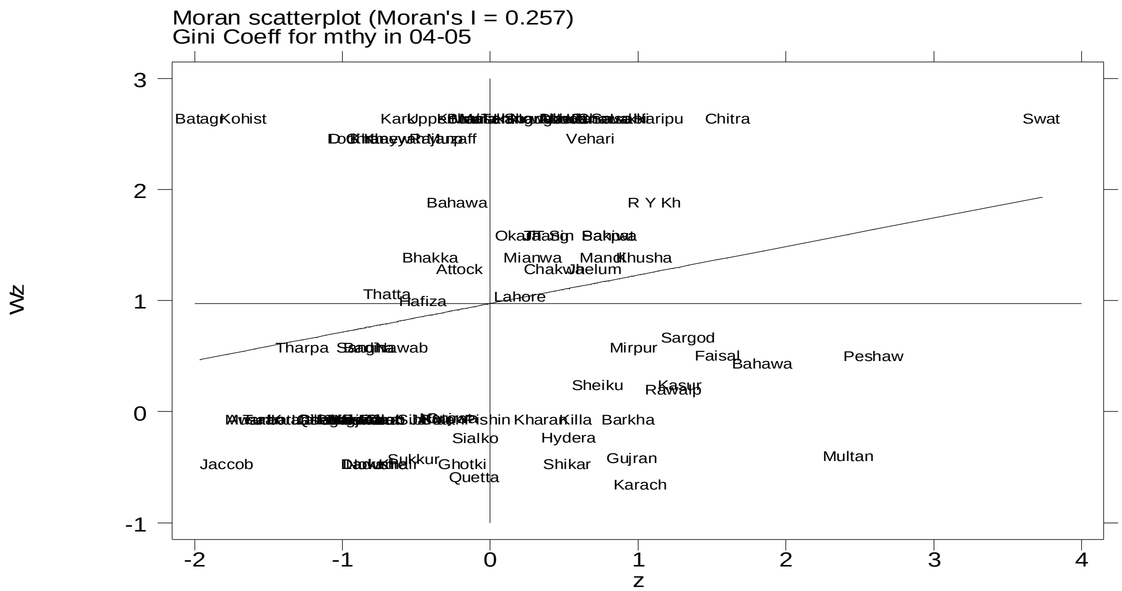


Figure A8: Spatial autocorrelation of district-wise income inequality using the ID matrix



BIBLIOGRAPHY

Anselin, L. (1988a). “Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeniety,” *Geographical Analysis*, 20:1-23.

Anselin, L. (1988b). *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer Academic Press.

Anselin, L. (1995a). SpaceStat. A Software Program for the Analysis of Spatial Data (version 1.80), Morgantown: Regional Research Institute, West Virginia University.

Anselin, L. (1995b). “Local Indicators of Spatial Association—LISA”, *Geographical Analysis,* 27: 93-115.

Anselin, L. (1992a). “Space and applied econometrics.” Special Issue, *Regional Science and Urban Economics*, 22.

Anselin, L. (1992b). SpaceStat, a Software Program for the Analysis of Spatial Data. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA.

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

Arbia, G., R. Benedetti, & G. Espa. (1996). ‘‘Effects of the MAUP on Image Classification.’’ *Geographical Systems,* 3: 123–41.

Arif, G.M., N Iqbal & S, Farooq. (2010). “The 2010 flood and poverty in Pakistan: A Preliminary District Level Analysis” Conference Paper for *The Environments of the Poor,* New Dehli.

Bailey, T.C. & A. C. Gatrell. (1995). *Interactive Spatial Data Analysis*. Addison Wesley Longman.

Baldwin, R., R. Forslid., P. Martin., G. Ottaviano & F. Robert-Nicoud. (2003a). *Economic Geography and Public Policy*, Princeton University Press.

Balisacan, A. M & Nobuhiko, F. (2004). "Changes in Spatial Income Inequality in the Philippines: An Exploratory Analysis," *Working Papers UNU-WIDER Research Paper*, World Institute for Development Economic Research (UNU-WIDER).

Beck, N., K. Gleditsch, & K. Beardsley. (2006). “Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy.” *International Studies Quarterly*, 50: 27-44.

Burki, A., A. K. Munir., M. Khan., U. Khan., A. Faheem., A. Khalid & S.T. Hussain. (2010). “Industrial Policy, Its Spatial Aspects and Cluster Development in Pakistan.” Analysis Report for the Industrial Policy 2010, Lahore University of Management Sciences.

Burki, A. A & Khan, M. A (2010). “Spatial Inequality and Geographic Concentration of Manufacturing Industries in Pakistan” 26th Annual General Meeting, Pakistan Institute of Development Economics.

Celebioglu, F *& S*. Dall'erba. *(*2010). "Spatial Disparities across the Regions of Turkey: An Exploratory Spatial Data Analysis," *The Annals of Regional Science*, 45(2): 379-400.

Cliff, A.D & J.K. Ord. (1981). *Spatial Processes*. London: Pion.

