Title: Is Labour Pooling Foundation for Agglomeration? The Nexus between Agglomeration and Workers' skills

Authors:

1. **Ms. Farah Atiq (Corresponding Author)**
M.Phil. Scholar
Applied Economics Research Centre
University of Karachi, Pakistan.
Email: farah.atiq95@gmail.com

2. **Dr. Ambreen Fatima**
Associate Professor
Applied Economics Research Centre
University of Karachi, Pakistan.

3. **Dr. Faisal Sultan Qadri**
Assistant Professor
Applied Economics Research Centre
University of Karachi, Pakistan.

Abstract

Industrial Agglomeration induce agglomeration economies and offers cost advantages /productivity gains to firms and labours involve. Literature so far suggests three founding factors for agglomeration i.e., labour pooling, knowledge spillover and input sharing. The labour pooling is considered related to the productivity gain of firm through utilization of labour market. In this research work, two types of labour pooling measures have been adopted. First adopted measure of labour pooling is built on firm to industry pooling, while the second adopted measure focuses on educational attainment of labour involved. Moreover, other relevant variables such as knowledge spillover and percentage of different skill levels and multiple controlling variables are also considered to verify their relevance as determinant of agglomeration. Lastly, to compute the dependent variable, location quotient is used. For data, manufacturing sector data of 21 industries in 14 cities is extracted from Labor Force Survey of Pakistan of 2014-15 and 2017-18. For empirical analysis, due to lack of longitudinal data Pseudo-Panel is formed and fixed-effect econometric technique is employed. As far results are concerned, it shows positive and significant coefficients of firm to industry pooling, percentage of tertiary level education, low-skill workers with on-the-job- training, percentage of low-skilled and skilled workers. It indicates all these factors contribute as founding factors of industrial agglomeration. As for recommendation, policymakers can boost up economic activity and utilize large labour market in the less privileged areas and form urban centre by providing cost incentives that matches agglomeration economies gains.

Keywords: Industrial Agglomeration, Labour Pooling, Urban Centre, Knowledge Spillover, Manufacturing Sector.

1. Introduction:

The geographic concentration of firms and labour brings immense productivity gain. Many empirical pieces of evidence reassure this stylized fact by quantifying them. The evaluation of productivity gain resulting from the doubling of agglomeration is lying between 2 to 8 percent, this factor is conditional on the estimation technique and sector under analysis (Rosenthal and Strange, 2004; Combes et al. 2008). There are many studies in the existing literature that provides sound theoretical micro-foundations for agglomerating economies. However, the main difficulty that arises with these theories is, they all predict the increase in productivity as spatial concentration rise, but the process of that mechanism is hard to trace. When firms agglomerate on a particular location, gradually that location evolves into industrial zones then cities or economic districts, it pulls the significant quantity of population than other locations. The concise question in this regard was raised by Fujita and Thisse (1996) i.e., why the firms or industries concentrate spatially. In their work, Fujita and Thisse (1996) identified three factors that cause the spatial concentration of firms or industries. First, firms avoid dispersing their plant-level activities to attain increasing returns to scale. Second, competition on a spatial level for market area always results in firms locating near to each other. Third, externality (positive spillover) also works as a source of spatial concentration. Furthermore, though there are many potential sources of industrial agglomeration, Alfred Marshall (1890) particularly highlighted one of the sources as labor market pooling. Though, labor market pooling as a foundation for agglomeration can be interpreted in several ways, Marshall stressed a specific one, that is *"a localized industry gain a great advantage from the fact that it offers a constant market for skill"* (Marshall 1890, 271).

So far, the evidence of industrial agglomeration has explored for various countries, such as Brazil (Sobrinho and Azzoni, 2016) and UK (Overman and Puga, 2010). It can be assumed that the spread pattern of industries can be a triggering factor for the growth in an economy, so far the number of studies have shown the effect of concentration of industries. There are several works that include the construction of spatial concentration indices such as Ellison and Glaeser (1997), Maurel and Sédillot (1999), and Duranton and Overman (2005). Meanwhile, there are also studies that show the empirical support for the factor behind the concentration of production agents (can be industries), these studies include the work of Rosenthal and Strange (2001), Resende and Wyllie (2004), Ellison et al. (2010), and Burki and Khan (2013). Though, the mentioned studies have not succeeded in differentiating the causing factor of productivity gains resulting from industrial agglomeration (Overman and Puga, 2010). Meanwhile, there are studies that provide the fundamental micro-based such as the works of Krugman (1991) Fujita and Thisse (2013) and Duranton and Puga (2004), but from the perspective of empirical construct, differentiating among these various mechanisms is complicated. The cause of this complexity is basically the result of those theories' anticipation about the productivity gain from agglomeration but that is hard to trace down with actual data (Duranton and Puga, 2004; Overman and Puga, 2010). From all these mechanisms, specifically, the labour pooling can be evident in Marshall's (2009) work, as an attracting force for industries to concentrate. The main pulling force is these mechanisms are advantages that firms can obtain when they concentrate on a particular location that has an abundant labour (Marshall, 2009). Further elaboration of this claim was given by Krugman (1991), in his

work, namely “labor market pooling model”. According to it, the adjustment of production level and employment by the establishment in response to idiosyncratic productivity shocks bring gains in case of labour pooling.

The theoretical section of this study used the microeconomic fundamentals of labor pooling as the foundation for agglomerating economies the version specified by Krugman (1991). It takes a sequence of sectors in which firms go through from idiosyncratic shocks. Profits are convex in the shocks that are firm-specific, since every firm manages its production along with its employment as a result of shock. Though, the change in the level of employment affects the wages locally, and more disperse the firms in the location from another (especially same sector or same), the intensive the effect. If local wages increase with the expansion in firm production, resulting from positive idiosyncratic shock, and decrease with the contraction of firm production as a result of negative idiosyncratic shock, these responses limit the firm’s ability to adjust its employment. Hence, firms which inclined to experience considerable fluctuations in their employment level as compared to other firms, find it more favorable to locate where there are more workers with similar skills. As a result, this model anticipates that sector whose firms go through more idiosyncratic fluctuations will tend to be more spatially concentrated.

The factors which determined industrial agglomeration has largely explored in the empirical literature for advanced countries. Though, in developing countries, such studies still need some work mostly due to limited and inconsistent data availability. This thesis has used data for Pakistan and specifically has attempted to answer one of the fundamental questions i.e., is labour pooling foundation for agglomeration?

The next two subsections of this intro explore the actual concept building of Agglomeration Economies, Industrial Agglomeration and Labour Pooling.

1.1 Agglomeration Economies and Industrial Agglomeration

The geographic nearness of people and firm forms cities and industrial clustering /agglomerations respectively. Moreover, the advantages or benefits that this spatial nearness brings are called Agglomeration Economies. To have a better understanding of where these ideas originate from, this section elaborates the origin of important terminologies in this study.

External economies refer to the conditions when a rise in the scale of the urban environment aids productivity, generally. There are three types of externality scope. First is the industrial scope, it is the degree to which agglomeration economies are spread-out across industries, both within and across all industries in a region. These two distinctions are also known by their separate term, for instance, economies of scale that results from a surge in the activity’s geographic concentration within a given industry is also called the localization. Whereas the externalities that result due to concentration of economic activity overall, are called urbanization economies. Hence, the studies that have conducted so far provide evidence that when economic agents come together in an industrial space, it gives more potential for externalities to arise. The second is the geographic scope. Due to the popularity of the idea of the city, almost all literature, whether it is books or research studies tries to explain why cities exist. The answer to this phenomenon is simple that nearness brings benefits. The component of geographic

proximity that this geographic scope is based on is the degree of agglomeration economies that depend on geographic distance. Therefore, it suggests if agents or economic agents are closer then there is more chance of interaction among them, and eventually more possibility of externality to arise. Lastly, the third scope is the temporal scope. As it clear from the term, the temporal scope is related to time. For instance, the interaction of agents in the past continues to have an effect on the productivity of one or all agents in present. Some of the examples of it include the learning which may take place in the past, but it could continue to affect present also, the possibility of the location of the supply chain. Though the learning can deteriorate over time it is static. But there is way more possibility that agents at different temporal states can continue to affect the other agencies like the other geographic and industrial scope interact which defines the temporal scope of agglomeration economies.

As explained above agglomeration economies has three dimensions. Moreover, the agglomeration economies, we refer to in this study will be related to industrial agglomeration. Though this work focuses on the industrial scope (as explained above), this study analyzes the industrial agglomeration which is the spatial concentration of industries, but the geographic and temporal scope will remain relevant to the concept from underneath the surface. The advantages of proximity initially arise from the cost-saving on transportation that is the actual difference in the distance of firm in the neighborhood from the firm in other regions, that it is convenient to connect the neighbor firm. Undeniably, the term transportation cost is used in the broader context, and it covers the transportation cost of people, goods, and ideas exchange. This kind of association in agglomeration economies and transport costs proposes that the importance of agglomeration economies declines as the costs linked to the transportation ideas, people, and goods go down. Moreover, since this concept came out, the world is transformed and has become more globalized. Thus, the paradox which still holds is that in the present world industrial agglomeration has to turn out to be as important if not more, even though the transfer of goods and ideas is more convenient across the regions.

The pile of studies has taken place over time to address the agglomeration economies relate to industrial scope. The literature caries the debate whether the agglomeration economies are associated with concentration within industry or to the concentration of other industries and the size of the city itself. Particularly, it is the argument that whether localization or urbanization economies hold more significance. The micro-foundation of agglomeration initially originates from Marshall's (1920) work emphasis on localization. Surprisingly, this narrative holds much importance after century:

"When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from the neighborhood to one another. The mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously ... Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require ... The advantages of a variety of employment are combined with those of localized industries in some of our manufacturing towns and this is a chief cause of their continued economic growth." [Marshall (1920, 271)]

The well-known example of localization industries includes the computer industry (Silicon Valley), ship breaking industry (Karachi Port) and Football industry (Sialkot).

In contrast, there is another micro-foundation that are provided by Jacobs (1969) which is based on the importance of the urban environment in the industrial scope. She proposed that diversity in industrial space promotes the interaction of varied environment and help industries grow. This argument perceives as a challenge to Marshall's claim, though the Marshall himself acknowledges the relative importance of urban diversity:

"On the other hand, a localized industry has some disadvantages as a market for labour if the work done in it is chiefly of one kind, such for instance as can be done only by strong men. In those iron districts in which there are no textile or other factories to give employment to women and children, wages are high and the cost of labour dear to the employer, while the average money earnings of each family are low. But the remedy for this evil is obvious and is found in the growth in the same neighborhood of industries of a supplementary character. Thus, textile industries are constantly found congregated in the neighborhood of mining and engineering industries, in some cases having been attracted by almost imperceptible steps; in others, as for the instance at Barrow, having been started deliberately on a large scale in order to give a variety of employment in a place where previously there had been but little demand for the work of women and children ... A district which is dependent chiefly on one industry is liable to extreme depression, in case of a falling-off in the demand for its products, or of a failure in the supply of the raw material which it uses. This evil again is in a great measure avoided by those large towns or large industrial districts in which several distinct industries are strongly developed. If one of them fails for a time, the others are likely to support it indirectly, and they enable local shopkeepers to continue their assistance to workpeople in it." [Marshall (1920, pp. 273-274)]

The literature in past shows over the years there are many pieces of evidence of urban economies are present, which shows how the increase in the presence of other industries and size of city surge the productivity (Shefer, 1973; Segal, 1976; Moomaw, 1981). Moreover, other research works have analyzed localization and urbanization simultaneously. Moomaw's (1983) results come up with evidence of both. Henderson (2003a) discovers that localization effects are strong, same as Rosenthal and Strange (2003) outcomes.

This detailed analysis of agglomeration economies and industrial agglomeration clarifies the basic stance of the underlying concept that this study follows. The next section discusses the Labour Pooling concept.

