Labor Pooling As a Determinant of Industrial Agglomeration

By

Najam uz Zehra Gardezi

Thesis submitted to the Lahore School of Economics in partial fulfillment of the requirements of the degree of MPhil in Economics

Supervised by: Dr. Azam Chaudhry

LAHORE SCHOOL OF ECONOMICS
LAHORE, PAKISTAN
2012

Abstract

This paper analyses agglomeration behavior exhibited by the manufacturing firms in Punjab. It employs a unique data set to construct a distance based measure of agglomeration to test the existence of localization economies. The M function which is 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 labour pooling is used to examine whether firms that experience greater fluctuations in employment are likely to be more concentrated. The results provide evidence as to the importance of labour pooling working through the stated mechanism in explaining the high level of concentration within industries.

Contents

1. Introduction

The geographical concentration of economic activity is a widely observed phenomenon in most economies. It is a consequence of such clustering that certain locations evolve into becoming cities or business districts attracting significant concentration of population while other areas remain relatively less developed. The pertinent question as to why the activities of firms as well as individuals within an economy are spatially concentrated was raised by Fujita and Thisse (1996). The authors identified three distinct motives behind 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 inevitably causes firms to be located in close proximity to their competitors. The third explanation is that of externalities or positive spillovers and has received significant attention as a source of industrial agglomeration.

Externalities are an essential aspect of localization economies - the benefits to firms of locating near others in the same industry. Marshall (1920) highlighted three major sources of intra-industry gains that induce clustering of plants within a geographical boundary. These include the diffusion of information regarding effective processes and efficient production techniques; the ability to support the specialized production of inputs; and the benefits of sharing a similar labour mix. The productivity benefits owing to these externalities are believed to strongly influence the location decisions of firms (Ellison & Glaesar, 2010). 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 province. In particular, a measure of each industry's potential for labour pooling as constructed by Overman and Puga (2009) is used to examine whether firms that experience greater fluctuations in employment are likely to be more concentrated. It is argued that the advantages of labour pooling should be more pronounced for establishments that need to adjust their level of production and employment frequently and are able to do so without affecting wages.

The forces of industrial agglomeration have received a great deal of attention in the empirical literature for developed countries. There is however, a dearth of such studies for developing economies, owing mostly to data restrictions. This paper adds to our current understanding of the micro foundations of agglomeration economies by measuring the impact of various industry characteristics on the spatial concentration of each industry. The study makes use of the establishment level information reported in the Census of Manufacturing Industries for Punjab to measure the industry characteristics that reflect the presence of externalities. The presence of industrial clusters in this province has become a stylized fact but any empirical analysis is severely lacking. Given that spatial concentration boosts productivity, agglomerations appear to be an important area for research. It is pertinent 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.

1.1 Spatial Distribution of Industries in Punjab

The motivation for this paper is derived from an interest in investigating the dispersion of manufacturing activity in Punjab. It is a widely held belief that economic activity in this province is highly concentrated without serious conjecture as to why this might be so. The aim of this study is therefore to identify whether industrial concentration in Punjab can be explained by the presence of the Marshalian externalities and in particular by the phenomenon of labour pooling. An essential prerequisite in carrying out such a study is to employ a consistent measure for agglomeration. A widely used measure is the Ellison and Glaesar index which makes use of the district's share of total employment and the district's share of industry employment to determine whether the distribution of a particular industry's employment follows that of total employment. The Ellison and Glaesar index has been computed for manufacturing industries in Pakistan by Burki and Khan (2010) and shows there is considerable industry level variation in the extent of concentration exhibited by these firms. Recent research has found however, that a bias may arise from the use of any such measure that is based on a discrete framework and which relies on political boundaries. This research therefore makes use of an extensive and unique data set comprising of the spatial addresses of the manufacturing firms in Punjab to construct a measure for the level of agglomeration in each industry. The industry specific measure of agglomeration is computed using geographical distances between pairs of firms. Such a distance based framework has not been used so far in Pakistan and recent literature has declared it to be superior to 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 as to the extent of agglomeration exhibited by each industry. This measure is then regressed using an Ordinary Least Square estimation on the industry specific characteristics that measure the presence of the three externalities i.e. knowledge spillovers, input sharing and labour pooling. To prevent any bias resulting from endogeniety, the lagged values of industry characteristics have been used. As a preliminary contribution however, this paper provides evidence regarding the location pattern of industries within Punjab. It makes use of data containing the names and addresses of all manufacturing firms in the 32 districts of Punjab, to compute the geographical coordinates of these firms. This allows for a mapping of the firms within an industry and thus shows the cross industry variation in the level of agglomeration. The mapping of a few industries based on the 2 digit ISIC (International Standard Industrial Classification) is shown in Figure 1 (a – f) as evidence of the variation of the level of dispersion of firms within each industry.

