

Spatial Inequality and Geographic Concentration of Manufacturing Industries in Pakistan

Abid A. Burki
Professor of Economics
Lahore University of Management Sciences (LUMS)
Opposite Sector U, DHA, Lahore Cantt.54792
Lahore, Pakistan
E-mail: burki@lums.edu.pk

AND

Mushtaq A. Khan
Associate Professor of Economics
Lahore University of Management Sciences (LUMS)
Opposite Sector U, DHA, Lahore Cantt.54792
Lahore, Pakistan
E-mail: mushtaq@lums.edu.pk

This version: November 15, 2010

Abstract:

This paper examines the nature of spatial inequality and causes of geographic concentration of manufacturing industries in Pakistan. The mapping of districts as spatial units suggests that firms are not uniformly distributed across space. They are mostly clustered in districts where there is large market, high road density and pooling of educated and skilled labor force. We use plant level data from CMI supplemented by external information to analyze geographic concentration of manufacturing plants. In general, our results suggest that strong and moderate concentration of manufacturing industries is widespread and that there is a declining trend in the dynamic industrial concentration. The econometric results confirm that increase in population size, higher road density, and pooling of technically trained workers promote geographic concentration of manufacturing industries. Moreover, localization economies (or within-industry externalities) are much more important in Pakistan than inter-industry learning or technological spillovers. Industries that offer highest local scale economies are also the most agglomerated. This paper also shows that productivity growth from 1995-96 to 2005-06 has remained stagnant in all industries, except food, beverage and tobacco industry.

Spatial Inequality and Geographic Concentration of Manufacturing Industries in Pakistan¹

1. Introduction

Geographic concentration or agglomeration of industries is one of the most striking features of economic activity in developed and developing countries including Pakistan. A vast empirical literature from developed countries suggests that firms and workers are unevenly distributed across spatial units; they agglomerate in some regions more than others [e.g., Ellison and Glaser (1997), Maurel and Sedilot (1999), Alonso-Villar et al. (2004), Bertinelli and Decrop (2005)].

The contributions of Paul Krugman and others, in what has been called the New Economic Geography, provide insights into the factors influencing these location decisions [e.g., Krugman (1991), Kim (1995), Fujita et al. (1999), Fujita and Thisse (2002)]. They argue that increasing returns to scale explain why economic activities are geographically concentrated. The benefits of agglomeration are associated with localization and urbanization economies. Localization economies make it beneficial for firms in the same industry to locate close to each other, which provides intra-industry benefits accruing through knowledge-diffusion, buyer-supplier networks, subcontracting facilities and labour pooling. Urbanization economies explain why firms of different industries locate in close proximity to each other. Their geographic proximity provides rather diverse benefits associated with size of cities which offer across industry spill-over from complementary services, e.g., financial institutions, marketing and advertising agencies, and other cheap infrastructure, etc. While there are a number of theories that explain why geographical concentration of economic activity takes place², little is known about the factors that drive internal economic geography of developing countries and what course this inequality takes. Since Krugman's (1991b) seminal paper, significant progress has been made in theoretical models on the new economic geography, but empirical studies, especially from the developing countries are limited [Lall et al. (2004), Lall et al. (2005)].

As part of a broad empirical evaluation of the industry agglomeration in developing countries, this paper contributes by presenting evidence from the manufacturing sector of Pakistan. Economic revival in Pakistan largely hinges on the performance of its industry and its forward and backward linkages. In the past three years the country has seen a dramatic retardation of economic activity characterized in particular by a stagnating manufacturing sector. Given the fact that the potential of growth and development of a country is inextricably linked to the extent of investment, urbanization, and industrialization, the continued dismal performance in industrial growth in Pakistan does not augur well for future growth. Therefore it is imperative

¹ This paper draws heavily from chapters 4 and 5 of Burki et al. (2010).

² For early works, see Marshall (1890), Weber (1909), Hotelling (1929), Florence (1948), Hoover (1948), Fuchs (1962), Henderson (1974), among others. For more recent contributions, see among others, Krugman (1991a, 1991b), Krugman and Venables (1995), Kim (1995), Krugman and Livas (1996), Ellison and Glaeser (1997), Fujita et al. (1999), Puga (1999), Fujita and Thisse (2002), Hanson (2005).

to investigate the agglomeration economies, their causes and trends to identify the barriers to firm growth and to suggest remedial measures.

In this paper, we examine what factors cause agglomeration of manufacturing industries in Pakistan and what is the nature of scale economies.³ Our goal is to explore whether manufacturing industries are agglomerated, and if so, which ones. Moreover, we also present evidence on how geographic concentration emerges from the dynamic process over time. We take districts as spatial units, which represent the third-level of administrative jurisdiction after provinces and administrative divisions. Because the area boundary changes for the districts are quite common, the number of administrative districts has increased from 43 in 1951 to 106 in 1998. To avoid distortions and inconsistency overtime, we freeze the district boundaries at the time of 1981 population census when the number of administrative districts was 65. However, we select 56 districts for the analysis to allow consistency in all the data. Therefore, we identify spatial units in Balochistan by administrative divisions instead of districts. We adopt a consistent definition and select only 56 districts where four administrative divisions of Balochistan are also treated as districts while Islamabad and Rawalpindi are merged to form a single district.

The paper is organized as follows. Section 2 presents detailed evidence on spatial disparity in Pakistan and mapping of districts by employment share in manufacturing industries, road infrastructure and human capital infrastructure. Section 3 describes the Ellison-Glaser (EG) measures of geographic concentration that are sensitive to the number of districts and the number of plants in the sample. Then we go on to present the results on the basis of the agglomeration index followed by the results on the dynamic trends. Section 4 examines the factors that cause agglomeration of manufacturing industries in Pakistan, while Section 5 presents evidence on the nature of scale economies and patterns of industry agglomeration. Section 6 summarizes the main findings of this paper.

2. Evidence on Spatial Inequality in Pakistan

We begin by presenting evidence on spatial distribution of manufacturing employment in Pakistan followed by mapping measures for road infrastructure and human capital infrastructure across spatial units. This evidence is also used in the empirical models tested to examine the sources of agglomeration and the nature of scale economies.

2.1 Spatial distribution of manufacturing employment

We begin by presenting a map of spatial distribution of manufacturing activity by taking district level employment shares of all the industries from CMI 2005-06 data. Figure 1 displays spatial employment shares for 2005-06. The average employment share of districts is 0.018 with SD of 0.0319. Note that highly concentrated districts are mostly clustered around metropolitan cities

³ For a recent literature review of Pakistan's manufacturing sector, see Burki et al. (1997) and Burki and Khan (2004), among others.

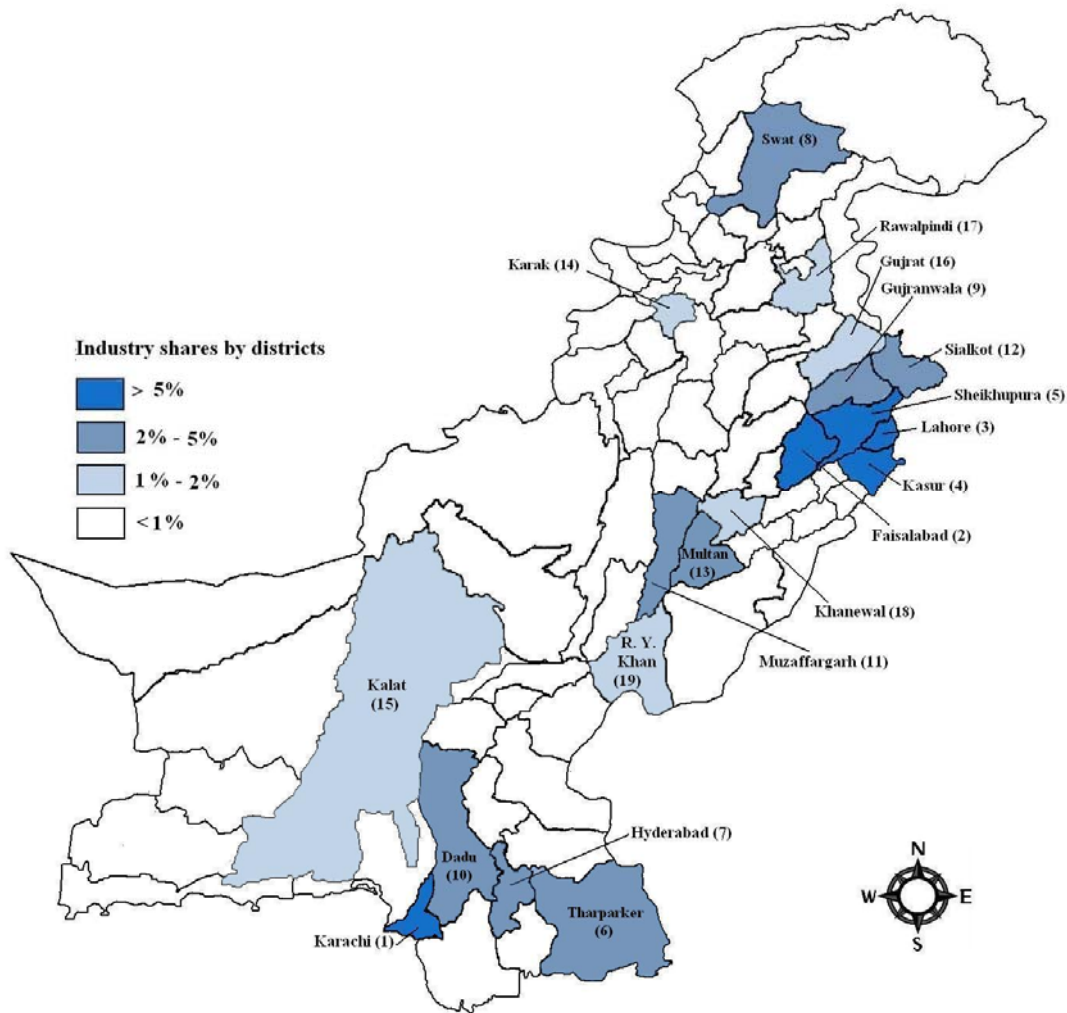


Fig 1: Distribution of manufacturing employment in Pakistan, 2005-06

of Karachi and Lahore. Even medium-concentrated districts are also clustered near two big cities. Only exceptions are Muzaffargarh and Swat districts. Muzaffargarh is located at the centre of the cotton growing belt in Punjab and due to natural advantage in cotton there is a large concentration of cotton ginning and textile manufacturing plants. Moreover, there is also a huge cluster of petroleum refining industry. Similarly, Swat is home to marble and mineral industries due to its natural advantage. In addition, there is a large concentration of pharmaceutical and plastic industries, which explain its large share.

