Nature & Effects of Agglomeration Economies and Manufacturing Firms in Punjab

PhD Thesis

By

Mutee Ul Rehman

PhD Student

School of Economics,

Quaid-i-Azam University, Islamabad

2024

Nature & Effects of Agglomeration Economies and Manufacturing Firms in Punjab

By

Mutee Ul Rehman

PhD Student

Supervisor:

Dr. Muhammad Tariq Majeed
Professor
School of Economics
Quaid-i-Azam University, Islamabad

Submitted in partial fulfillment of the requirement for the
Doctor of Philosophy in Economics
at the School of Economics, Faculty of Social Sciences,
Quaid-i-Azam University, Islamabad
May 2nd , 2024

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, titled **"Nature & Effects of Agglomeration Economies and Manufacturing Firms in Punjab"** was conducted by **Mr. Mutee ul Rehman** under the supervision of **Dr. M. Tariq Majeed,** Professor, Director School of Economics.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the school of Economics, Quaid-i-Azam University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the field of Economics, School of Economics, Quaid-i-Azam University, Islamabad.

Student Name: Mutee ul Rehman **Signature:** ____________

Examination Committee:

a) External Examiner

Dr. Eatzaz Ahmad
Professor
Iqra University, Islamabad **Signature:** ____________

b) External Examiner

Dr. Muhammad Iqbal
Professor (R)
Pakistan Institute of Development Economics (PIDE), Islamabad **Signature:** ____________

Supervisor:
Prof. Dr. M. Tariq Majeed
Professor
School of Economics
Quaid-i-Azam University, Islamabad **Signature:** ____________

Director:
Prof. Dr. M. Tariq Majeed
Professor & Director
School of Economics
Quaid-i-Azam University, Islamabad **Signature:** ____________

Declaration

I, Mutee Ul Rehman, S/o Hafiz Abdul Rehman, registration no. 03091411011, candidate for PhD Economics at School of Economics, Quaid-i-Azam University, hereby declare that this thesis, titled "**Nature & Effects of Agglomeration Economies and Manufacturing Firms in Punjab"**, submitted for the partial fulfillment of Doctor of Philosophy (PhD) degree in Economics, represents my original work. This research work has not been formerly included in any dissertation submitted to this or any other institution for a degree.

Mutee Ul Rehman

Reg No. 03091411011

Dedicated to my sons, Amar & Hussain

Acknowledgement

This thesis is a cumulative outcome of a few human traits such as resilience, dedication, support, hard work and professional experience. Honestly, I might not have completed it if any of these traits were not available to me.

My supervisor, Prof. Dr. Tariq Majeed, played a pivotal role during the entire process of this thesis. His level of sincerity, integrity and patience makes him an ideal supervisor and mentor. I must acknowledge his unrelenting guidance and continuous support despite my quirkiness. He taught me the basics of how to write academic research and helped me transform the idea of this thesis into a feasible research study.

I must acknowledge the support and technical inputs I received from Dr. Eatzaz Ahmad. His approachability and willingness to guide makes him a pillar of strength for those under his apprenticeship. Through his guidance, students not only gain academic prowess but also develop critical thinking and advance data analytics skills. Dr. Eatzaz's commitment to his students' success and his passion for teaching makes him a national asset for the academic community of Economics in Pakistan. He taught me the correct use of econometrics, especially spatial econometrics, for my theoretical model of this thesis.

I express my sincere gratitude to all the administrative staff of the School of Economics of the Quaid-i-Azam University. It is me and them who know how we made the finalization of this thesis possible. Mr. Asif Khan, Mr. Sajid Mehmood, Mr. Zahid and Mr. Ikram Ullah. They were all really very helpful throughout the PhD.

Table of Contents

List of Tables

List of Figures

Abstract

This research quantifies the nature of agglomeration economies and its effects on the annual turnover of the manufacturing firms located in Punjab, a province of Pakistan. Using firm level cross-section data, this research first explains the nature of agglomeration within manufacturing firms in Punjab by calculating the agglomeration indicators such as urbanization and localization. We used the firm's distance from urban centre as a proxy for urbanization. We used spatial autocorrelation as a measure to show the localization of firms. We constructed the translog production function to measure the effects of urbanization and localization on firms' annual turnover. This research has used spline regression to measure the impact of distance of the firm from urban centre on the annual turnover of the firm.

The overall impact of both agglomeration economies is positive on firms' turnover. However, the effect of urbanization economies is more prevalent than the effects of localization economies in Punjab. At two-digit sector level, thirteen out of twenty-three manufacturing sectors showed positive effect of urbanization on firm's turnover. For localization, seven sectors showed positive effect on firm's turnover. Spline regression shows there is a negative relation between distance of the firm from urban centre and the annual turnover. Firms' annual turnover attenuates with the increase in its distance from urban centre. The results are consistent with the literature and the existing status of manufacturing sector of Punjab.

CHAPTER 1

1. Introduction

In the recent decades, economic activity concentration has increased considerably all over the world. The concentration or clustering refers to a group of similar or related firms or business that account for a substantial proportion of economic activity and play an active role in the success of regional development. The economic activity concentration is of great importance as it helps the potential firms/businesses to decide where to locate themselves to gain the competitive advantage. This concentration in a particular region tends to enhance the economic activity. This relation is bi-directional. It is intriguing to know the reasons and outcome of the concentration of economic activity at a specific location. Enormous literature has suggested various reasons for concentration of economic activities in a specific region (Henderson 2003, Duranton 2005, Krugman 1991, Baldwin 2008, Overman and Puga 2010).

Literature suggests that economic activity concentration is a productive phenomenon. It tends to increase the advantage of the related firms in that particular cluster through positive externalities i.e., labor pooling, knowledge spillover and input sharing (Hill & Brennan 2000). The importance of the concentration of economic activity in a particular region cannot be overemphasized, as it is known to contribute significantly to the growth of cities and different industrial sectors.

Historically, the industrial clusters have evolved differently in different countries with specific indigenous development characteristics. In China, many industrial clusters have been formed and the government had a lot to do with their development process. Many of the industrial clusters in China have been formed after the open-door policies and reforms that allowed for the foreign investment to enter and private sector to grow in China (Zeng 2013). High technology clusters like those in U.S. (Silicon Valley in California; Austin, Texas; and Research Triangle Park in North Carolina) and Europe (technological parks in Sophia-Antipolis, France) are the example of how different clusters have emerged based on their positive externalities. As pointed out by (Markusen, 1996), Silicon Valley hosts all four kinds of clusters which she identified in her study i.e., Marshallian Industrial Districts, Hub-and-Spoke, Satellite Industrial Platforms and State-Anchored industrial Districts.

We have explained the development of industrial clusters on the basis of specific development characteristics i.e., they might be formed because of the conducive national policies or to take the advantage of the surrounding externalities. However, the industrial clusters can also evolve over time such that the type of industries within that cluster may be alteterd over time reflecting the technological development (Montana & Nenide, 2008).

The above discussion on the clustering/concentration of economic activities explains the broader and bigger phenomenon famously known as agglomeration economies. The concentration of economic activities in a particular region/area with the intention of taking competitive advantage of resources, input sharing and labor pooling is known as agglomeration economies. The idea of agglomeration economies has been studied extensively over the last century. Alfred Marshall (1920), though he did not explicitly used the term “Agglomeration Economies” first presented the idea in which he discussed the factors of concentration of economic activities. Building upon the Marshall’s work many researchers have contributed to the knowledge of regional development. Other notable researchers include Jacobs, 1969) and (Krugman, 1991) that introduced another dimension into the agglomeration literature which states that regional development can be achieved through cities and core-periphery. The benefits of agglomeration economies have been discussed extensively in literature (Kukalis 2010, Glaeser and Mare 2001, Holmes 1999). Agglomeration economies offers a lot of benefits to the firms located in that particular region that are not limited to only gaining productive advantages (Jaffe et al, 1993).

Literature suggests that the benefits of agglomeration economies can be accrued because of one of the two reasons i.e., localization economies and urbanization economies. (Alfred Marshall, 1920) and (Jacobs, 1969) presented the ideas of localization economies and urbanization economies. While localization economies provide with the idea that firms benefit from the cluster of similar firms based on labor pooling, input sharing and knowledge spillover, urbanization economies focus on the benefits accrued because of proximity to the urban centres. Enormous literature has tested and found significant and positive results on the types of agglomeration economies. Different studies have studied different scenarios to test for the effects of localization and urbanization on agglomeration economies (Gardezi 2013, Henderson et al 1995, Costa & Kahn 2003, Overman & Puga 2010).