1.2 The Concept Building of Labour Pooling

Labour pooling actually refers to the cost of moving people. As Marshall (1920) has explained the concept of labour pooling in his work in this context:

"[A] Localized industry gains a great advantage from the fact that it offers a constant market for skill. Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skill

as theirs and where therefore it is likely to find a good market." [Marshall (1920, Book IV, Chapter X, para. 9)]

In the literature, there are two concepts of labour pooling. First, emphasizes on the matching that the process of labour matching with respective jobs takes place efficiently in cities and industrial concentrations. Though it is vague that how one can determine the matching quality. One way is to assess the rates of layoff or termination in jobs. This measure implies that if the matching procedure is not good enough the termination rates will be high, it could work as labour market pooling direct estimate. However, in an environment where bad matching is common, firms or establishments would less likely terminate employees because it could increase the recruitment cost with no certain outcome. Hence, the practicality of this method is uncertain.

The second approach to capture labour pooling through the assessment of turnover of firms. It implies that it is more convenient first for workers to switch work second for employers to change employee. But it gives the same kind of problem that termination gives. Alternatively, labour pooling can be explained through specialization. This concept is consistent with Adam Smith's division of labour theory along with Marshall's explanation of labour pooling. Particularly, it shows the agglomeration can foster specialization.

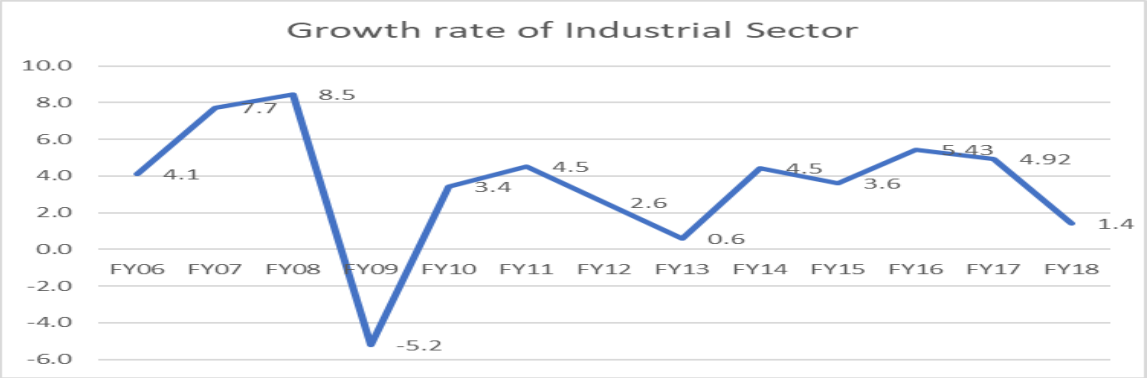
To understand the labour pooling it is vital to fully understand the mechanism that makes it more productive to agglomerate. This study follows the explanation of the classic Krugman (1991) model that basically stressed the benefits of agglomeration that results from a decline in the cost of moving goods over space. Whereas the cost related to good transport has reduced largely both in monetary and time context, the moving of people is still costly. Even though the advancement over the years makes it convenient to transport goods, but there is still benefit in clusters to reduce the cost related to the movement of people over geographic space. Overman and Puga in 2010 examine the labour pooling, which originates more explanation by Marshall (1920). Their basic concept follows the second approach defined above that is if there are enough employers in a geographic location, then it will be possible for workers to switch employers. The concept surrounds that the argument is, during productivity shocks, labour pooling allows efficiency without changing residences. This job hoping will give positive outcomes if it is unknown to workers where their work will give more productivity or if the firm's productivity varies with time. Because on same geographic location worker can switch from less productive firm to the efficient ones. Basically, their work is the extension of Krugman's (1991) simple model of the labour market. They broaden their spectrum over the various sectors and industries. This kind of extension gives this concept more relevance to the current world and helps in building anticipation regarding how different kinds of firms' co-location will be beneficial. One of the crucial findings that come out is, when the shocks among firms of different sectors are heterogeneous the agglomeration is more advantageous. Though this result largely depends on the assumption that sectors are similar enough to share the labour market pool. Moreover, Overman and Puga empirically analyze this phenomenon over the sectors in the United Kingdom. Their result shows that geographically concentrated sector are those which have more plant-level employment shocks. Following the path of Overman and Puga (2010), this study tries to test century-old ideas, but unlike them, it will be tested on a developing country dataset.

After concept building of key features, the next section provides the significance of the study.

1.3 Stylized Facts about Pakistan's Manufacturing Sector; Highlights to Significance

So far literature has a number of growth and development theories, each of them on some level emphasizes on the importance of the industrial sector in the development process. Unfortunately, the industrial sector of Pakistan for few years' remains neglected, due to many factors, some of which include the Energy Crisis, Political instability, and Security issues. Along with these aspects, the lack of interest and attempt for the betterment can also be seen in past years.

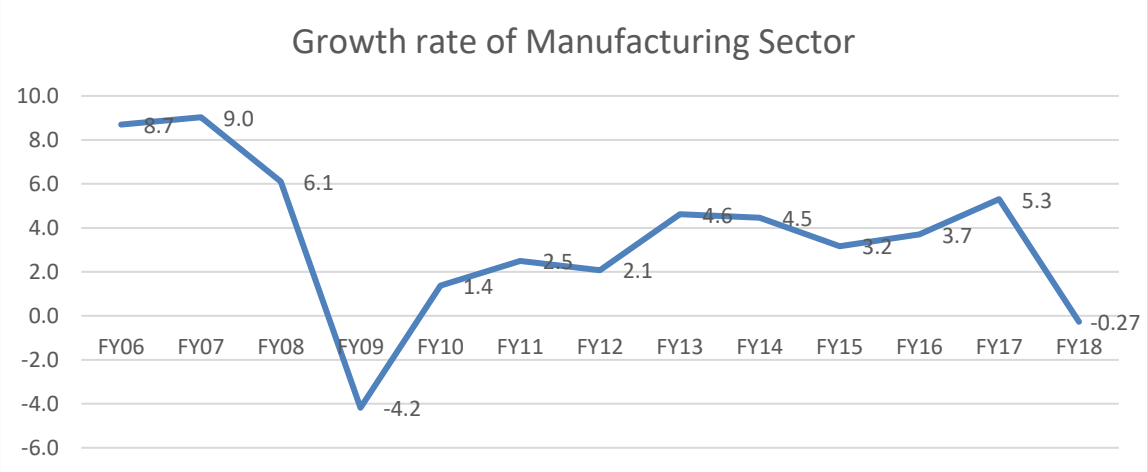
Figure 1: Annual Growth Rate of Industrial Sector in Pakistan



Source: Author illustration (Pakistan Economic Survey 2018-19)

The graph above shows the fluctuating trend of industrial sector growth. Except for fiscal year 2007-08 and 2008-09, the value of growth rate remain below 5.5%. At the fiscal year 2009-10, the growth rate become negative -5.2, the energy and gas's production, which affects other components in the industrial sector has a negative growth rate, which is -12.1, and the manufacturing sector, which has the largest share in the industrial sector is also negative in value -4.2, the negative growth of two important components of industrial growth pretty much explains the overall negative growth. Lastly, the growth rate in the fiscal year 2018-19 is also very low which 1.4 in both absolute and comparative terms.

Figure 2: Annual Growth Rate of Manufacturing Sector in Pakistan



Source: Author illustration (Pakistan Economic Survey 2018-19)

The graph above shows the fluctuating trend of manufacturing sector growth, which nearly is imitating the shape of the industrial sector growth graph which again is justified because the manufacturing sector has the largest share in industrial sector production. Hence, it is justified to examine the manufacturing sector of Pakistan is this study.

This study is trying to trace the contribution of labor pooling as a source of industrial agglomeration and agglomeration economies, especially in the context of Pakistan. The reason for focusing on Pakistan in this study is simple, Pakistan is labour abundant country. It is beneficial if this abundance can be intentionally utilized for the arrival of new firms to form economic zones and lead the country towards efficient utilization of resources. Major business and economic activities concentrate in cities. By analyzing the cities as per Pakistan Bureau of Statistics classification, this study will help to understand that how the manpower which Pakistan has in abundance (mostly semi-skilled and unskilled) can contribute to the agglomeration of industries in the manufacturing sector. Additionally, this study also highlights the regional and labour market factors that affect the agglomeration of industries. Focuses on determining what factors that are relevant in the case of Pakistan. But at this point, the question is why agglomeration is so important and emphasized in this study. The answer is many empirical studies so far have proved that an increase in the agglomeration of industries increases the productivity of firms. Some examples of work in this regard is as follows: Rosenthal and Strange (2004) and Combes et al. (2008). Furthermore, by calculating the industrial spatial concentration index and regional specialization index this study can help policymakers to set objectives in changing economic structure. Also, this study assists to identify the concentration pattern in the manufacturing sector and brings the attention of the researcher in further exploring the effectiveness and policies that can be suggested for further production growth.

This particular study is aiming to assess the labour pooling as a foundation for industrial agglomeration with the help of two major measures of labor pooling. Also, trying to evaluate the developing country based proxy of knowledge spillover. Lastly, explore whether the low-skilled workers and skilled workers serves as the source to attract industrial agglomeration.

In the subsequent section, the literature related to agglomeration and its determinant specially labour is discussed.

2. Theoretical Background & Review of Related Literature

This section discusses the basic theoretical base of this study and all sort of literature available in this regard.

2.1 The Theoretical Framework

The following section provides the theoretical base that theoretically supports this study, explore the labour pooling as a foundation for agglomeration. This section begins with the general explanation of founding factors of agglomeration and move toward more specific the determinant under focus i.e., the labour pooling.

Notably, Marshall (1920) had one of the founding works in regard to agglomeration economies. In his works, he explained the three main determinants that could work as

a foundation for firms to locate a particular area. These fundamentals included intermediate input sharing, a common skilled labor pool, and information spillovers. Later, Krugman (1991a) elaborated the Marshall's findings. Also, Hoover (1948) assembled agglomeration into the firm, sector, and city-specific categories to decline those externalities as sector-specific. Jacobs (1969) expanded the city-specific aspects of agglomeration and put forward the concept that economic activity's concentration on geographical level boosts the region's development, mainly because of knowledge spread. Contrary to this Hoover emphasized that city size is the factor of agglomeration. Jacobs (1969) basically proposed the concept that urban areas got benefit from an economy that was more diversified. In today's literature, the localization and urbanization economies are actually Marshall-Arrow-Romer (MAR) externalities and Jacob's externalities respectively.

Krugman (1991b) formed an economic model to describe the production agglomeration mechanism which gave birth to "new economic geography" theories. Krugman's and the work of other new theorists were grounded on the modeling of imperfect competition and increasing returns to scale, as established by Dixit and Stiglitz (1977). Thus, the new theory of economic geography described the first category as per Hoover's provided classification i.e., firm-specific agglomeration economies. As per these models, if internal economies of scale were present significantly and transportation costs would be low, then production activities inclined towards concentration, as well as sustainability at certain places. These new economic geography theories had revived focus in research on agglomeration economies, and more work had been done based on previous frameworks (Rosenthal and Strange 2001, 2003).

As per the Duranton and Puga (2004) analysis, there was three micro-foundations for the agglomeration economies e.g., learning, market linkages, and labor market efficiency. In their study, we explored the foundation factor results from the pool of worker in an area, with the help of labour pooling concept presented by Marshall (2009) i.e, "*the concentration of specialized labor generates gains for firms and workers, such as by reducing the time of hiring and increasing the ability of firms to adjust to productivity shocks.*"

Krugman (1991) suggested a theoretical model, which clarified how this force would be working on the distribution of productive activity. Also, this work of Krugman (1991) later followed by the Overman and Puga (2009) and Gardezi (2013).