2.2 Spatial disparity in road infrastructure

Market access is determined by the ease of connectivity with the market centers in spatial vicinity of the firm, which in turn depends on the availability of good road infrastructure, firm's

distance from the market, size of the market and the availability of quality transport networks. Absence of all or some of these factors limits the extent of the market for a firm because the firm would be unable to connect to a wider market area, i.e., other cities and districts, other provinces or the rest of the world. Therefore, spatial inequality in road infrastructure may constrain market efficiency and promote market failures by creating factor scarcities and distort factor prices, which in turn may prevent these spatial units to specialize in production by comparative advantage or dynamic comparative advantage. The recent literature also suggests that improvements in roads at the regional level can significantly contribute to the pursuit of socially inclusive growth [e.g., Khandker et al. (2009), Jacoby and Minten (2009)]

Even though there are popular concerns about the sufferings of regional and spatial units, there is no systematic documentation of what has happened to spatial development of road infrastructure in Pakistan over the last two decades. We document spatial concentration and the changes taking place in the road infrastructure from 1990-91 to 2005-06. Much of this material is summarized from Burki (2010) and Burki et al. (2010).

A consistent time-series data on road density at the district level is not available from any published source. Part of the problem is that more than one dozen government institutions at the federal, provincial, district and municipal government level, besides armed forces and state-owned corporations, are involved in construction and maintenance of road infrastructure in the country that makes the data collection exercise extremely cumbersome.

We obtain road density data of the districts of Punjab from the Punjab Highway Department for the period 1992 to 2006 while the data for KP was collected from the *Provincial Development Statistics*, which occasionally report district level data on road density. The road density is based on national highway roads, farm to market roads and district government roads. However, they do not cover the road network maintained by cantonment boards and defense housing authorities located in bigger cities. We assume that the omission of data on cantonment and DHA roads would not directly affect firm location in other districts.

The most striking pattern that emerges from the data is that while spatial inequality in road density has been large, it has widely fluctuated over time. Figure 2 displays mapping of the districts of Punjab on the basis of road density for 2005-06 as well as their relative rank from most dense to least dense. It appears from there that before recent floods, districts in southern Punjab were most deprived in road density while districts in northern Punjab were the ones with highest road density.

Figure 3 plots road density of 29 districts relative to the road density of Lahore in 1992-93 and 2005-06.⁴ We have chosen Lahore district because it had the highest road density in both 1992-93 and 2005-06 due to which the units of relative road density can be readily interpreted. Changes in road density across districts are illustrated by departures from the 45-degree line. The share of Lahore in road density relative to Lahore is 100%. Districts above the 45-degree line experienced improvement in relative shares over time while those below the line

⁴ We consider road density of each district divided by the road density of Lahore district.

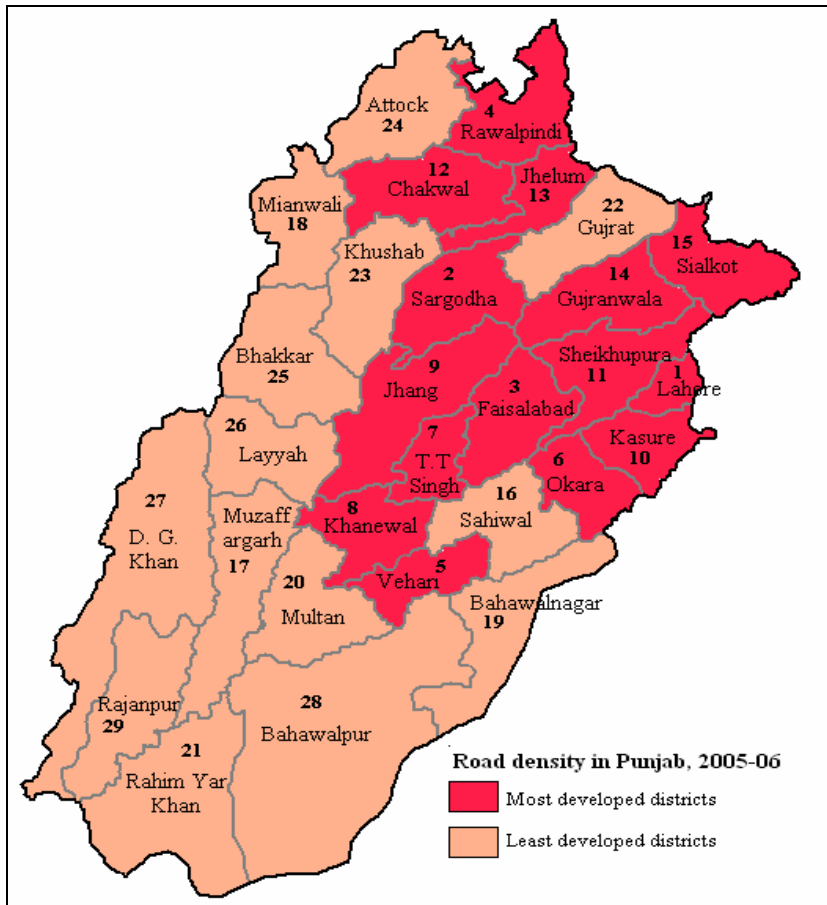


Fig 2: Spatial inequality in road density in Punjab, 2005-06

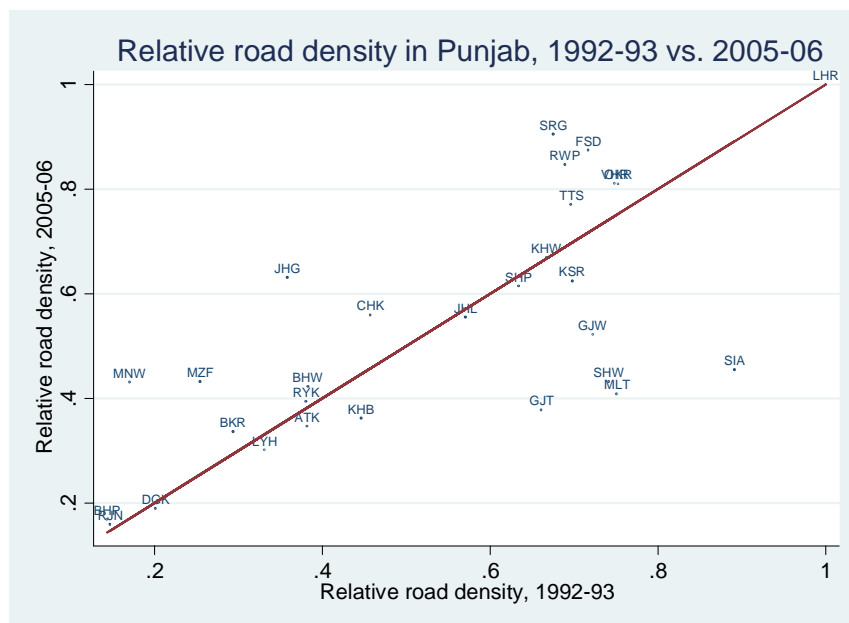


Fig 3: Relative road density of Punjab's districts with Lahore district

experienced decline. Some Southern districts, e.g., Rajanpur, D.G. Khan, Bhakkar, Layyah and Rahimyar Khan, with relative road density of less than 40% of Lahore district in 1992-93 experienced no change in relative road density by 2005-06. Sargodha, Faisalabad and Rawalpindi districts improved their relative shares from around 70% of Lahore to more than 80% mainly due to construction of 400 km motorway and other ancillary roads. Jhang, Chakwal, Muzaffargarh and Mianwali districts are other noticeable exceptions where relative road density has substantially increased during this 13-year period. Sialkot, Gujranwala, Sahiwal, Multan and Gujrat districts are some of the districts that have suffered significant decline in relative road density. Sialkot had the second highest road density after Lahore in 1992-93 with road density equal to 90% of Lahore. But, by 2005-06, Sialkot's road density had fallen to below 50% of Lahore.

Figure 4 displays spatial inequality in road density in KP before the August 2010 floods. Note that Peshawar, Abbottabad, Bannu and Kohat districts had highest road density. Figure 5 displays that Mardan, Karak, Mansehra and Dir had below 40% of the road density of Peshawar. On the other hand, Kohat and Bannu significantly increased their relative share in road density from 1992-93 to 2005-06, Abbottabad significantly lost its share from more than 100% of Peshawar in 1993-94 to around 70% level in 2005-06.

2.3 Spatial distribution of human capital infrastructure

Emphasis of recent studies in the industrial organization literature on human capital accumulation as key to firm survival makes sense [Mata and Portugal (2002), Karlsson (1997)]. However, social spending in Pakistan has always suffered due to competition with other development heads and stagnant revenue generation leaving little capacity in the hands of policy makers to meet increasing backlog of investment in education and health in the country. While spatial deprivation and rank ordering of districts in Pakistan has been attempted before [see, among others, Jamal (2003), Wasti and Siddiqui (2008)], but their adverse impact on industrial development has not been studied before.

We construct education and skill endowment variables as indicators of labor pooling in each district. We apply a multivariate statistical weighting approach known as the method of principal components [see, Greene (1997)] to select the principal components that account for the largest variance.⁵ The retained principal components are uncorrelated with each other, which is useful in principal component regression analysis. The selected principal components are used as mapping measures of spatial concentration and as education and skill endowment variables in the regression analysis discussed below.