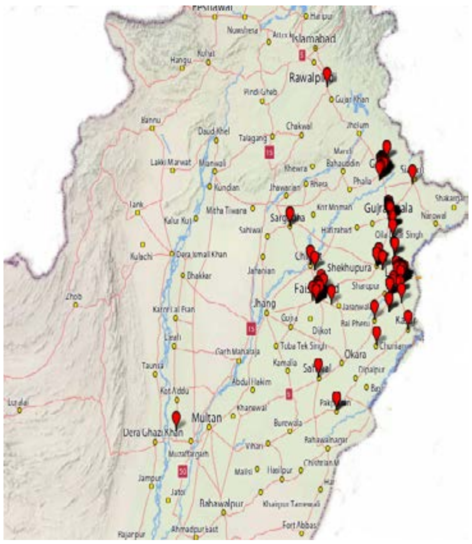

Fig 1a: ISIC 24 - Manufacture of Basic Metal

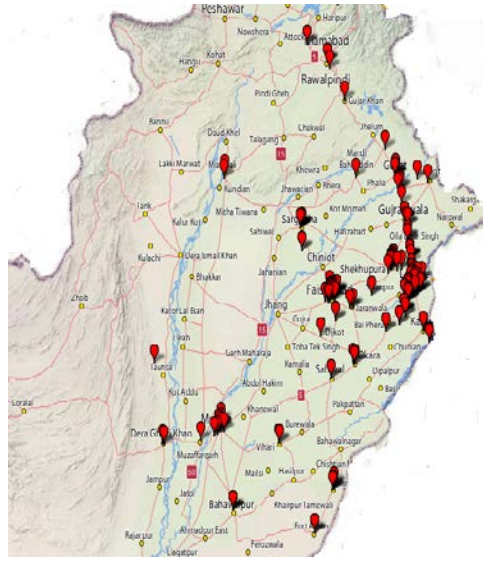

Fig 1b: ISIC 20 - Manufacture of Chemical and Chemical Products

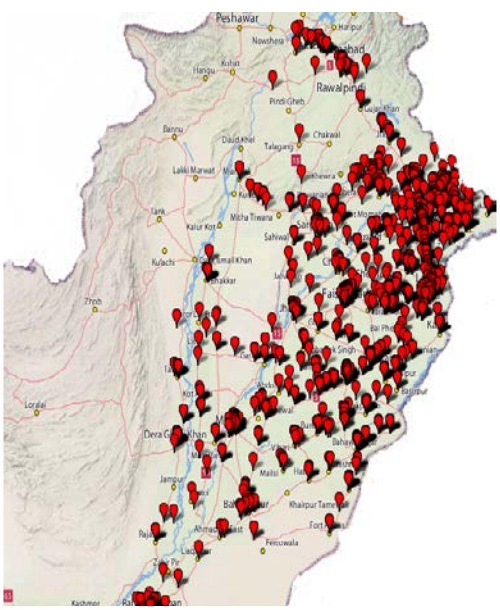

Fig 1c: ISIC 10 - Manufacture of Food Products

Fig 1d: ISIC 15 – Manufacture of leather Products

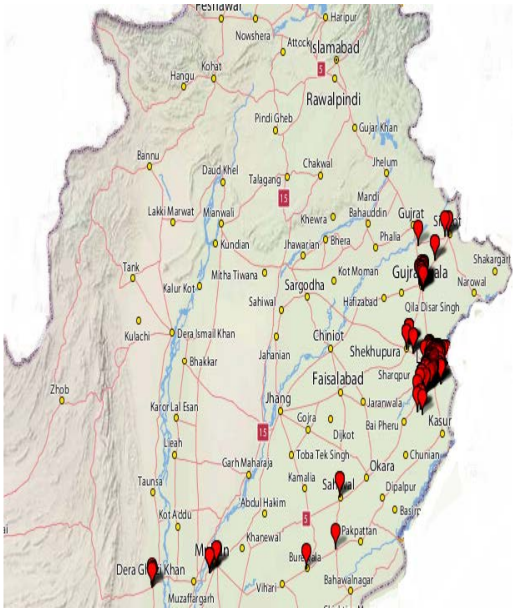

Fig 1e: ISIC 29 - Manufacture of Transport Equipment, motor vehicles and trailers

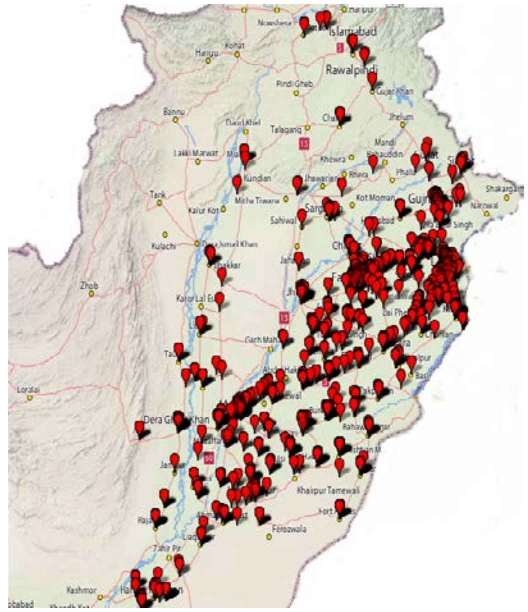

Fig 1f: ISIC 13 - Manufacture of Textiles

It is apparent that the degree of concentration/ dispersion of firms varies with the industry to which they belong. Figure 1c and 1f show that industries and firms engaged in the manufacture of Food products and Textiles are spread out across the province. Other industries such as those involved in the Manufacture of Basic Metal (Fig, 1a) seem relatively concentrated. The map depicts the location pattern of individual plants but for the purpose of empirical estimation carried out in this paper, concentration is measured in terms of the density of employment within an area.

The rest of the paper is organized as follows. A review of literature is presented in the next section followed by the theoretical framework for the measure of the potential for labour pooling within each industry. A discussion of the measure for agglomeration and the empirical estimation are presented in Section 5 and 6 respectively. Section 7 gives the description of the data. The results and their discussion are given in the subsequent two sections. The last section contains some concluding thoughts.

2. Literature Review

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

One of the earliest explanations of agglomeration economies was given by Marshall (1920). He discussed the various advantages to an industry of being geographically concentrated. The sources of these benefits, which subsequently came to be known as the Marshallian externalities; can be summed into three broad categories: input sharing, labour pooling and 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. Similar firms are believed to locate their production activities near each other because the positive spillovers from which they benefit outweigh other considerations such as those of transportation costs (Fujita & Thisse, 1996). There is also, considerable literature that focuses on the benefits of clustering attributed to the size of a geographical unit. The impact of these urbanization economies on productivity however, does not appear as strong as that of localization in explaining agglomerations (Rosenthal and Strange, 2003). Regardless, it cannot be said that the location decisions are independent of the considerations of infrastructure or market access. Indeed the natural characteristics of an area are often the strongest forces in the formation of clusters. Thus, studies focusing on such aspects have enough significance to merit a discussion in any paper focusing on industrial agglomeration.