As explained in the literature, the nature and type of agglomeration economies can be studied for a specific region. Punjab is one of the most important and populous provinces of Pakistan, consisting of more than half of country's population (52.9% of the country's population). The province contributes significantly to the agriculture sector of the country in fact, the contribution of agriculture sector of this province is greater than other provinces of the country. On the other hand, the industrial sector of the province cannot be ignored as Punjab is amongst the most industrialized provinces of Pakistan (detailed discussion in chapter 3). The manufacturing industries of the province are producing textiles, electrical appliances, surgical instruments, sports goods, machinery, processed food and many other essential items. The industrial sector of the province plays an integral role in the economy of the country as it contributes significantly to the national income of the country and employs a large number of labor force.

As classified by Pakistan Bureau of Statistics, there are about 23 major sectors under the manufacturing activities (see table 5.1). However, in case of Punjab province, only 6 of those 23 majors sectors dominate and constitutes about 69% of the total manufacturing units. The manufacturing firms in Punjab are known to be highly clustered in different districts of the province and are enjoying the benefits of larger markets, large pool of skilled and professional labor and high road density. The regions where clusters are located have an enhanced market potential as compared to the regions where firms are far apart. One of the well-known clusters of manufacturing firms includes the manufacturing firms in Sialkot, Gujrat and Gujranwala districts, called the golden triangle. These districts are positioned in such a way that it looks like a geometrically shaped triangle. This clustered region is also known as the industrial hub of Punjab. Other famous cluster of manufacturing firms includes the agglomeration of textile industry in Faisalabad. Major known clusters of Punjab are located in Sialkot, Faisalabad, Gujrat, Gujranwala, Lahore, Rawalpindi, Sheikhupura, and Wazirabad. Since, the decision of where the newly entering manufacturing firms will locate themselves is based on the availability of resources like labor, market potential of the region etc.

This concentration of manufacturing firms intrigues many policy questions, e.g., why firms are so closely concentrated in specific regions? What kind of benefits these firms are accruing through this agglomeration? Are the firms taking benefits of urbanization aspect of agglomeration more or the localization? Is it the huge market size and population of the province which attracts the firms' agglomeration or is it the specialization of manufacturing sectors which has caused this? Which indicator of agglomeration (urbanization/localization) is dominant in contributing to the annual turnover of the firms? The problem is, so far, the existing literature on agglomeration of manufacturing in Punjab have not been able to answer these questions. They either use proxy variables to ascertain the nature of agglomeration or have used district level aggregated data (Burki and Khan, 2010). This research intends to fill that gap and will provide answers to these question by using advanced datasets and analytical methods.

Based on the above discussed theory and the agglomeration of manufacturing firms in the Punjab province, this research uniquely contributes to the existing literature and policy work for the Punjab province. First, there has not been any study which has explored the nature and effect of agglomeration economies and its impact on any dimension of manufacturing firms in Punjab using microdata. The existing research (Nasir, 2017) have used either district level aggregated data and indicators or have used proxies. However, this research used microdata set of manufacturing firms in Punjab disaggregated at sectoral level with agglomeration indicators. Secondly the existing research have used standard methods to measure agglomeration indicators which are heavily dependent on proxies. But this research has used geo-spatial point data and spatial econometric methods to calculate the agglomeration indicators. Hence, this research fills in the gaps in the literature and policy by using advanced tools, new econometric methods and geo-spatial point dataset, which makes this research unique and robust.

The significance of this research is that it spatially explore the impact of agglomeration economies on the annual turnover of the manufacturing firms in Punjab. This study will be analyzing the spatial concentration of manufacturing firms through spatial analysis i.e., spatial autocorrelation.

Previous studies have emphasized on estimating the spatial autocorrelation in terms of productivity and GDP (Rey and Montouri, 1999, Wang, Chang and Wang, 2019 & Stankov et al, 2017) where this study will be spatially analyzing the manufacturing firms in Punjab using spatial autocorrelation in order to understand the nature of agglomeration. In addition to that, a firm's level production function with agglomeration indicators (localization and urbanization) is employed in this research as proposed by (Nakamura, 1983). Such advance tools have not been used previously by any researcher who has worked on manufacturing firms of Punjab.

1.1. Objectives of the study

Manufacturing firms form clusters to take benefits from the similar type of industries or from the various types of industries. The question arises that why firms agglomerate in particular city or region? Does it matter to be agglomerated? What is the nature and type of agglomeration? If the firms agglomerate in one city or region on the basis of different characteristics, then it is important to know what is the impact of agglomeration on the annual turnover of the firms. The firms or industries also agglomerate on the basis of the sectors, so it is also required to evaluate the impact of agglomeration on the annual turnover of firms at disaggregated sector level.

This thesis will answer above-mentioned questions. The literature suggests that industries agglomerate to share common market, promote skills of workers and enable the exchange of suppliers. This study will evaluate the impact of agglomeration on the annual turnover of manufacturing firms in Punjab province. The spatial concentration of firms plays very critical role in the growth and development of cities and regional economies. The uniqueness of this research is the employment of such geo-spatial firm level data based analysis that has not be conducted previously in Pakistan. Therefore, this study will provide detailed empirical analysis on the nature and impact of agglomeration on the annual turnover of the manufacturing firms in Punjab.

There are two major objectives of this study.

1. First, this study will identify the nature and type of agglomeration in manufacturing sector of Punjab. This study will also identify the extent of urbanization and localization in the province.
2. Secondly, this study will explore the effect of agglomeration economies on the annual turnover of firms in Punjab, both at overall and at sector level.

Keeping in view the research objective, the following three research questions are addressed in this study:

1. What is the nature of agglomeration in the manufacturing firms of Punjab?
2. What is the effect of agglomeration economies on the annual turnover of the manufacturing firms in Punjab?
3. Which manufacturing sectors accrue more benefit from the agglomeration economies in Punjab?

The sequence of the thesis is as follows. This research will be critically reviewing the previous studies in Chapter 2 where the underlying economic theory behind the study will be discussed. In addition to that, a critical review of studies based on agglomeration economies and their effects on firms will be conducted. In Chapter 3, the study will be briefly discussing the economy of Pakistan and Punjab province in the context of manufacturing sector. Following that, a brief description on the sector-wise concentration of manufacturing firms in Punjab will be discussed. Chapter 4 discusses the theoretical framework and detailed methodology which have been employed in this research. The chapter then discusses the methodologies that this study will opt to achieve the stated objectives and research questions. In Chapter 5, the data source and the variables for agglomeration economies along with other independent variables are discussed. Chapter 6 provides with the results and discussions of the study. The chapter first provides the descriptive statistics then discusses the results of localization and urbanization indicators. Finally, in Chapter 7, the study concludes by explaining the extent and effects of localization and urbanization on the annual turnover of the manufacturing firms.

CHAPTER 2

2. Literature Review

This chapter provides extensive literature about the concept of agglomeration economies. As per the topic of this research thesis, this chapter will provide detailed literature review about the nature and effects of agglomeration economies. The following literature review starts with the seminal theories of agglomeration economies. It is important to start the literature with the seminal work and hardcore theories which provide the theoretical foundation to the research. Once the main theories have been explained, this chapter explains literature of other studies related to agglomeration economies. Since agglomeration economies are broadly categorized into two indicators i.e., localization and urbanization, the following literature review also provides the details on both indicators extensively. Finally, this chapter provides the literature about the effects of agglomeration economies and manufacturing firms in Pakistan.

2.1. Review of Theoretical Studies

The oldest literature which presents something seminal on agglomeration economies is the idea developed by Alfred Marshal in 1920. In his classical book "The Principles of Economics", Marshall identified three sources of industrial agglomeration i.e., input sharing, labor pooling and knowledge spill overs. He presented the idea that the productivity of the firms can be increased due to agglomeration economies, such as.

> "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 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 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. (1920, pp.271).

Thus, Marshall's idea of concentration of economic activity in a specific region/city becomes the foundation theory of the concept of agglomeration economies. Although he did not directly use the term "agglomeration economies", his concept of the concentration of the economic activity was later coined as agglomeration economies by other researchers. Marshall presented three factors which causes the concentration of economic activity. These are labor pooling, input sharing and knowledge spillovers. These factors are famously known as Marshallian externalities in the literature (Ellison & Glaeser, 1997, Rosenthal & Strange, 2004).

However, the Marshallian externalities only point out towards a specific kind of concentration of economic activities. It only explains the concentration of similar kind of economic activity in region/city. Marshall's idea of concentration of economic activity provided justification for the concentration of similar kind of economic activities or industrial clusters. Hence, according to Marshall, Marshallian externalities (labor pooling, input sharing and knowledge spillovers) encourage agglomeration of specific clusters. Such concentration of economic activities promotes a specific kind of clusters are called localization.