⁵ Only small number of principal components from large number of variables is chosen that contribute highly in explaining the variance. Components associated with smallest eigenvalues are discarded because they are least informative. We adopt Kaiser-Gutman Rule whereby only principal components associated with eigenvalues greater than one are retained.

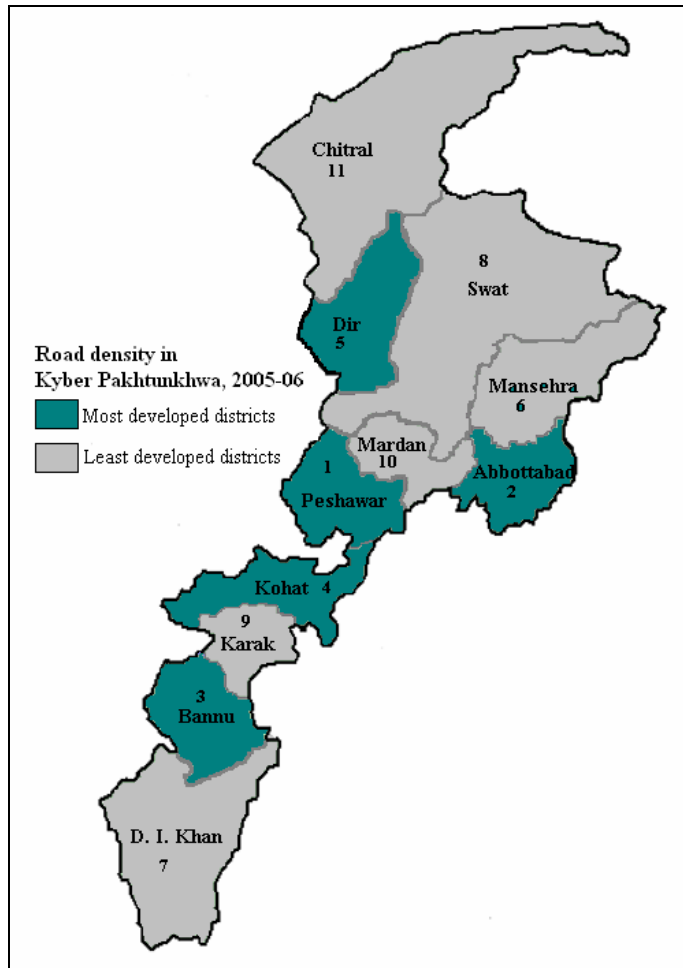


Fig 4: Spatial inequality in road density in KP

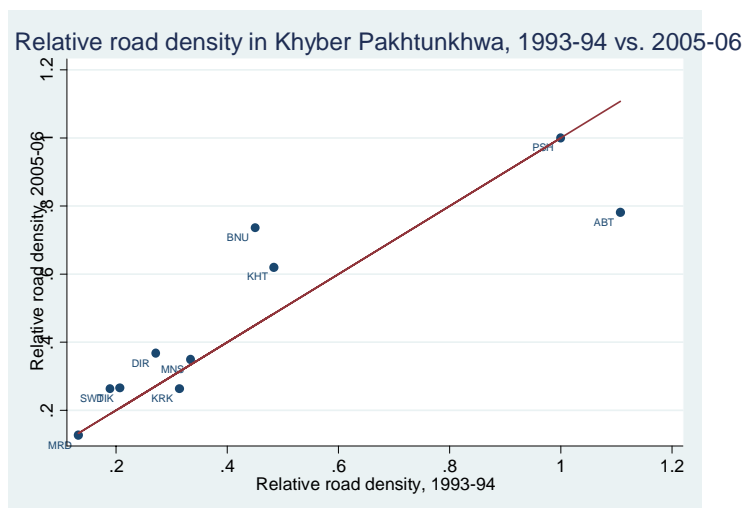


Fig 5: Relative road density of KP districts with Peshawar district

To construct district level formal and technical education variables we employ external information. We take corresponding data on 29 district level formal education and skill-specific indicators from Pakistan's *Labor Force Survey* 1995-96, 2000-01 and 2005-06. After varimax rotation, we retain four principal components using Kaiser eigenvalue criterion that account for 82.2% variation in the total variance. However, 73.5% of the variance was explained by first two factors. Therefore, we select factor 1 (F1) and factor 2 (F2) where factor 1 accounts for 51% variance and is characterized by high factor loadings on formal education (e.g., primary, secondary, intermediate, under-graduate and graduate degrees, and professional degrees), while factor 2 has high factor loadings on technical skills (e.g., vocational training, technician, garment making, leather works, polishing and soldering, interior decoration and carpentry, etc.).

Appendix Table A1 lists the most and least developed districts for selected years. Note that the most developed districts are also most urbanized, e.g., Karachi, Lahore, Faisalabad, etc. Figure 6 displays complete mapping of districts for 2005-06 indicating that most highly ranked districts are in the northeast (Lahore, Faisalabad, Gujranwala, and Rawalpindi) and Karachi. Districts located away from these urban demand centers (e.g., southern Punjab, interior of Sindh and remote districts of northern KP and Balochistan) have lowest education endowments.

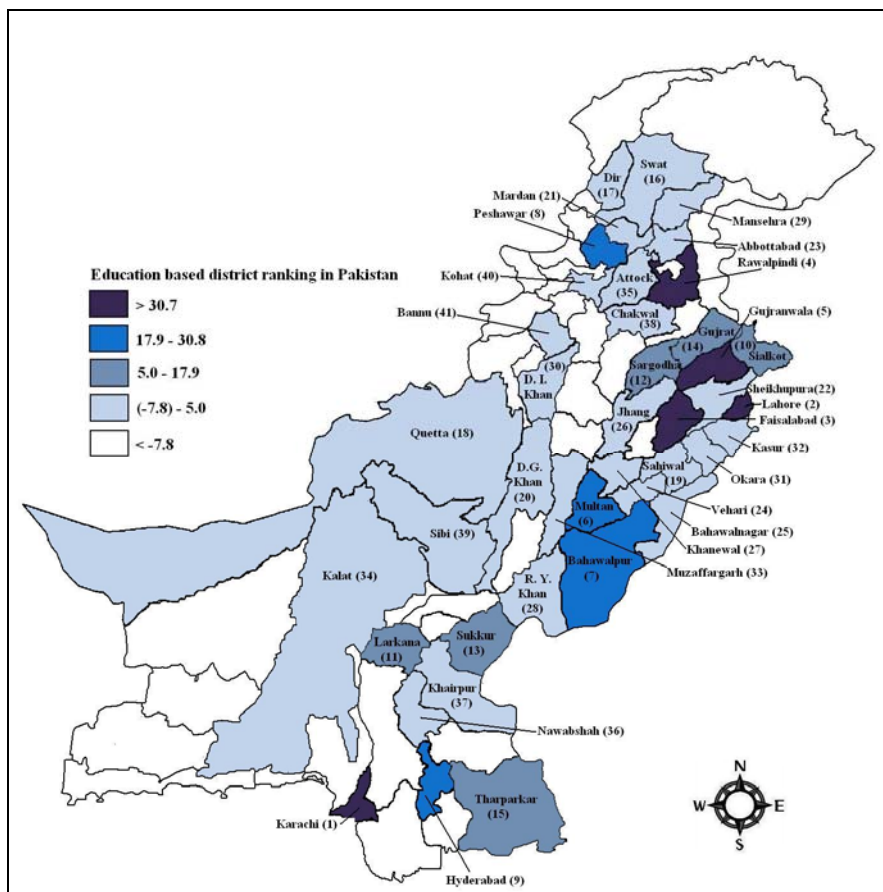


Fig 6: Education endowments, 2005-06

In sum, manufacturing firms in Pakistan are not uniformly distributed across space. Most firms are concentrated in districts where they have access to a large market. Higher road density and labor pooling in and around these cities/districts allows them reduction in transport costs and lower wages. However, the agglomeration process and trends overtime are not obvious, which require more rigorous and formal approach to which we turn to below.

3. Measuring Industry Agglomeration by Ellison and Glaeser index

It may be that the simple employment shares discussed above are not sufficient to completely capture the concentration of industry in each district. In what follows we present measures of geographic concentration that are sensitive to the number of districts and the number of plants in the sample. We then present distribution of industries that are most highly concentrated in comparison with low concentration industries and how concentration of industries has evolved overtime.

To measure the extent of geographic concentration we follow the method proposed by Ellison and Glaeser (1997). Their index is based on a rigorous statistical model that takes random distribution of plants across spatial units as a threshold to compare observed geographic distribution of plants. Ellison and Glaeser (1997) assume that plants make location decisions to gain from internal and external economies peculiar to a particular location. Because the industrial structure in Pakistan consists of many small, medium and large plants, proper weights are required to correct for the diverse sizes of plants and this is taken care of in the Ellison and Glaeser index. They present the following estimator to measure the agglomeration of industries

$$\gamma_j = \frac{\sum_{i=1}^M (s_{ij} - x_i)^2 - \left(1 - \sum_i x_i^2\right) H_j}{\left(1 - \sum_i x_i^2\right) (1 - H_j)}$$

where s_{ij} is the share of industry j 's employment located in district i ; x_i is the share of industry's overall manufacturing employment in district i ; $\sum_i (s_{ij} - x_i)^2$ is an index of raw geographic concentration given by the sum of squared deviations of employment shares of the industry j known as Gini-coefficient; $H_j = \sum_k Z_{kj}^2$ is a Herfindahl-style measure of the industry j 's plant level concentration of employment, where Z_{kj} is the k th plant's share in industry j 's employment.