The study by Ellison and Glaesar (1997) highlighted the fact that the level of agglomeration is not the same across industries. This indicated that the benefits derived from external economies are more pronounced for some industries than for others. The authors made a significant contribution to the literature on agglomeration by creating an index that measured geographical concentration. Up until then, the most common measure had been the Gini index which the authors criticized for producing misleading results due to the inability to account for differences in plant size. The index proposed in this paper accounted for the geographical distribution of employment and the intra industry allocation of employment to each establishment. Based on the results of their indices, Ellison and Glaesar concluded that the extent of localization exhibited by all industries is greater than that which could be accounted for by a random distribution of plants. Their measure of agglomeration became quite popular and was widely used in many subsequent studies. (Rosenthal & Strange,2001; Bertinelli & Decrop 2005; Ellison et al., 2010). However, there has now emerged, a new strand of literature which questions the underlying assumption of discrete spatial units in the computation of Ellison and Glaesar 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 that highlighted the shortcomings of using indices based on the assumption of discrete spatial units is that by Duranton and Overman (2005). In their study, an industry level measure of agglomeration based on a model of continuous space is derived using distribution of geographical distances between pairs of firms in an industry. While their measure reaffirmed that there is significant variation in the level of agglomeration across industries, its results differ from the findings of Ellison and Glaesar(1997) as to the extent of concentration exhibited by industries in the United Kingdom. While the findings of Duranton and Overman were accepted as the more accurate, the use of such a measure of concentration as suggested by the authors has generally been limited. This is primarily due to the extensive data requirements – in particular, the need for the exact spatial addresses or geographical coordinates of each firm.

The work of Marcon and Puech (2009) is one of the most recent research towards introducing a comprehensive measure of spatial concentration based on geo-distances. The M-function introduced by the authors can be used to identify both inter and intra industrial geographical concentration and fulfills the required criterion set forth by Duranton and Overman(2005) for what constitutes a sound measure of agglomeration. The paper presents a valid theoretical model and establishes the primacy of the 'M-function' over other distance based measures but the work is recent enough to not have seen any application on actual data.

There is significant literature devoted to measuring the level of agglomeration using different discrete as well as continuous models and to building a consensus regarding 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 result in clustering of firms across space. In particular, these studies have looked at the industry specific characteristics that would induce one industry to be more prone to agglomeration than the other.

Rosenthal & Strange (2001) examine the sources of differences in the extent of geographical concentration exhibited by manufacturing industries in the United States. Their paper on the determinants of agglomeration has direct significance for the current study which proposes a similar empirical estimation as that used by the authors. Rosenthal and Strange use proxies for the three prominent causes of localization: labour pooling, input sharing and knowledge spillovers and measure the impact that these have on agglomeration at different levels of geographical aggregation. The results of their analysis provide evidence to support all three theories but found that the labour pooling motive exerts the strongest influence on the firm's decision to concentrate its production activities near others within the same industry. The analysis used three measures of labour pooling: labour productivity, the ratio of management workers to production workers and the percentage of industry's workers with more than a Doctorate, Master's degree, and Bachelors degree. Their results are consistent with the findings of Dumais et al. (2002) that patterns of co-agglomeration are dominated by labour pooling considerations. The incentive to locate near establishments using a similar labour mix appears to be a strong force of agglomeration.

Further evidence to support the forces of agglomeration as pointed out by Marshall comes from Ellison et al. (2010). They construct co-agglomeration indices using the U.S. plant level employment data, and regress these on industry attributes such as the extent to which two industries buy and sell to each other, employ the same type of labour, and share similar technology. Input-output commodity level data was used, along with the data on occupations within each industry. The third covariate was measured using data on patent citations. Interestingly, the co-agglomeration indices are matched with the industry characteristics of establishments in the United Kingdom and this effectively eliminates the endogeneity bias inherent in the model.

As mentioned earlier, there are numerous studies that focus on the probability of a firm locating in a specific region based on the various characteristics of that area rather than emphasizing the localized, industry specific spillovers. Adopting such an approach Ellison and Glaesar (1999) looked at 'natural advantages' as the possible explanation of geographic concentration of employment. Their empirical analysis however is refined to incorporate the sensitivity of input cost based on the use of that input by each industry. Each industry's employment share in the state is shown to be affected by the potential advantages in natural advantages, labour and transportation costs. Other studies have measured the impact of geographical characteristics on firm level measures of profitability. Lall & Chakravorty (2005) used such an approach to identify the key factors in industrial location decisions. Rather than using profit measures however, their analysis involves estimating a cost function based on various economic geography variables (in addition to the input prices of capital, labour, energy and materials). These include market access (measured by indicators for accessibility and inter industry linkages), own industry employment in the district, buyer-supplier linkages and degree of economic diversity as measured by the Herfindahl[1] index. The elasticity of costs with respect to these variables was computed for each industry as well as for small, medium and large firms respectively. The importance of each of these variables differed by firm size and across industries. This paper presented strong support for urbanization economies of scale and showed that the most significant reduction in costs occurred as a consequence of locating in a region where there is industrial diversity.

In Pakistan, Burki and Khan (2010) have studied the trend in agglomeration exhibited by the manufacturing sector in Pakistan. Their analysis also employs the Ellison and Glaesar index

[1] The Herfindahl index is the sum of squared employment shares of all industries in a particular region.

and provides evidence that district characteristics and infrastructure are vital in explaining the high levels of concentration. In particular, the authors have looked at road density, the concentration of demand and the technical expertise of the workforce within each district to conclude that metropolitan cities attract the greatest levels of economic activity.

Deichmann et al. (2008) have also studied the determinants of industrial location in a developing country. They present a survey of a number of earlier empirical studies to highlight various factors that affect the profitability of activities in different locations. These include the prices/quality of inputs, market size and transport quality, access to labour, clustering effects (own industry concentration, supplier-customer linkages and industrial diversity), knowledge spillovers and regulation. While the direction of impact of these variables is consistent throughout the studies, the magnitude of influence is not the same across industries/countries. The empirical analysis by the authors, involves regressing the share of employment in each industry relative to the share of employment for that industry in the country as a whole, on various industry characteristics. Their results also show that labour pooling as measured by skill intensity (ratio of production to non production worker employment) exerts a strong influence on concentration of industry within a specific district.