On the contrary, many researchers have argued that the agglomeration of economic activities may be because of the obtaining the benefits of environment and big cities. Hence the city size or its diversity may also contribute towards the concentration of economic activities. Almost 50 years after the Marshall's theory, (Jane Jacobs, 1969) presented another dimension of agglomeration economies. Jacobs defends the benefits of urbanization over localization. The urbanization agglomeration describes the benefits accrued because of the presence of other economic activities such as services sector, diversified labor skill set, consumption and population. Hence, whereas Marshall advocated for localization agglomeration and Marshallian externalities, Jacobs emphasized the role of urbanization agglomeration and city size as core factor of agglomeration economies.

Thus, agglomeration economies are defined as the concentration of economic activity in a region/city where people and firms benefit from each other through labor pooling, input sharing and knowledge spillover, city size, population, consumption and other economic activities. It can be divided into two broader types, i.e., localization agglomeration and urbanization agglomeration. Third major stream of work on agglomeration economies was presented by (Krugman, 1991). He wrote in his seminal work on economic geography about the concentration of economic activity in one region. He presented his theory of "core" and "periphery" and provides economic rationale for regional divergence. Based on the geographic concentration of manufacturing activities, Paul Krugman suggested that when transportation costs decline and the importance of economies of scale becomes more significant, the manufacturing firms tend to cluster in one or few "core" areas, while other regions take on the role of "periphery" or suppliers of agriculture products to the manufacturing core. The concentration of manufacturing occurs near the areas where there are large demand markets since it minimizes the transportation costs, and the markets will be large where the manufacturing is concentrated. Thus, Krugman advocated in favor of both localization and urbanization of agglomeration.

2.2. Agglomeration Economies: Impact of Urbanization Indicators

This part of the chapter provides extensive literature on the one indicator of agglomeration economies i.e., urbanization economies. In this part, detailed explanation on urbanization indicators of agglomeration economies including geographical proximity to the urban centres and affiliated positive externalities will be provided. We will review the studies explaining the effect of distance of the firms from the urban centres on various characteristics of manufacturing process (productivity, growth, wages, employment). Moreover, we will review the literature on country specific case studies on the impact of urbanization indicators.

Physical proximity is an important aspect of agglomeration as (Rosenthal and Strange, 2003) estimated the influence of distance on effect of agglomeration using firm level micro data. Although the impact of distance varies across sectors, overall, the benefits of agglomeration economies attenuate with distance. (Cainelli and Ganau, 2018) conducted a study to examine the role of neighboring firms characterized in agglomerated areas. The study aimed to examine whether the firms with varying characteristics have varying abilities to generate local externalities. The results of this study emphasized to understand the nature and type of the manufacturing firm in order to measure the urbanization agglomeration. This study has an interesting aspect that the closely located firms may not take the advantage of localization agglomeration. It says that benefits of urbanization agglomeration decrease with the increase in distance of the firms from the urban centres, however if the firms are closely located their productivity growth may decrease. Hence this research advocates that lesser the urban distance, more will be the benefits of urbanization agglomeration and the benefits of localization agglomeration will be decreased.

The empirical results indicated that characteristics of neighboring firms matter for generation of externalities. Both in case of inter and intra industrial clusters. Positive externalities affiliated with localization economies are dependent on distance. As with the increase in distance, positive localization economies increase, accounting for specific characteristics of neighboring firms. It shows that at short distance the weighted and un-weighted diversified forces have negative impact on productivity growth. On the other hand, longer distance has positive effect regardless of weighing scheme.

For an analysis on how distance influence the effect of agglomeration spline regression models can be helpful as different knots indicating the distance will explain its influence on the effect of agglomeration. Though, the spline regression models are rare in practice since its application is really limited but it does not mean that it hasn't been used entirely. For example, (Greiner & Kauermann, 2007) using the smoothing spline regressions confirmed the Bohn's conjecture, (Bechard, 2020) used the spline regression approach to indicate that the impact on monthly taxable sales in tourism-related sectors due to slight increase in number of days is negative. (Bokemeier & Greiner, 2015) used spline regression model in their estimation of finding the relation between public debt and economic growth, (Zareipour, Bhattacharya, & Canizares, 2006) used multivariate adaptive regression splines approach for forecasting the short-term behavior of hourly Ontario energy price.

(Geambasu, Jianu, & Geambasu, 2010) used the spline linear regressions for evaluating the financial assets which proved to better follow the trend from the empirical evidence as compared to the other regression models, (Alkan et al, 2014) have used three types of spline regression i.e., linear, squared and cubic spline regression models for the estimation of the export-import ratio distribution in Turkey from the year 1923 till 2010, (Oktay et al, 2012) used the linear spline regression models to study the demand of tourism in Turkey from 2000 till 2010, (Geambaşu & Stancu, 2011) used the spline regression models in order to describe that the spline regressions are simple solutions when explaining the historical values and when we are forecasting the new ones.

Regardless of strong growth during the 1990s, high poverty persists in remote rural areas, with significant reductions in US poverty measures. Consequently, a study by (Partridge & Rickman, 2008) has employed database of urban proximity measures for the new geographic information system to examine the relationship between poverty and their remoteness in rural American counties, especially their geographical proximity to large urban centers. It has been found that the poverty rate is positively correlated with large (high) metropolitan areas (citrus paribus) with large rural distances. The conclusion explained that the outcomes arise with the elimination of urbanization agglomeration over long distances and the incomplete labor supply adjustment in the case of transportation and relocation to remote rural areas. The results explain the shortcomings in terms of benefits to urban aggregate economies, remote rural areas may benefit from location-based economic development policies, especially in terms of their impact on poverty.

Despite the focus on distance in regional economics and new economic geography models in general, to the best of author's knowledge, no study has been conducted that has experienced the relationship between poverty in rural areas and their geographical proximity in urban classification. Distances can affect poverty by affecting both the demand and supply of rural labor. Short distances between firms have many economic benefits for metropolitan areas, leading to increased economic activity (Rosenthal and Strange, 2001), and higher wages. Nevertheless, the effects of consensus-related wages diminish with distance from the core area. Distance is also an important factor, which is a key factor in increasing employment and population. People adjacent to metropolitan areas grew the fastest during the 1990s and during this period the growth of non-metropolitan areas has declined due to greater distance from the urban core (Partridge et al., 2008). Furthermore, they found that the favorable effects of development in nearby urban areas extend to rural areas up to about 180 km.

The scope of local agglomeration economies is very complex in nature as explained in the agglomeration literature. The researchers have explained this issue through analyzing individual wages with the panel data. The findings of the study show that agglomeration less than the 5km have no significant effect on wages while it has positive and significant affect between the distance of 5 -10 km. This effect decreases rapidly in geographical location and is not significant after 40-80 km. However, these results do not mean that agglomeration is irrelevant to production. At further distances, there must be a significant threshold in the regions for the nearby accumulation and to accumulate benefits from agglomeration. Furthermore, on studying the impact on wages in the Netherlands, no significant evidence was found indicating the large-scale impact of foreign economies (FDI). The findings of the study are an indication towards the national borders that comes out as a major obstacle in the interaction between economic agents (Verstraten, Verweij & Zwaneveld, 2018).

(Ascani, et al., 2012) stated that, in a global world, distance and geography continue to be significant factors. It was noted that industrial production, skilled labor, and higher wages tend to concentrate where proximity between economic agents fosters communication and creates an environment conducive to regular interaction and exchange of ideas.

(Glaeser and Mare, 2001) estimated that wages are higher in large cities, which show the impact of urbanization economies. Input sharing involves the local outsourcing by the firms. Yarn, textile and garments sector in district Faisalabad provides such examples. The presence of yarn factories and garments improve the scale economies and lower the transportation costs. (Holmes, 1999) measured the input intensity showing the benefits of input sharing.

The urban areas that are densely populated are known to be rich in the sources of knowledge as they offer the opportunities of interaction among various stakeholders and the process of knowledge accumulation and spillover starts through the interaction among businesses, universities and customers. Furthermore, the flow of workers and technical staff between companies is high in densely populated areas, which stimulates the spread of embodied skills and knowledge among the people (Almeida and Kogut 1999). Other researchers have found that the flow of knowledge is due to the civic interactions, innovations, patents, and its registrations (Audretsch and Feldman 1996, Jeffe et al. 1993).