In practice, the value of the EG index indicates the strength of agglomeration externalities in an industry. Usually a γ score of more than 0.05 indicates highly agglomerated industry; a score of between 0.05 and 0.02 suggests moderate agglomeration and a score of less than 0.02 shows randomly dispersed industry. The present study uses plant level data taken from the *Census of Manufacturing Industries* (CMI), 2005-06 provided by the Federal Bureau of Statistics, Government of Pakistan. The CMI provides data for 2-digit, 3-digit, 4-digit and 5-digit

classifications under the Pakistan Standard Industrial Classification (PSIC) according to geographic subdivision at the district, province and national levels.

3.1 Results on the agglomeration index

Next, we discuss the geographic concentration of 3-digit industries performed at the district level by the Ellison-Glaeser index. As suggested by Ellison and Glaeser (1997), we take values higher than 0.05 as high concentration, values in the range of 0.02 to 0.05 show intermediate concentration and values lower than 0.02 represent low concentration. Results of the EG index are also compared with raw geographic concentrations known as Gini index and Herfindahl index.

From the results summarized below we find that agglomeration of the 3-digit manufacturing industries is widespread in Pakistan while a small number of industries fall in the category of low concentration industries. For example, Table 1 shows that 35.3% of the industries are highly agglomerated, 38.2% are moderately concentrated and only 26.5% industries are not agglomerated. This is further corroborated in Figure 7 which plots the frequency distribution of the EG index across 3-digit industries, where each bar is for the number of industries for which the index is in that interval. The tallest bar is at the EG value of 0.03 and 0.04 and a large number of industries are in the range that corresponds to highly concentrated industries.⁶

Turning to specific 3-digit industries, we find that the most highly concentrated industry is other manufacturing (PSIC 394) with the EG index of 1.04 and raw concentration of 0.97 indicating that the industry is located in only one district. This is not a surprising result since this industry classification relates to ship-breaking industry, located only in Gadani (Kalat division) of Balochistan. Sports and athletic goods (PSIC 392) is the second most highly agglomerated industry with the EG index value of 0.90, where high value for the Gini index (0.842) indicates that the industry is located in few districts (5 districts) and the low value for Herfindahl index (0.077) shows that the employment is distributed across many plants. Located in Sialkot and its surrounding districts, the driving force for concentration of sports and athletic goods industry is natural advantage of specialized labor, inter-industry spillovers, local transfer of knowledge and market access in export markets due to close ties with international sports brands.

Other most concentrated industries represent sectors where it is critical for the industry to spread out to reach to the final consumers or the suppliers, e.g., furniture and fixtures, scientific instruments, pharmaceutical industry, wearing apparel, handicrafts and office supplies, printing and publishing, pottery and china products, paper and paper products, etc. On the other hand, the demand for least concentrated industries is diversified across many districts due to which they have substantial raw concentration, e.g., iron and steel (0.426), footwear (0.247) and tobacco (0.207), but their employment is distributed across few large plants, e.g., iron and steel (0.463), footwear (0.255), tobacco (0.232) and rubber (0.198).

⁶ The most highly concentrated industries indicate that they are more concentrated than would be expected if the industries were randomly located across spatial units.

Table 1. Agglomeration of 3-digit manufacturing industries in Pakistan, 2005-06

3-digit PSIC	Industry	Rank	Number of Plants	Number of Districts	Ellison-Glaeser index	Gini coefficient	Herfindahl index
394	Other manufacturing industries	1	30	1	1.043	0.965	0.035
392	Sports and athletics goods	2	51	5	0.901	0.842	0.078
332	Furniture and fixtures	3	34	8	0.231	0.268	0.076
385	Scientific instruments	4	95	7	0.193	0.218	0.052
350	Pharmaceutical industry	5	213	22	0.171	0.171	0.017
322	Wearing apparel	6	236	12	0.156	0.164	0.025
393	Handicrafts and office supplies	7	43	12	0.151	0.177	0.048
342	Printing and publishing	8	43	5	0.141	0.271	0.176
361	Pottery and china products, etc.	9	97	7	0.139	0.175	0.058
341	Paper and paper products	10	131	22	0.124	0.210	0.117
325	Ginning and bailing of fibers	11	540	27	0.099	0.096	0.004
383	Electrical machinery	12	240	19	0.068	0.125	0.072
372	Non-ferrous metals	13	41	7	0.047	0.108	0.073
380	Fabricated metal, cutlery and aluminum products	14	75	15	0.044	0.092	0.058
382	Non-electrical machinery	15	206	25	0.042	0.075	0.041
354	Petroleum refining, petroleum products and coal	16	30	12	0.041	0.164	0.142
369	Other non-metallic mineral products	17	311	32	0.039	0.059	0.026
352	Other chemical products	18	150	24	0.038	0.081	0.051
356	Plastic products	19	141	20	0.035	0.048	0.017
381	Copper and brass industrial products	20	111	17	0.033	0.067	0.041
321	Made-up textiles, knitting mills, carpets and rugs	21	261	23	0.030	0.049	0.024
323	Leather and leather products	22	227	22	0.029	0.052	0.028
311	Dairy products and processed food	23	1190	55	0.026	0.033	0.010
384	Transport equipment	24	186	15	0.021	0.057	0.041
362	Glass and glass products	25	34	13	0.020	0.121	0.113
320	Spinning and weaving of cotton & wool	26	1081	44	0.019	0.023	0.006
324	Footwear manufacturing	27	35	10	0.015	0.247	0.255
331	Wood and cork products	28	15	10	0.014	0.139	0.138
351	Industrial chemicals	29	111	21	0.003	0.029	0.028
371	Iron and steel industries	30	198	18	-0.005	0.426	0.463
314	Tobacco industry	31	13	6	-0.012	0.207	0.232
313	Beverage industry	32	36	16	-0.020	0.080	0.105
355	Rubber products	33	30	10	-0.030	0.161	0.198
312	Animal feed & ice factories	34	65	20	-0.035	0.067	0.104

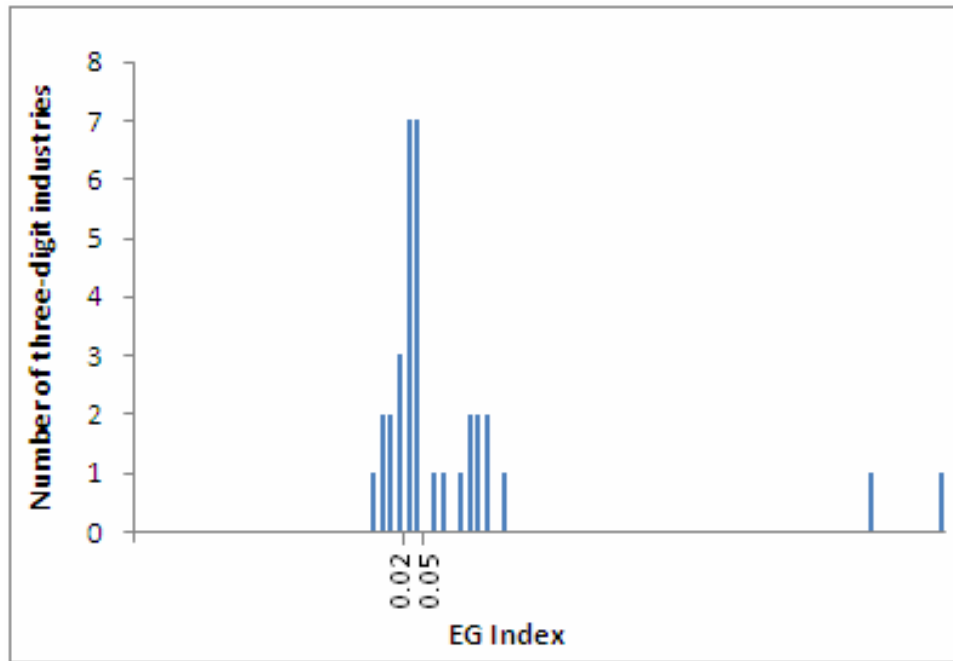


Fig 7: Distribution of 3-digit Ellison-Glaeser index

3.2 Agglomeration in the period 1995-96 to 2005-06

Our earlier discussion has suggested that agglomeration of manufacturing industries is widespread. To analyze this further we ask how industrial concentration has evolved overtime. More specifically, we consider whether the most agglomerated industries in 2005-06 were also agglomerated in previous years. We perform this analysis on the data from Punjab province where most of the manufacturing industry in located.⁷ To make this dynamic analysis possible we select three data points over the 10-year period from 1995-96 to 2005-06. This supplementary data was provided by the Punjab Bureau of Statistics.⁸ As before, we use the 1981 district boundaries to identify our spatial units and focus on 3-digit industries to measure industrial concentration.

Table 2 shows that the rankings resulting from Punjab’s data do not drastically differ from the national rankings, except in industry 393 Handicrafts and office supplies and 354 Petroleum refining, petroleum products and coal, which have lower agglomeration index value in Punjab.

⁷ The plant-level CMI data of 1995-96 and 2000-01 for Sindh, KP and Balochistan provinces was not available from any source.

⁸ The CMI 1996-97 and 2000-01 were conducted on the basis of same PSIC, but the classification was changed for CMI 2005-06. To produce consistent and comparable estimates, we regrouped CMI 2005-06 data according to CMI 1995-96 classification.