In addition to the empirical evidence provided by the above studies, labour pooling had also received attention in theoretical literature as being an important source of agglomeration. Helsley and Strange (1990) for example, discusses the potential for better matching between jobs and skills in large labour markets as conducive to the formation of agglomeration economies. David and Rosenbloom (1990) developed a model that shows the interaction between labour pooling as a source of externality and industrial location. In particular, they focus on the "risk pooling" benefit owing to a large labour market. They argue that both firms and workers are able to minimize fluctuations in their earnings by locating near firms in the same industry but for which individual shocks are not positively correlated. Based on the significance of this particular source of externality, Overman and Puga (2009) examine its influence on the spatial concentration of manufacturing establishments in the United Kingdom. The study is unique in that it provides a direct measure of labour pooling rather than using indirect proxies of the extent to which workers possess industry specific skills. Following the model of Krugman (1991), they authors describe a model of labour pooling in which the individual establishments have the benefit of altering their level of employment in response to possible demand shocks that they might face. A measure of firm level employment shocks is constructed by computing the differences between the percentage change in the employment at the establishment and the percentage change in employment experienced by the industry. This measure is then averaged across all plants to obtain an industry level measure of the extent of volatility experienced by firms. It is argued that firms will have a tendency to agglomerate within the industry if they are prone to experiencing fluctuations in employment. Based on this measure and using the Ellison and Glaesar index of concentration, the authors tested the relative importance of labour pooling across industries in the United Kingdom. The empirical analysis supports their hypothesis that indeed, labour pooling as operating through such mechanisms do result in greater concentration of economic activity.

3. Theoretical Framework

The model presented in this section is closely related to the work of Krugman (1991). It builds the rationale for the study of labour pooling as a source of agglomeration economies by highlighting the microeconomic foundations through which this mechanism affects the location decisions of firms across industries. Furthermore, Overman and Puga (2009) have used it to develop their key hypothesis regarding the industry characteristics that explain the variations in agglomeration behavior.

3.1 Set up

We begin by considering a number of industries indexed $j = 1,\ldots, J$. In each industry there will be a discrete number of firms sub-indexed $i = 1, .., n$, each employing workers with a skill specific to that industry. Each firm operates under decreasing returns to scale and is prone to experiencing a productivity shock that is uncorrelated across firms i.e. it is specific to the establishment. The firm has the following production function:

$$y_i = [\beta + \epsilon_i]l_i - \frac{1}{2}\gamma {l_i}^2 \qquad (1)$$

The intensity of decreasing returns is denoted by γ and ϵ_i represents the firm specific idiosyncratic shock. The firm decides on a location at time period zero and in the subsequent period faces a productivity shock (ϵ_i). After observing this shock, the firm decides on the amount of labour it must hire from the available labour pool in the industry. Thus the profit π_i, of an individual establishment that chooses an employment level l_i is given by:

$$\pi_i = [\beta + \epsilon_i]l_i - \frac{1}{2}\gamma {l_i}^2 - \mathrm{w}l_\mathrm{i} \qquad (2)$$

3.2 Wages

It is assumed for the simplicity of the model that market wages are predetermined. Each firm after observing the shock, adjusts its labour such that the marginal value product equals the wage. The labour demand for firm i is:

$$l_i = \frac{\beta + \epsilon - w}{\gamma} \qquad (3)$$

We can use the labor market condition that $\sum_{i=1}^{n} l_h = L$ and aggregate the hours of labour across the industry in a given district to get the labour market clearing equilibrium:

$$\mathrm{L} = \sum_{i=1}^{N} (l_\mathrm{i}) = \frac{\beta + \mathrm{w} - \sum_{i=1}^{N} \epsilon_i}{\gamma} \qquad (4)$$

The market clearing wage can subsequently be derived from the above equation as:

$$w = \beta - \gamma \frac{L}{N} + \frac{1}{N} \sum_{i=1}^{N} \epsilon_i \qquad (5)$$

The expected wage is thus:

$$E(w) = \beta - \gamma \frac{L}{N} \qquad (6)$$

3.3 Profits

Substituting the firm’s labour demand into its profit equation:

$$\pi_i = \frac{[\beta - w + \epsilon_i]^2}{2\gamma} \qquad (7)$$

The profits are convex in the individual specific shock and in wages. The firm that experiences an idiosyncratic shock is required to adjust its production and employment levels. We can derive the firms expected profits using the above equation. Recall that the expected value of a square of a random variable equals the square of its mean plus its variance.

Taking expectations thus gives:

$$E(\pi_i) = \frac{[\beta - E(w)]^2 \ + var[\epsilon_i - \ w]}{2\gamma} \qquad (8)$$

Plugging in expected wage and simplifying:

$$E(\pi_i) = \frac{\gamma}{2}(\frac{L}{N})^2 \ + \ \frac{[var[\epsilon_i] + var\ (w) - 2cov[\epsilon_i, w]}{2\gamma} \qquad (9)$$

In the absence of any shock, the expected profits would be explained entirely by first expression on the right hand side. An increase in the ratio of workers per firm results in a lowering of expected wage and consequently raises expected profits. The labour pooling effect is captured by the latter term. It shows that the variance of the productivity shock as well as the variance of local wage work to increase expected profits whereas the covariance of the two decreases the expectation of profit. This is because when the firm needs to expand production after facing a positive shock, the expected profits will decline from a simultaneous rise in wages. Likewise, profits will fall in expectations if the wages are lower when the firm wishes to contract production. This provides the rationale that highlights the micro economics foundations of labour pooling as a source of agglomerations. Firms that are prone to fluctuations will not wish to be located in isolation from other establishments where productivity shocks get reflected in local wages.

An important assumption of this model is that shocks are identically distributed over [-ϵ, ϵ] with zero mean and variance equal to σ_s. Thus we can get from equation (5) the variance of market wages var (w) = cov[ϵi, w] = σ_s / N. We can simplify the equation for expected profits to yield:

$$E(\pi_i) = \frac{\gamma}{2}\ (\frac{L}{N})^2 + \left[1 - \frac{1}{N}\right]\frac{\sigma_S}{2\gamma} \qquad (10)$$

The greater the variance of the shock, the more pronounced becomes the labour pooling effect $[(1 - 1/N)\ \sigma_s / 2y]$. The expected profit depicted as in equation (10) will be equal for all firms in the same location and sector. The model predicts that the benefit of labour pooling is greater for firms that experience more individual specific shocks relative to the industry. The critical assumption regarding these shocks is that they should heterogeneous across firms. At a time when an individual establishment is faced with a positive shock and requires more labour, it must be that other firms in the industry are contracting employment so that wages would not be positively correlated with the shock. Industries will exhibit greater agglomeration if the firms in that industry face heterogeneous fluctuations in employment. The expected profits for such firms will be higher if they are agglomerated than if they were to locate in isolation because in the latter case, these shocks would be transmitted into local wages.