Transportation investment through agglomeration economies can encourage positive productive benefits with the increase in the efficiency and scale of local economic interactions. In order to estimate the gross benefits of transportation investment, there is need to understand the local level at which they are distributed externally (Graham, Gibbons & Martin, 2009). This study explored the relationship between urbanization agglomeration and productivity and have estimated how the benefits of urbanization decrease with the increase in distance of the firms from the urban centres. At the sectoral level, the benefits of urbanization agglomeration attenuates more rapidly due to increase in distance from the urban centres for services sector as compared to the manufacturing sector.

In addition to the urban distance, (Rosenthal and Strange, 2004) emphasize two other benefits which may provide benefits of agglomeration economies. Hence three types of benefits i.e., geographical, industrial and temporal can accrued from agglomeration economies. The research concluded that doubling the size of city may increase the productivity by roughly 3-8%.

Similarly, other recent literature suggests that the benefits of agglomeration economies need to be studied beyond urbanization and localization indicators. For example, to save the transportation cost can be a strong incentive to look in large agglomerations regardless of the agglomeration economy (Baldwin and Okubo 2006).

Duranton and Puga (2004) have explored the benefits of increasing return of urbanization economies in a unique way. The researchers have categorized the urbanization benefits into three types i.e., sharing, matching, and learning. Sharing refers to the incentives accumulated from the ability to sustain large areas, it includes diverse input suppliers and deep labor division to improve labor productivity. Matching, like models proposed by Helsely and Strange (1990) and Kim (1990), refers to the fact that more mediators try matching the quality expected. Learning refers to the provision of opportunities by civic bodies for generation, spread, and accumulation of knowledge. The fore mentioned mechanisms provide explanations on how productivity can be increased in the urban consensus and why it remains so.

Firms tend to locate themselves in the urban areas in order to take benefits from the different outcomes of urban areas. An empirical analysis of Milan, using the trans-log production function, has proved that urbanization economies positively affect the productivity of manufacturing firms (Capello, 1999). These results are similar with the study of (Audretch & Feldman, 1996) that stated that diversified knowledge spillover in urban areas is the main reason for the technological change and innovative activities as compared to the specialized knowledge spillover. Another study examined the growth of productivity in urbanization economies. The study has estimated productivity in services and industrial sector as a function of competence, diversity and density. The findings of the study show that diversity has strong and positive impact on the productivity of services and industrial sector (Gutierrez & Ordóñez, 2019). Urbanization economies played their role in productivity enhancement and these results are in line with other studies of developed countries (Cingano, Schivardi, 2004 & Guevara et al., 2019).

The relationship of productivity and agglomeration in Sweden has been examined by a number of studies. The studies by (Aberg, 1973, Klaesson and Larsson, 2009) found that the productive capacity of labor in the Swedish manufacturing industry is a growing work of regional size and the highest productivity of manufacturing firms is in the metropolitan areas. (Braunerjelm and Borgmann, 2004) analyze the productivity of the service and manufacturing sectors in the Swedish regions using the panel data for 1975-1999. The results of this study showed that the sectors which are more concentrated in specific regions have more productivity. (Anderson and karlsson, 2007) also showed the similar results and stated that there is a positive correlation between regional size and number of employees.

2.3. Agglomeration economies: Impact of localization indicators

In this part, we review the second major indicator of agglomeration economies i.e., localization economies. Here detailed explanation on localization indicators of agglomeration economies including Marshallian externalities (knowledge spillover, labor pooling, input sharing) will be discussed. We review the studies explaining the various dimensions of advantages of Marshallian externalities. Finally, we will provide the literature on country specific case studies on the impact of localization indicator.

The emergence of cities and industrial clusters is the cause and effect for regional concentration of economic activity. A lot of literature suggests various dimensions of advantages of the concentration of economic activity. Main sources of agglomeration economies are, according to Marshall, knowledge spillovers, input sharing, and labor pooling. Knowledge spillovers include the knowledge sharing by firms as (Audretsch and Feldman, 1996) showed that innovative industries are geographically concentrated because location of industry is important for the economic knowledge in any industry. (Jaffe et al, 1993) calculated that an innovator is 5-10 times more likely to cite patent from the firm in the same metropolitan as compared to the firms from elsewhere in the country.

In addition, Frenken et al., 2007 highlighted that spillover of knowledge within the region are expected to occur mainly among related sectors and only to a limited extent among unrelated sectors. Research also presents distance as an important aspect of agglomeration.

The industrial cluster proposes that complementary industries co-locate, compete against and share their resources with each other. Hill & Brennan (2000) stated that a competitive industrial cluster is one in which competitive firms are concentrated in the same industry having close link with other industries exists in the same region, sharing common technology and specialized workforce. The positive externalities give benefits to clustered firms that result from the geographical proximity of firms. (Kukalis, 2010) analyzed the association between financial performance and agglomeration economies. The data of 194 firms for the period of 31 years has been used which show no significant difference between non clustered and clustered firms. The results also show that at early stages of industry life cycle the firms have no significant difference in financial performance during the economic contraction. (Gardezi, 2013) favor these Marshall externalities. While investigating the factors of localization economies for the case of manufacturing firms, the author observed that the labor pooling and input sharing are the determinants of Localization economies. For mature industries, (Henderson et al, 1995) found that localization agglomeration is more effective.

Krugman (1991) argues that labor pooling reduces the search cost, implying that the agglomeration offer labor a kind of insurance. Further, (Costa and Kahn, 2003) exemplified the labor pooling case that the married couples who have high educational degrees, will be more likely to locate in large metropolitan areas.

Huge Portion of literature considers agglomeration economies more generally. We now investigate the case studies on the effects of agglomeration economies and whether Marshallian externalities effect the agglomeration of manufacturing industries. Henderson (1986) estimated the effect of agglomeration in US and Brazil. His research found no evidence of urbanization economies but found evidence of impact of localization economies. For mature industries, (Henderson et al, 1995) found that localization agglomeration is more effective. Henderson (2003), using the firm level data of US, finds the evidence of positive effect of localization economies on high-tech industries but not in machinery industries. However, the study does not find any evidence of presence of urbanization agglomeration for two industries.

Lu & Tao, (2009) measured the impact of Marshallian externalities on agglomeration in the manufacturing industries in China. Using the panel data of manufacturing industries, from 1998-2003, the researchers first measured the presence of agglomeration by calculating the Ellison

Glacier (EG) index. The results show that during the said period, agglomeration has increased in China; however, the speed of agglomeration is sluggish due to the national protectionist policies. Moreover, using the dynamic GMM model, this research concluded that the presence of Marshallian externalities affects the agglomeration of manufacturing industries in China i.e., labor pooling, knowledge spillovers and input sharing all contribute to the spatial concentration of industries.

Further, in UK the localization of industries has been tested by using micro data (Duranton and Overman, 2005). The researchers used the geo-spatial data and have tried to understand the localization of industries with various aspects. The very first thing that the researchers did was to challenge the Ellison-Glaeser (EG) Index of concentration and provided their own measure of concentration. This distance-based measure was calculated through the Euclidean space between the industries and gauged the dispersion between the industries. There are certain merits of the distance-based approach that these researchers have adopted. They constructed this measure by incorporating the five crucial elements which were previously absent in the Ellison-Glaeser Index. These five elements are such that the formula that was constructed to measure the concentration should be comparable across the industries, the measure should control the uneven distribution of manufacturing, control the industrial concentration and there must not be any aggregate bias and that the measure should be statistically significant. Based on these assumptions, the formula was constructed; the k-function which is basically used to quantify the geographic existence. The results are very interesting to a level that it was found that 52% of the industries exhibit the localization and since they have calculated the distance, the localization is between zero to 50 kms. The industries which were beyond 50 kms, the impact of localization was diminishing.

Huge Portion of literature on labor market pooling considers agglomeration economies more generally. Researchers have used a unique source of Italian data to provide an inclusive approach to labor market polling. It mutually considers different aspects of labor market relations, including business, matching, learning and hold-up. It also focuses on the expansion of the labor market with the perspective of firm, workers and various industries. It reports on the general positive relationship of business to the density of the local population, which is in line with theories of an uncertainty and agglomeration. It also provides evidence of job learning which is in line with the theories of labor poaching, labor pooling, and hold up. In addition, this research provides evidence according to the combination of job match improvement (Andini et al., 2013).

Overman and Puga, (2010) determined the role of pooling of labor in determining spatial concentration of manufacturing in UK. The study calculated the fluctuations in the employment of individual establishment in relation to their sectors and averaging them across sector over time. The results of the study showed that sectors those were more experiencing idiosyncratic volatility were more spatially concentrated. (Combes and Duranton, 2006) explored the tradeoff between benefits of labor pooling and cost of labor poaching in a duopoly game. The study characterizes the strategic choices of firms with respect to wages, prices, poaching and locations. The results of the study showed that co-location is not the equilibrium outcome whereas it was efficient in study framework.