Table 2. Geographic concentration of 3-digit industries in Punjab, 1995-96 – 2005-06

3-digit				
PSIC	Industry	1995-96	2000-01	2005-06
Fifteen most concentrated industries in 1995-96				
392	Sports and athletics goods	1.047087	1.068879	0.913328
385	Scientific instruments	0.8610103	0.589197	0.280299
362	Glass and glass products	0.5038906	0.30704	0.191059
361	Pottery and china products, etc.	0.415917	0.327417	0.164882
371	Iron and steel industries	0.4151095	0.391224	0.276167
322	Wearing apparel	0.3793302	0.208751	0.077318
350	Pharmaceutical industry	0.3547366	0.351063	0.339121
384	Transport equipment	0.2308846	0.364937	0.145223
355	Rubber products	0.2160143	0.089129	-0.00429
380	Fabricated metal, cutlery, aluminum and products	0.2126588	0.108821	0.173101
321	Made-up textiles, knitting mills, carpets and rugs	0.1896874	0.214717	0.037884
381	Copper and brass industrial products	0.169055	0.184687	0.249403
331	Wood and cork products	0.1624373	0.164167	0.045256
383	Electrical machinery	0.1582761	0.193224	0.20588
332	Furniture and fixtures	0.1406263	0.229663	-0.15062
Fifteen least concentrated industries in 1995-96				
324	Footwear manufacturing	-0.1127579	-0.17016	-0.06546
312	Animal feed & ice factories	-0.106044	3.36E-05	-0.07352
372	Non-ferrous metals	-0.0379526	0.216773	-0.00275
354	Petroleum refining, petroleum products and coal	-0.0311225	0.054006	0.192322
313	Beverage industry	-0.0057071	-0.00215	-0.02943
314	Tobacco industry	0.0068856	0.4708	0.064257
351	Industrial chemicals	0.0151802	0.010317	-0.00265
311	Dairy products and processed food	0.0179265	0.04199	0.021544
356	Plastic products	0.0179869	0.042642	0.055971
382	Non-electrical machinery	0.0185093	0.041723	0.108921
393	Handicrafts and office supplies	0.0209904	0.044206	0.077944
369	Other non-metallic mineral products	0.0213563	0.057317	0.098664
352	Other chemical products	0.0277417	0.009313	-0.02643
320	Spinning and weaving of cotton & wool	0.0340134	0.049388	0.031043
323	Leather and leather products	0.0953683	0.050376	0.0288

These industries are distributed across few districts, but within them they are mostly located in Sindh province.

The dynamic process in 3-digit industries in Punjab demonstrates declining agglomeration levels in the most concentrated industries in 1995-96. Except for industry 381 and 383, the industrial concentration generally falls between 1995-96 and 2005-06 when more than 50% of the most concentrated industries experienced a sharp decline in the concentration index while five of these industries experienced increase in the first five years but a decline in the next five

years. Entry of more firms in an industry explains steady decline in concentration index as we note that the industry specific Gini index of these industries gradually falls during the study period. Of the fifteen least concentrated industries in 1995-96, in general, the moderate to high concentration industries again show a consistent decline in concentrations. Most other industries experience a moderate increase in concentration index, but the magnitude of increase in least concentrated industries is much smaller than the decline in most concentrated industries. Whereas dramatic movements in the concentration index are rare, there are only two noticeable increases in the concentration index that are worth mentioning. We notice a dramatic increase in the concentration index of tobacco and non-ferrous metal industries between 1995-96 and 2000-01, which is the result of exit of few plants that makes a dramatic impact on their Gini index as well as on the absolute value of the EG index.

These results are further corroborated by the declining mean values of the geographic concentration levels. In Table 3 we report mean concentration level of 3-digit industries in Punjab. The table shows that the mean value of EG index remains roughly constant from 1995-96 to 2000-01, but the index drastically falls by about 33% in the next five years. The decline in industrial concentration is mainly explained by the decline in raw concentration measure (Gini index), and the change in average plant size measured by the Herfindahl index is less important in explaining the decline in industrial concentration. The correlation coefficient of the EG index further corroborates these results, which indicates a declining trend in the dynamic industrial concentration levels in Punjab across industries. It reveals that the correlation coefficient of EG index between 1995-96 and 2000-01 was 0.86, which fell to 0.79 between 2000-01 and 2005-06. It suggests that the agglomeration of industries followed a declining trend overtime.

Table 3. Mean values of industrial concentration measures overtime

Concentration measure	Concentration in Punjab		
	1995-96	2000-01	2005-06
Ellison-Glaeser index	0.1833	0.1789	0.1189
Gini index	0.2794	0.2656	0.2186
Herfindahl index	0.1465	0.1406	0.1374
Survey Year	Correlation of the EG index		
2000-01	0.8626	1	~
2005-06	0.7820	0.7873	~

4. What Factors Cause Agglomeration of Manufacturing Industries?

The results of the previous section have suggested that there is evidence of widespread, but declining, agglomeration of manufacturing industries. In what follows, we would like to know what forces drive agglomeration of manufacturing industries in Pakistan.

From the literature we come to know that there are three types of transport costs that play an important role in “moving goods”, “moving people” and “moving ideas” or knowledge

spillovers [Marshall (1920)]. Firstly, the firms would like to locate near the demand centers and input suppliers to save shipping cost. Secondly, agglomeration of industries also offers the advantage of scale economies on account of labor market pooling due to inherent benefits of pooling, which also allow labor to optimally allocate time to maximizes productivity. Finally, agglomeration allows firms to gain from the free flow of ideas or technology spillovers. In this regard, Ellison and Glaeser (1999) describe the finance industry in urban centers where density speeds up the flow of new ideas.⁹ However, these models cannot be easily translated into variables that could be used in the empirical models trying to find out the determinants of agglomeration.

Natural advantage is another key factor that may be motivating firms to locate at particular regions where the considerations of Marshall's agglomeration forces may be weak or non-existent. To illustrate, some regions offer natural environments that are suited to certain industries due to significant natural cost advantages. For example, milk processing industry has a natural advantage to locate in regions where milk supply is surplus and farm gate price of milk is relatively low. Similarly, petrochemical industry may save shipping costs by locating near a port. One may like to offer some anecdotal examples of industries that have otherwise agglomerated in a particular region to reap benefits of natural advantage or any of these transport costs. However, what is desirable from an empirical standpoint is to examine the relative significance of these factors in explaining the causes of industry agglomeration.

We analyze the factors that help explain the causes of agglomeration by using an empirical specification that allows us to relate the Ellison-Glaeser index of industry concentration to industry characteristics and agglomeration forces put-forward by the theory. The empirical specification used is

$$\gamma = \alpha + \beta X + \delta + \lambda + \varepsilon$$

where γ depicts the EG industry agglomeration index, X is a vector of industry characteristics that explain agglomeration, δ and λ are year and industry fixed-effects where the industry effects refer to 2-digit industries in each 3-digit industry, and ε is a random error term.

As before, we use pooled data of 3-digit EG index of Punjab based on CMI 1995-96, 2000-01 and 2005-06. The New Economic Geography literature views locating near demand centers and major input suppliers as a major driver behind industry agglomeration aimed at saving transportation cost. Therefore, we take district population as an indicator of market access and district level data on road density to proxy for market access and transportation cost. However, we expect a strong correlation between the two. An industry would decide to locate in areas where the supply of educated and skilled labor force is high so that they find industry specific skilled labor force easily at market wage rates in its area of production activity.

⁹ Similarly, Arzaghi and Henderson (2008) draw our attention to the benefits of networking to marketing firms in Manhattan.

Table 4 shows descriptive statistics on agglomeration and sources of agglomeration variables. The district level industry agglomeration index (the EG index) is worked out by taking the product of each industry's agglomeration index, γ , and the industry's share of manufacturing in each district, i.e., $\gamma \times s$. Road density and education and skill endowment variables are described above. The principal component formal education and technical education indexes have high standard deviations, which indicate spatial inequality across district.

Table 4. Descriptive statistics for agglomeration regressions

Variable	Mean	Std. Dev.	Min	Max
Ellison and Glaeser index (γ)	0.0179	0.076	-0.098	1.068
Road density	0.3420	0.146	0.044	0.697
Population (millions)	3.5329	1.632	0.852	7.419
Principal component formal education index (F1)	0.5459	1.284	-0.725	4.575
Principal component technical education index (F2)	0.4313	2.129	-0.815	10.381
Year 2000-01 (yes=1, no=0)	0.3125	0.463	0	1
Year 2005-06 (yes=1, no=0)	0.3637	0.481	0	1
Industry 31 (yes=1, no=0)	0.1714	0.377	0	1
Industry 32 (yes=1, no=0)	0.2513	0.434	0	1
Industry 33 (yes=1, no=0)	0.0389	0.193	0	1
Industry 34 (yes=1, no=0)	0.0422	0.201	0	1
Industry 35 (yes=1, no=0)	0.1690	0.375	0	1
Industry 36 (yes=1, no=0)	0.0745	0.262	0	1
Industry 37 (yes=1, no=0)	0.0355	0.185	0	1
Industry 38 (yes=1, no=0)	0.1868	0.390	0	1
Industry 39 (yes=1, no=0)	0.0300	0.170	0	1
N	899	--	--	--

Note: Industry 31=food beverage and tobacco; 32= textile and leather industry; 33= wood and wood products; 34= paper and paper products; 35= chemicals, rubber and plastic; 36= mineral products; 37=basic metal; 38 = metal products; 39= other industry & handicrafts.

Table 5 shows Pearson's correlation matrix between the agglomeration index and the sources of agglomeration. Note that simple correlation between the EG index and its correlates is always positive. However, caution is warranted since the correlation coefficient between road density, population and F1 is also very high, which is expected to create robustness issues in the full regression models.

Table 5. Pearson's correlation

	EG (γ)	Road density	Population	F1	F2
Ellison and Glaeser index (γ)	1				
Road density	0.0827	1			
Population (millions)	0.1303	0.6169	1		
Principal component formal education index (F1)	0.0861	0.6757	0.849	1	
Principal component technical education index (F2)	0.0716	-0.0537	0.2404	-0.0001	1

We estimate the model explaining the causes of industry agglomeration by using pooled data consisting of 899 observations (Table 6). We present four models for different industry characteristics: in column (1) we present the full model, in column (2) we include only population variable, in column (3) we include only road density variable, and in column (4) we include the two principal component variables, F1 and F2. All models include a complete set of year and 2-digit industry fixed effects. Overall, the empirical estimates are quite robust. Total variation explained by these models is about 9%.