3.4. Equilibrium with Simultaneous Relocation by Firms and Workers

Overman and Puga (2009) presented a formalization of the risk sharing explanation of labour pooling by defining a situation of urban equilibrium. Production and location decisions are modeled in a dynamic setup. In the first stage, firms and an exogenously determined number of workers choose to locate in a particular area. The equilibrium in the first period is such that wages are equalized across locations and neither the worker nor the establishment has any incentive to deviate i.e. locate elsewhere. The ratio of workers per establishment is the same across locations. Moreover, relocation by an individual worker has no effect on wages or profits; an individual firm however can influence wages at the origin and the place of destination if it decides to relocate.

In the second time period, each firm experiences a productivity shock ϵ_i. The firm is thus faced with the prospect of deviating profitably. There are two aspects that it will consider. If the firm decides to relocate, the ratio of workers to establishment will decrease in the destination location. This will increase the expected wage and put a downward pressure on expected profits. The first term in equation (9) explains this *labour market tightness* effect. On the other hand, if the destination has a large number of firms, the local wages will not be greatly affected by the firm's productivity shock. This allows the firm to adjust employment shocks and obtain higher expected profits as explained by the *labour pooling* effect in the preceding discussion.

The *labour market tightness* effect favours dispersion of economic activity but with a greater variance in firm specific shock, the advantages of *labour pooling* become more pronounced for the specific industry to which such firms belong. This is because firms will be able to make frequent adjustments in employment without having the impact of shock reflected in local wages. The equilibrium in the model occurs when workers and establishments are allocated such that each location has at least $\sigma_s / 2(y^2R^2 + \sigma_s)$ as many establishments as those in the district where the industry is most concentrated. The industries most prone to heterogeneous productivity shocks will tend to be most agglomerated in order to have higher expected profits.

4. Measure of an Industry's Potential for Labour Pooling

The model presented in the previous section highlights one particular mechanism through which labour pooling might result in agglomeration behavior. Firms in an industry will have higher expected profits by being spatially concentrated because fluctuations in employment in case of productivity shocks will not be reflected in wages. Thus the potential for labour pooling can be measured by the extent of heterogeneous shocks faced by an industry. This measure for labour pooling affects agglomeration behavior through a risk sharing channel.

There are in addition, other theoretical arguments that highlight the role of a large labour market as a source of positive externality. These include the improved prospects for firms to find the right quality worker for a job, just as workers are able to find better matching of their skills Helsley and Strange (1990). Moreover, a large market promotes the specialization of labour which can generate increasing returns for firms that agglomerate geographically. (Duraton, 1998). For the purpose of this study however, the measure for labour pooling is based on the ability of firms to adjust employment levels in response to productivity shocks.

Earlier studies that have looked to measure the impact of labour pooling as a source of positive externalities have often relied on indirect proxies. This is because it is difficult to measure the extent to which workers might possess industry specific skills. By limiting the analysis of the role of labour pooling to that discussed in the theoretically framework section, it is possible to compute a measure of employment shocks at the firm level which can be averaged to reflect the industry's tendency to experience heterogeneous shocks.

The advantage of locating near other establishments in the industry arises from the need of an individual firm to adjust its level of employment at a time when other firms are not faced with similar productivity shocks. The wage level should therefore remains 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 fluctuations. This in turn reflects the potential for labour pooling within each industry.

Percentage change in firm's employment:

$$\Delta E_i = \frac{E_2 - E_1}{E_1} * 100$$

Percentage change in industry's employment:

$$\Delta E_j = \frac{E_2 - E_1}{E_1} * 100$$

$$Industry\ level\ fluctutaions = \frac{\sum_i^n |(\Delta E_i - \Delta E_j)|}{n} \quad \text{(i)}$$

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

5. Measure of Geographical Concentration of Industries

The importance of agglomeration varies with the type of industry. High tech sectors for example, may exhibit a higher degree of geographical concentration because of a greater potential for knowledge spillovers. Most studies on agglomeration have relied on employment shares to measure the density of economic activity. Recent literature however has criticized the crude measures that are used 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 (2002) have argued that any index for spatial concentration should satisfy a number of requirements. The employed measure should allow a comparison of geographical concentration across industries and also should control for the overall aggregation across industries. It must also distinguish and separate industrial concentration from spatial concentration. The index that fulfills these requirements and that has consequently been used in a large number of studies is the Ellison and Glaesar (1997) measure of agglomeration. There is however an additional requirement that remains unfulfilled by the use of this and other such agglomeration indices that rely on the assumption of discrete spatial units. A consistent measure must produce results that are unbiased with respect to the degree of spatial aggregation. The latter condition is violated if we make the assumption 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 in the agglomeration literature, recent studies have moved away from employing cluster based methods and have placed an emphasis on 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 however, are quite extensive and thus their application in the literature on economic geography is rather limited.

The measure for localization employed in this study is the ‘M function’ proposed by Marcon and Puech(2009). It makes use of the average number of neighbours of plants 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’ to 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 from each other.

The calculation of this measure is such that 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 is computed. This is then averaged across the industry and is compared to the ratio of all firms in industry ‘i’ over the total number of firms in Punjab. To control for the concentration within plants and accounting for the size of each firm, it is feasible to attach weights and use the number of employees rather than the number of firms. Thus, for each industry, the average proportion of employees for that industry within a radius ‘r’ can be obtained. 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 gives an industry level measure for the intra-industrial geographical concentration.

$$M_i(r) = \left. \frac{1}{N_i} \sum_{s=1}^{N_i} \frac{\sum_{n=1,s\neq n}^{N_i} c_i(s,n,r)w_n}{\sum_{n=1,s\neq n}^{N} c(s,n,r)w_n} \middle/ \sum_{s=1}^{N_i} \frac{Wi - w_s}{W - w_s} \right.$$

In the above term[2], the numerator depicts the share of industry ‘i’ in the 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 there is relatively a greater proportion of employees near firms belonging to industry ‘i’ 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. Thus, this measure is comprehensive and fulfills the five criterions put forth by Duranton and Overman (2002).

