

CREB Working Paper No. 04-13

Labor Pooling as a Determinant of Industrial Agglomeration

Najam uz Zehra Gardezi



**Centre for Research in Economics and Business
Lahore School of Economics**

Centre for Research in Economics and Business (CREB)

Naved Hamid
Director CREB

CREB Advisory Board

Shahid Amjad Chaudhry
Rector
Lahore School of Economics

Sohail Zafar
Dean
Faculty of
Business Administration

Azam Chaudhry
Dean
Faculty of Economics

Muneer Ahmed
Director
Centre for Policy and
Environmental Studies

Shahid Siddiqui
Director
Centre for Humanities
and Social Sciences

Rana Wajid
Director
Centre for Mathematics
and Statistical Sciences

Iqbal M. Khan
Editor
Lahore School
Case Study Journal



Lahore School of Economics

Intersection Main Boulevard Phase VI, DHA and Burki Road
Lahore 53200, Pakistan
Tel: 042-36561230; 042-36560936
Email: creb@lahoreschool.edu.pk

Labor Pooling as a Determinant of Industrial Agglomeration

Najam uz Zehra Gardezi

Teaching Fellow

Lahore School of Economics

© 2013 Centre for Research in Economics and Business
Lahore School of Economics
All rights reserved.

First printing September 2013.

The views expressed in this document are those of the author and do not necessarily reflect the views of the Centre for Research in Economics and Business or the Lahore School of Economics.

Lahore School of Economics
Intersection of Main Boulevard, Phase VI, DHA and Burki Road
Lahore 53200, Pakistan
Tel.: +92 42 3656 1230
creb@lahoreschool.edu.pk
www.creb.org.pk

Price: Rs100

Preface

The Centre for Research in Economics and Business (CREB) was established in 2007 to conduct policy-oriented research with a rigorous academic perspective on key development issues facing Pakistan. In addition, CREB (i) facilitates and coordinates research by faculty at the Lahore School of Economics, (ii) hosts visiting international scholars undertaking research on Pakistan, and (iii) administers the Lahore School's postgraduate program leading to the MPhil and PhD degrees.

An important goal of CREB is to promote public debate on policy issues through conferences, seminars, and publications. In this connection, CREB organizes the Lahore School's Annual Conference on the Management of the Pakistan Economy, the proceedings of which are published in a special issue of the *Lahore Journal of Economics*.

The CREB Working Paper Series was initiated in 2008 to bring to a wider audience the research being carried out at the Centre. It is hoped that these papers will promote discussion on the subject and contribute to a better understanding of economic and business processes and development issues in Pakistan. Comments and feedback on these papers are welcome.

Abstract

This paper analyzes the agglomeration behavior exhibited by manufacturing firms in Punjab. Employing a unique dataset, it constructs a distance-based measure of agglomeration to verify the existence of localization economies. The M function—the industry-level measure of concentration—is regressed on a number of industry characteristics that measure the presence of positive externalities. In particular, a measure of each industry's potential for labor pooling is used to determine whether firms that experience greater fluctuations in employment are likely to be more concentrated. The results provide evidence of the importance of labor pooling in explaining the high level of concentration within industries.

Labor Pooling as a Determinant of Industrial Agglomeration

1. Introduction

The geographical concentration of economic activity is widely observed in most economies. As a result of such clustering, certain locations evolve into cities or business districts, attracting a significant concentration of population while other areas remain relatively less developed. Fujita and Thisse (1996) raise the pertinent question as to why the activities of firms and individuals in an economy are spatially concentrated. The authors identify three distinct motives for the formation of economic agglomerations: (i) increasing returns to scale, (ii) spatial competition, and (iii) externalities. Increasing returns to scale prevent firms from dispersing their plant-level activities, while spatial competition for market area invariably leads firms to locate in proximity to their competitors. Externalities or positive spillovers have received significant attention as a source of industrial agglomeration.

Externalities, an essential aspect of localization economies, refer to the benefits to firms of locating near others in the same industry. Marshall (1920) has highlighted three major sources of intra-industry gains that induce plant clustering within a geographical boundary. These include the diffusion of information on effective processes and efficient production techniques, the ability to support the specialized production of inputs, and the benefits of sharing a similar labor mix. The productivity benefits owing to these externalities are believed to strongly influence firms' location decision (Ellison & Glaeser, 1997).

This study examines the impact of various industry characteristics that measure the presence of these externalities, on the extent of agglomeration exhibited by various manufacturing industries in the Punjab. Given that spatial concentration boosts productivity, agglomeration appears to be an important area of research. It is important to note, however, that the purpose of this study is to examine the sources of differences in agglomeration across industries and not to draw inferences regarding the productivity or growth within these clusters.

Although the presence of industrial clusters in the Punjab is now a stylized fact, the subject lacks substantive empirical analysis. The motivation for this paper stems from an interest in investigating the dispersion of manufacturing activity within the province. Economic activity in the province is believed to be highly concentrated, but without serious conjecture as to why this might be so. Our aim is to identify whether industrial concentration in the Punjab can be explained by the presence of Marshallian externalities and, in particular, by the phenomenon of labor pooling.

A prerequisite for carrying out such a study is to use a consistent measure of agglomeration. Drawing on an extensive and unique dataset comprising information on the location of manufacturing firms in the Punjab, we construct an industry-specific measure of agglomeration by computing the geographical distance between pairs of firms. Such a distance-based framework has not been used in Pakistan and the recent literature confirms its superiority over measures based on discrete spatial units. The M-function computed in this study is based on the theoretical model proposed by Marcon and Puech (2009), and provides empirical evidence on the extent of agglomeration exhibited by each industry.

As a preliminary contribution, however, this paper provides evidence on the location pattern of industries in the Punjab. Using data on the names and addresses of all manufacturing firms in the 32 districts of Punjab, we compute the geographical coordinates of each firm. This allows us to map the firms in an industry and thus show cross-industry variations in the level of agglomeration or dispersion. Figures 1 to 4 map firms in selected industries based on the two-digit International Standard Industrial Classification (ISIC).