The theoretical building of this framework begin by considering the number of industries indexed as j = 1,…, J. In every industry there is particular number of firms sub-indexed i = 1, .., n, and each of them have a pool of workers with specific skills and expertise. After located at particular place each establishment go through a particular shock $\boldsymbol{\epsilon_i}$. Productivity shock that experience by each firm is heterogynous in nature i.e., the productivity shock will be establishment specific, as well as, has no correlation across firm. Establishment decides the quantity of labor it chooses to hire from the locally available labor pool. Suppose if the level of employment chooses by establishment is li, and $\boldsymbol{\lambda}$ shows intensity of decreasing return to scale of a firm's production function, the profit will be as following:

$$\pi_i = [\beta + \epsilon_i]l_i - \frac{1}{2}\lambda l_i^2 - w l_i^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

As in work of Krugman (1991) the local wage is considered as given by firms. Hence, the firm after undertaking shocks set its quantity of workers where, the marginal productivity of worker will be equal to wages. The demand for labor in firm i is:

$$l_i = \frac{\beta + \epsilon_i - w}{\lambda} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Aggregated supply of hours of the work by labors is symbolized as $L_{i,j}$ in a given city and sector. Equilibrium in labor market together form in equation (3) as follows:

$$L_i = \sum_{i=1}^{n} l_i = \frac{\beta + \sum_{i=1}^{n} \epsilon_i - w}{\lambda} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Equilibrium wage can be determined by above equations:

$$w = \beta - \lambda\frac{L}{N} + \frac{1}{N}\sum_{i=1}^{n} \epsilon_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

Thus, the expected wage is:

$$E(w) = \beta - \lambda\frac{L}{N} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Solving equation (1) by putting equation (2) in it leads to:

$$\pi_i = \frac{[\beta + \epsilon_i - w]^2}{2\lambda} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

Firms adjust their production in response to productivity shocks therefore; firm's profit is a convex function of production shocks and wages. Incorporating expectation in profit function yields:

$$E(\pi_i) = \frac{[\beta - E(w)]^2 + \mathrm{var}[\epsilon_i - w]}{2\lambda} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

In equation (7), Substitute equation (5) and var [$\boldsymbol{\epsilon i}$,–w] = var [$\boldsymbol{\epsilon i}$] + var [w] – 2cov[$\boldsymbol{\epsilon i}$, w], which results in

$$E(\pi_i) = \frac{\lambda}{2}\left[\frac{L}{N}\right]^2 + \frac{\mathrm{var}[w] + \mathrm{var}\,[\epsilon i] - 2\,\mathrm{cov}[\epsilon_i, w]}{2\lambda} \ldots\ldots\ldots \quad (8)$$

When there is no productivity shock, the firm's profit can be represented by the first term of the right-hand side of the equation. The increase in labor to firm ratio increases the expected profit of a firm because from equation (5), it can be seen that an increase in this ratio reduces expected wages. Moreover, the second term of the right-hand side part of the equation captures the effect of labor pooling. It reflects the expected profit is likely to increase with the increase in productivity shock's variance and local wage's variance, however, decreases with an increase in productivity shock and wage's covariance. This can be explained as if firms increase their production due to productivity shock but simultaneously it also increases the wages or vice versa in negative productivity shocks, this will reduce the firm's expected profit. The explanation above provides the micro-level base for the framework of labor pooling as a foundation for agglomeration. So, firms that

face the productivity shocks which affect the local wage prefers to agglomerate with other firms to lessen the impact of wage fluctuation.

To further comprehend the equation (8), equation (4) and (5) can be used to determine var[w]=$\frac{\sigma}{N}$ cov[ϵ_i, w] = $\frac{\sigma}{N}$. After substitution of these variances the equation (8) will become:

$$E(\pi) = \frac{\lambda}{2}\left[\frac{L}{N}\right]^2 + \left[1 - \frac{1}{N}\right]\frac{\sigma}{2\lambda} \quad \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

To generalize, equation (9) does not include the sub-index i, so the expected profit of the firms is equal that have the same location and sector. The more heterogeneity in firm-specific shock leads to the greater effect of labor pooling. An explanation of the statement earlier is when a firm goes through a positive shock and increases its employment, to neutralize this effect other firms must be reducing their employment. Industries are likely to agglomerate more if the employment pattern is heterogeneous across firms. Otherwise, the local wage will exhibit shocks which result in reducing expected profit.

Overman and Puga (2009) have modeled production and relocation in a dynamic setup. Initially firms explicitly choose the location along with the number of workers. At equilibrium, neither firm nor worker has an incentive in relocation because wages are equalized. The labor per firm ratio is the same across locations. Furthermore, the relocation of a worker does not affect the wages or profit; though, the wages at origin and final location can get affected if a firm decides to relocate. Afterward, firms face the productivity shock ϵi which creates the deviation in profitability. The destination at which the firm decides to relocate, the worker-to-firm ratio decreases over there. This will put upward pressure on expected wages and as a result expected to decline profit. Moreover, if there is a large number of firms at the destination locations, productivity shock will not largely affect local wages? Because of this firms are able to lighten down the shocks and gain higher expected profit. If an industry's shocks are more heterogeneous in nature, it will tend to agglomerate more to gain more expected profit.

This is the theoretical base used for empirical the model building. This potential labour pooling along with another proxy and the measure of geographical concentration used in this study, as described in the following sections, were computed from micro-data the LFS, which is a comprehensive survey on employment data.

2.2 Literature Regarding Agglomeration and Labour Pooling

This section discusses the review of literature both international and national respectively.

2.2.1 International Literature

Agglomeration economies are broadly considered in describing the geographical concentration of industries. There is a vast literature on agglomeration economies, some of the fundamental founders include Smith (1776), Marshall (1920), and Hoover (1948) and later economists revived this issue to give it a descriptive and more practical approach such as Krugman (1991b), Fujita and Thisse (1996), and Puga (2010). One of the highlighting contributions in this area was made by Ellison and Glaser (1999), while elaborating on the geographic concentration of industries pointed out that the decision where the firm locate is sensitive to the cost of inputs, specifically, the production process

is used intensively. They performed empirical work by using 1987 data of four-digit manufacturing industries and apply nonlinear least squares to illustrate the relations intended to reflect the advantages in natural resources, transportation costs, and labor. The paper concluded that the firms that were concentrated geographically were because they had a cost advantage. In their study, they formed Ellison Glaser index, which latter was widely used (Abid A. Burki 2010, and Overman and Puga 2010).

Silicon Valley came to the attention of economists, because of the way that the valley had been agglomerated. D P Angel (1991) conducted a survey-based case study in Silicon Valley. The focus of that study was to analyze the pattern in which firm cluster in Silicon Valley and the way labor market contributed to the growth of high technology firms that cluster. In this research, the assessment of local labor-market dynamics was conducted along with the labour market link to the organization of high-technology production in the agglomeration. It was a case study that conducted on clustered semiconductor industry in Silicon Valley, covering aspects such as labour market and production activity. The study examined the pattern of wage by regression analysis, where the dependent variable consists of an average hourly wage. The independent variable included were whether the workforce was unionized, whether the institution was an integrated captive facility of a large firm, employment level in electronics, and whether they made integrated circuits. Moreover, the study also tabulated the survey findings. As per the implications, this study presented the significance of knowledge and experience flows among firms in Silicon Valley. Moreover, the effectiveness of Silicon Valley was more as a production hub than as an individual firm, as per conclusion (Ellison and Glaser, 1999).

Whether the growth and productivity could result from the concentration of manufacturing industries is the subject that has been researched multiple times in literature so far (Moomaw, 1981) (Nakamura,1985) and (Henderson, 1995). As per these studies, localization, and urbanization affect the productivity of manufacturing industries which creates concentration, however, localization turns out more efficient than urbanization according to empirical investigation. Rosenthal and Strange (2001) studied agglomerating economies' micro-foundations for the manufacturing sector in the U.S. Their work measured spatial concentration by the EG index as a function of spillover of knowledge, labour pooling, cost of shipping, natural pulling factors, and input sharing. Their study broke down this analysis on zip code, state, and country level. As a matter of focus in their study was labour pooling, as Rosenthal and Strange (2001) represented it through the percentage of degree attained by workers. Usually, the degree attainment and its specialization determine the skill the labour possesses. (However, as far as our study is concerned the quality of education normally in developing countries is so low that it does not capture the effect). As the theory and develop world practices implied labour market pooling came out as positively related to concentration on all levels.

The cost of big cities can be seen nowadays, more than ever and it can be seen in terms of environmental degradation, traffic, and intense unplanned concentration of masses. However, it has its own benefits, both revealed hidden. Some of the hidden benefits were pointed out by Marshall (1920), which presented the analysis of agglomerating economies, claiming cities improve productivity by letting the labour pool, technical knowledge spread, and input to share among industries. Rosenthal and Strange (2003) in their study provided two notable contributions in economies of agglomeration. First, using a rich

database and mapping software technique they identified the extent of agglomeration externalities geographically. Second, they provided empirical evidence on agglomerating economies that was organizational, close to the theoretical study of Saxenian (1994). Their outcome showed that the industries they analyzed, gave the sign of localization economies varies across the different distance in the location of industries. Due to the study's limited scope, it only provided the theoretical justification of the presence of all three sources, namely labour pooling, knowledge spillover, and input sharing as causing factors of agglomeration economies for the case under analysis.

The location decision of firms frequently used as an important factor in the agglomeration model. So, it is essential to differentiate the agglomeration sources to fully understand these decisions. La Fountain (2005) examined that firms that choose to locate near each other can be because of "natural advantage models" that is some firms used the raw material that was only available on location or production of specific goods need a certain degree of climate. The other models that cause this phenomenon included "market access models" according to which producers, as well as consumers, preferred to concentrate to avoid heavy transportation costs. Another factor that contributed to this nearness of firms was externalities i.e., when one firm closely located to other firms, it resulted in producing positive externality. Every model described above, according to his results anticipated different associations among land prices where industry was located, the productivity of industry, industrial contribution in employment in the region, and degree of diversion in labour. To perform empirical work, he used the US's county data for 48 states for the years 1980 to 1990. Also, the industries consist of 2-digit level 18 manufacturing industries. In his result, he broke down the industry category-wise and elaborated categorically on the firm's decision to locate. As per his results, he provided an industry-wise assessment of the reason stated above, such as electronics, paper, chemicals, petroleum and coal, instruments and textile industries' firm gave prime importance to primary input location while making a location decision.

When firms cluster and create the spatial concentration of firms and industries, this phenomenon has its own cost and benefit. Combes and Duranton (2006) emphasized that when firms and industries cluster, the two main costs and benefit arise in the context of the labor market. First, the benefit, according to their work was the labor pooling, which makes access to a pool of labor, who have skills and expertise possible (helps to reduce the cost). Second, the cost was, the labor poaching that loss of competitive workers due to competition and higher wage (which in result becomes costly to retain worker). According to Combes and Duranton, there was a trade between labor pooling and labor poaching, which could be explored better in a duopoly game. Firms made choices strategically in this setting, for instance, where to locate, prices and wages to charge, and about poaching, depending on the degree to which goods was differentiated, specific market size, heterogeneous nature of worker in term of their knowledge and expertise.

Firms throughout their existence keeps on changing the level of input use due to their strategic planning of a business or shocks its faces. Overman and Puga (2010) drew attention to the sort of agglomeration that occurs due to idiosyncratic shocks that firm was facing, which was a factor that intensify the procedure of labor pooling (due to variation in production and employment that the firm went through because of idiosyncratic shock). They took the data from the Annual Respondents Database (ARD) since 1994 to 2003.