6. Empirical Estimation

The empirical estimation involves measuring the effect of agglomerative externalities on the degree of spatial concentration exhibited by various industries. In particular, we wish to observe the impact on agglomeration of the potential for labour pooling within an industry. A linear model can thus be estimated using the Ordinary Least Squares which is the estimation technique conventionally used in literature. The first equation to be estimated is as follows:

[2] For a detailed explanation, refer to the original work of Marcon and Puech (2009).

$$M_i = \beta o + \beta L_j + \beta \mathbf{X}_j + \beta Z_j + e_j \qquad \text{(iii)}$$

M_i in the above equation is the M function that measures the extent of geographical concentration of an industry within a certain radius. For this study, the M function has been computed for a2 mile and a 10 mile radius. It is to be regressed on a number of industry characteristics: L_j denotes our measure of the potential for labour pooling in the industry, $\boldsymbol{X}$ 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 is added to control for transport costs in an industry and is denoted by Z.

The regression equation specified above is close to that employed by Rosenthal and Strange (2001) in their investigation of the various sources of agglomeration. The primary hypothesis that is to be tested is that those establishments which are more prone to experiencing productivity shocks will exhibit greater concentration because they have a need to adjust employment levels without affecting their profitability. It is thus stipulated that the potential for labour pooling will differ for each industry based on the firm specific fluctuations that are experienced.

While our key interest is to the measure the impact of labour pooling as a source of agglomeration, the significance of the remaining Marshallian externalities cannot be undermined. The industries sensitive to costs of input and knowledge sharing are more likely to exhibit agglomeration behavior. These have been added as controls in the proposed model. As regards to the proxy for input sharing, it is suggested that sharing the 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 exhibit a tendency to cluster near that plant which is the source of their inputs. Accordingly, agglomeration should be positively related to this measure of input sharing. The other externality, that of knowledge sharing, is relatively difficult to measure given the existing data constraints. This externality is generally measured using the R&D expenditure by firms 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 which import spare parts are in some way engaged in adding to the existing knowledge of 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 (iii) measures transport costs. Transport costs are a significant determinant of the level of agglomeration or dispersion of economic activity. This fact has received much attention in the theoretical literature (Fujita & Thisse, 1996). High transport costs of supplying to the consumer market will deter agglomeration in manufacturing industries. The CMI provides information regarding 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 is required to be more frequently transported to the market, firms manufacturing the good will locate closer to the consumer and hence be relatively dispersed. A negative sign is thus hypothesized.

6.1 Hypothesis:

The primary hypothesis that we wish to test is that firms that experience greater heterogeneous productivity shocks have a greater potential to benefit from labour pooling within the industry and are consequently are more likely to be spatially concentrated.

7. Data

There are two distinct objectives of this paper. The first of these is the computation of 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, the geographical coordinates of all firms are required so to compute the distances between pairs of firms. This data is extracted from the Directory of industries which contains in addition to other information; the names, addresses and employment of over 17000 plants operating in Punjab. The most recent data is available for the year 2010 and it has been used to compute the geographical coordinates for independent firms which then allowed us to compute the distances between all firms, thus allowing the computation of the M function. These coordinates have also been used to include the geographical mapping of industries presented earlier. Table 1a is a brief summary of the M function computed for the two distances.

Table 1a

Summary Measures of Agglomeration among Manufacturing Industries

M function	Mean	SD	Min	Max	Correlation
10 miles	1.656979	1.332022	0	5.960157	1.000
2 miles	1.137627	0.4069265	0	2.852808	0.4687

The second objective entails the application of this measure for agglomeration. The data source employed to investigate the relationship between labour pooling and spatial concentration is the Census of Manufacturing Industries (CMI) for the province of Punjab. It provides plant level data on the quantities and value of inputs and outputs as well as details of employment and employment costs for 3528 establishments. Table 1 provides selected summary statistics of the explanatory variables that are used to measure the presence of externalities. A limitation arises in constructing the measure for labour pooling for which we require the percentage change in employment at the firm level. While the census is carried out for 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 the all firms. Another limitation is that there is no information on the expenditure on Research and Development (R&D). The expenditure on imported spare parts has consequently been used as a proxy for the presence of knowledge spillovers in this paper. The other controls for industry characteristics have been computed using the input and output data reported in the Census.

Table 1b

4 digit Manufacturing Industries: Definitions and selected summary statistics for key Explanatory Variables

Variable Name	Definition	No. of Obs.	Mean	SD	Min	Max
Labour Pooling (plant to industry)	The difference between the percentage change in plant's employment and the percentage change in industry employment (idiosyncratic shock)	3505	0.196	0.247	0.006	3.407
Labour Pooling (plant to total)	Idiosyncratic shock to a firm relative to manufacturing sector as a whole	3505	0.236	0.281	0.02	1.363
Labour Pooling (industry to total)	Idiosyncratic shock experienced by an industry relative to manufacturing sector as a whole	3505	0.129	0.244	0.001	3.046
Input Sharing	Share of manufactured inputs as a cost of total inputs	3528	0.141	0.124	0.007	0.62
Knowledge Spillovers	Share of expenditure on imported spare parts	3528	9E-04	0.003	0	0.056

To carry out the empirical estimation discussed in the previous section, the observations in the Directory of Industries were to be matched with those in the Census of Manufacturing. The two data sets were thus coded on the basis of the 4 digit International Standard Industrial Classification so to allow an industry level analysis. Additionally, since the CMI includes only medium sized and large scale firms with a minimum of 10 employees, thus in computing the 'M-function 9577 firms from the Directory of Industries have been included while the rest of the observations have been dropped. Lastly, the most recent Census of Manufacturing was available for the fiscal year June 2005 – June 2006 but the measure for agglomeration is computed using data from 2010. The rationale for this lies in the observation made by Overman and Puga (2009) that a lag between characteristics and outcome allow for a more sound economic interpretation. In addition to this, the lag prevents the possible problem of endogeneity between certain industry characteristics and the measure of intra-industry geographical concentration.

8. Results

The M function that is the distance based measure of agglomeration employed in this study has been computed for a 10 mile and a 2 mile radius. The results depicting the employment density of industries classified at the 4digit ISIC are somewhat similar within the two distances. Table 2a provides a list of the most agglomerated and the least agglomerated industries in the province. The most agglomerated firms exist within the industries that manufacture fabricated metal products (2 digit ISIC 25) and those involved in the manufacturing of non metallic mineral products (2 digit ISIC 23). Based on the 4 digit industrial classification, it appears firms that manufacture sports goods and bodies for motor vehicles appear to be highly concentrated. The manufacture of the machinery for textile and leather is also a concentrated activity within the province. Among the least agglomerated industries are those engaged in the production of tobacco and in the manufacturing of weapons and ammunition. The M function attaches a value of zero to such industries because these plants do not have neighboring firms within a 2 mile or at10 mile distance.