Figure 1: ISIC 24 – manufacture of basic metal

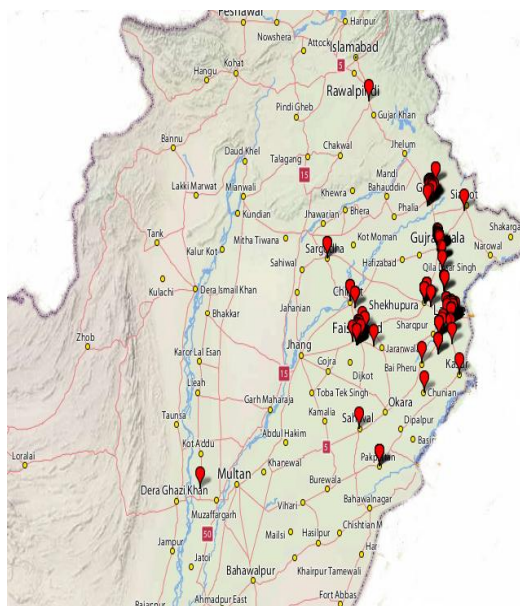


Figure 2: ISIC 10 – manufacture of food products

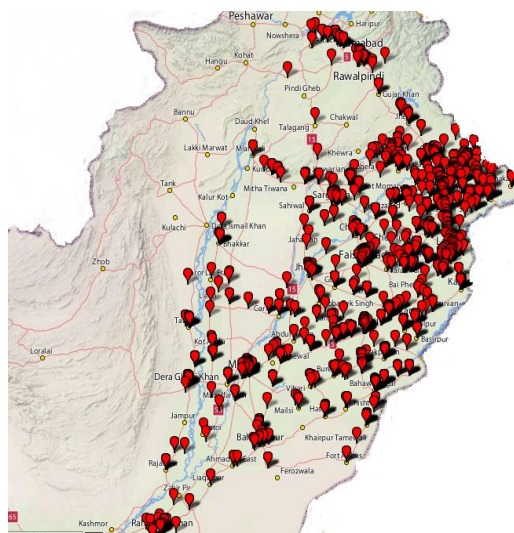


Figure 3: ISIC 29 – manufacture of transport equipment, motor vehicles, and trailers

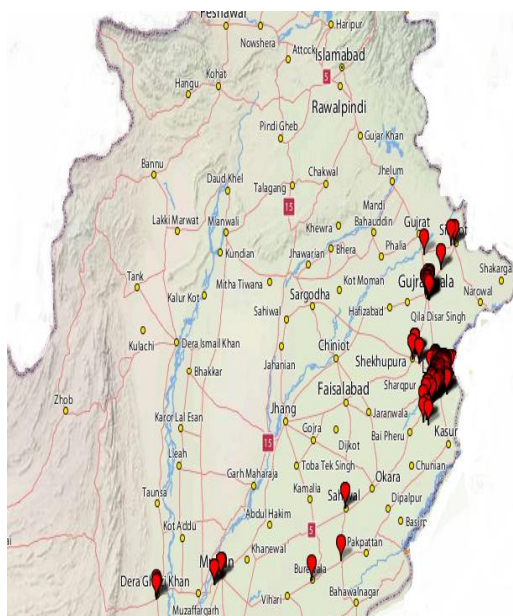
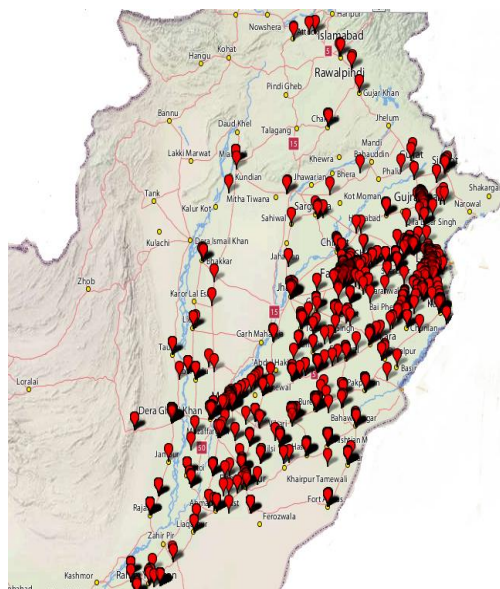


Figure 4: ISIC 13 – manufacture of textiles



The degree of concentration or dispersion of firms evidently varies with the industry to which they belong. Figures 2 and 4 show that industries and firms engaged in the manufacture of food products and textiles, respectively, are spread out across the province. Other industries, such as those involved in the manufacture of basic metal products (Figure 1) seem relatively concentrated. The maps depict the location pattern of individual plants, but for the purpose of empirical estimation, we measure concentration in terms of the density of employment within an area.

The rest of the paper is organized as follows. Section 2 presents a review of the literature, followed by the theoretical framework used to measure the potential for labor pooling in each industry in Section 3 and the measure of agglomeration in Section 4. Section 5 presents our empirical estimation, and Section 6 describes the data used. Section 7 presents the study's results, and Section 8 provides some concluding remarks.

2. Literature Review

The geographical organization of an economy presents an interesting area for research. Fujita and Thisse (1996) have discussed the existence of various centripetal and centrifugal forces that determine the extent of agglomeration or dispersion in economic activity. Most subsequent research, however, has paid closer attention to the reasons that firms locate in proximity to each other, i.e., the centripetal forces. Numerous studies examine the relative importance of factors influencing the location decisions of firms, and derive their premise largely from the classical theories of Marshall (1920) and, subsequently, Krugman (1991).