Darlauf, S.N & D.T., Quah. (1999). "The New Empirics of Economic Growth," in Taylor, J.B. and Woodford, M. (Eds), *Handbook of Macroeconomics*, Vol. IA, Chap 4, North-Holland, Amsterdam.

Dominicis, L., G, Arbia & H. L.F, de Groot. (2007). "The Spatial Distribution of Economic Activities in Italy." *Tinbergen Institute Discussion Papers* 07-094/3, Tinbergen Institute.

Fotheringham, A., C. Bunsden & Charlton M. (2000). *Quantitative Geography*. London: Sage.

Franzese, R. & J. Hays. (2007b). "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data," *Political Analysis*, 15(2):140-64.

Galor, O., & D. Tsiddon. (1997). "Technological Progress, Mobility, and Economic Growth." *American Economic Review, 87*: 363–382.

Glomm, Gerhard., & B. Ravikumar. (1992). "Public Versus Private Investment in Human capital: Endogenous Growth and Income Inequality." *Journal of Political Economy,* 100(4): 818-834.

Glaeser, E., Kallal, H., Sheinkman, J., & A, Shliefer (1992). "Growth in Cities," *Journal of Political Economy*, 100: 1126-1152.

Haining, R. (2003). *Spatial Data Analysis. Theory and Practice*. Cambridge, Cambridge University Press

Henderson, J. V (2003). "Marshall's scale economies," *Journal of Urban Economics*, 53(1): 1-28.

Jamal, H & A.J. Khan. (2003). "The Changing Profile of Regional Inequality," *The Pakistan Development Review*, Pakistan Institute of Development Economics, 42(2): 113-123.

Jamal, H., A. J. Khan., I. A. Toor & N. Amir. (2003). "Mapping the Spatial Deprivation of Pakistan," *The Pakistan Development Review*, Pakistan Institute of Development Economics, 42(2): 91-11.

Jamal, H & A. J. Khan. (2005). "The Knowledge Divide: Education Inequality in Pakistan" *The Lahore Journal of Economics*, 10(1): 83-104.

Jamal, H & A. J. Khan. (2008). "Trends in Regional Human Development Indices", Research Report 73.

Jamal, H & A. J. Khan. (2008). "Education Status of Districts: An Exploration of Inter-Temporal Changes", Research Report 71.

Kanbur,R & T.Venables.(2005). "Introduction: Spatial Inequality and Development," *Journal of Economic Geography*, 5(1).

Khan, R. I. A. (2003). *Spatial Distribution of Population with Special Refernce to 1998 Population Census*. In A.R.Kemal, M.Irfan and N.Mahmood (eds.*). Population of Pakistan: An Analysis of 1998 Population and Housing Census*. Islamabad: Pakistan Institute of Development Economics.

Krugman, P. (1991). *Geography and Trade*. MIT Press, Cambridge.

Krugman, P & A. J. Venables. (1995). "Globalization and Inequality of Nations." *Quarterly Journal of Economics,* 11: 857-880.

LeSage, J.P. (1999). “The Theory and Practice of Spatial Econometrics.” Unpublished, Dept. of Econ., University of Toledo.

Le Gallo, G & C. Ertur. (2003) "Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995." *Journal of Economics,* 82(2): 201.

Mion, Giordano (2004). "Essays in spatial economics," Open Access publications from Université catholique de Louvain.

Naqvi, S. A. A (2007). "A Look at Spatial Inequality in Pakistan - Case Study of District Sargodha." *Center for Global, International and Regional Studies. Mapping Global Inequalities - conference papers.* Paper mgi-9.

Pose, A.r & V, Tselios (2007). "Mapping the European regional educational distribution: Educational attainment and inequality," Working Papers, 2007-18, Instituto Madrileño de Estudios Avanzados (IMDEA) Ciencias Sociales.

Planning Commision of Pakistan. (2011). Framework for Economic Growth Pakistan. Government of Pakistan. Islamabad.

Quah, D. (1996). “Regional convergence clusters in Europe.” *European Economic Review,* 35: 951-958.

Rey, S. J & B.D. Montouri. (1999). “US Regional Income Convergence: A Spatial Econometric Perspective.” *Regional Studies, 33*(2): 143-156.

Saint-Paul, G & T, Verdier. (1993). "Education, Democracy and Growth." Journal *of Development Economics,* 42(2): 399-407.

Siddiqui, R. (2008) . "Income, Public Social Services, and Capability Development: A Cross-district Analysis of Pakistan," PIDE-Working Papers 2008:43.

Sokal R. R., N.L. Oden N. L., & B.A. Thomson. (1998). “Local Spatial Autocorrelation in a Biological Model.” *Geographical Analysis,* 30: 331- 354.

Tobler, W.(1970). "A computer movie simulating urban growth in the Detroit region". *Economic Geography,* 46(2): 234-240.

United National Development Program (2003). *Pakistan National Human Development Report*. Oxford University Press, Karachi.