Table 2: Geographical Concentration of 4 digit industries in Punjab – 2010

ISIC	Description	M 2miles	ISIC	Description	M 10miles
	10 most concentrated industries				
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 bodies for motor vehicles	5.77063
2394	Manufacture of cement, lime and plaster	2.4151	2393	Manufacture of porcelain and ceramic products	4.987743
1311	Preparation of spinning of textile fibres	1.3258	2593	Manufacture of cutlery, hand tools & general hardware	4.628244
3230	Manufacture of sports goods	1.2887	2815	Man. 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 bodies for motor vehicles	1.2730	2394	Manufacture of cement, lime and plaster	3.444931
2826	Man. Of machinery for textile, apparel, & leather production	1.2724	3230	Manufacture of sports goods	3.40682
2511	Manufacture of structural metal products	1.2615	2826	Man. Of machinery for textile, apparel, & leather production	3.357219
2599	Manufacture of other fabricated metal n.e.c.	1.2595	3250	Manufacture of medical and dental instruments	3.239103
	10 least concentrated industries				
1200	Manufacture of Tobacco products	0	1200	Manufacture of Tobacco products	0
2520	Manufacture of weapons and amunition	0	2520	Manufacture of weapons and ammunition	0
2813	Manufacture of other pumps, compressors, taps & valves	0.3153	2029	manufacture of other chemical products n.e.c.	0.106498
2310	Manufacture of glass and glass products	0.5757	1311	Preparation of 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	weaving of textiles	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	Man. of soft drinks, mineral waters & other bottled waters	0.532559
1312	weaving of textiles	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 feeds	0.601087

Firms that are involved in the manufacturing of food products, leather products and consumer items such as lubricants, detergents, cosmetics and perfumes (ISIC 20: Manufacture of chemical and chemical products) are found to be dispersed. These findings are consistent with the findings of Burki and Khan[3] (2010) and support the conjecture that theleast 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 firms.

A more empirical explanation for the variation in industrial concentration as depicted by the M function may lie in the Marshallian externalities. The results of regressions based on the Ordinary Least Squares are presented in Table 3a and 3b. The specification is such that the measure of agglomeration is regressed on the lagged values of industry characteristics. As observed by Overman and Puga, this gives a lag from characteristics to outcomes such that firms are able to observe industry characteristics prior to making a decision regarding location. The primary reason for taking lagged values, however, was to avoid possible problems of endogeneity.

The measure for labour pooling has been computed as idiosyncratic shocks relative to the industry (column 1) and then changing the reference category to manufacturing sector as a whole (column 2 & 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 is that firms that experience frequent changes in employment, wish to locate closer to other firms in the same industry so that wages and in turn the firms profitability is not affected. The impact of labour pooling is also significant for industries that experience shocks relative to the changes in the manufacturing sector as a whole (column 3). It can be stipulated that such industries would be geographically concentrated but might also choose to locate in more urbanized areas where they can benefit from labour pooling across industries.

[3] Burki and Khan (2010) have used the Census of Manufacturing for the year 2006 to measure agglomeration using the Ellison and Glaesar index. The results of the M function in this study are based on data from the Directory of Industries 2010.

Table 3a. OLS Regression Dependant variable: M function - 10miles			
	(1)	(2)	(3)
Labour Pooling (plant to industry)	1.69 *		
	(0.96)		
Labour Pooling (plant to manufacturing)		1.78	
		(4.06)	
Labour pooling (industry to manufacturing)			1.94**
			(0.86)
Input sharing	21.00*	21.08*	21.03*
	(11.90)	(11.85)	(11.87)
Knowledge spillover	-57.97	-59.78	-49.96
	(162.21)	(162.87)	(164.06)
Transport cost	-122.46*	-125.76*	-121.77
	(72.79)	(71.50)	(73.78)
Constant	9.22***	9.27***	9.29***
	(2.24)	(2.24)	(2.28)
R^2	0.10	0.09	0.10
n=70			

Robust Standard Errors in paranthesis
****,** and * indicate statistical significance at 1%, 5% and 10% respectively*

The employment data for firms within the most concentrated industries show that agglomeration of these firms exists largely in the industrialized districts of Gujranwala and Sialkot. Since the regression estimates are based on the the 4-digit classification, it is possible that the impact of labour pooling (firm to industry) is undermined due to labour 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 to this, we have accounted for the transport cost of industries which 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 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 labour pooling.The exception is column (2) in Table 3b. These results which include the measure for labour 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 measure of labour pooling itself.

Table 3b. OLS Regression
Dependant variable: M function - 2miles

	(1)	(2)	(3)
Labour Pooling (plant to industry)	0.52***		
	(0.06)		
Labour Pooling (plant to manufacturing)		0.82*	
		(0.41)	
Labour pooling (industry to manufacturing)			0.63***
			(0.05)
Input sharing	0. 68*	0.71	0.79**
	(0.39)	(0.44)	(0.36)
Knowledge spillover	12.67	1.99	13.71
	(13.98)	(15.42)	(14.66)
Transport cost	-0.86*	-0.73	-0.84*
	(0.49)	(0.53)	(0.50)
Constant	0.95***	0.94***	0.98***
	(0.09)	(0.11)	0.08
R^2	0.45	0.29	0.48
n=70			

Robust Standard Errors in paranthesis
****,** and * indicate statistical significance at 1%, 5% and 10% respectively*

Excluding these results, firms that employ manufactured inputs in their production tend to be more closely located, most likely near the supplier of these manufactured inputs. It may be however that within the rather small radius of 2 miles, the effect is undermined because other factors take precedence. The last Marshallian externality that we have tested for is that of knowledge spillovers and it is not found to be significant in any of the regressions.

Table 4a. 2digit Fixed effects

Dependant variable: M function - 10 miles

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

Robust Standard Errors in paranthesis
****,** and * indicate statistical significance at 1%, 5% and 10% respectively*

The proxy for this externality was based on the expenditure on imported spare parts. The agglomeration literature commonly measures knowledge spillovers using the expenditure on R and D. Such data was not available for the manufacturing industries in Pakistan and therefore a more indirect proxy was opted for. 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 existing literature, transport costs are an important determinant of agglomeration[4]. The effect of transport costs on agglomeration might however be stronger higher levels of geography than those accounted for in this analysis.