However, the growing profits in the labor market that we measure are largely small, and it looks like workers and firms will have a significant share of the agglomeration profits. Another study has explained industrial agglomeration with the use of micro data at firm level in the Northeast of Brazil. The study has used theoretical model suggested by Krugman (1991) with some changes proposed by Overman and Puga (2010) to investigate the reaction of firms to shocks in labor market that affect their productivity. The paper regressed EG index as a proxy function for labor polling as well as observed external shocks in the labor market with the controlled sector features that vary in time and sector. The results of the study revealed that industrial agglomeration decline during the period of 2002 and 2014. The role of labor pooling variable is significant and positive as it is expected (Almeida & Kogut, 1999). Moreover, the plants of industries face more passionate shocks and are more locally concentrated as compared to industries.

2.4. Literature on Collective Effect of Localization and Urbanization

So far, we have reviewed the literature on localization and urbanization where details of their indicators have been discussed extensively. Further to this discussion, the empirical evidence has proved that the existence of localization and urbanization economies is not mutually exclusive. It can be the case where a specific manufacturing industry is benefited from both i.e., localization and urbanization economies. This section will extensively review the studies that have investigated for both localization and urbanization economies. In addition to that, country specific case studies will be reviewed where collective effect of localization and urbanization is studied.

Moomaw (1998) found the evidence for both urbanization and localization of agglomeration. The results were also similar for the 3-digit data and 2-digit data of industries. The study also concluded that data of 2-digit industries does not enlarge the prominence of agglomeration economies. Further, Lambert et al (2006) finds that the urban areas attract the manufacturers to select plant locations and the main factors affecting the location decision are population, skilled labor, and infrastructure. Martin et al (2008) used the micro data of firms and provided very insightful results that agglomeration economies in France benefit due to urbanization in short-run, however agglomeration economies are more beneficial due to localization in long-term. The benefits they achieved were because of increased sectoral clustering. The gains which firms received from clusters are also internalized by firms.

In addition, Mahmood, et al. (2016) argues that urban growth has significant positive impact on economic and employment sectors of local area as well as the areas surrounding it. The influence of agglomeration varies according to the nature or type of industries too. Research shows that this impact is different for mature and new entrant industries. Glaeser and Mare (2001) found that wages are higher in larger cities which shows the impact of urbanization economies. In metropolitan area of US, the micro foundation of occupational agglomeration has been examined and labor market pooling was measured by specialized workers who have knowledge on wide range of topics. The results of the study show that level of geographic concentration is higher with unique knowledge base as compared to the generic knowledge base workers. Further the results of co agglomerate patterns depict that similar knowledge base firms will co agglomerate. The cities are not completely diversified and not completely unique in nature. In a model the intermediate case of cities for some industrial co agglomeration can be generated. In some cases, industries will co agglomerate inefficiently and in other situation it could enhance efficiency. Both agglomeration and co agglomeration occur when beneficial (Helsley and Strange, 2014).

Theory of agglomeration economies link two concepts that economies agglomerate when firms share inputs for the production in urban areas. The shareable resources: proximity of labor and businesses that produce positive externalities which decrease the cost of production of one business and enhance the output of other business. The business sharing generates externalities that are technical expertise, common labor pool, personal contacts and general knowledge. The review of theoretical and empirical studies suggests that sharing of tangible inputs like water treatment facilities, highways, public capital stock and communication systems directly affect the workers' productivity and operation of cities through the facilitation of business activities. Some studies have shown positive effect of infrastructure on productivity whereas in some cases it is insignificant. Only few studies have examined infrastructure and agglomeration economies effect instantaneously. Studies that have incorporated both shared inputs physical infrastructure and spatial proximity have positive impact on the productivity of firms (Eberts &McMillen, 1999).

In another study, (Baldwin et al, 2008) tested the Marshallian externalities (knowledge spillover, input sharing and labor pooling) for labor productivity on Canadian manufacturing firms. It was indicated that the effect of labor productivity on agglomeration within the manufacturing sector is positive. It was found that all three types of externalities exist, and they have a positive relation with plant productivity. Plant productivity tends to be higher in the cities where the specialization of upstream industries was present which indicates the presence of input sharing. Also, the presence of input sharing in upstream industries is linked with higher productivity. The cities having a mix labor pool had firms with higher productivity. The plants surrounded by large number of other establishments also showed a higher productivity and since they are surrounded by other plants is an indication of knowledge sharing among the firms. They also calculated the Marshallian externalities to a specific level. Marshallian externalities attenuate after some point which indicates that within the 10km radius the plants have strong and positive influence in terms of knowledge spillover.

There are certain ways to measure the impact of urbanization agglomeration on industries. Mitra (1999) in his paper discussed two methods. First the direct impact of agglomeration on manufacturing industries. The researcher took labor, capital and indicator of agglomeration i.e., population to perform econometric regression. The author took the proxy of population and city work force to measure the impact of agglomeration. In second method the researcher used stochastic frontier model to measure the impact of agglomeration on efficiency of the manufacturing industries. This paper has provided very interesting and pertinent examples. This study has used two industrial sectors of Indian industries i.e., 1) electrical machinery and 2) cotton and cotton textiles to check for the agglomeration. Econometric model of regression was used, to estimate the impact of agglomeration. Results showed that the city population has a significant impact on cotton and cotton textiles. However, there is no empirical indication of any significant impact of city population or agglomeration on electrical machinery sector.

Econometric studies show that agglomeration economies (localization or urbanization) are major determinants of disparities in productivity between cities and regions (Aberg 1973; Moomaw 1998 and 1998; and Nakamura 1985). Further, (Moomaw 1998; Yezer and Goldfarb 1978, Krugman 1991, and Glaeser et al. 1992) extended the previous analysis and concluded that agglomeration economies are not only important to firm location and city-size distribution theories, but they also provide a possible explanation of economy-wide increasing returns and thus endogenous growth.

Many studies have used different variables like population density and market access, location quotients and inverse of Hirschman-Herfindahl index, city population for the purpose of measuring the impact of agglomeration economies (Otsuka et al. 2010; Otsuka and Goto, 2015; Agovino and Rapposelli, 2015; Widodo, Salim, and Bloch,2015; McCoy and Moomaw, 1995; Fukao, Kravtsova, and Nakajima, 2014; Kim et al. 2009; Lakner et al.2012). Urbanization and localization forces affect labor productivity due to diversity of knowledge, skills and market closeness which shows that metropolitan areas are the major cause of boosted productivity of labor (Brunow & Blien, 2015).

The agglomeration economies influence total factor productivity of firms, but few studies have calculated this effect. A study has contributed to literature by estimating impact of agglomeration economies on the total factor productivity of firms. This analysis has been conducted on four different industry groups including agglomeration economies effect on firms based on their technology, research and development of firms to improve productivity, radii and administrative boundaries for outlining regions. The results of the study show that urbanization has greater effect on the total productivity of high-tech firms and no effect on low-tech firms. The localization economies and research and development have significant and positive effect on the TFP, and these effects rise with the technological intensity (Gornig and Schiersch, 2019).

A study has examined the relationship between agglomeration economies and productivity at firm level with the use of dynamic panel model. The findings of the study show that at individual level firms have positive relationships between labor productivity and size of the region. This relationship exists when the attributes of individual level firms like ownership structure, human capital, industry classification, physical capital and time trend are controlled. Secondly, results also explained that learning effect of larger agglomeration economies improve firms' productivity. Moreover, this study did not show clear picture of role of agglomeration coupling with firm size (Andersson & Loof, 2011).

The impact of agglomeration on sectors has been measured by using trans-log production function by (Graham and Kim, 2008). They used a standard trans-log production function which had the elements of labor, capital and 'U' as a measure of agglomeration of industries and services sectors in UK. First, an analytical framework for agglomeration of economies using the UK data of manufacturing and services sector was defined. The trans-log specification showed that all nine industries had a positive effect of agglomeration on the sectors. Secondly, the researchers calculated the elasticities of different variables with respect to agglomeration. The elasticities were found to be negative for construction sector, business services and public services and positive for manufacturing, distribution, transport, and real-estate sectors. When agglomeration elasticities were positive and have a significant impact in terms of UK industries, it meant that increase in agglomeration leads to an increase in industrial concentration and wages.