The result of their study showed firms and sectors that went through more idiosyncratic shocks was more spatially and geographically concentrated. As per their argument firms got benefit from locating in proximity because idiosyncratic shocks that firms faced, was heterogeneous and it was what benefited firms i.e., when one firm increased its employment level other firms in the observed period was contracting in terms of employment and diminished the labour market wage impact for all and increased profit margins. Similarly, the claim that labour pooling works as an agglomerating factor has also been tested on a number of data sets worldwide, for instance, the USA, Brazil, Turkey, Italy, Pakistan, etc. The empirical evidence of Brazil was analyzed by Almeida and Rocha (2018) for the period 2002-14, particularly the northeast region of Brazil. They measured manufacturing industry's concentration through the Ellison Glaeser index. Their results showed an overall index decreased in the given period showing the de-concentrating of industries. Though, if those results segregated by activity in the manufacturing industry, roughly few of these on segregated level showed high agglomeration, for example, technology-intensive industries, which showed more concentration than other analyzed. Their study supported the argument that the geographic concentration of industries gave productive gains from the accessibility of specialized labor in comparison to firms that located in separated areas. More importantly, their study followed the same measure of labour pooling as in the study of Overman and Puga (2010). Also, the study of Gardezi (2013) followed the same labour pooling measure with different agglomeration to assess the agglomeration source.[1]

One of the phenomena that we observe in today's world are firms and people tend to be concentrated in industrial zones and cities respectively. However, these concentrations bid up prices and cause crowding and congestion in day-to-day life. Puga (2010) analyzed theoretically, over the 30 years, economists could identify 3 approaches of the productivity advantages of cities and urban clusters. These approaches argued that whether it was firms or people, they were more productive in a dense urban environment rather than in any other area. He also discussed the causes of agglomeration in detail. According to his examination, there was a broad range of literature with which the mechanism of urban agglomerative economies could be explained. First, the large market or a big cluster of firms made the efficient sharing of the labor pool, local infrastructure, provision of input supply, possible. Second, again if there was a large market the efficient matching among the buyer, supplier, employer, employee, and partner took place. Finally, the large market or firms' clusters helped in the diffusion of knowledge and business practices. As concluding remarks, Puga emphasized the need for a microeconomics foundation of agglomeration magnitude causes and matching as a source of it.

The pair of words ''agglomeration economies'' is used to signify the processes that lead labour and firms to locate near each other on a geographical level. Monseny, López, and Marsal (2011) investigated the agglomeration processes that was explained by Marshall, through analyzing the new manufacturing firm location, particularly in Spain's 477 cities. In their main function, they took the number of new firms classified by their industry and location as a function of the use of workers who were alike in their skills (labour pooling), input sharing and the usage of technologies that was alike by firms in similar industries (spillovers of knowledge). They estimated the variation in the geographic location of new

[1] The paper of Gardezi (2013) discussed in much detail in the national review of literature section.

firms across the area and how each of these mechanisms of agglomeration caused concentration. Also, notably, they generated the labor and technology similarity index to identify Marshall's claim and count of new firms considered as a dependent variable. As far as the results were concerned. They supported the existence of all agglomeration mechanism which is sensitive to the extent of area under analysis.

The spatial concentration of industries is also affected by the transportation cost and international trading atmosphere. Behrens, Brown, and Bougna (2013) examined the agglomeration determinant of manufacturing industries for the period 1990-2019. A long panel of localization was used in this study to reassess the work of Rosenthal and Strange (2001). It particularly emphasized the factors such as input sharing, cost of transportation, and access to the international market by using micro-data. Their results suggested for the data of two decades that there was a continuous decline in localization of industry. Localization fall could be because of competition in the international market, the reason could be cost-efficient countries in term of wages or from the increased in transportation cost etc. The outcome showed high responsiveness in localization because of trade, however, knowledge spillover and labor pooling were not that much affecting the localization.

As it is discussed above (Puga, 2010) concentration pattern could be explained through the difference in spatial natural advantages. Ellison, Glaeser and Kerr (2010) assessed the co-agglomerating pattern of industries and evaluated the significance of the different theories of geographic concentration. They pointed out, there was different theory made different inferences about patterns and tendencies with which industries co-agglomerated. For the measurement, they used the longitudinal data in manufacturing industries of the US, from 1972 to 1997. Using OLS, as well as IV techniques, concluded that all the three Marshall Theories regarding industry pair of industrial agglomeration were relevant and significant.

Many prior studies directly or indirectly undertake the same assumption that all industries agglomerate for similar reasoning which is majorly based on Marshall's provided traditional explanation (reasons include labour pooling, input sharing, and knowledge spillover). Another important aspect of a major portion of studies in the context of industrial agglomeration included the generalization of the results originated from the very setup of Silicon Valley. Faggio et. al (2017) in their work considered the industrial heterogeneity that existed in the micro-based foundation of agglomeration economies. Their study focused on the UK and data was taken from the Business Structure Database and the Quarterly UK Labour Force Survey. Moreover, their empirical investigation concentrated on the association among the marshal's provided industrial links and industrial pair co-agglomeration. The main finding of their study explicitly showed that the different industries react in a different way to the forces of agglomeration. So, even though a specific scenario's study was essential, but it was equally important not to directly generalize the finding without further consideration. Their study highlighted a very significant factor that remain often left unseen by the policymakers i.e., the heterogeneous nature of agglomeration economies.

As far as the measure of spatial concentration of industry is concerned throughout our literature review different measure has been used so far such as EG index and MS index

etc. The papers using these indices have vast data resources in term of the comprehensive detail survey on a large sample size. However, the developing countries usually fall behind in terms of data availability and sample scale. So, the available literature also contains the use of location quotient index as a measure of spatial concentration which is quite convenient in use. First, Guimarães et al. (2009) explored whether location quotient in more depth as s relative concentration index. Their study followed the footsteps of Ellison and Gleaser (1997) dartboard method. In concluding remarks, their study justified that firms take location decisions to capture the benefits of Marshall's specified economies, clusters on the industry in an area emerge, and this fundamental occurrence of concentration by a high LQ in a relative term. Second, Strotebeck (2010) decomposed the LQ index and analyzed its effectiveness as a measure of geographic concentration of industries. In his work, he suggested various methods with the LQ relevance in terms of interpretation and accuracy could be raised by using biotechnology industry data of Germany. Also, Boasson and Boasson (2011) examined the spatial concentration in the health industry and services sector with help of LQ. Their study used tools such as LQ and Moran's I statistics to examine the geographic and time-based pattern of clustering of heath and services industries across the 62 counties in New York, USA. Lastly, Liu (2014) measured the geographical concentration of manufacturing data by using multiple indices including LQ. The data used in his study was China's county-level data. Maps resulting from the local Moran statistic were likely to pinpoint fewer agglomeration than found in the LQ.

As far as the determinants of agglomeration explained by Marshall are concerned, various studies have mentioned above supports the claim that labour pooling works as a foundation factor for agglomeration along with other factors. Though, there are shreds of evidence that show some causes of agglomeration can be more effective than others depending on the regions and industries under analysis. Moreover, most studies use their own perceived and derived implication of Marshall's provided explanation of labour pooling, knowledge spillover, and input sharing over the course of time. Also, this study for labour pooing follows the measure suggested by Overman and Puga (2010). Since it is more recent and relevant in the current literature as a measure of labour pooling in agglomeration sources. This study aims to reassess the significance of labour pooling as a foundation for agglomeration in Pakistan. Also, the study discussed above mostly were using developed country data. So, whether this holds as much as relevancy in Pakistan city-level's case, as different countries have different external factors affecting the outcomes.

2.2.2 Literature Related to Pakistan

This subsection reviews the national studies conducted in this regard. In the case of Pakistan, Burki, and Khan (2010) with the help of CMI 2005-06 data explored whether there was agglomeration in the manufacturing industry exists and factors that cause agglomeration in the manufacturing industry, and the dynamic process with which agglomeration took place over time. Their paper concluded that concentration could be seen in large manufacturing industries and large manufacturing city district such as Lahore, Karachi was highly clustered in term of employment. As their results suggested the agglomeration could be observed in 3-digit industries. Furthermore, the paper pointed factors that caused agglomeration in manufacturing industries, for instance, the population

of the district, road density, and skilled pool of the labor market. These district characteristics were one of the prime inspirations in our analysis to include a variable that might be varied across the region and affect the agglomeration. Their paper emphasized that Pakistan's localization plays rather more important than spillovers of technology.

Is industry agglomeration beneficial for the firms? And if it is what is the extent of the benefits? These questions mentioned above are sensitive to the presence of technical inefficiency in the production process. Burki and Khan (2013) analyzed how the technical inefficiency in the production process affected a firm's productivity by using plant-level cross-sectional data for 1995-96, 2000-01, and 2005-06. It also included the over-the-year Pakistan-level industry agglomeration assessment through the EG index. According to this paper, the concentration of industries could be evident, but with a declining trend. Also, their results suggested that industrial agglomeration had vitally benefited the firms, showed by a negative relationship between the firms' agglomeration index and technical inefficiency. Moreover, it also concluded that factors such as district population size, road density of districts and most importantly the technically trained pool of worker works as causes of manufacturing industry agglomeration in Punjab.

Punjab is the largest province of Pakistan in terms of population and serves as a home for large agriculture, as well as industrial activities. Gardezi (2013) examined the manufacturing industries' pattern of agglomeration in Punjab with the CMI 2005-06 32 districts' data. Unlike, other studies available before this study employed a varied form of agglomeration measure called the M-function, proposed by Marcon and Puech (2009). She analyzed manufacturing industries' agglomeration behavior with area characteristics and distance-based measure. It pointed out industry-specific features that encourage firms to locate in proximity. It was the first study in Pakistan that addressed Marshall suggested positive externalities that give cost advantages if firms locate near each other. This study followed the labour pooling measure based on idiosyncratic shocks explained in Overman and Puga (2010) and found a significant positive impact of labour pooling on agglomeration.

Agglomeration is highly important when the pattern of formation and scale of economies of new firms is determined. Chaudhry, Nasir, and Haroon (2012) examined the different aspects that influence the entry and scale of operation of new firms in Punjab's export clusters. Their results suggested that that firm entry increased considerably as a result of a decline in the trade-weighted real exchange rate, though the impact of variations in trade partner tariffs was not significant. Chaudhry (2005, 2010) showed the benefits of agglomeration in Pakistani clusters while Chaudhry (2011) looked at the disadvantages of agglomeration and showed that firms in clusters might be reluctant to try new suppliers when their current ones was linked to them socially.

Lastly, there are various studies that explored industrial agglomeration in Pakistan, though they are not very close to the core concern of this study but provide valuable insight into national-level industrial agglomeration. Some of this work includes studies such as Zhao and Mudassar (2013) analyzed the industrial agglomeration of the manufacturing sector in Pakistan with the help of CMI data from 2000 to 2012. Nasir (2013) examined the geographic, as well as industrial concentration of firm affects the turnover of the firm in

Punjab, Pakistan. The study's outcome showed that entry, as well as exit, was likely to happen more in highly agglomerated industries because first firms was attracted due to agglomeration economies. Nonetheless, already existed large firms with a sizable market share made other firms reluctant to enter or cause them to shut down.

In the subsequent section, the methodology adopted in this work is discussed.

3. Methodology

In the urban economics literature, Industrial agglomeration is the focus of economists in current decades, due to the significant benefits it provides e.g., industrial sector growth, economic activity acceleration, etc. For estimation purposes of the agglomeration measure, labour pooling, and other determinants, this research study uses the Labor Force Survey (LFS) for Pakistan in 2014-15 and 2017-18. Since independence due to discontinuous policy pattern industrial sector is suffering. Though there were some hopeful moments and growth evident in the past eras, sudden authoritative decisions create the wave of uncertainty. The manufacturing sector has the biggest share in the industrial sector. Furthermore, there is much evidence that presents that the urbanization of remote areas can be promoted through more employment opportunities with the help of industrial agglomeration. Side by side, it will also increase the industrial activity overall which has its own benefit. Also, the establishment of new urban regions with the help of industrial agglomeration will too release the pressure from the existed congested ones. Therefore, in this particular study, to assess the agglomeration and its sources, 14 cities data for Pakistan is analyzed. However, LFS has one major drawback of not having firm-level data. To cope up with this lacking, this section took 2 digit industry-level data as a proxy for industry and 4 digit level data as firm's employment. This step has also been taken by Gabe and Abel (2010) while facing similar data limitations, during the calculation of industrial agglomeration index.