Marshall (1920), who discusses the various advantages to an industry of being geographically concentrated, provides one of the earliest explanations for agglomeration economies. The sources of these benefits, which subsequently became known as Marshallian externalities, can be divided into three broad categories: (i) input sharing, (ii) labor pooling, and (iii) knowledge spillovers. The phenomenon of localization economies, which constitute the benefits to an industry of being spatially concentrated, is in large part explained by these three forces.

Ellison and Glaeser (1997) show that the level of agglomeration is not consistent across industries, indicating that the benefits derived from external economies are more pronounced for some industries than for

others. A significant contribution of their study to the literature on agglomeration is an index that measures geographical concentration. Based on its results, Ellison and Glaeser conclude that the extent of localization exhibited by all industries is greater than that accounted for by a random distribution of plants. Their measure of agglomeration has been used widely in many subsequent studies (see, for example, Rosenthal & Strange, 2001; Bertinelli & Decrop, 2005; Ellison, Glaeser, & Kerr, 2010).

However, a new strand of literature has now emerged that questions the underlying assumption of discrete spatial units in the computation of the Ellison–Glaeser (E-G) index and other similar indices. Research emphasizing the primacy of distance-based measures of agglomeration over discrete space models has recently gained momentum in the literature on economic geography. One of the earliest studies to highlight the shortcomings of indices that assume discrete spatial units is that of Duranton and Overman (2005). They derive an industry-level measure of agglomeration based on a model of continuous space, using the distribution of geographical distances between pairs of firms in an industry. While their findings are accepted to be more accurate, the use of such a measure of concentration has generally been limited. This is primarily due to its extensive data requirements—in particular, the need for the exact spatial address or geographical coordinates of each firm included in the index.

More recently, Marcon and Puech (2009) have introduced a comprehensive measure of spatial concentration based on geo-distances. The “M-function” they use helps identify both inter- and intra-industrial geographical concentration and fulfills the criteria set by Duranton and Overman (2005) for what constitutes a sound measure of agglomeration. Although the study presents a valid theoretical model and establishes the primacy of the M-function over other distance-based measures, the work is too recent to have been applied to actual data.

A significant body of literature is devoted to measuring the level of agglomeration, using different discrete as well as continuous models, and to building a consensus on an ideal measure. At the same time, the application of such a measure has also received considerable attention and many researchers have attempted to identify the factors that cause firms to cluster across space. In particular, these studies have looked at the industry-specific characteristics that may induce one industry to be more prone to agglomeration than another.

Rosenthal and Strange (2001) examine the sources of differences in the extent of geographical concentration exhibited by manufacturing industries in the US. Their work on the determinants of agglomeration has direct significance for our study, which proposes a similar empirical estimation. Rosenthal and Strange use proxies for the three prominent causes of localization—labor pooling, input sharing, and knowledge spillovers—and measure their impact on agglomeration at different levels of geographical aggregation. Their results provide evidence to support all three theories, but indicate that labor pooling is the strongest motive driving a firm's decision to concentrate its production activities near others within the same industry. Their results are consistent with the findings of Dumais, Ellison, and Glaeser (2002) that patterns of co-agglomeration are dominated by labor pooling considerations.

The incentive to locate near establishments that use a similar labor mix appears to be a strong force of agglomeration. In Pakistan, Burki and Khan (2010) have studied the trend in agglomeration exhibited by the country's manufacturing sector. Their analysis also employs the E-G index and provides evidence that district characteristics and infrastructure are vital in explaining the high levels of concentration.

In addition to the empirical evidence provided by the above studies, labor pooling has also received attention in the theoretical literature as an important source of agglomeration. Overman and Puga (2009) examine its influence on the spatial concentration of manufacturing establishments in the UK. Their study is unique in that it provides a direct measure of labor pooling rather than using indirect proxies of the extent to which workers possess industry-specific skills. Following Krugman's (1991) model, the authors present a model of labor pooling in which individual establishments have the benefit of altering their level of employment in response to possible demand shocks. The study argues that firms will tend to agglomerate within the industry if they are prone to fluctuations in employment. Based on this measure and using the E-G index of concentration, the authors test the relative importance of labor pooling across industries in the UK. The empirical analysis supports their hypothesis that labor pooling that operates through such mechanisms does indeed result in a greater concentration of economic activity.

3. Measure of an Industry's Potential for Labor Pooling

The potential for labor pooling is measured by the extent of heterogeneous shocks faced by an industry. Spatially concentrated firms in an industry will have higher expected profits because fluctuations in employment in case of productivity shocks will not be reflected in wages. Another theoretical argument that highlights the role of a large labor market as a source of positive externalities centers on a firm's improved prospects for finding the right worker for a job, just as workers are able to find jobs that better match their skills (Helsley & Strange, 1990). Moreover, a large market promotes the specialization of labor, which can generate increasing returns for firms that agglomerate geographically (Duranton, 1998). For the purpose of our study, however, the measure of labor pooling is based on firms' ability to adjust their employment levels in response to productivity shocks.

An industry will exhibit greater agglomeration if its constituent firms face heterogeneous fluctuations in employment. The expected profits for such firms will be higher if they are agglomerated than if they were located in isolation because in the latter case, these shocks would be transmitted to local wages. The advantage of locating near other establishments in the industry arises from the firm's need to adjust its level of employment at a time when other firms do not face similar productivity shocks. The wage level should, therefore, remain largely unaffected.

Following Overman and Puga (2009), we measure the idiosyncratic shock to a single firm by calculating the difference between the absolute value of percentage change in that firm's employment and the percentage change in the industry's employment. These are then averaged across all the firms in the industry to obtain an industry-level measure of fluctuation. This, in turn, reflects the potential for labor pooling within the industry.

A slight adaptation of this measure would be to take the difference between the change in industry-level employment and the change in total manufacturing employment in Punjab. Taking manufacturing employment as a whole rather than merely that of the industry allows us to observe the benefits of labor pooling across sectors. Similarly, it is possible to construct this measure of labor pooling by taking the change in firm employment relative to the change in the manufacturing sector as a whole.