The above results substantiate that the presence of the potential labour pooling in an industry has an important impact on the agglomerative behavior shown by firms in a industry. The R squared in each of the Ordinary Least Square regressions presented in 2a and 2b however are low thereby indicating that the variation in agglomeration across industries is not sufficiently explained by the externalities accounted for in this analysis. There can be other, industry specific characteristics that could result in a cross industry variation in the level of agglomeration. There are a number of demand side factors specific to the industry that could affect the location decision of firms as to whether they are dispersed or agglomerated. To address this concern, we have changed the specification to include industry fixed affects. These fixed effects are based on two digit International Industrial Classification. The results are shown in table 4a and 4b for the 10 and the 2 mile radius respectively.

The evidence that the two Marshalian externalities i.e labour 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 labour pooling as measured by the idiosyncratic shocks to firms relative to manufacturing sector are insignificant for the larger radius of 10 miles. This may be because individual plants faced with productivity shocks might not be able to attract labour from different sectors if it requires greater mobility on the part of the workers. Hence while such a measure of

[4] Natural advantage was a control variable included in the original study by Rosenthal and Strange but has been omitted in these regressions because of a strong correlation between the proxy for this variable and that of Input sharing.

labour pooling is positively associated with agglomeration for a smaller distance, it becomes insignificant as the distance is increased.

Table 4b. 2 digit Fixed Effects
Dependant variable: M function - 2 miles

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

Robust Standard Errors in paranthesis
****,** and * indicate statistical significance at 1%, 5% and 10% respectively*

Transport costs become insignificant in all regressions at the 10 mile radius once the fixed effects are incorporated. On the other hand, it is highly significant at the lower level of geography i.e. at 2 miles. The results shown in table 3b provide evidence that the higher the cost of transport involved in making the good available to the consumer, the lesser will be the level of geographical concentration of an industry. This supports the argument presented earlier that firms will be dispersed if the demand for their goods is well spread out.

9. Conclusion:

The paper aims to add to the current understanding of the micro foundations of agglomeration economies in developing countries. It attempts to identify the factors that can account for the concentration of industries beyond that which can be explained by the area characteristics. The computation of a distance based measure and its application is a contribution unique to this study and allows for greater accuracy in accessing the extent of concentration exhibited by manufacturing firms. The use of such a measure can be of particular use to researchers aiming to study the impact of industrial concentration on growth or productivity. The distance based measure allows for an assessment of the impact of agglomeration over various levels of geography without encountering the bias associated with political state boundaries.

The current paper provides evidence that there are certain industry specific characteristics because of which manufacturing firms opt to locate near other establishments in the same industry. These characteristics were highlighted by Marshall(1920) as the positive externalities that allow firms within the same industry to experience cost advantages if they choose to locate in close proximity to other firms. Two particular mechanisms through which these cost advantages can occur have been identified in this paper. The first being the potential for labour 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 other is that of input sharing whereby firms can have the advantage of attracting the production of specialized inputs. An important implication that stems from such findings is that policy makers can facilitate economic activity in the less developed areas through cost incentives that match the benefits arising from such externalities.

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

References

Burki, A.A, & Khan, M.A. (2010) Spatial Inequality and Geographic Concentration of Manufacturing Industries in Pakistan. Pakistan Society of Development Economists (PIDE), 27th Annual General Meeting and Conference.

Bertinelli, L., & Decrop, J. (2005). Geographical Agglomeration: Ellison and Glaesar's index applied to the case of Belgian manufacturing industry. *The Journal of Regional Studies Association*. Vol. 39(5), 567 – 583.

David, P. A. & Rosenbloom, J. L. (1990). Marshallian Factor Market Externalities and the Dynamics of Industrial Localization. *Journal of Urban Economics.*Vol. 28, 349-370.

Deichmann, U., Lall, S.V., Redding, S. J., & Venables, A. J. (2008). Industrial Location in Developing Countries. *World Bank Research Observer*. Vol. 23(2), 219-246.

Dumais, G., Ellison, G. & Glaesar, E. L. (2002). Geographic Concentration as a Dynamic Process, *Review of Economics and Statistics*, Vol. 84(2), 193-204.

Duranton, G. (1998). Labor Specialization, Transport Costs, and City Size. *Journal of Regional Science*. Vol. 38(4), 553 – 573.

Duranton, G. & Overman, H.G. (2005). Testing for Localization using Micro-Geographic Data. *Review of Economic Studies*, Vol 72, 1077-1106.

Ellison, G., Glaeser, E. L., & Kerr, W. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*. Vol. 100(3), 1195-1213.

Ellison, G., &Glaeser, E. L., (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?. *The American Economic Review*, Vol. 89 (2), 311-316

Ellison, G. & Glaeser, E. L. (1997). Geographic concentration in U.S. manufacturing Industries: A dartboard approach. *Journal of Political Economy*. Vol. 105(5), 889- 927.

Fujita, M. & Thisse, J.F. (1996). Economics of Agglomeration. *Journal of the Japanese and International Economies*. Vol. 10, 339-378.

Hensley, R. W. & Strange W. C. (1990). Matching and Agglomeration Economies in a System of Cities. *Regional Science and Urban Economics*. Vol. 20, 189-212.

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

Lall, S. V. & Chakravorty, S. (2005). Industrial location and Spatial Inequality: Theory and Evidence from India. *Review of Development Economics*. Vol. 9, 47 – 68.

Marcon, E. & Puech, F. (2009). Measures of the Geographical concentration of Industries: Improving distance based methods. *Journal of Economic Geography*. Vol 10(5), 745-762.

Marshall, A. (1920). Principles of Economics. London, U.K.: MacMillan and Co.

Overman, H. G. & Puga, D. (2009). Labour pooling as a source of agglomeration: An empirical investigation. CEPR Discussion Paper No. DP7174.

Rosenthal, S.S. & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*. Vol. 50(2), 191–229.

Rosenthal, S.S. & Strange, W. C. (2003). Geography, Industrial location, and Agglomeration. *Review of Economics and Statistics*. Vol. 85(2), 377-393.

Rosenthal, S.S. & Strange, W. C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. *Handbook of Urban And Regional Economics*. Vol. 4, 2119-2171.