In another paper related to the previous, (Garaham, 2009) tried to identify the urbanization and localization agglomeration in manufacturing and services sector in UK. This paper is an extension of the previous paper. Researcher extended the previous analysis based on trans-log production function with input demand system. To measure the agglomeration, the paper used the factors such as distance-based measure for localization and urbanization. The study identified that both effects (localization and urbanization) are present for industrial agglomeration. However, the localization effect is relatively less or small as compared to the urbanization effect. This effect is being calculated through elasticities for the different sectors. The elasticities of manufacturing industries for localization are less as compared to the elasticities of different manufacturing sectors for urbanization. They have used the distance as a proxy for measuring the agglomeration effect. With the increase in distance the effect of localization attenuates. The agglomeration externalities also affect hourly earnings of the worker and increased number of labor employment will increase the wages of workers. The findings of the study provide evidence that within 5 kilometers the 100,000 raises in jobs will cause 1.19% increase in hourly wages of workers and after this distance the affect will start decline (Graham and Melo, 2009).

In another study conducted by (Duranton and Overman, 2005) same approach was utilized, only their understanding of localization was distinct from the previous approach. The study examined the extent to which localization exists in the manufacturing sector by analyzing the various features such as the status of firms as either entrants or exiters, ownership of firms; whether they are foreign or domestically owned, size of establishments as large or small, and the level of vertical linkages among industries. It was found that the phenomena of localization were dominant between these firms. The study revealed that the firms that are foreign owned do not appear to locate differently from the domestic owned ones in terms of localization. Furthermore, the analysis indicated that large firms tend to cluster closely together than small firms. There is evidence of co-localization for vertically linked industries as well. Another study has been conducted to analyze impact of urbanization and localization on performance of exports. The results of the study showed that localization economies which are formed due to the similar firms are not that much effective on industrial spaces and innovations whereas urbanization economies and traditional economies mutually have large effects on the performance of exports (Malmberg et al, 2000). These results are consistent with the findings of the study of (Harrison et al., 1996).

There has been somewhat mixed empirical evidence about the existence of agglomeration economies in the context of productive efficiency. (Lakner, et al., 2012) examined the efficiency of agricultural farming in Germany by means of a stochastic frontier analysis (SFA) using data from 1994-95 to 2005-06. They observed that localization and urbanization economies have substantial impact on the technological efficiency of those farms. Same findings were observed by

(Kim et al., 2009), by using Data Envelopment Analysis to calculate regional efficiency in the US biotech industries. Similarly, in order to analyze the role that agglomeration economies play in influencing the technological efficiency of the production process, using the SFA technique, Agovino & Rapposelli (2015) estimated the aggregate production function for the 20 Italian regions for the period 1970-1993.They noted that the localization and urbanization economies impact the production efficiency positively in various Italian regions.

(Driffield & Munday, 2001), using the same methodology, observed the same finding for the case of UK industries. They concluded that UK industries are moving closer to the technological efficiency frontier due to agglomeration. The technical efficiency of Estonian firms increased due to foreign ownership as compared to state ownership and employee. The budget constraints badly affect technical efficiency whereas labor quality and firm size improve efficiency. Moreover, the size of Estonian firms increase with the time due to operating at high level of efficiency (Sinain, Jones & Mygind, 2007). In addition, Fukao et al., 2006) explores productivity spillovers from Japan's efficient factories by using the Data Envelopment Analysis (DEA) and factory-level data of Japan's Census of Manufactures. They found that industrial clustering takes place in Japan in each sector and the efficient firms tend to locate in specific regions to boost their efficiency.

In contrast to these findings, (Widodo, Salim, and Bloch, 2015) identified that localization economies had a positive impact on technical performance, while in case of diversity (linked with externalities of Jacob urbanization) opposite is true. Moreover, McCoy and Moomaw (1995) used panel data of 50 Canadian cities to estimate SFA. Regression analysis of the determinants of the inefficiency of a city suggests that both population size and density improve efficiency. These findings indicate that urbanization has significant impact on nation’s growth. Similarly, (Arif,

2012) argued that in developing countries, although industrial clusters are expanded but their role in improving the efficiency in term of quality and production is not much satisfactory or up to their potential.

Similarly, there is another country specific research which presents an important aspect of the nature and causes of agglomeration. The research work by (Baldwin and Okubo, 2006) on the manufacturing firms of Canada provide in depth analysis of the agglomeration literature. Baldwin breaks the assumptions of homogenous firms while trying to understand their spatial location pattern and their agglomeration and geography. Previously many studies have assumed that the firms are homogenous, but author explained that if study allows this assumption to relax and assume that the firms are heterogeneous in their nature, then there are two types of firms which include productive and unproductive firms. According to Baldwin, the productive firms tend to agglomerate in the areas which are at the center, the core of any city or the core of any industrial hub. Whereas the non-productive firms tend to locate away from the center. There are two factors for such type of agglomeration i.e., selection effect and sorting effect. Sorting effect is the phenomenon when the regional policies attract the firms into some main hub. The phenomena of selection effect occur with the concentration of firms in an area where the regional productivity is higher.

Further, in another paper, Baldwin et al (2010) has extended the previous analysis onto the panel data of Canadian industries. The researchers took two-year data of 1989 and 1999, constructed a panel and applied the GMM and 2SLS econometric model to identify the presence of Marshallian characteristics and other agglomeration indicators with respect to productivity of the firm. The results confirmed the existence of Marshallian externalities i.e., labor pooling, input sharing and knowledge spillover. This research measured the effect of three sources of agglomeration on productivity using plant level data. Only this time, their research showed that, among all sources, labor pooling is the most important source of agglomeration in manufacturing industries. Another empirical analysis showed impact of agglomeration externalities on the regional growth with the sample of 259 Europe regions and 2-digit industries from the period of 1990-2007. The findings of the study explained the agglomeration externalities have strong and positive effect of competitive growth and sectoral diversity. The effect of externalities is different for different regions and impact of Marshallian and Jacob externalities is different on employment growth and is dependent on the density of regional economy. The employment growth has larger effect of Marshallian externalities in less dense areas and high dense urban areas the sectoral diversity has positive effect, the results are in line with (Jacobs, 1969).

The impact of agglomeration economies on the manufacturing industries has been explored very profoundly by (Nakamura, 1985). Researcher employed cross-sectional data of manufacturing industries for Japan and estimated the impact of agglomeration on the manufacturing industries. For agglomeration parameters, the researcher took the proxy variable of city population and city climate along with the standard factors of production. The research employed the trans-log production function in order to extensity analyze the behavior of the variables under consideration. The results showed that the agglomeration economies affect the output of the firms. However, when compared, the effect of urbanization economies came out dominant over localization economies. 19 out of 20 sectors showed that there is positive effect of agglomeration economies in terms of urbanization. And 9 out of 20 sectors show that there is positive effect of localization on the manufacturing sector. In another paper, (Nakamura, 2008) extended the same analysis to a panel of manufacturing industries. He took a panel of 1990-2000 data of manufacturing industries in Japan. The research extended the previous work and tried to explore the impact of agglomeration on the Marshallian externalities. The research concluded that agglomeration affects all the Marshallian externalities. It implied that the impact of agglomeration is evident on Marshallian externalities in the manufacturing sector of Japan. The results showed that doubling the size of industry scale leads to 4.5% increase in productivity, whereas, doubling the size of city population leads to a 3.4% increase in productivity.

Similarly, multi-stage estimation for Japan has explored industrial agglomeration with the agglomeration degree of each industry. The researchers first identified the clusters through cluster detection methodology, scalar measure approach and distance-based approach. Once the industrial clusters are identified, the researchers gathered the possible determinants of agglomeration. Market access, transactions access, labor access, labor cost advantage and natural conditions are used as determinants of agglomeration. Afterwards, the study used simple regression model to measure the effects of these determinants and categorized the clusters as less "concentrated" and more "concentrated industries". Secondly, cluster level spatial autocorrelation was checked. The results of spatial autocorrelation show that certain clusters are highly correlated (metal products) and some are negatively correlated (paper and paper products). At the third stage, this research has used Cobb-Douglas production function to measure productivity using value added, labor and capital. Residual term of the regression is used as total factor productivity for each cluster. Spatial autocorrelation of the residual term was measured with the help of Moran's I test to see the effect of agglomeration on productivity (Tomoya & Smith, 2013).

Agglomeration economies also played their role in improving product quality, profits and productivity of firms. They reduce the marginal cost of the products which enables firms to use more inputs and produce quality products which ultimately enhances the profits of the firms. A study has used plant product level data to find the effect of agglomeration on the quality of products and productivity for the manufacturing in Japan. The results of the study show that with the expansion in market size, the quality of product enhances, and suggested that agglomeration is very effective for the enhancement of firms profit with the improvement in productivity and product quality. The contribution of agglomeration in quality upgrading and profit improvement has been ignored in literature.