4. Measure of Geographical Concentration of Industries

The importance of agglomeration varies with the type of industry. High-technology sectors, for example, may exhibit greater geographical concentration because of their higher potential for knowledge spillovers. Most studies on agglomeration have relied on employment shares to measure the density of economic activity, but the recent literature has criticized the use of crude measures to determine whether the distribution of a particular industry's employment follows that of total employment in a specific geographical unit. In particular, Duranton and Overman (2005) argue that any index of spatial concentration should satisfy a number of requirements. The employed measure should allow a comparison of geographical concentration across industries and control for the overall aggregation across industries. It must also distinguish between industrial concentration and spatial concentration.

The E-G index fulfills these requirements and, consequently, has been widely used. However, along with other such agglomeration indices that rely on the assumption of discrete spatial units, it fails to meet the criterion that a consistent measure must produce results that are unbiased with respect to the degree of spatial aggregation. The latter condition is violated if we assume that firms locate in discrete states because the position, size, and shape of the chosen spatial unit (tehsil, district, or province) can affect the results of the analysis.

To address this shortcoming, recent studies on agglomeration have moved away from employing cluster-based methods and instead emphasized distance-based methods, which are independent of politically assigned boundaries and thus do not suffer from a statistical bias induced by the choice of spatial unit. The data requirements for such measures of agglomeration are, however, quite extensive and their application in the literature on economic geography is, therefore, rather limited.

The measure for localization that we employ in this study is the M function proposed by Marcon and Puech (2009). It makes use of the average number of a plant's neighbors in a given radius (r) and compares the location patterns of an industry within that area to that of the entire manufacturing sector. The intuition behind this model is that we can compare the number of plants belonging to the same industry within a distance r with a benchmark distribution of the industry. This

should allow us to detect whether firms are more or less concentrated than if they were distributed randomly and independently of each other.

The measure is calculated by computing for each firm the ratio of neighboring firms belonging to the industry i within a radius r over the number of all firms within that distance. This is then averaged across the industry and compared to the ratio of all firms in industry i over the total number of firms in Punjab. We use weights to control for the concentration within plants and account for the size of each firm, using the number of employees rather than the number of firms. Thus, for each industry, we can obtain the average proportion of employees for that industry within a radius r . Similarly, the ratio of all employees in industry i to the total employment by the manufacturing sector can be computed. The M function uses the ratio of these two quantities averaged across all firms in an industry and yields an industry-level measure of intra-industrial geographical concentration.

In equation 1 below,¹ the numerator depicts the share of industry i in industrial activity in circles of radius r while the denominator represents the share of the industry in all manufacturing activities in Punjab. The benchmark for the M function is 1 so that any value greater than 1 indicates that a greater proportion of employees near firms belonging to industry i are located within a distance r relative to the industry's share of employees over the entire province. The statistical significance of this measure can be tested using confidence intervals. The measure is thus comprehensive and fulfills the five criteria put forth by Duranton and Overman (2005).

$$M_i(r) = \frac{1}{N_i} \sum_{s=1}^{N_i} \frac{\sum_{n=1, s \neq n}^{N_i} c_i(s, n, r) w_n}{\sum_{n=1, s \neq n}^N c(s, n, r) w_n} \bigg/ \sum_{s=1}^{N_i} \frac{W_i - w_s}{W - w_s} \quad (1)$$

5. Empirical Estimation

Our empirical estimation involves measuring the effect of agglomerative externalities on the degree of spatial concentration exhibited by various industries. In particular, we want to observe the impact on agglomeration of the potential for labor pooling within an industry. The conventional estimation technique involves estimating a linear model using ordinary least squares (OLS). The first equation to be estimated is as follows:

¹ For a detailed explanation, see Marcon and Puech (2009).

$$M_i = \beta_0 + \beta L_j + \beta X_j + \beta Z_j + e_j \quad (2)$$

where M_i is the M function that measures the extent of the geographical concentration of an industry within a certain radius. For this study, the M function is computed for a two-mile and ten-mile radius. It is regressed on several industry characteristics: L_j denotes our measure of the potential for labor pooling in the industry and X represents the vector of industry characteristics that proxy for the presence of the remaining two Marshallian externalities—knowledge spillovers and input sharing. An additional variable, denoted by Z , is added to control for transport costs in the industry.

The regression equation specified above is close to that employed by Rosenthal and Strange (2001). The primary hypothesis being tested is that establishments more prone to productivity shocks will exhibit greater concentration because they will need to adjust employment levels without affecting their profitability. The potential for labor pooling will, therefore, differ for each industry based on the firm-specific fluctuations experienced.

While our key interest is measuring the impact of labor pooling as a source of agglomeration, the significance of the remaining Marshallian externalities should not be undermined. Industries sensitive to the costs of input and knowledge sharing are more likely to exhibit agglomeration behavior—we add these to the proposed model as controls. As regards the proxy for input sharing, we suggest that sharing suppliers of intermediate goods is a motive for the concentration of manufacturing activity.

Input sharing is measured using the ratio of manufactured inputs as a share of total inputs. This reflects the presence of vertical linkages such that industries that make greater use of the output of other plants will tend to cluster near the plant that provides their inputs. Accordingly, agglomeration should be positively related to this measure of input sharing. The other externality, knowledge sharing, is relatively difficult to measure given the existing data constraints. It is generally measured using firms' research and development (R&D) expenditure or some proxy for innovation such as new products. Such data could not be found for manufacturing firms in Punjab.

The share of expenditure on imported spare parts is used as a crude proxy for the presence of knowledge sharing. The rationale for this measure lies in the argument that firms that import spare parts are in some way engaged in adding to the existing knowledge on techniques and processes. It is also presumed that such firms are more aware of the existing technology and thereby more likely to generate spillovers for other firms in the same industry.