Table 6. OLS specifications for agglomeration regressions

Variable	Full model	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
Road density	0.0365* (1.79)	--	0.0549** (2.14)	--
Population (millions)	0.0046 (0.93)	0.0047** (2.06)	--	--
Principal component formal education index (F1)	-0.0029 (-0.44)	--	--	0.0044 (1.40)
Principal component technical education index (F2)	0.00050 (0.37)	--	--	0.0014* (1.70)
Year 2000-01 (yes=1, no=0)	-0.0008 (-0.31)	-0.0016 (-0.62)	-0.0036 (-1.33)	-0.0001 (-0.05)
Year 2005-06 (yes=1, no=0)	-0.0129** (-2.75)	-0.0117** (-2.86)	-0.0174** (-2.81)	-0.0107** (-2.07)
Industry 32 (yes=1, no=0)	0.0038 (1.04)	0.00380 (1.08)	0.0044 (1.24)	0.0043 (1.21)
Industry 33 (yes=1, no=0)	0.0086* (1.82)	0.0090* (1.83)	0.0101** (2.18)	0.0100** (2.13)
Industry 34 (yes=1, no=0)	0.0193 (1.32)	0.0194 (1.34)	0.02121 (1.39)	0.0204 (1.40)
Industry 35 (yes=1, no=0)	0.0062 (1.35)	0.0061 (1.40)	0.0078 (1.61)	0.0073 (1.60)
Industry 36 (yes=1, no=0)	0.0284** (2.35)	0.0273** (2.26)	0.0294** (2.51)	0.0273** (2.25)
Industry 37 (yes=1, no=0)	0.0334 (1.60)	0.0328 (1.60)	0.0360* (1.68)	0.0343* (1.67)
Industry 38 (yes=1, no=0)	0.0194** (2.15)	0.0196** (2.12)	0.0210** (2.17)	0.0210** (2.11)
Industry 39 (yes=1, no=0)	0.1129 (1.26)	0.1134 (1.26)	0.1159 (1.29)	0.1158 (1.26)
Constant	-0.0179 (-1.19)	-0.0075 (-0.94)	-0.0080 (-1.06)	0.0045 (1.29)
R ²	0.093	0.0904	0.0896	0.0877
N	899	899	899	899

Note: All the models are estimated by the OLS. Numbers in parenthesis are *t*-values obtained from robust standard errors corrected for clustering at the district level. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively. Industry 31=food beverage and tobacco; 32= textile and leather industry; 33= wood and wood products; 34= paper and paper products; 35= chemicals, rubber and plastic; 36= mineral products; 37=basic metal; 38 = metal products; 39= other industry & handicrafts.

The results tell quite a consistent story. Note a declining dynamic concentration levels in 3-digit industries in Punjab, which are in line with the results reported in Tables 2 and 3. For example, in the first five year period (i.e., 1995-96 to 2000-01), the agglomeration of manufacturing industries in Punjab shows no change as implied by a statistically insignificant coefficient on dummy variable for 2000-01. However, in the second five year period, industry agglomeration significantly declined as revealed by a statistically significant coefficient on year 2005-06. This is consistent with a 33% fall in the EG index in the second five year period (see Table 3). Due to high correlations between some explanatory variables, the coefficients of population, F1 and F2 are all statistically equal to zero in the full model, while the coefficient on road density is marginally significant. Our estimates in model 2 imply that the size of district population increases agglomeration; in model 3 roads density variable is positive and statistically significant indicating that increased road density promotes more agglomeration. The estimates using F1 and F2 in model 4 are also positive, but statistically significant only for the technical education index at the 10% level, this suggesting that formal education is not a constraint for industry concentration in Punjab, decrease in technical education index does lead to a significant decrease in concentration levels.

5. Localization vs. Urbanization Externalities

Localization and urbanization externalities are related to local scale externalities that arise from local information spillovers linked with input and output markets and local technological developments. If firms in a district learn from local firms in their own industry, this is called localization externalities; if firms learn from all firms in the district, it is termed as urbanization economies [Henderson et al. (2001)]. While several urban areas dominate in Pakistan, the relative strength of “localization economies” versus “urbanization economies” is not known. In the present age, there is a greater role for information spillovers leading to greater cross-industry learning economies. These relationships and bonding may explain presence of mega industries around large cities. Moreover, if these externalities are dynamic, then one may expect that presence of industry at a location in the past would affect productivity today. In these cases, industry would be unwilling to move to far off regions where there is no “built-up stock of knowledge” that makes diversification of industries more and more difficult.

We follow Henderson et al. (2001) to measure the scale of local externalities by a diversity index written as

$$g_i^s = \sum_{n=1}^N [(E_{in}/E_i) - (E_n/E)]^2$$

where i and n index district and industry, respectively, g_i^s is the index of localization and urbanization economies in i th district, E_n is for employment in industry n , E is for total national manufacturing employment, and E_i is for employment in i th district and E_{in} is for employment in i th district in n th industry. A lower value of g_i^s (minimum value is zero) implies that the city is non-specialized (high diversity) while a higher value (approaching two) indicates

complete specialization. In other words, as g_i^s goes up specialization increases and the diversity tends to fall. Henderson et al. (2001) find a negative relationship between g_i^s and productivity in Korea.

As before, we use CMI data of Punjab for 1995-96, 2000-01 and 2005-06 and calculate localization and urbanization index of 3-digit industries for 29 districts of Punjab. In Table 7 we show localization versus urbanization externalities across districts. Our results suggest that localization economies are much greater in magnitude than urbanization economies since the diversity index has higher values for most of the districts with a mean of 0.253 in 1995-96. It shows that there is a greater role for within-industry externalities and that technological spillovers and inter-industry learning has little role in the manufacturing industries in Punjab.

Table 7. Localization versus urbanization externalities in Punjab, 1995-95 to 2005-06

	g_i^s , 1995-96	g_i^s , 2000-01	g_i^s , 2005-06
Sheikhupura	0.0169	0.0539	0.0312
Khushab	0.0407	0.0385	0.0623
Khanewal	0.0493	0.0448	0.1891
Multan	0.0529	0.0452	0.0811
D.G. Khan	0.0625	0.2092	0.0843
Attock	0.0806	0.0735	0.0937
Faisalabad	0.0892	0.076	0.0489
Jhang	0.0919	0.1259	0.2931
T.T. Singh	0.1013	0.2746	0.2907
Gujranwala	0.1196	0.1620	0.0373
Chakwal	0.1381	0.1149	0.1386
Bahawalpur	0.1504	0.2804	0.4388
Kasur	0.1612	0.1286	0.1142
Muzaffargarh	0.1718	0.1975	0.0427
Lahore	0.1724	0.2279	0.1372
Sargodha	0.1967	0.1762	0.2775
Rawalpindi	0.2057	0.0249	0.0280
R.Y. Khan	0.2256	0.2566	0.4002
Sialkot	0.2574	0.5092	0.3548
Gujrat	0.2963	0.3342	0.2082
Mianwali	0.3126	0.4481	0.4431
Vehari	0.3465	0.1553	0.1695
Jhelum	0.3471	0.2824	0.1634
Bahawalnagar	0.3955	0.4615	0.2429
Sahiwal	0.4233	0.0532	0.2364
Okara	0.4896	0.4567	0.2252
Layyah	0.6518	0.7641	0.7094
Bhakkar	0.7365	0.1244	0.1238
Rajanpur	0.9589	0.9778	0.3583
Mean	0.253	0.244	0.208

The on-going dynamic process further reveals that the raw diversity index, g_i^s , fell overtime indicating that the forces of agglomeration were encouraging the diversity of local industries at an increasing rate. For example, while the index fell moderately (3.6%) in the first five years, it drastically fell (15%) in the second five year period indicating that the forces of diversity were increasing that may have led to increased productivity. We regress the specialization index, $g_i^s(t)$, on district population along with controls for survey years. By pooling data of 3-digit industries in three surveys, our results reported below (standard errors in parentheses) clearly indicate that district population is negatively correlated with the index of specialization.

$$g_i^s(t) = 1.09 - 0.061 \ln(Dist_{pop}) + 0.004 Year_{00-01} - 0.021 Year_{05-06}$$

$$(0.128) \quad (0.0086) \quad (0.0113) \quad (0.0109)$$

N = 901; $R^2 = 0.062$.

The time dummy for the second five year period is statistically significant, which confirms the results that the index declined in 2005-06, compared with 1995-96.

6. Nature of Scale Economies and Patterns of Industry Agglomeration

A natural extension to the analysis of agglomeration of industries is the question of the nature of scale economies and the pattern of agglomeration in different manufacturing industries. To understand this pattern, we examine the nature of local scale externalities (localization versus urbanization) in major 2-digit industries by taking CMI data of Punjab from the 1995-96, 2000-01 and 2005-06 census years.

We estimate a value-added production function relating industry factors of production by controlling for variables that capture local industry spillovers. To investigate these relationships, we follow Henderson et al. (2001) to specify the following value-added production function

$$y_{in} = A(W_{in})f(k_{in})$$

where y_{in} is the value added output per production worker in i th district in n th industry; k_{in} is capital per worker; $f(\cdot)$ represents production technology in n th industry; W_{in} represents a vector of shift factors including variables for externalities associated with localization and urbanization economies (spillover effects) and employment in each district in respective industries. The basic econometric model is given by

$$\ln(V_{in}(t)) = \alpha_j + \alpha_j \ln(k_{in}(t)) + \sigma_j \ln(E_{in}(t)) + \kappa_j g_i^s(t) + \rho_n + \mu_{ij} + \delta(t) + u_{in}(t)$$

where $V_{in}(t)$ is value-added output per production worker in i th district in n th 3-digit industry in each 2-digit industry, $j=1, \dots, 7$; $k_{in}(t)$ is capital per worker; $E_{in}(t)$ is total employment in

own industry to measure localization economies; $g_i^s(t)$ measures urbanization economies discussed above; ρ_n is a control variable sub-industry fixed effects given intercept term α_j ; μ_{ij} controls for district fixed-effects; and $\delta(t)$ is a dummy variable to capture time-effects. We assume that the technology is same across all 2-digit industries. We use three-year pooled data of CMI 1995-96, 2000-01 and 2005-06 for the Punjab province, consisting of 901 observations from 29 districts with an average of about 10 industries per district in each year of the census.