Fernandes et al (2017) measured the effect of agglomeration on productivity using total factor productivity and trans-log specification for estimation. (Faggio et al., 2017) used data from UK to document heterogeneity among industries in micro foundation of agglomerated economy. The results obtained were in consistent with the traditional theories of Marshal. There was noticeable heterogeneity among industries. The findings of the research are in accordance with (Jacobs, 1969) idea about unplanned interactions being an important part of agglomeration process.

2.5. Review of Literature on Agglomeration Economies in Pakistan

The review of literature on agglomeration economies in Pakistan indicates that there are quite few studies that has been done in this domain. This part reviews the studies that have been previously conducted on the two indicators of agglomeration i.e., urbanization and localization economies in Pakistan. Here, we review the studies explaining the agglomeration behavior of the manufacturing industries in Pakistan.

Gardezi (2013), in another study used a distinctive dataset in order to develop a distance-based measure of agglomeration. The study investigated manufacturing firms in the region of Punjab on the existence of localization economies and analyzed their agglomeration behavior. An industrial level measure of concentration known as the “M” function was used in this study. In addition to that, several industrial characteristics were examined that are known for measuring the presence of positive externalities. In order to determine the probability of the concentration of firms that experience employment fluctuations, the study used the potential of each industry for labor pooling. It was revealed that the role of labor pooling in explaining the concentration of firms within an industry is significant.

The concept of industrial district can be regarded as a geographic and spatial concentration of firms that have a compact network of local inter-firm relations, characterized by the organization of their products. Evidence on how agglomeration impacts the entry and exit in the domestic industries of Punjab has been presented by (Nasir, 2017) by conducting a firm-level analysis. It also sheds light on how some industries tend to locate themselves in clusters while others are dispersed. The results indicate that industries with a high degree of agglomeration tend to have higher rates of firm entry and exit. On the other hand, industries that are operating with a dispersed geography tend to have more stable firm populations. The aspect of agglomeration economies explained that firms contribute to positive externalities which are generated from the geographical clustering of any industry as well as firms also gain benefits from these positive externalities.

Azhar and Adil (2019) conducted a study to analyze the variation of agglomeration across districts over time in Punjab and investigated the impact of agglomeration on socioeconomic outcomes specifically in terms of social inclusion and efficiency of firms. The Principal Component Analysis

(PCA) was employed by the authors in order to analyze the social inclusion variable using control variables, and Data Envelopment Analysis (DEA) with bootstrap technique to estimate the district-wise efficiency of firms. The use of nonparametric techniques was aimed at ensuring reliable results, as it is free from specification bias. The findings indicates that average efficiency of firms has a positive relation with district agglomeration. Also, a significant positive relationship was reported when examining social inclusion. In addition to that, the study recommended that setting up clusters in urbanized areas, rather than highly urbanized ones, under the China Pakistan Economic Corridor (CPEC) to have a significant positive impact on the economy of Punjab and the country as whole.

To examine the nature of spatial inequality among manufacturing industries in Pakistan and the causes of their geographic concentration, a study was conducted by Burki and Khan (2010). The study used the plant-level data. By mapping districts as the spatial units, the study claimed that across space the firms are not evenly distributed. Clusters were found in the districts with large markets, high road density, and pool of educated and skilled labor. The study's econometric results confirmed this evidence. The study indicated that strong and medium clusters of manufacturing industries were indeed present however, declining trend was observed in the dynamic concentration of industries. The study suggested that for Pakistan the importance of localization economies is dominant over urbanization economies. Furthermore, it was revealed that from 1995-96 to 2005-06 for all industries except food, beverage, and tobacco the productivity growth remained stagnant.

In another study, Burki and Khan (2010) presented the evidence on increased technical efficiency due to an increase in agglomeration. Using cross-sectional data, the study employed a trans-log stochastic frontier and technical inefficiency models. The study revealed that the clustering of industries was widespread but declining overtime. It was observed that the value of technological spillovers is not realized by the firms at first, because of the benefits received from localization economies. The trend has however changed in the recent years because of the increasing competition among the regions which firms are unable to deal with. The study shoes that agglomeration index and technical inefficiency of firms have a strong negative relationship indicating towards the benefits that the firms have amassed through industrial agglomeration. The study also indicated that localization in leather and textile industry whereas urbanization in beverage and tobacco, food, chemical, rubber and plastic industry is highly beneficial.

Siddiqi (2013), examined the spatial agglomeration and productivity of leather and textile manufacturing firms in Punjab, Pakistan. The study suggests that the spatial agglomeration of the industries under observation has a significant role to play when determining the respective productivities. It was revealed that the impact of localization is dominant over urbanization, indicating that if the manufacturing firms of leather and textile are located within an industry of particular district, the results in terms of the productivity of manufacturing firms will show positive results. The study recommends that policies should be focused on promoting localization specifically in the industries of textile and leather.

The literature provides evidence on the two indicators of agglomeration economies i.e., localization economies and urbanization economies where in some industrial settings the impact of one indicator is stronger than the other while in other industrial settings the effect of both indicators is dominant. It suggests that agglomeration economies provide with the positive externalities whether it's from geographical proximity to urban centres or industrial clusters (Marshallian externalities). Such externalities result in increased productivity of the manufacturing industries. The literature suggests that competitive industrial clusters have the property of sharing common technology, region and a specialized workforce. The externalities have a huge role to play in the formation of agglomeration economies. The agglomeration behavior of the manufacturing firms has been reviewed in detail using different case studies of several countries including Pakistan. It was indicated that the impact of the indicators of agglomeration economies on the manufacturing firms varies according to the nature and type of the industries.

CHAPTER 3

3. Overview of Pakistan Economy

Like many other countries, Pakistan also calculates its GDP on the basis of its different sectors. As explained in Pakistan Economic Survey, the economy of Pakistan is divided into agriculture, industry and services. During 2017-18, agriculture contributed 18.86% in total to the GDP of country. On the other hand, the industrial sector contributed 20.91% to the GDP of country. As compared to these sectors, the services sector contributes significantly (60.23%) to the GDP of country. The industrial sector has been categorized into four sub-sectors i.e., Mining, Manufacturing, Electricity & Gas Production and Distribution and Construction. Within the industrial sector, the manufacturing sector contributes significantly to the GDP and is an integral part of the industrial sector. The manufacturing sector can be sub-categorized into three sub-sectors including Large Scale Manufacturing, Small Scale Manufacturing and Slaughtering. The manufacturing sector contributes about 13.6% to the GDP of the country and from that 80% comes from the Large-Scale Manufacturing of the Manufacturing Sector. In this study, our unit of analysis are the manufacturing firms which falls in the category of small- and large-scale manufacturing activities.

According to the Population Census 2017, the total population of Pakistan was estimated to be around 207.77 million and the population growth rate was recorded to be around 2.4%. It has been observed that more than half of total population (132. 19 million) lives in rural area and 75.58 million lives in urban area. The account for the distribution of the population it has been categorized into male, females and transgenders. The large segment of population consists of young people. It has been noted that 59% of the population falls in the 15-59 age group where, 27% of population falls in the 15-29 age group. The total labor force of the country has been estimated to be around 65.5 million during the period of 2017-18 (Pakistan Bureau of Statistics 2017-18). From the total, around 61.7 million of the labor force is employed where 3.8 million labor force is unemployed. The unemployment rate is calculated to be around 5.8%. About 16.1% of the labor force has been employed in manufacturing sector of the country as shown in Table 3.1.

Table 3.1: Percentage Share of Sectors in GDP

Sectors	**Percentage share in GDP**	**Percentage of employed persons**
Agriculture	18.86	38.49
Industry • Manufacturing	20.91 • 13.6	24.62 • 16.05
Services	60.23	36.88

Source: Pakistan Economic Survey 2017-18& Labor Force Survey 2017-18

The contribution of manufacturing sector to the GDP and total employment of the country is 13.6% and 16.05% respectively. The Large-Scale Manufacturing (LSM) contributed 10.8% to the GDP and it dominates the whole sector. Where, the Small-Scale Manufacturing contributes around 1.88% of GDP. A growing labor force and rapid urbanization have created an ideal opportunity for the manufacturing sector to benefit from “agglomeration economies” the concept of economies of urban scale, higher economic efficiency resulting from clustering of firms in a given industry or related industries, and a higher demand for goods and services.

The Census of Manufacturing Industries (CMI) that is conducted after every five years but the latest census was conducted after 10 years. The census of 2015-2016 covered manufacturing establishments according to Pakistan Industrial classification and the current census shows that manufacturing establishment is 390% more than in 2005-06.