An additional control represented by Z in equation 3 measures transport costs, which are a significant determinant of the level of agglomeration or dispersion of economic activity. This idea has received much attention in the theoretical literature (see Fujita & Thisse, 1996). High transport costs of supplying to the consumer market will deter agglomeration in manufacturing industries. The Census of Manufacturing Industries (CMI) provides information on the cost of transporting finished goods to the market. The ratio of transport cost to sales is thus computed to measure the industry-level variation in transport cost. If the transport cost is high because a good needs to be more frequently transported to the market, firms manufacturing that good will locate nearer the consumer and, hence, become relatively dispersed. We therefore expect the variable to carry a negative sign.

6. Data

This study has two distinct objectives. The first is to compute a measure of the geographical distribution of firms that is not subject to the statistical bias resulting from the choice of discrete spatial units. In order to employ a continuous approach to space, we require the geographical coordinates of all firms to compute the distance between pairs of firms.

This data is taken from the Directory of Industries, which contains in addition to other information the names, addresses, and employment figures of over 17,000 plants operating in Punjab. The most recent data is available for 2010 and is used to compute the geographical coordinates for independent firms. This, in turn, allows us to calculate the distances between all firms, and thus determine the M function. These coordinates are also used to include the geographical mapping of industries presented earlier. Table 1 summarizes the M function computed for the two distances.

Table 1: Summary measures of agglomeration among manufacturing industries

M function	Mean	SD	Min.	Max.	Correlation
10 miles	1.656979	1.3320220	0	5.960157	1.0000
2 miles	1.137627	0.4069265	0	2.852808	0.4687

Source: Author’s calculations.

The second objective entails the application of this measure of agglomeration. The data source used to investigate the relationship between labor pooling and spatial concentration is the CMI for Punjab, which provides plant-level data on the quantities and value of inputs and outputs as well as details of employment and employment costs for 3,528 establishments. Table 2 provides selected summary statistics for the explanatory variables used to measure the presence of externalities.

Table 2: Four-digit manufacturing industries: Definitions and selected summary statistics for key explanatory variables

Variable	Definition	Obs.	Mean	SD	Min.	Max.
Labor pooling (plant to industry)	Difference between percentage change in plant’s employment and percentage change in industry employment (idiosyncratic shock)	3,505	0.196	0.247	0.006	3.407
Labor pooling (plant to total)	Idiosyncratic shock to a firm relative to the manufacturing sector as a whole	3,505	0.236	0.281	0.020	1.363
Labor pooling (industry to total)	Idiosyncratic shock experienced by an industry relative to the manufacturing sector as a whole	3,505	0.129	0.244	0.001	3.046
Input sharing	Share of manufactured inputs as the cost of total inputs	3,528	0.141	0.124	0.007	0.620
Knowledge spillovers	Share of expenditure on imported spare parts	3,528	9E-04	0.003	0.000	0.056

Source: Author’s calculations.

One limitation that arises in constructing the measure of labor pooling is that we require the percentage change in employment at the firm level. While the census spans three years, it is not possible to match the establishments across years. The change in employment for this analysis is, therefore, the change across quarters within the year; this information

is recorded for all surveyed firms. Another limitation is that there is no information on firms' expenditure on R&D. Consequently, we use firms' expenditure on imported spare parts as a proxy for the presence of knowledge spillovers. The other controls for industry characteristics are computed using the input and output data reported in the CMI.

To carry out the empirical estimation discussed in the previous section, we match the observations in the Directory of Industries with those in the CMI. The two datasets are thus coded on the basis of the four-digit ISIC to allow an industry-level analysis. Additionally, since the CMI includes only medium-sized and large firms with a minimum of 10 employees, in computing the M-function we include 9,577 firms from the Directory of Industries and discard the remaining observations.

Finally, the most recent CMI is available for the fiscal year 2006, but the measure of agglomeration is computed using data for 2010. The rationale for this lies in Overman and Puga's (2009) observation that a lag between characteristics and outcome allows for a more sound economic interpretation. In addition, the lag prevents the possible problem of endogeneity between certain industry characteristics and the measure of intra-industry geographical concentration.

7. Results

The M function, i.e., the distance-based measure of agglomeration employed in this study, is computed for a ten-mile and two-mile radius. The results depicting the employment density of industries classified according to the four-digit ISIC are somewhat similar within the two distances. Table 3 lists the most and the least agglomerated industries in the province.

The most agglomerated firms exist within industries that manufacture fabricated metal products (two-digit ISIC 25) and those involved in the manufacture of nonmetallic mineral products (two-digit ISIC 23). Based on the four-digit industrial classification, firms that manufacture sports goods and bodies for motor vehicles appear to be highly concentrated. The manufacture of machinery for textiles and leather is also a concentrated activity in Punjab.

Among the least agglomerated industries are those engaged in the production of tobacco and the manufacture of weapons and ammunition. The M function assigns a value of 0 to such industries because these plants do not have neighboring firms within a two-mile or ten-mile radius.

Firms involved in the manufacture of food products, leather products, and consumer items such as lubricants, detergents, cosmetics, and perfumes (ISIC 20: manufacture of chemical and chemical products) appear to be dispersed. These findings are consistent with Burki and Khan (2010)² and support the conjecture that the least concentrated industries are those for which demand is well spread out. Hence, the need to be closer to the consumer could be a strong determinant in the location choice of such firms.