We estimate the model by the OLS and present basic results for seven 2-digit industries in Table 8. Two industries, paper products and printing (PSIC 34) and basic metal industry (PSIC 37), had less than 32 observations, therefore, to increase the degrees of freedom, we merge PSIC 34 (paper and paper products) with PSIC 39 (other industry and handicrafts), and PSIC 37 (basic metal industry) with PSIC 38 (metal products industry). We run OLS regressions for each 2-digit industry by including in all models a complete set of sub-industry (3-digit industries), year and district fixed effects. We also run the model and report the results of the regression for all industries, where the coefficients of all n industries are restricted to be the same which represent the average effects across all industries. The explained variation in each OLS model is more than 70%.

Table 8. Scale externalities and productivity in manufacturing industries in Punjab, 1995-96 – 2005-06

Variable	Industry (31)	Industry (32)	Industry (33)	Industry (35)	Industry (36)	Industry (37& 38)	Industry (34&39)	All industries
lnK	0.406*** (4.52)	0.450*** (7.83)	0.301*** (3.95)	0.417*** (5.77)	0.349** (2.54)	0.437*** (8.05)	0.357*** (3.80)	0.432*** (16.55)
Localization (lnE)	0.338*** (3.76)	0.424*** (7.30)	0.409*** (3.58)	0.385*** (5.06)	0.288** (2.12)	0.365*** (6.29)	0.435*** (3.89)	0.407*** (15.30)
Urbanization, g_i^s	1.284 (1.42)	0.164 (0.22)	-1.121 (-0.82)	-0.579 (-0.38)	2.570 (0.88)	1.838* (1.98)	1.167 (0.61)	0.628 (1.45)
Year 2000-01	0.346* (1.84)	0.126 (0.80)	-0.165 (-0.94)	0.145 (0.66)	-0.100 (-0.34)	0.135 (0.86)	-0.003 (-0.10)	0.161** (2.04)
Year 2000-01	0.434** (2.16)	-0.162 (-1.04)	-0.473 (-1.65)	-0.386* (-1.73)	0.162 (0.51)	0.253 (1.58)	0.013 (0.05)	-0.039 (-0.48)
R ²	0.784	0.826	0.822	0.772	0.798	0.819	0.861	0.773
N	154	226	35	152	67	200	67	901

Note: Industry 31=food beverage and tobacco; 32= textile and leather industry; 33= wood and wood products; 35= chemicals, rubber and plastic; 36= mineral products; 37 & 38 = basic metal and metal products; 34 & 39= paper and paper products and other industry & handicrafts. All models include a constant term and a complete set of sub-industry and district fixed effects. ***, ** and * indicates statistically significant at the 1%, 5% and 10% levels, respectively.

The impact of externalities due to localization economies (ln E) is always positive and always statistically significant, including the coefficient for all industries, at least at the 5% level. Since in our model we have logs on both sides, the coefficient on (ln E) indicates that a 1% increase in own-industry employment leads to a σ -% increase in value added output per worker. On the

basis of the coefficient of all industries, our results predict that, holding all else as constant, a 10% increase in own industry employment increases the value-added output per worker by an estimated about 4%. At sample means, our estimates imply that local scale economies are present in all industries, but, localization economies are highest in industry 34 & 39 (paper and paper products and other industry & handicrafts), followed by industry 32 (textile and leather), and then industry 33 (wood and wood products).¹⁰ In general, industries where local scale economies are present are expected to be most agglomerated and concentrated across districts. Likewise, these industries are also most highly agglomerated and concentrated industries in Pakistan across districts (e.g., see industry 341, 342; 392, 393; 322, 325, 321, 323; and 332 in Table 1). Localization economies are lowest in industry 36 (mineral products), which is relatively least agglomerated industry (see industry 362 in Table 1). Therefore, a key feature of the industrial geography of Pakistan is that industries that offer highest local scale economies are also the most agglomerated industries across districts. Thus, higher magnitudes of localization economies correspond well with inter-district agglomeration and concentration of industry, which is in line with expectations.

Based on the prevailing evidence in other countries, we expect that urbanization or Jacob economies are likely to be highest in industries that are regarded as high-tech and vice-versa. But, the urbanization economies are positive and statistically significant in only one industry, i.e., basic metal and metal products industry (37 & 38). An increase in urbanization index by one standard deviation (0.139) increases productivity in basic metal and metal products industry by 29%. These results seem to suggest that even though there is a greater role for inter-industry learning from technological spillovers, basic metal and metal products industry would greatly benefit from increase in the urbanization index.

The average share of capital in value added is 0.432 in all industries, which is not far from the average share of 0.37 in South Korea [Henderson et al. (2001)]. The highest share of capital is in the textile and leather industry (0.45), followed by basic metal and metal products (0.44) and petrochemical industry (0.42). The lowest share of capital is in wood and wood products (0.30) followed by mineral products (0.35).

The two time dummies reflect productivity growth, which averages at only 1.6% per annum from 1995-96 to 2000-01 for all industry and remains stagnant in the later period. This pattern hides high productivity industry behind the stagnant industries. Looking at the individual industry effects, we note that average productivity growth in food, beverage and tobacco industry over the 10-year period was 4.4% per annum. Unfortunately, this is the only 2-digit industry where the average productivity consistently increased over the study period. However, productivity growth was stagnant in all other industries (see Henderson et al. (2001)).

¹⁰ It is possible that some of the industries may be under-reporting contractual labor to avoid negative fall-outs of the labor laws, which may be leading to over-stating of the magnitudes of localization economies.

6. Concluding Remarks

This paper examines the nature of spatial inequality and scope of geographic concentration of manufacturing industries in Pakistan. We use plant level data from CMI supplemented by external information to analyze geographic concentration of manufacturing plants. While there is little doubt that increasing returns to scale associated with agglomeration externalities do exist at a wider scale in Pakistan, it is much more difficult to identify factors that cause industry agglomeration. This paper also explored how geographic concentration of manufacturing industries emerges from the dynamic process and what is the nature of agglomeration economies. Moreover, we note that scale externalities generate localization and urbanization economies that arise from local information spillovers with input and output markets and local technological developments in spatial units. The conclusions and policy implications of our investigation are summarized below.

Firstly, we find that the economic geography does matter in Pakistan. With few exceptions, the most highly concentrated districts in large-scale manufacturing employment are clustered around metropolitan cities of Karachi and Lahore, and their surrounding districts. However, despite many empirical studies on convergence and divergence in many countries, we cannot offer a concrete guide on how the policy makers in Pakistan can move to achieve the objective of faster convergence. Nor do the existing literature offer any benchmarks of remoteness below which the chances of viable development activity disappear.

Secondly, we find that agglomeration of 3-digit manufacturing industries is widespread where only a small proportion of industries fall in the category of low concentration industries. Our results suggest that 35% of the industries are highly agglomerated (EG index > 0.05) and 38% of the industries are moderately concentrated (EG index between 0.02 and 0.05), while only about 27% of the industries are not agglomerated. Ship-breaking and sports and athletic goods industries are two most highly agglomerated industries where the driving force is their natural advantage of specialized labor, inter-industry spillovers, local transfer of knowledge and access to international supplier and buyer networks. Other most concentrated industries represent sectors where it is critical for the industry to spread out to reach to the final consumers or the suppliers. On the other hand, the demand for least concentrated industries is diversified across many districts due to which they have substantial raw concentration, e.g., iron and steel, footwear and tobacco, but their employment is distributed across few large plants, e.g., iron and steel, footwear, tobacco and rubber.

Thirdly, we asked whether the most agglomerated industries in 2005-06 were also agglomerated in previous years. The evidence shows that the industry concentration dramatically fell between 1995-96 and 2005-06. The mean value of EG index remained roughly constant from 1995-96 to 2000-01, but the index drastically fell by about 33% in the next five years. This is further corroborated by the correlation coefficient of EG index and the econometric results. Entry of more firms in an industry explains the steady decline in concentration.

Fourthly, the literature tells us that the benefits of agglomeration of industries are often associated with reduction of three types of transport costs, e.g., “moving goods”, “moving people” and “moving ideas” or knowledge spillovers. We examine the factors that help explain the causes of agglomeration of manufacturing industries in Pakistan and the results tell quite a consistent story. We find that the size of district level population, increase in road density in a district, and increase in the pool of technically trained workers in a district all help promote agglomeration of manufacturing industries. The determinants of industry agglomeration guide us on the causes of dense economic activity across spatial units, and the difficulties faced in attracting manufacturing activities in remote districts. However, a range of policy instruments have already been tried in Pakistan as well as other developing countries, e.g., tax holidays, building infrastructure in industrial estates, free trade zones, export processing zones, etc. However, it goes without saying that there is no evidence on their systematic success or failure from any country. Therefore, it is advised that more empirical research need to be conducted to evaluate the effectiveness of the past policies so that we are able to conclude that under what circumstances these programs and policies are likely to succeed.

Fifthly, we find that in Pakistan, as in some developed countries, localization economies or within-industry externalities are much more important than urbanization economies, or inter-industry spillovers, which shows that there is much less role for technological spillovers and inter-industry learning. By implication, policy makers should be able to make a dent on spatial inequality by focusing on “industry-specific subsidies or infrastructural investments”. Industries that offer highest local scale economies are also the most agglomerated industries across districts. Thus, higher magnitudes of localization economies correspond well with inter-district agglomeration and concentration of industry, which is in line with expectations. The urbanization economies are found in only basic metal and metal products industry where an increase in urbanization index by one standard deviation increases productivity by 29%. These results seem to suggest that even though there is little role for inter-industry learning from technological spillovers, basic metal and metal products industry would greatly benefit from increase in the urbanization index.

Finally, the average share of capital in value added is 43.2% in all industries, which is not far from the average share of 37% in the manufacturing sector in South Korea. The highest share of capital is in the textile and leather industry, followed by basic metal and metal products and petrochemical industry. The lowest share of capital is in wood and wood products followed by mineral products. We find that from 1995 to 2005, productivity growth in the large scale manufacturing sector of Pakistan has remained stagnant in all industries, except food, beverage and tobacco industry where productivity increased at 4.4% per annum. In all other industries, productivity remained stagnant during this period.