There are reasons due to which this study is employing LQ Index as a proxy for agglomeration; empirical model. Firstly, LFS does not provide the sort of depth in the data that allows calculation of EG Index for the industries by cities separately. Second, like any other agglomeration indices, the EG measure of agglomeration depends on the discreteness of spatial units. EG Index's consistency and unbiasedness could vary with the extent of spatial aggregation. Thus, in this study, we assume firms located in the discrete states because, the size, location, and form of the chosen spatial unit (city, district, or province) can affect the results of the analysis. Furthermore, location quotient as a measure of agglomeration is also estimated to avoid the data limitations. The estimator of LQ further help in quantifying the agglomeration in Pakistan and in the overall robustness of the results. Specifically for estimating the impact of labour pooling on agglomeration, this study uses the LQ as a proxy of agglomeration.

Hence, to perform detailed empirical work which requires data by industry in the respective 14 cities. This study uses Location Quotient (LQ) as a measure of agglomeration. Many research studies mentioned as well as used LQ as an index for spatial concentration of industries, such as Guimarães et al.'s (2009), Boasson and Boasson, (2011), and Fracasso and Marzetti, (2017). The computation of LQ for assessing the industrial concentration is accomplished through the division of share of employment in an industry in a specific region (in our case city) by share of employment in the industry

under consideration on large reference groups such as a country/nation. As mentioned earlier, there is a severe lack of data in Pakistan for the analysis under consideration, so the LQ is a very conveniently useful measure to capture the industrial geographic concentration.

Furthermore, throughout the literature various explanations and measures of labour pooling are proposed. This study employee two different measures of Labour Pooling; one based on educational attainment and second based on quality of matching. Both measures are explained in the descriptive part in detail and used in empirical estimation.

In the following section, the methodology for evaluating the linkage between agglomeration and labour pooling is discussed. The first part begins by discussing the estimation of Location Quotient and Labour Pooling. Furthermore, after discussion on index formation. The second part discusses the model for the empirical estimation along with the econometric specification of the empirical model.

3.1 Development of Indices and Variables

The detailed description of variables explanation is explained in these headings below. These variables used both in descriptive and empirical part.

3.1.1 Location Quotient Index

The index of location quotient (LQ) primarily proposed by Florence (1939) as an indicator of relative concentration. As per Florence's explanation the LQ of an industry can be written in term of the formula as.

$$LQ_{ic} = \frac{\frac{E_{ic}}{E_c}}{\frac{E_i}{E}}$$

E_{ic}= employment in industry i in city c

E_c = total employment in city c

E_i = total employment in industry i on the aggregate level

E = total employment at the aggregate level

Where the index of concentration of the industry i in the city c is benchmarked against a similar ratio, but on a larger context or aggregated level. Due to the benchmarking method, the LQ index can be used as a concentration index (Fracasso and Marzetti, 2017). The location quotient index is used as a proxy measure for spatial concentration. Location Quotient (LQ) computes the within regional concentration of industries and also known as the Hoover-Balassa coefficient. It is in-depth the assessment suggests that the value of LQ higher than 1 takes place when an industry has a large share of employment in a region as compared to the industry share on a broader/ national level, it signifies the relatively high industrial specialization. Industries having high value on regional LQ are most often considered geographically concentrated industries. For what should be the value of LQ to be considered as geographically concentrated remain arbitrary by studies. However, the

implication can be derived from the discussion above that as LQ increases, it leans toward the indication of industrial geographic concentration.

3.1.2a Labour Pooling Based on Firm to Industry Pooling

With references to the title of the study, evaluation of labour pooling as the foundation for industrial agglomeration, labour pooling is crucial in our analysis.

The first measure pooling that is used in this study is suggested by Overman and Puga (2010), though in little altered form due to data limitation but it serves the purpose of tracing the labour pooling. N_f be the number of firms in the industry

$$LP_i = \sum_f |\Delta EMF - \Delta EMI| / N_f$$

Where Δ EMF is a change in the plant employment; and ΔEMI is a change in the industry employment. Furthermore, productivity shocks are heterogeneous, and it is computed by taking the difference in employment variation.[2] The term |Δ EMF - ΔEMI | captures the effect of heterogeneous productivity shocks across plants in the industry and it will take large value for plants that raise their employment while the other plant in an industry is shrinking or vice versa. It is crucial to subtract between the plant's employment variation from the industry's variation in employment because there is little or no gain from labour pooling if each time plant increases their employment, various other plants in the similar industry also grow.

The is vital to mention that this study is taking this variable as by industry by city, so for instance, change employment at a plant of a specific industry belongs to a specific city. Also, to avoid high standard deviation difference in endogenous and exogenous variables which could lead to an exceptionally small coefficient, this variable is used in log form.

3.1.2b Labour Pooling Based on Educational Attainment

As per the importance of labour pooling defined in the above text. This measure is proposed by the Ellison and Glaeser (1999) and Rosenthal and Strange (2001), according to it, the percentage of workers hold different education attainment can be used as a proxy of the quality of the labor pool. Though, practically due to the low quality and attainment of education, there is a low correlation in the degree and type of skill requirement available in the job market. But it is crucial to include and test this hypothesis.

The attainment of education among labour force is really low in Pakistan and also the quality of education is low. So, unlike the referred research's reference, this study will divide into 3 shares. First, the percentage of workers (related to a specific industry) with primary education which covers attainment educational attainment till 8 grades. Second, the percentage of workers (related to a specific industry) with secondary education which covers educational attainment of grade 9 to 12. Lastly, the percentage of workers (related to a specific industry) with tertiary education which covers all degrees equal to and more than a bachelor degree.

3.1.3 Knowledge Spillover/ Learning

[2] Throughout the study the terms "firm" and "plant" are used alternatively.

Knowledge spillover is one of the vital and multi-aspect micro-foundations of agglomeration, due to its connectivity to theories of growth as well human capital economics. Though, it is a concept that is hard to identify to perform empirics. Knowledge primarily often transferred without being bought and sold and there is important significance to the word "spillover" (Helsley and Strange, 2002). In developing countries like Pakistan, where invention and innovation are not practiced widely, spillover through experience (practical knowledge) is suitable for consideration.

In the previous studies, the proxy for knowledge spillover is often regarded as expenditure on R&D (Overman and Puga, 2010). This study used very direct measure to knowledge spillover which is very relevant for the case of Pakistan due to its low spending on R&D overall more specifically, It is the "on the job training" on workers in a specific industry. It attracts the industries to agglomerate because it brings down their own cost of on the job training or gives the advantage of the possibility of getting the skilled labour force (Rosenthal and Strange, 2004). The other proxies in the models for knowledge spillover are "low-skilled workers with on the job training" and "low-skilled workers with on the job training", which also can work as a source of agglomeration because low-skilled worker's standard wage will be less than the skilled ones.

3.1.4 Skilled Level Workers

As mentioned earlier in this work, the literature conducted on this regard largely focuses on developed country's data set and thus representing determinants of industrial agglomeration relevant to developed world. As most works suggest skilled workers serves as the pulling factor in agglomeration source. This study assess whether it is applicable in developing country such as Pakistan or not.

Usually, the labour market in developing countries such as Pakistan has large unskilled workers as well as semi-skilled workers. Also, the type of industries that exist in these also require low-skilled workers in a large sum.

The variable of an unskilled worker in this study captures with unskilled and medium-skilled workers, so it's acceptable that both unskilled and medium-skilled workers require low wages which work as an agglomeration factor for industries. Moreover, these workers may have some set of skills that differ from region to region and also that skill set works as prime input for some industries. Hence provide the agglomeration source.

3.1.5 Construction of Some Important Variables

Unlike previous studies in this regard and due to the inclusion of pseudo-panel cohorts in this study, it is vital to control the model for it. So, there is some controlling variable present in this study to capture the regional (city-level) and industrial characteristics. This controlling variable includes region and city wise variable such as the percentage of migrants, the percentage of married workers, the percentage of workers literate, the size of the informal sector, and age of workers.

3.1.6 Regional, Industrial and Time Dummies

To trace the effect of cities, industries and time period involved, the relevant dummies are used are per econometric model requirement.

3.2 Econometric Model

This study as a quantitative approach empirically analyzes the Krugman (1991) reasoning by considering a series of sectors where firms experience idiosyncratic shocks (more detailed description in theoretical foundation). Further, this quantitative analysis will help to trace the existence of labor pooling as the foundation for agglomeration in industries of major cities in Pakistan. The included major cities specified by LFS are

"Karachi, Lahore, Islamabad, Faisalabad, Rawalpindi, Multan, Gujranwala, Sargodha, Sialkot, Bahawalpur, Hyderabad, Sukkur, Peshawar, and Quetta".

Also, industries under analysis contains

"Motor vehicles, trailers and semi-trailers, Manufacture of Basic Metals, Machinery and Equipment n.e.c, Wood and Related, Products Manufacture, Electrical Equipment Manufacture, Paper and Related Products Manufacture, Chemicals and Related Products Manufacture, Other Manufacturing, Other Transport Equipment, Textiles Manufacture, Food Products Manufacturing, Repair and installation of Machinery and Equipment, Wearing Apparel Manufacture, Other Non-metallic Mineral Products Manufactures, Fabricated Metal Products, Leather and Related Products Manufacture, Rubber and Plastics Products Manufacture, Beverages Manufacture, Printing and Record Media, Basic Pharmaceutical Manufacture and Furniture Manufacture".

Thus, the model is specified as

Model 1:

$$AG_{ct} = LP_{ct}\,\beta_1 + SP_{ct}\beta_2 + SS_{ct}\beta_3 + ST_{ct}\beta_4 + OJT_{ct}\beta_5 + UNS_{ct}\beta_6 + CV_{ct}\beta_7 + ID + CD + TD + \varepsilon_{ct}$$

Model 2:

$$AG_{ct} = LP_{ct}\,\beta_1 + SP_{ct}\beta_2 + SS_{ct}\beta_3 + ST_{ct}\beta_4 + OJT_{ct}\beta_5 + SK_{ct}\beta_6 + +CV_{ct}\beta_7 + ID + CD + TD + \varepsilon_{ct}$$

The dependent variable shows the geographical concentration of industry measured through the location quotient index, in a cohort. In the model explained above, LP, SP, SS, ST,OJT, UNS, SK, and CV are exogenous variable having the factor of time variance in them, and also they respectively stand for labour pooling, percentage of workers with primary education, percentage of workers with secondary education, percentage of workers with tertiary education, percentage of workers with on the job training, percentage of low-skilled workers, percentage of skilled workers, and control variables for region and industries. Furthermore, as specified, the model above is Least Square Dummy Variable, thus it includes industrial, regional, and time dummies as ID, CD, and TD.

3.3 Methodology of Empirical Model

In this study, we assess the labour pooling as the foundation for agglomeration in manufacturing industries in 14 cities of Pakistan as per LFS classification. As the location decision of firms and tracing the pattern of industrial agglomeration in cities is dynamic analysis. Thus, this quasi-longitudinal study will help in understanding the pattern of

industrial agglomeration in major cities over time and the degree to which labour pooling is determining it.

One of the major data limitations mentioned in the above text is that every year different households are surveyed in LFS, Pakistan. The main issue of using LFS in the dynamic analysis is the unobserved individual heterogeneity that could disturb the outcome. To avoid this issue of unobserved individual heterogeneity, the work of Deaton (1985) has proposed, that instead of using individual observations in empirical estimation the 'cohorts mean' can be used. As per Deaton (1985) outcome shows, the use of pseudo panel can limit the actual panel problems and allow the more significant over the year analysis, despite the different cross-sections. Hence, this study is using 2014-15 and 2017-18 surveys of LFS Pakistan for a pseudo-panel. Deaton (1985) claimed that the use of pseudo-panel for cross-sectional data is beneficial because it allows defining past patterns and side by side anticipating the future course. Cross-sectional data which is repeated and large in term of the household or individual covered in it. As per classification and distribution provided by PBS (LFS design (2008-09) and the PSIC (2010)), this research study has taken the City and the industry as a cohort. As far as the sampling size distribution by PBS is concerned, it begins by setting the sample in the provinces then its breakdown in the urban-rural region. After it, major cities are formed from the sample of urban regions. As for the industrial classification, Pakistan Standard Industrial Classification (2010) provided classification is used.