Table 3: Geographical concentration of four-digit industries in Punjab, 2010

ISIC	Description	M 2 miles	ISIC	Description	M 10 miles
3092	Manufacture of bicycles and invalid carriages	2.8528	2396	Cutting, shaping, and finishing of stone	5.960157
2393	Manufacture of porcelain and ceramic products	2.4151	2920	Manufacture of motor vehicle bodies	5.770630
2394	Manufacture of cement, lime, and plaster	2.4151	2393	Manufacture of porcelain and ceramic products	4.987743
1311	Spinning of textile fibers	1.3258	2593	Manufacture of cutlery, handheld tools and general hardware	4.628244
3230	Manufacture of sports goods	1.2887	2815	Manufacture of ovens, furnaces, and furnace burners	4.144241
3250	Manufacture of medical and dental instruments	1.2870	2511	Manufacture of structural metal products	4.010656
2920	Manufacture of motor vehicle bodies	1.2730	2394	Manufacture of cement, lime, and plaster	3.444931
2826	Manufacture of machinery for textile, apparel, and leather production	1.2724	3230	Manufacture of sports goods	3.406820
2511	Manufacture of structural metal products	1.2615	2826	Manufacture of machinery for textile, apparel, and leather production	3.357219
2599	Manufacture of other fabricated metal n.e.c.	1.2595	3250	Manufacture of medical and dental instruments	3.239103

² Burki and Khan (2010) have used the CMI for 2006 to measure agglomeration using the E-G index. The results of the M function in this study are based on data from the Directory of Industries for 2010.

ISIC	Description	M 2 miles	ISIC	Description	M 10 miles
10 least concentrated industries					
1200	Manufacture of tobacco products	0	1200	Manufacture of tobacco products	0
2520	Manufacture of weapons and ammunition	0	2520	Manufacture of weapons and ammunition	0
2813	Manufacture of other pumps, compressors, taps, and valves	0.3153	2029	Manufacture of other chemical products n.e.c.	0.106498
2310	Manufacture of glass and glass products	0.5757	1311	Spinning of textile fibers	0.163311
2029	Manufacture of other chemical products n.e.c.	0.7619	2910	Manufacture of motor vehicles	0.224832
2011	Manufacture of basic chemicals	0.7768	1312	Textile weaving	0.384844
1399	Manufacture of other textiles n.e.c.	0.8883	2011	Manufacture of basic chemicals	0.389968
2023	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations	0.9371	1072	Manufacture of soft drinks, mineral water, and other bottled water	0.532559
1312	Textile weaving	0.9525	1104	Manufacture of other food products n.e.c.	0.576307
2910	Manufacture of motor vehicles	0.9604	1079	Manufacture of prepared animal feed	0.601087

Source: Author's calculations.

A more empirical explanation for the variation in industrial concentration as depicted by the M function may lie in the Marshallian externalities. Tables 4 and 5 present the results of OLS regressions. The specification is such that the measure of agglomeration is regressed on the lagged values of industry characteristics. As Overman and Puga (2009) observe, this produces a lag between characteristics to outcomes such that firms are able to observe industry characteristics prior to making a decision regarding location. The primary reason for using lagged values, however, is to avoid possible problems of endogeneity.

Table 4: OLS regression
(Dependant variable: M function – 10 miles)

	(1)	(2)	(3)
Labor pooling (plant to industry)	1.69 * (0.96)		
Labor pooling (plant to manufacturing)		1.78 (4.06)	
Labor pooling (industry to manufacturing)			1.94** (0.86)
Input sharing	21.00* (11.90)	21.08* (11.85)	21.03* (11.87)
Knowledge spillovers	-57.97 (162.21)	-59.78 (162.87)	-49.96 (164.06)
Transport cost	-122.46* (72.79)	-125.76* (71.50)	-121.77 (73.78)
Constant	9.22*** (2.24)	9.27*** (2.24)	9.29*** (2.28)
R2	0.10	0.09	0.10
n = 70			

Note: Robust standard errors given in parentheses. ***, **, and * indicate statistical significance at 1, 5 and 10 percent, respectively.

Source: Author's calculations.

The measure of labor pooling is computed as idiosyncratic shocks relative to the industry (column 1) and then by changing the reference category to the manufacturing sector as a whole (columns 2 and 3). The results show that industries in which individual firms experience greater idiosyncratic employment shocks are more likely to be geographically concentrated. As stated earlier, the rationale for this is that firms prone to frequent changes in employment choose to locate closer to other firms in the same industry so that wages and, in turn, their profitability is not affected.

The impact of labor pooling is also significant for industries that experience shocks relative to changes in the manufacturing sector as a whole (column 3). Such industries may be geographically concentrated but might also choose to locate in more urbanized areas where they can benefit from labor pooling across industries.

The employment data for firms within the most concentrated industries show that agglomeration occurs largely in the industrialized districts of Gujranwala and Sialkot. Since the regression estimates are based on the four-digit classification, it is possible that the impact of labor pooling (from firm to industry) is undermined by labor mobility within a broader reference group.

The other agglomeration forces that we have controlled for are the two Marshallian externalities: input sharing and knowledge spillovers. In addition, we have accounted for the transport cost of industries, which the literature shows to be an important determinant of location. There is evidence that the ability to benefit from the production of specialized inputs is positively associated with the level of concentration exhibited by industries. Input sharing is found to be significant for both levels of geography and the coefficients have a higher magnitude than that of labor pooling.

The exception is column 2 in Table 5. These results, which include the measure of labor pooling computed as the idiosyncratic shocks faced by individual plants relative to the manufacturing sector as a whole, are insignificant for all industry characteristics and only marginally significant for the labor pooling measure itself.

Excluding these results, firms that use manufactured inputs in their production tend to be located closer together, most likely near the supplier of these manufactured inputs. It may be, however, that within the rather small radius of two miles, the effect is undermined because other factors take precedence. The third Marshallian externality that we test for is knowledge spillovers, which is not found to be significant in any of the regressions.

The proxy for this externality is based on the expenditure on imported spare parts. The agglomeration literature commonly measures knowledge spillovers using firm expenditure on R&D, but this data is not available for manufacturing industries in Pakistan and, therefore, we opt for an indirect proxy. It may be that any knowledge spillovers generated by the firms that import spare parts are not sufficient to affect the level of industrial agglomeration. The control for transport costs is negative but at a low significance level. According to the literature,

transport costs are an important determinant of agglomeration.³ The effect of transport costs on agglomeration might, however, be stronger for higher levels of geography than those accounted for in this analysis.