References

- Alonso-Villar, O., J.-M. Chamorro-Rivas, X. Gonzalez-Cerdeira (2004). Agglomeration Economies in Manufacturing Industries: The Case of Spain, *Applied Economics*, 36, 2103 - 2116.
- Arzaghi, M., J.V. Henderson (2008). Networking off Madison Avenue. *Review of Economic Studies*, 35(4), 1011-1038.
- Bertinelli, L., J. Decrop (2005). Geographic Agglomeration: Ellison and Glaeser's Index Applied to the Case of Belgian Manufacturing Industries, *Regional Studies*, 39, 567-583.
- Burki, Abid A. (2010). Exploring the Links between Inequality, Polarization and Poverty: Does Public Sector Infrastructural Investments Produce Uniform Spatial Effects?. Report prepared for the South Asia Network of Economic Research Institutes (SANEI), Dhaka.
- Burki, Abid A., Kamal Munir, Mushtaq A. Khan, Usman Khan, Adeel Faheem, Ayesha Khalid, S. Turab Hussain (2010). Industrial Policy, its Spatial Aspects and Cluster Development in Pakistan. A report prepared for the Ministry of Industries and Production, Government of Pakistan in association with the World Bank, Washington, D.C. October.
- Burki, Abid A., Mahmood-ul-Hasan Khan (2004). Effects of Allocative Inefficiency on Resource Allocation and Energy Substitution in Pakistan's Manufacturing, *Energy Economics*, 26, 371-388.
- Burki, Abid A., Mushtaq A. Khan, Bernt Bratsberg (1997). Parametric Test of Allocative Efficiency in the Manufacturing Sectors of India and Pakistan, *Applied Economics*, 29, 11-22.
- Ellison, Glenn, Edward Glaeser (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach, *Journal of Political Economy*, 105, 889-927.
- Ellison, Glenn, Edward Glaeser (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review*, 89(2), 311-316.
- Florence, P. Sargant (1948). *Investment, Location and Size of Plant*. London: Cambridge University Press.
- Fuchs, V. (1962). *Changes in the Location of Manufacturing in the US Since 1929*. New Heaven: Yale University Press.
- Fujita, M., Jacques-François Thisse (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge University Press.
- Fujita, Masahisa, Paul Krugman, and A.J. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press.
- Greene, W.H. (1997). *Econometric Analysis*, New Jersey: Prentice Hall International.
- Hanson, Gordon H. (2005). Market Potential, Increasing Returns, and Geographic Concentration, *Journal of International Economics* 67, 1-24.
- Henderson, J. Vernon (1974). The Sizes and Types of Cities, *American Economic Review*, 64, 640-656.
- Henderson, J. Vernon, Todd Lee, Yung J. Lee (2001). Scale Externalities in Korea, *Journal of Urban Economics*, 49, 479-504.
- Henderson, J.V., Z. Shalizi, A.J. Venables (2000). Geography and Development, Policy Research Working Paper No.2456, The World Bank, Washington, D.C.

- Hoover, E.M. (1948). *The Location of Economic Activity*. New York: McGraw Hill.
- Hotelling, H. (1929). Stability in Competition, *Economic Journal*, 39, 41-57.
- Jacoby, H.G., B. Minten (2009). On Measuring the Benefits of Lower Transport Costs, *Journal of Development Economics*, 89, 28-38.
- Jamal, H., A.J. Khan, I.A. Toor, N. Amir (2003). Mapping the Spatial Deprivation of Pakistan, *Pakistan Development Review*, 42(2), 91-111.
- Khandker, S.R., Z. Bakht, G. B. Koolwal (2009). The Poverty Impact of Rural Roads: Evidence from Bangladesh, *Economic Development and Cultural Change*, 57, 685-722.
- Kim, S. (1995). Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987, *Quarterly Journal of Economics*, 110, 881-908.
- Krugman, Paul (1991a). *Geography and Trade*. Cambridge: MIT Press.
- Krugman, Paul (1991b). Increasing Returns and Economic Geography, *Journal of Political Economy*, 99, 483-99.
- Krugman, Paul, and A.J. Venables (1995). Globalization and the Inequality of Income, *Quarterly Journal of Economics*, 110, 857-80.
- Krugman, Paul, Raul E. Livas (1996). Trade Policy and the Third World Metropolis, *Journal of Development Economics*, 49, 137-50.
- Lall, Somik Vinay, Sanjoy Chakravorty (2005). Industrial Location and Spatial Inequality: Theory and Evidence from India. *Review of Development Economics*, 9, 47-68.
- Lall, Somik Vinay, Z. Shalizi, U. Deichmann (2004). Agglomeration Economies and Productivity in Indian Industry, *Journal of Development Economics*, 73, 643 - 673.
- Marshall, A. (1890), *Principles of Economics*, MacMillan, London.
- Maurel, F., B. Sedillot (1999). A Measure of the Geographic Concentration in French Manufacturing Industries, *Regional Science and Urban Economics*, 29, 575-604
- Puga, D. (1999). The Rise and fall of Regional Inequalities, *European Economic Review*, 43, 303-34.
- Wasti, S.A., M.U. Siddiqui (2008). Development Rank Ordering of Districts of Pakistan: Revisited, *Pakistan Journal of Applied Economics*, 18(1&2), 1-29.
- Weber, A. (1909). *Theory of the Location of Industries*. University of Chicago Press, Chicago, Ill.

APPENDIX

Table A1. Most and least developed district based on principal component education endowments

1996-97		2001-02		2005-06		
S.No.	District	PC value	District	PC value	District	PC value
1	karachi	79.20	karachi	105.37	karachi	125.78
2	lahore	28.56	lahore	72.28	lahore	99.99
3	rawalpindi	23.62	faisalabad	41.32	faisalabad	52.56
4	faisalabad	5.10	rawalpindi	29.94	rawalpindi	46.56
5	hyderabad	3.50	gujranwala	27.80	gujranwala	44.82
6	peshawar	0.28	multan	26.33	multan	30.10
7	sargodha	-1.38	sialkot	13.13	bahawalpur	22.17
8	bahawalpur	-2.15	peshawar	13.00	peshawar	21.32
9	multan	-2.88	hyderabad	10.99	hyderabad	20.89
10	sukkur	-3.37	bahawalpur	10.57	sialkot	16.82
11	vehari	-4.00	sargodha	9.32	larkana	11.57
12	jhelum	-4.70	sukkur	7.93	sargodha	10.87
13	okara	-5.14	larkana	7.388	sukkur	8.234
14	gujranwala	-6.10	swat	6.88	gujrat	6.13
15	D.G khan	-6.65	gujrat	2.69	tharparker	5.21
16	mardan	-7.10	D.G khan	1.10	swat	3.63
17	layyah	-7.16	tharparker	0.63	dir	2.23
18	T.T singh	-7.77	dir	0.16	quetta	1.67
19	kasur	-8.02	quetta	-0.49	sahiwal	1.51
20	quetta	-8.36	sheikhupura	-1.23	D.G khan	0.98
21	abbottabad	-8.39	mardan	-1.26	mardan	-0.81
22	tharparker	-8.68	sahiwal	-2.54	sheikhupura	-1.02
23	sanghar	-8.77	R.Y. Khan	-3.33	abbottabad	-1.65
24	attock	-9.48	D I Khan	-4.59	vehari	-2.19
25	mansehara	-9.54	abbottabad	-4.66	bahawalnagar	-3.01
26	badin	-9.55	jhang	-4.69	jhang	-3.46
27	rajanpur	-9.70	okara	-5.39	khanewal	-3.59
28	khairpur	-9.93	vehari	-5.53	R.Y. Khan	-3.69
29	dir	-9.99	kasur	-5.71	mansehara	-3.79
30	dadu	-10.7	T.T singh	-5.82	D I Khan	-4.38
31	nawabshah	-10.59	khanewal	-5.93	okara	-5.19
32	shikarpur	-11.32	muzzafargarh	-6.05	kasur	-5.82
33	jhang	-11.46	bahawalnagar	-7.04	muzzafargarh	-6.14
34	D I Khan	-11.71	mansehara	-7.62	kalat	-6.17
35	sibi	-11.77	jacobabad	-8.26	attock	-6.29
36	chakwal	-11.80	kohat	-8.54	nawabshah	-6.41
37	kohat	-11.86	attock	-8.55	khairpur	-6.51
38	bhakkar	-12.04	chakwal	-9.01	chakwal	-6.52
39	larkana	-12.13	kalat	-9.18	sibi	-7.42
40	R.Y. Khan	-12.21	khairpur	-9.77	kohat	-7.43
41	swat	-12.6	bannu	-9.92	bannu	-7.84
42	kalat	-12.36	dadu	-9.94	T.T singh	-7.92
43	sialkot	-12.68	sibi	-10.18	dadu	-7.99
44	gujrat	-12.69	thatta	-10.18	thatta	-8.36
45	khushab	-12.82	mekran	-10.41	bhakkar	-8.57
46	khanewal	-12.92	rajanpur	-10.67	sanghar	-8.72

47	sheikhupura	-13.21	nawabshah	-10.86	jhelum	-9.17
48	thatta	-13.27	layyah	-10.95	badin	-9.69
49	karak	-13.38	khushab	-10.98	layyah	-9.71
50	mekran	-13.45	jhelum	-11.36	khushab	-10.75
51	jacobabad	-13.45	bhakkar	-12.03	jacobabad	-11.06
52	muzzafargarh	-13.49	mianwali	-12.55	rajanpur	-11.21
53	bannu	-13.61	shikarpur	-12.66	shikarpur	-11.68
54	sahiwal	-13.84	badin	-12.78	mianwali	-11.90
55	bahawalnagar	-14.78	sanghar	-13.37	mekran	-12.36
56	mianwali	-15.86	karak	-14.02	karak	-12.96