3.3.1 Applying Pseudo-Panel Approach

The core concept for the application of pseudo panel on cross-sectional surveys is the compilation of the individual/household into groups on the basis of common attributes among them and form cohorts (Deaton, 1985). This study forms two cohorts e.g. city and industry, after ensuring that each cohort holds adequate individuals. In the final pseudo-panel dataset, the observations consist of the mean value of these cohorts i.e. each observation in variables is formed from the average value of cohort c and time t.

$$y_{it} = \beta + x_{it}\varphi + \alpha_i + \epsilon_{it} \quad t= 1\ldots T$$

In above simple model represents the exogenous and endogenous variables respectively. More elaborately, is a vector of exogenous variables, where i stands for individuals observed (LFS Pakistan) and t stands for the time period involved and β is the scalar. Furthermore, shows the individual-level effects which are unobservable and constant over time and indicates the unobserved residual effects the change over time and individuals.

On the basis of the explanation provided by Deaton (1985), city and industry are perceived as cohorts such that each individual in a city belongs to their own industry, exclusive from any other individual for every respective time period involved.

In the following equation, Least Squares Dummy Variable (LSDV) model is assessed while using pseudo panel data. Averaging over the cohorts gives

$$\bar{y}_{ct} = \bar{x}_{ct}\varphi + \bar{\alpha}_{ct} + \epsilon_{ct} \quad c=1\ldots C \quad (2)$$

Let the size of the group c at time t, be z_{ct} such that the average observation of all included x_{it} in cohort c at time t can be written as

$$\bar{x}_{ct} = z_{ct}^{-1} \sum_{i \in c}^{s=t} x_{is} \quad (3)$$

This result in pseudo-panel data set consist of repeated observations over time periods T and cohorts C. Moreover, it includes explanatory variables with cohorts' dummy set. It is assumed that each cohort comprises a measurable component, also this component is fixed over the time period. The core issue in the equation (2) estimation is the possible correlation among which is dependent on t and. Treatment of term as a random error would raise the problem of estimator's inconsistency, though consideration of it in term of fixed unknown parameters would give the identification issue in estimation unless the ignorance of time variation is possible. This assumed ignorance only holds when the cohorts are average from our large observation. Verbeek and Nijman (1993) have altered Deaton's estimator to attain consistency for a fixed t and i per cohort. If the individuals in each cohort are large enough, then the size of the cohort approach to infinity, and as result the measurement error approach to 0. This assumption results in the estimator β of cohort, which is similar to Deaton (1985) provided estimator

$$\hat{\beta} = \left(\sum_{c=1}^{C} \sum_{t=1}^{T} z_{ct} \, (\bar{x}_{ct} - \bar{x}_c)^2 \right)^{-1} \left(\sum_{c=1}^{C} \sum_{t=1}^{T} [z_{ct} \, (\bar{x}_{ct} - \bar{x}_c)(\bar{y}_{ct} - \bar{y}_c)] \right)$$

Where $\bar{x}_c = (\sum_{t=1}^{T} (z_{ct})^{-1} \sum_{t=1}^{T} \bar{x}_{ct} z_{ct})$ the average that is fixed over time for the cohort c. The β estimator properties are reliant on the asymptotic properties under consideration. According to the explanation, Moffitt (1993) has suggested the number of cohorts that should be fixed, meanwhile the size of cohorts should grow with the increase in the number of individuals. This will deal with the variables error problem as per Moffitt.

In contrast to the actual panel, (where the same individual is observed in cross-section over time) which normally consists of time and individual, the segregation of pseudo panel data is further expanded into the observation size in each cohort and the number cohorts. Moreover, whether the number of observations in the cohort should be large or the number of cohorts, is debatable. Because each choice has some opportunity cost. In case the choice is large observation in each cohort, but a small number of cohorts could lead to individual-level heterogeneity. In the case of other selection i.e., a few observations and large cohorts could result in an inefficient estimator. In the best-case scenario, the selection should be taken place in such a manner that provides the homogeneity in cohorts along with its adequate size. Providentially, with the LFS observation size and selection of industries and cities as cohorts, this standard is fulfilled. The selection of years 2014-15 and 2017-18 is done due to two reasons, first, the alteration in the pattern of agglomeration (firm location decision) is gradually realized over time, and second, the availability of 4 digit level industrial classifications which is crucial for proxy in variables, is limited to these years. Also, the 2 digit level industries are considered as an industrial cohort, which is 21 in number, and the city cohort is 14 in number as per the classification of LFS Pakistan. Furthermore, the number of industries in each city is different across cities, but constant over time which results in strongly balanced data. The pseudo panel data building is accomplished by calculating the cohort mean in each available cross-section. Thus, the resulting cells after cohort mean computation gives by city industries for two years, where cells are defined as industry codes.

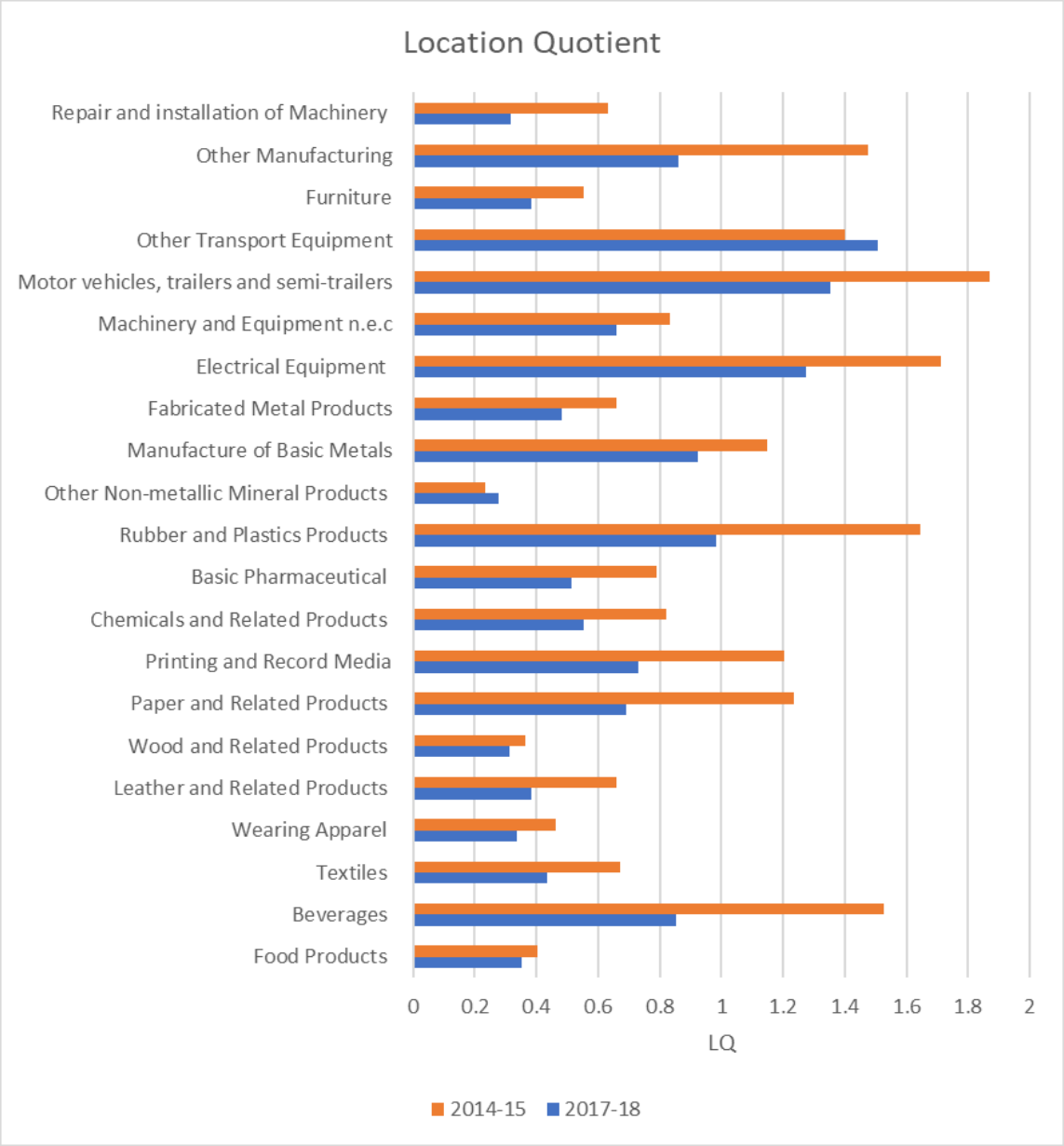
4. Results

This section begins with the descriptive statistics of the prevailing condition of agglomeration, labour pooling and variables affecting agglomeration which is followed by the models' estimation.

4.1 Evaluation of Location Quotient and Labour Pooling across Manufacturing Sector/Industries for the year 2014-15 and 2017-18 - Average across Cities

In this subsection, the average values of the location quotient and labour pooling for manufacturing sector industries are presented.

Figure 3: Location Quotient Values of Manufacturing Sector Industries for 2014-15 and 2017-18

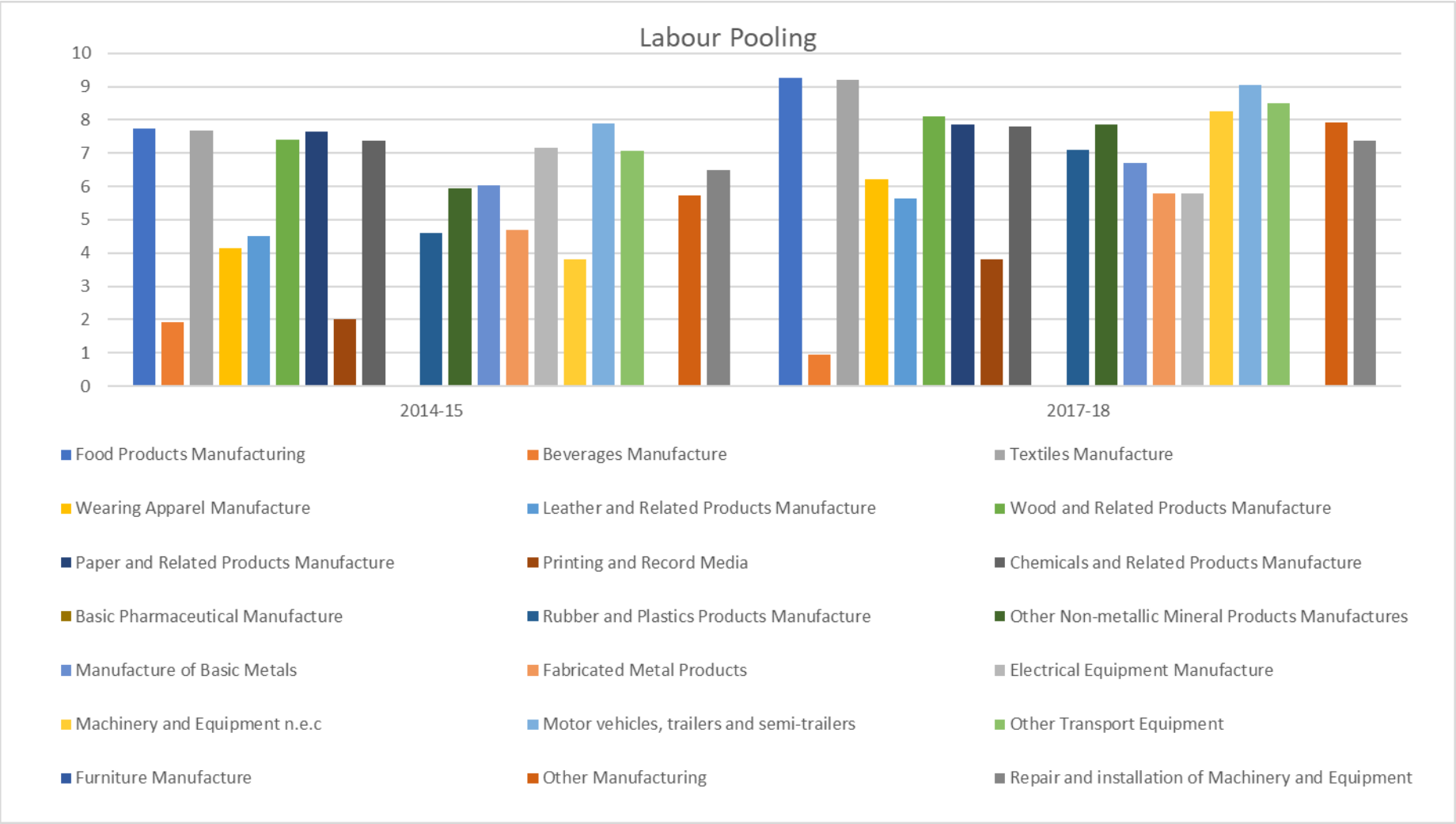


Source: Author's Estimation

The figure above shows the Location Quotient values for each industry. It is actually the industry-wise LQ, average across cities. The sole purpose of averaging out the LQ value