Table 5: OLS regression
(Dependant variable: M function – 2 miles)

	(1)	(2)	(3)
Labor pooling (plant to industry)	0.52*** (0.06)		
Labor pooling (plant to manufacturing)		0.82* (0.41)	
Labor pooling (industry to manufacturing)			0.63*** (0.05)
Input sharing	0.68* (0.39)	0.71 (0.44)	0.79** (0.36)
Knowledge spillovers	12.67 (13.98)	1.99 (15.42)	13.71 (14.66)
Transport cost	-0.86* (0.49)	-0.73 (0.53)	-0.84* (0.50)
Constant	0.95*** (0.09)	0.94*** (0.11)	0.98*** 0.08
R2	0.45	0.29	0.48
n = 70			

Note: Robust standard errors given in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Source: Author's calculations.

The results above show that the presence of potential labor pooling in an industry has an important impact on the agglomerative behavior shown by firms in that industry. The R-squared terms in both the OLS regressions presented in Tables 2 and 3 are, however, low, thereby indicating that the variation in agglomeration across industries is not sufficiently explained by the externalities accounted for in this study.

Other, industry-specific characteristics could result in cross-industry variation in the level of agglomeration. A number of demand-side factors

³ Rosenthal and Strange include natural advantage as a control variable in their original study but we have omitted it in these regressions because of a strong correlation between the proxy for this variable and that of input sharing.

specific to the industry can affect the location decision of firms in terms of whether they are dispersed or agglomerated. To address this concern, we change the specification to include industry fixed effects that are based on the two-digit ISIC. The results are shown in Tables 6 and 7 for a ten- and two-mile radius, respectively.

Table 6: Two-digit fixed effects
(Dependant variable: M function – 10 miles)

	(1)	(2)	(3)
Labor pooling (plant to industry)	0.96** (0.36)		
Labor pooling (plant to manufacturing)		0.26 (0.78)	
Labor pooling (industry to manufacturing)			0.77*** (0.28)
Input sharing	3.23** (1.39)	3.53** (1.53)	3.54** (1.47)
Knowledge spillovers	9.63 (47.00)	6.44 (51.50)	11.39 (50.84)
Transport cost	-0.20 (1.91)	0.06 (1.85)	-0.03 (1.87)
Constant	0.68* (0.34)	0.83** (0.37)	0.77** (-0.33)
R2	0.51	0.45	0.48
n = 70			

Note: Robust standard errors given in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Source: Author's calculations.

The evidence that the two Marshallian externalities (labor pooling and input sharing) affect agglomeration is stronger when we account for industry fixed effects. The proxies for these externalities are significant for both levels of geography. As with the OLS regressions, however, the results for labor pooling as measured by the idiosyncratic shocks to firms relative to the manufacturing sector as a whole are insignificant for the larger ten-mile radius. This could be because individual plants faced with productivity shocks might be unable to attract labor from different sectors if it requires greater mobility on the part of the workers. Hence, while such a measure of labor pooling is positively associated with

agglomeration for a smaller distance, it becomes insignificant as the distance increases.

Table 7: Two-digit fixed effects
(Dependant variable: M function – 2 miles)

	(1)	(2)	(3)
Labor pooling (plant to industry)	0.47*** (0.08)		
Labor pooling (plant to manufacturing)		0.51** (0.43)	
Labor pooling (industry to manufacturing)			0.65*** (0.11)
Input sharing	0.76* (0.40)	0.85** (0.44)	0.90** (0.36)
Knowledge spillovers	18.29 (17.11)	12.05 (17.15)	19.73 (16.73)
Transport cost	-1.25*** (0.41)	-1.07** (0.42)	-1.18*** (0.39)
Constant	0.96*** (0.09)	0.94*** (0.12)	0.93*** (0.07)
R2	0.64	0.34	0.70
n = 70			
20 fixed effects			

Note: Robust standard errors given in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Source: Author's calculations.

Transport costs become insignificant in all regressions for the ten-mile radius once fixed effects are incorporated. On the other hand, the variable is highly significant at the lower level of geography, i.e., at two miles. The results shown in Table 5 provide evidence that the higher the cost of transport involved in making a good available to the consumer, the lower will be an industry's level of geographical concentration. This supports the argument presented earlier that firms will be dispersed if the demand for their goods is well spread out.

8. Conclusion

This paper has aimed to add to the current understanding of the micro-foundations of agglomeration economies in developing countries. It has attempted to identify the factors that can account for the concentration of industries beyond that explained by area characteristics. Our computation of a distance-based measure and its application is a unique contribution to the literature, and allows for greater accuracy in assessing the extent of concentration exhibited by manufacturing firms. Such a measure may be particularly useful to researchers aiming to study the impact of industrial concentration on growth or productivity. The distance-based measure allows one to assess the impact of agglomeration over various levels of geography without encountering the bias associated with political state boundaries.

We have shown that certain industry-specific characteristics lead manufacturing firms to choose to locate near other establishments in the same industry. Marshall (1920) has highlighted these characteristics as the positive externalities that allow firms within the same industry to experience cost advantages if they choose to locate in proximity to other firms.

Two particular mechanisms through which these cost advantages can occur have been identified. The first is the potential for labor pooling that may arise for firms that need to alter their employment levels, i.e., those firms that are more susceptible to productivity shocks. The second is input sharing whereby firms can have the advantage of attracting the production of specialized inputs. An important implication of these findings is that policymakers can facilitate economic activity in less developed areas through cost incentives that match the benefits arising from such externalities.