is to get a brief representation of the detailed information. For most of the industries the value of LQ has decreased during the time period of 2014-15 and 2017-18, as indicated earlier. The industries show the high value of LQ in results, representing the high spatially concentrated industries include “Beverages”, “Motor Vehicles, Trailers and Semi-trailers”, “Electrical Equipment” and “Rubber and Plastics products”. Among all the LQ values that have declined over time, mostly industries showed a minor decline, only the industries such as “Other Manufacturing”, “Beverages”, “Motor Vehicles, Trailers and Semi-trailers”, and “Rubber and Plastics Products” showed a sharp decline. This finding, as mentioned above, reassure the prevailing conditions as “Motor Vehicles, Trailers, and Semi-trailers” is concentrated in few cities only. Though two industries show a slight increase in the LQ value such as “Other Transport Equipment” and “Other Non-metallic Mineral Product”. As stated above, each value of LQ by industry, is actually the average of 14 cities, hence, the industrial value might vary. For having the slightest idea of the average labour pooling by industry, the next figure depicts its visual representation.

Figure 4: Labour Pooling in the Manufacturing Sector/Industries for 2014-15 and 2017-18



Source: Author's Estimation

The table above shows the labour pooling values by industry. In 2014-15, “Motor vehicles, trailers and semi-trailers” industry has the highest value of labour pooling. Also, it is evident from the finding above that “Motor vehicles, trailers and semi-trailers” is the most spatially concentrated industry. Hence, this is pointing out that one of the reasons for its spatial concentration is labour pooling. Similarly, industries like “Electrical Equipment Manufacture”, “Rubber and Plastics Products Manufacture”, “Other Transport Equipment” and “Other Manufacturing” have both one of the highest LQ values along with large values of labour pooling. Also in 2017-18, “Food Products Manufacturing” has the highest labour pooling values than other industries. Though the LQ value of “Food Products Manufacturing” is not quite high, supporting the view that for this particular industry dispersion is more beneficial despite having high labour pooling. For both years, the industries such as “Basic Pharmaceutical Manufacture” and “Furniture Manufacture” depict no sign of labour pooling. Likewise, “Beverages Manufacture” and “Electrical Equipment Manufacture” industries indicate a decline in labour pooling over the years. All other industries show various degrees of increase in labour pooling, except for 4 industries. Furthermore, the considerable rise can be evident in the “Machinery and Equipment n.e.c” industry. While “Paper and Related Products Manufacture” industry shows a very small increase relative to other industries.

The next heading shows the first model regressed in this research work.

4.2 Empirical Model Employing Low-skilled Workers and Labour Pooling Based on both, Firm to Industry Pooling and Tertiary Level Educational Attainment

Table 1: Regression Results of Model 1

	Model (4)	
Dependent Variables	Coefficient	Std Dev
Labour Pooling	0.074621	(0.01806)***
Percentage of Workers with Tertiary Education	0.602116	(0.45353)*
Percentage of the Workers Low-skilled	1.392155	(0.40670)***
Percentage of Workers with On the Job Training	0.169742	(0.27570) *
Average Age	0.246445	(0.05990)***
Square of Average Age	-0.00336	(0.00081)***
Percentage of Married Workers	-0.44782	(0.24820)*
Percentage of Migrants	-0.45887	(0.21400)**
Size of Informal Sector (Average)	0.128177	(0.27313)
Industry Dummies	Yes	
City Dummies	Yes	
Year Dummy	Yes	
Constant	-4.93031	(1.23343)***

Number of obs	316
Prob > F	0
R-squared	0.4005

Source: Author's Estimation
"Note: * = 1% significance level, ** = 5% significance level and *** = 10% significance level"
Dependent variable: Location Quotient

The outcome from the regression of the industrial and city-specific characteristics on the Location Quotient (used as agglomeration measure) of industries in the specific cities is presented in model 1 of Table 5.8. The core concern of this analysis is about the relationship between labour pooling and spatial concentration of industries. One of the complex things in micro-foundations illustrate by Marshall, is to find a proxy for labour pooling. If the industry in the region experience labour pooling then it can reap benefit through agglomeration, because it will be in their reach to able to hire labour with industry-specific skill. The difficulty with using a proxy to trace pooling in the labour market that it is hard to trace down. In the model above, the two types of proxies of labour pooling have been used. First, the firm to industry labour pooling proxy suggested in the work of Overmen and Puga (2010). The results come out as per the anticipation of the theoretical framework of this model. The coefficient of labour pooling variable turns out positive and significant at the 1 % level significance, following the path of previous studies (Overmen and Puga, 2010). Hence, industries in cities where the firms on average experience more idiosyncratic shocks than others, happen to be more spatially concentrated than other industries. The second proxy used in this model, is proposed by Rosenthal and Strange (2001) i.e. the percentage of workers with different education level and professional qualification (Burki and Khan, 2010). Though, this model only includes education levels with degrees only that is graduation and above. Resultantly, the table above shows the positive relationship among workers with tertiary education and industry's spatial concentration at the 10 % significance. This is according to the conceptual point of view that graduation level and above degree provide field-specific knowledge and add skills and as a result, promoting agglomeration. The industries that are located in 14 cities under analysis, and in general in Pakistan, are not highly specialized and technology-based and usually use the abundant Low-skilled labour in the labour market which also efficient on the cost side. Thus, the percentage of low-skilled workers' coefficient turns out positive and significant. The other variable of on-the-job training provides a proxy for the knowledge spillover. Industries spatially concentrate more to gain the benefit of knowledge spillover. And table also shows the percentage of on-the-job training has a positive and significant coefficient. For controlling variables, age, age square, migrants and married women turns out significant.

The year, regional and industrial dummies included in the model above (Kuncoro and Wahyuni, 2009). After the inclusion of industrial dummies in the model, the overall model upgrades as it has better the R square.

5.3 Empirical Model Employing Skilled Workers and Labour Pooling Based on both, Firm to Industry Pooling and Tertiary-Level Educational Attainment

Table 2: Regression Results of Model 2

Dependent Variables	**Model (3)**

	Coefficient	Std Dev
Labour Pooling	0.0793933	(0.01902)***
Percentage of Workers with Tertiary Education	0.6364229	(0.33162)**
Percentage of the Workers Skilled	1.056899	(0.47971)**
Percentage of Workers with On the Job Training	0.1692964	(0.26361) *
Average Age	0.2386646	(0.05573)***
Square of Average Age	-0.0032609	(0.00078)***
Percentage of Married Workers	-0.6907477	(0.26416)***
Percentage of Migrants	0.6776337	(0.28632)**
Size of Informal Sector (Average)	-0.0275786	(0.25567)
Percentage of Workers with Literacy	-0.098	(0.5031)
Industry Dummies	Yes	
City Dummies	Yes	
Year Dummy	Yes	
Constant	-3.89135	(1.00769)***
Number of obs	316	
Prob > F	0	
R-squared	0.4009	

Source: Author's Estimation
"Note: * = 1% significance level, ** = 5% significance level and *** = 10% significance level"
Dependent Variable: Location Quotient

The table above shows the second model of this study. It follows the same variable setting as first model but instead of the percentage of low-skilled workers, this model employs the percentage of skilled workers. This helps in enlarging model understanding in different settings and the degree of result consistency. The main variables of labour pooling which is based on firms to industry pooling turns out positive, as well as significant at 1 %. Furthermore, the coefficient of second proxy based on educational attainment turns out positive and significant at 5 %. Also, the proxy for knowledge spillover used in this work i.e., on-the-job-training shows direct relation with industrial agglomeration. The coefficient of the percentage of skilled workers has a positive sign with significance level of 5 %. Pointing out the fact that for Pakistani industries though low-skilled workers have considerable importance as the type of industries country has but skilled labour also serves as foundational factor for industrial agglomeration (T. Choi). Controlling variables result remain as all the models above. Finally, the industrial dummies mostly give significant results and regional dummies outcome remains more or less as model 1.

5. Conclusion and Policy Recommendations

In this section of the study, the conclusion has drawn regarding various results discussed above. Also, the recommendation that this study implies, is also discussed along with guidelines for the future studies.

5.1 Conclusion

This research study focuses on exploring the labour pooling as foundation for agglomeration in Pakistan and also how workers with different skill level works as source of industrial agglomeration in the manufacturing industries of Pakistan. In this study, the Labour Force Survey of Pakistan for the year 2014-15 and 2017-18 is used. However, due to the unavailability of data at firm level, 4 digit industry codes are used as a proxy for plants/firms and 2 digit industry classification as the industry.

Also, as per the title suggests this study is elaborating the determinant of industrial agglomeration. As highlighted by Marshal (1920), firms that chose to locate close to each other reap the benefits of positive externalities which is basically a cost advantage attained when firms agglomerate. For empirically assessing geographic concentration on a more elaborative and large scale with available data i.e., Location Quotient, which is used in this study as suggested by Kuncoro and Wahyuni (2009) and Szakálné, Kanó and Vas (2013). The main emphasis among determinants, as per the title is on labour pooling. From 1920 when Alfred Marshall addressed labour pooling as the foundation for agglomeration, the urban literature has made various attempts to trace labour pooling through proxies. In the literature of industrial agglomeration, the various researchers try to trace down the effect of labour pooling through the different measure. If a labour market in an industry in a city goes through pooling, that industry can benefit by agglomerating due to better job skill matching. To get a better understanding of labour pooling, two variant proxies are used from the notable work of Overmen and Puga (2010) and Rosenthal and Strange work (2001). First proxy is from Overmen and Puga (2010), it proposes that firm to industry labour pooling according to which in particular industry and location when a firm raises employment simultaneously other firms there shrinking employment and vice versa. The second proxy of labour pooling is adopted from Rosenthal and Strange (2001) work which is the percentage of workers with different educational qualifications. The empirical models result of the proxy given by Overmen and Puga (2010) shows positive and significant in all the models. Moreover, proxy given by Rosenthal and Strange (2001) gives positive significant outcomes in both models as well. Pakistan as a whole has an abundance in the labour force, but mostly it consists of low-skilled or unqualified workers. Also, the locating industries are not high technology, invention, and innovation-based. Lack of technological advancement in industries is a very much result of a low-skilled and unqualified worker in high percentage. However, this abundance of low-skilled labour drives down wages and as a result increase profit margin. Also, the type of industries that exist right now in Pakistan is highly inclusive of these types of labour due to their availability and abundance. Hence, the variable of the percentage of low-skilled workers turns out positive and significant. However, this analysis suggest skilled workers also hold as much importance as low-skilled workers in the process of explaining industrial agglomeration's sources. Further, the variable of the percentage of on-the-job training is used as a proxy for the knowledge spillover. As per Marshall (1920) the explanation, the industries spatially concentrate more to avail knowledge spillover advantages. Thus, the resulting coefficient's

positivity gets lined up with the prior statement regarding knowledge spillover. The year, regional and industrial dummies included in both models (Kuncoro and Wahyuni, 2009). After the addition of industrial dummies in the model, overall models upgrade as it increases the R square. Moreover, the post estimation tests result supports the inclusion of dummies (Kuncoro and Wahyuni, 2009).

Finally, Exploration of the different determinants of industrial agglomeration gives a valuable understanding of possible sources of cost and productivity advantages that agglomeration offers. Also, it highlights the course of industrial dynamics that critical in explaining the firm pattern of decision making, however, the relevant sources of agglomeration in developed and developing countries might vary due to differences in the type of the industries as this study suggested. The labour pooling and knowledge spillover implication in empirical models indicates the relevancy of this concept. Though, the labour pooling, which is measured through educational attainment in this study, showed less significance except for the tertiary education share proxy.

5.2 Policy Recommendation

In the competing environment of the 21st century, any research study cannot emphasize enough on the manufacturing and industrial sector growth and development. Nonetheless, any theory new or old always emphasize the industrial base of the country as a way toward economic development. Because the spatial concentration gives a boost to "agglomeration economies" which allow firms to get production and cost advantages. These advantages consist of all the things discussed above like labour pooling, knowledge spillover, geographical characteristics like quality and quantity of inputs, available and explored chain of supplier and collaborators.

The promotion of industry agglomeration is rather larger in scope than it seems. Also, competitiveness triggers the knowledge flow and drive for betterment. Though, with benefits, the cost follows in terms of inflation (demand push inflation i.e., high land prices) and congestion, etc. Till the benefit remain higher than the cost the industrial agglomeration that also gives a boost to urban agglomeration (in term of city growth) will continue to grow, but the cost will be increasing more and more until it outweighs the benefit. Thus, after this worker and firms choose to locate in different locations.

In this context, this study recommends that:

- In Pakistan, it has been observed that in some cities due to the large pool of labour and better infrastructure than other cities firm choose to locate more. However, with time the no development in available cities and no growth of new cities and urban areas, the cost in these few cities/regions is outweighing the benefit such as Karachi, Lahore, and Gujranwala, etc. In the current state, Pakistan is in dire need of new urban-center/cities as well as the development of current ones. This motive can be achieved through better infrastructure in areas where the government wants to promote the agglomeration of workers and industries.

- The sources of manufacturing sector agglomeration result suggest that workers with tertiary education (tertiary education as described in this study) attract the agglomeration. Various levels of higher study incentives should be provided by the government to increase tertiary education share. These scholarships should be field-specific as per industry requirements. And internships as per educational year should

be given to the student, it will give a boost to experience and qualified labour availability.

- To increase employment and manufacturing sector growth, regions unemployment level should also be determined. After it, to assess the type of pool available various surveys should be conducted. Then, as per the available pool, horizontal industry promotion policies should implement. It will be favorable for both the region and industry. It will use the local pool of labour, also it increases region's employment, and industries will gain profit through low price labour available and attract other industries.

- Moreover, the authorities can themselves prompt type of the skill development program in a labour pool that is in demand by industries available. The government can also give incentives to industries for conducting the practical skill development program. This will work in generating labour pooling and knowledge spillover,

- Finally, another policy point for the recommendation that can conclude our study is government, authorities, and practitioners can boost the economic activity level in the underdeveloped or targeted area by providing the cost incentives to firms to locate there till the agglomeration benefit starts happening.

5.3 Future Directions

In regard to future recommendation of this study, as mentioned earlier, it is evolving concept in the urban literature and researchers are still trying to identify the correct or best measure depicting the theory. This work can be extended by expending regions, for instance district level analysis. Also, expansion of this work can be done through the inclusion of determinant other than labour pooling as a main focus and by enlarging the scope of industries that is covering the whole industrial sector. Though, there is a dire need for the expansion of data collection and publication regarding industries on micro-level, especially in Pakistan. The understanding and implication of this study can be broadened through the micro-level and time consistent data. Also, since the concept is still under major rebuilding over a century now, there is still a dire need of evaluation of the distinction in sources of agglomeration in the case of developed and developing countries.

References

- Angel, D. P. (1991). High-Technology Agglomeration and the Labor Market: The Case of Silicon Valley. Environment and Planning A: Economy and Space, 23(10), 1501–1516.
- Behrens, Brown & Bougna, (2013). The Determinants of Agglomeration Redux: Transportation, Trade, and Input-Output Linkages, ERSA conference papers ersa14p32, European Regional Science Association.
- Boasson, E., & Boasson, V. (2011). The spatial-temporal cluster development in the healthcare service industry–an integrated GIS approach. Environmental economics, (2, Iss. 3), 131-140.
- 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.
- Burki, A.A., Khan, M.A., (2013). Agglomeration economies and their effects on technical inefficiency of manufacturing firms: evidence from Pakistan. Sciences.
- Chaudhry, T. T. (2005). Industrial clusters in developing countries: A survey of the literature. The Lahore Journal of Economics, 10(2), 15-34.
- Chaudhry, T. T. (2010). Relational contracting in Pakistan's surgical instrument cluster: an empirical study. The Pakistan Development Review, 213-237.
- Chaudhry, T. T. (2011). Contracting and Efficiency in the Surgical Goods Cluster of Sialkot, Pakistan. South Asia Economic Journal, 12(1), 91-115.
- Chaudhry, A., Nasir, M., & Haroon, M. (2012). The Birth of Exporters: Entry and Scale of Firms in Punjab's Export Sectors. The Lahore Journal of Economics, 17, 191.
- Combes, Pierre-Philippe & Duranton, Gilles. (2006). Labour Pooling, Labour Poaching, and Spatial Clustering. Regional Science and Urban Economics.
- Combes, P. - P., G. Duranton, and L. Gobillon. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63 (2): 723– 42.
- Deaton, A. (1985). Panel data from time series of cross-sections. Journal of econometrics, 30(1-2), 109-126.
- Duranton, G., Puga, D., (2004). Micro-foundations of urban agglomeration economies. Handbook of Regional and Urban Economics, vol. 4. Elsevier,pp. 2063–2117.
- 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. (1997). Geographic concentration in U.S. manufacturing Industries: A dartboard approach. Journal of Political Economy. Vol. 105(5), 889- 927.
- Ellison, G., &Glaeser, E. L., (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?. The American Economic Review, Vol. 89 (2), 311-
- Ellison, G., Glaeser, E. L., & Kerr, W. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. American Economic Review. Vol. 100(3), 1195-1213.
- Faggio, G., Silva, O., & Strange, W. C. (2017). Heterogeneous agglomeration. Review of Economics and Statistics, 99(1), 80-94.
- Figueiredo, O., Guimaraes, P., & Woodward, D. (2009). Localization economies and establishment size: was Marshall right after all?. *Journal of Economic Geography*, *9*(6), 853-868.

- Florence, P. S. (1939). Report on the Location of Industry. Political and Economic Planning, London, UK.
- Fracasso, A., & Vittucci Marzetti, G. (2017). Estimating dynamic localization economies: the inadvertent success of the specialization index and the location quotient. Regional studies, 52(1), 119-132.
- Fujita, M. & Thisse, J.F. (1996). Economics of Agglomeration. Journal of the Japanese and International Economies. Vol. 10, 339-378.
- Gabe, T. M., & Abel, J. R. (2010). Labor market pooling and occupational agglomeration. FRB of New York Staff Report, (392).
- Gardezi, Z. G., (2013). "Labor Pooling as a Determinant of Industrial Agglomeration," CREB Working papers 4-2013, Centre for Research in Economics and Business, The Lahore School of Economics, revised 2013.
- Guimarães, P., Figueiredo, O., & Woodward, D. (2009). Dartboard tests for the location quotient. Regional Science and Urban Economics, 39(3), 360-364.
- Helsley, R. W., & Strange, W. C. (2002). Innovation and input sharing. Journal of Urban Economics, 51(1), 25-45.
- Henderson, (1995). J. V., A. Kuncoro, and M. Turner, "Industrial Developmenti n Cities," Journal of Political Economy 103, 1067-1085.
- Henderson, J. V. (2003a). Urbanization and economic development. Annals of Economics and Finance, 4, 275–341.
- Hoover, E. M. (1948). *Location of economic activity*. McGraw-Hill Book Company, Inc., New York.
- Jacobs, J. (2016). The economy of cities. Vintage.
- Kanó, I. S., & Vas, Z. (2013). Spatial distribution of knowledge-intensive industries in Hungary. *Transition Studies Review*, *19*(4), 431-444.
- Kim, S. (1995). Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987. *The Quarterly Journal of Economics*, *110*(4), 881-908.
- Krugman, P. R. (1991). *Geography and trade.* Cambridge, MA: MIT Press.
- Kuncoro, M., & Wahyuni, S. (2009). FDI impacts on industrial agglomeration: the case of Java, Indonesia. Journal of Asia Business Studies.
- La Fountain, Courtney. (2005). Where Do Firms Locate? Testing Competing Models of Agglomeration. Journal of Urban Economics.
- Liu, Z. (2014). Global and local: Measuring geographical concentration of China's manufacturing industries. The Professional Geographer, 66(2), 284-297.
- Haroon, M. and Chaudhry, A. (2014). Where Do New Firms Locate? The Effects of Agglomeration on the Formation and Scale of Operations of New Firms in Punjab. Economics Discussion Papers, No 2014-21, Kiel Institute for the World Economy.
- Marshall, A. (1890). *Principles of economics.* London: Macmillan.
- Marshall, A., (2009). Principles of Economics: Unabridged Eighth Edition. Cosimo, Inc.
- Marcon, E., & Puech, F. (2009). Measures of the geographical concentration of industries: Improving distance-based methods. Journal of Economic Geography, 10(5), 745–762.

- Moffitt, R. (1993). Identification and estimation of dynamic models with a time series of repeated cross-sections. Journal of Econometrics, 59(1-2), 99-123.
- Monseny, López, & Marsal,. (2011). The Mechanisms of Agglomeration: Evidence from the Effect of Inter-Industry Relations on the Location of New Firms. Journal of Urban Economics.
- Moomaw, R. L. (1983). Is population scale a worthless surrogate for business agglomeration economies?. Regional Science and Urban Economics, 13(4), 525-545.
- Moomaw,R . L., (1981), "Productivity City Size: A Critique of the Evidence," Quarterly Journal of Economics 96 675-688.
- Nakamura R,., (1985), "Agglomeration Economies in Urban Manufacturing industries :A Case of Japanese Cities," Journal of Urban Economics 17 108-124.
- Nasir, M. (2013). Agglomeration and Firm Turnover.
- Overman, H.G., Puga, D., (2010). Labor pooling as a source of agglomeration: an empirical investigation. Agglomeration Economics, vol. 2010.University of Chicago Press, pp. 133–150.
- Ozturk, Selcen & Kilic, Dilek. (2016). Do Firms Benefit From Agglomeration? A Productivity Analysis For Turkish Manufacturing Industry. Ekonomik Yaklasim. 27. 131. 10.5455/ey.35907.
- Puga, Diego. (2010). The Magnitude and Causes of Agglomeration Economies. Journal of Regional Science. 50. 203-219. 10.1111/j.1467-9787.2009.00657.x.
- 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 organization, and agglomeration. The Review of Economics and Statistics, 85(2), 377-393.fsax
- Rosenthal, S. S., and W. Strange. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, vol. 4, ed. V. Henderson and J.- F. Thisse, 2119– 71. Amsterdam: North- Holland.
- Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, MA: Harvard University Press.
- Segal, D. (1976). Are there returns to scale in city size?. The Review of Economics and Statistics, 339-350.
- Shefer, D. (1973). Localization economies in SMSA'S: a production function analysis. Journal of Regional Science, 13(1), 55-64.
- Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations: Volume One. London: printed for W. Strahan; and T. Cadell, 1776.
- Sobrinho, Ednaldo & Azzoni, C. (2016). Potencial inovativo da indústria nas regiões brasileiras. Revista Brasileira de Inovação. 15. 275. 10.20396/rbi.v15i2.8649131.
- Strotebeck, F. (2010). The location quotient–assembly and application of methodological enhancements.
- Verbeek, M., & Nijman, T. (1993). Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections. Journal of Econometrics, 59(1-2), 125-136.
- Zhao, Y., & Mudassar, M. (2013). A Research on Industrial Structure Evolution and Adjustment in Pakistan. *International Journal of Business and Social Science*, *4*(9).