As for any developing country, the manufacturing sector is vital to Pakistan's overall economic performance. Such an empirical analysis has not previously been carried out for the country and it is an emerging area of research in the international literature. Identifying the various sources of agglomeration provides valuable insight into potential sources of productivity advantages, and indicates that the industry dynamics that are vital in explaining firm behavior in developed countries might not hold as strongly for developing counterparts with different cost structures. This subject could be explored further and in greater detail if the current data limitations were overcome.

References

- Bertinelli, L., & Decrop, J. (2005). Geographical agglomeration: Ellison and Glaeser's index applied to the case of Belgian manufacturing industry. *Regional Studies*, 39(5), 567–583.
- Burki, A. A., & Khan, M. A. (2010). *Spatial inequality and geographic concentration of manufacturing industries in Pakistan*. Paper presented at the 27th Annual General Meeting and Conference of the Pakistan Society of Development Economists, Islamabad, Pakistan.
- Dumais, G., Ellison, G., & Glaeser, E. L. (2002). Geographic concentration as a dynamic process. *Review of Economics and Statistics*, 84(2), 193–204.
- Duranton, G. (1998). Labor specialization, transport costs, and city size. *Journal of Regional Science*, 38(4), 553–573.
- Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies*, 72, 1077–1106.
- Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing Industries: A dartboard approach. *Journal of Political Economy*, 105(5), 889–927.
- Ellison, G., & Glaeser, E. L. (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, 89(2), 311–316.
- Ellison, G., Glaeser, E. L., & Kerr, W. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, 100(3), 1195–1213.
- Fujita, M., & Thisse, J.-F. (1996). Economics of agglomeration. *Journal of the Japanese and International Economies*, 10(4), 339–378.
- Helsley, R. W., & Strange, W. C. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics*, 20(2), 189–212.
- Krugman, P. (1991). *Geography and trade*. Cambridge, MA: MIT Press.

- Marcon, E., & Puech, F. (2009). Measures of the geographical concentration of industries: Improving distance-based methods. *Journal of Economic Geography*, 10(5), 745–762.
- Marshall, A. (1920). *Principles of economics*. London, UK: Macmillan.
- Overman, H. G., & Puga, D. (2009). *Labour pooling as a source of agglomeration: An empirical investigation* (Discussion Paper No. 7174). London, UK: Centre for Economic Policy Research.
- Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191–229.
- Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial location, and agglomeration. *Review of Economics and Statistics*, 85(2), 377–393.
- Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson & J.-F. Thisse (Eds.), *Handbook of urban and regional economics* (vol. 4, pp. 2119–2171). Amsterdam, The Netherlands: North Holland.

Lahore School of Economics
Centre for Research in Economics & Business

Working Paper Series

No. 03-13

The Effects of Agglomeration on the Formation and Scale of Operation of New Firms
Maryiam Haroon

No. 02-13

Agglomeration and Firm Turnover
Marjan Nasir

No. 01-13

Determinants of School Choice: Evidence from Rural Punjab, Pakistan
Hamna Ahmed, Sahar Amjad, Masooma Habib, Syed Ahsan Shah

No. 03-12

The Effects of External Migration on Enrolments, Accumulated Schooling, and Dropouts in Punjab
Rabia Arif, Azam Chaudhry

No. 02-12

The Determinants of Child Health and Nutritional Status in Punjab: An Economic Analysis
Uzma Afzal

No. 01-12

Investigating the Proposed Changes to Pakistan's Corporate Bankruptcy Code
Ali Hasanain, Syed Ahsan Ahmad Shah

No. 02-11

Cross-Country Growth Spillovers: Separating the Impact of Cultural Distance from Geographical Distance
Azam Chaudhry, Rabia Ikram

No. 01-11

The Determinants of Female Labor Force Participation in Pakistan: An Instrumental Variable Approach
Mehak Ejaz

No. 01-10

The Determinants of Interest Rate Spreads in Pakistan's Commercial Banking Sector
Ayesha Afzal, Nawazish Mirza

No. 03-09

Evaluating the Impact of Microcredit on Women's Empowerment in Pakistan
Salman Asim

No. 02-09

Speculative Bubbles in Karachi Stock Exchange
Nawazish Mirza

No. 01-09

Economic Development: A View From the Provinces
Khalid Ikram

No. 02-08

Agricultural Growth in Irrigated Punjab: Some Issues and Policies
Mahmood Hassan Khan

No. 01-08

Size and Value Premium in Karachi Stock Exchange
Nawazish Mirza

Policy Paper Series

No. 01-11

Pakistan and Lessons from East Asia: Growth, Equity, and Governance
Khalid Ikram

No. 01-10

A Strategy for Reversing Pakistan's Dismal Export Performance
Hamna Ahmed, Mahreen Mahmud, Naved Hamid, Talal-Ur-Rahim

These papers can be accessed at: www.creb.org.pk

The Lahore School of Economics (established in 1993) is one of Pakistan's leading centres of learning for teaching and research in economics, finance and business administration. Its objectives are (i) to train young Pakistanis as professional economists, finance managers, accountants, financial analysts, bankers, and business executives, and (ii) to undertake research in economics, management, finance, and banking to deepen the understanding of major facts, issues, and policies.

The Centre for Research in Economics and Business (CREB) is an independent research centre at the Lahore School of Economics. CREB's mission is to conduct and facilitate research, coordinate and manage the Lahore School's postgraduate program, and promote discussion on policy issues facing Pakistan. The research focus at CREB is on the management of the Pakistan economy, income distribution and poverty, and the role of the modern services sector in the area of economics; and financial markets in the area of business management.

The Lahore School's publication program comprises the Lahore Journal of Economics, Lahore Journal of Policy Studies, Lahore Journal of Business, a Text Book Series, Lahore School Case Study Journal, the CREB Working Paper Series, and CREB Policy Paper Series. The program encourages both in-house and external contributors.



Lahore School of Economics

**Intersection Main Boulevard Phase VI, DHA and Burki Road
Burki Lahore 53200, Pakistan**