Table 3.2: Number of Manufacturing Establishments

Region	2015-16
Pakistan	42,578
Provinces	
Punjab	32,256
Sindh	6,299
KP	3,357
Baluchistan	342
Islamabad	324

Source: Census of Manufacturing Industries (CMI) 2015-16, PBS

The recent census reports that the GDP contribution at basic prices of manufacturing firms has increased as compared to previous census. At provisional level, the contribution of Punjab in GDP remains the highest at around Rs. 1,430,996 in comparison with Sindh, KP and Baluchistan. The table 3.3 shows the contribution of manufacturing firms in GDP at provincial level.

Table 3.3: Contribution to GDP at Basic Prices (PKR Millions)

Region	2015-16
Pakistan	2,946,644
Provinces	**2015-16**
Punjab	1,430,996
Sindh	1,255,895
KP	166,912
Baluchistan	63,160
Islamabad	29,683

Source: Census of Manufacturing Industries (CMI) 2015-16, PBS

The manufacturing firms has been classified by the Pakistan Standard Industrial Classification (PSIC) at 2 digits, 3 digits and 4 digits level. According to sectors at 2 digits level, there are 23 manufacturing sub-sectors. Table 3.4 shows the contribution of manufacturing firms in GDP at factor cost in 2015-16. The textile industry followed by food products, other non-metallic minerals and chemical and chemical products are the major contributors to the GDP of the country.

Table 3.4: Contribution in GDP at Factor Cost in '000' PKR

Manufacturing Firms		**Contribution in GDP**
PSIC Code	**Description**	**2015-16**
10	Food Products	491,927,862
11	Beverages	58,692,429
12	Tobacco products	60,864,622
13	Textile	585,560,911
14	Wearing Apparel	237,110,189
15	Leather and Related Products	40,528,010
16	Wood and Products	9,324,682
17	Paper and Paper products	53,441,688
18	Printing & reproduction of recorded media	20,168,643
19	Coke & Refined petroleum products	98,981,373
20	Chemical and Chemical products	255,278,258
21	Basic Pharmaceutical	158,056,499
22	Rubber and Plastic Products	58,521,405
23	Other Non-metallic mineral	317,037,674
24	Basic Metal	106,800,096
25	Fabricated Metal Products	34,515,460
26	Computer, Electronic and Optical Product	5,321,784
27	Electrical Equipment	54,210,886
28	Machinery and Equipment	29,084,841

29	Motor Vehicles and Trailers	91,367,024
30	Other Transport Equipment	21,834,223
31	Furniture	19,880,199
32	Other Manufacturing (surgical, sports)	39,858,300

Source: Census of Manufacturing Industries (CMI) 2015-16, PBS

Moreover, the employment in the manufacturing sector has increased during the 2015-16 period. Table 3.6 shows the contribution of manufacturing firms in the total employment. It can be observed that Textile is the leading contributor to the total employment followed by Non-Metallic Mineral Products and Wearing Apparel.

Table 3.5: Number of Persons Employed in Manufacturing Industries

Manufacturing firms		**No. Of persons employed**
PSIC Code	**Description**	**2015-16**
13	Textiles	706,146
23	Other non-metallic mineral products	453,942
14	Wearing apparel	273,415
10	Food products	253,934
21	Pharmaceutical's products	92,080
20	Chemicals and Chemical Products	73,288
15	Leather and Related Products	54,576
22	Rubber and Plastics Products	46,465
29	Motor vehicles, trailers etc.	45,912
27	Electric Equipment	42,852
24	Basic Metals	39,501
17	Paper and Paper Products	37,848
11	Beverages	24,948

28	Machinery and Equipment N.E.C.	24,654
19	Coke and Refined Petroleum Products	7,441
12	Tobacco Products	4,433
30	All other industries	159,531

3.1. Problems of Manufacturing in Pakistan

As it has been highlighted before that the contribution of industrial sector to the GDP of the country has been estimated to be about 20.9% during 2017-18 as compared to 2012-13 when it was estimated to be about 20.4% which shows only 0.5% increase in 5 years. Similarly, the share of manufacturing sector was estimated to be around 13.56% during the 2017-18 period as compared to 2012-13 when it was estimated to be around 13.4% showing an increase of 0.16% during the 5 years. Moving towards the manufacturing sector, the large-scale manufacturing sector haven't shown any improvement as its contribution towards GDP has remained stagnant from 2012-13. So, any contribution to industrial growth can be put into the account of small-scale manufacturing as the large-scale manufacturing has not experienced any growth.

The manufacturing sector of Pakistan has been facing different structural changes and issues such as low growth of investment, exports and output. The human skill development is very important for the technological advancement and Pakistan lacks in human resource development. Specifically, the educational institutions are not able to produce qualified scientific manpower because of certain reasons such as a smaller number of technical teaching staff and scholar students do not go for careers in science and technology. The establishment and productivity of different industries is affected adversely because of the deficiency of skilled workers. Therefore, in order to develop diversified industrial sector, there is need to pay attention to human resource development

(Kemal, 2006) & (Burki, 2010). In Pakistan social spending has always been in competition with other development heads and due to stagnation in revenues, policy makers have little ability to meet the growing backlog of investment in health and education sector of the country [(Jamal, 2003) & (Wasti and Siddiqui, 2008)].

Science and technology are imperative to sustain development in quickly changing economic environment internationally. In Pakistan the status of science and technology has not been satisfactory. Pakistan lacks strong foundations of engineering, and the shortage of qualified personnel has further hampered the development of the country's technical base. Low level of science and technology development may also contribute to weak link between, research institutions, industry and academia, lack of awareness about the technological capacity and needs of domestic industries, deficiency of resources for technological development and scientific research (Kemal, 2006). Clustering of manufacturing firms can be really helpful in enhancing the technological development as the knowledge spillover effect comes into play when many manufacturing firms agglomerate together.

The labor market of the country faces the issues of unemployment of educated people and scarcity of middle level skilled labors (Iqbal and Siddiqi, 2013). In order to enhance productivity level and boost industrial diversification the skilled manpower is very essential. The demand for skilled workers has declined due to the little attention that is paid to quality of products and negligence of technological based industries. Moreover, the supply of skilled workers is also inadequate to fulfill the demand. Resultantly, the producer will make informal decision due to absence of skilled workers which ultimately lead to loss of output and low level of productivity. The producers are the main beneficiaries of skilled labor force because their profits, output and productivity would go up. Irrespective of that, the producers have made very minute efforts to enhance and improve skills of workers (Hussain et al., 2012) & (Kemal, 2006). Agglomeration economies can be really helpful in improving the productivity of the workers by improving the skill level of each worker. Unskilled/Low Skilled workers in regions where manufacturing firms are agglomerated together will be keen on keeping up with the Joneses i.e., the skilled workers, increasing the productivity of the workers and in turn the whole cluster.

Further, the transport network for economic development is very important. In Pakistan, the transport sector is bearing heavy losses because of the absence of proper maintenance of existing facilities. The current state of road networks is also not satisfactory, and lack of maintenance and repair of roads has led to fast deterioration of the road networks. The poor quality of roads not only cause delays but also contributes to wear and tear of vehicles and consequently contributing to the increase in transportation cost [(Kemal, 2006) & (Burki, 2010)]. It has been noted that the transport investments induce positive productivity benefits by increasing the accessibility to the economic mass and a large number of suppliers. So, agglomeration of manufacturing firms can help improve the transport network and road densities through input sharing phenomenon.

However, the manufacturing sector has yet to tap into this potential advantage and has remained constrained by the poor business environment, low skilled labor and failure to diversify production and move up the value chain. Punjab in particular, is still largely dependent on agriculture. Punjab’s share in the agriculture has been fluctuating at around two-thirds of national value added in agriculture since 1999. In contrast, Punjab’s share of national value added in the industrial sector is much lower, at around 43-45% between 1999 and 2011.

The quality constraints of the manufacturing products have restricted the development of the industry and also limited the potential for growth (Afraz, Hussain and Khan, 2014). The exporting industries of the country are not able to compete with exports of the international market due to low quality products. In Pakistan, the manufacturing sector is lagging behind due to lack of technological advancement and also lacks the adaptation of advanced technology which ultimately cause low quality products for exports (Ahmed et al., 2017). Research reported that high tariffs and import substitutions are the main restraints that demoralize the efficiency of manufacturing sector (Mahmood, et al. 2016). The domestic producers do not go for expansion of market share in international market due to lack of competitiveness.

Another concern related to industrial development is the lack of diversification. The major portion of industrial value added is contributed by textile and food sector (Akhtar, 1955). On the other hand, the machinery and electronics has made minute progress whereas, the engineering sector is based on imported parts and only basic components are produced. Pakistan lags behind in developing capital goods because producers prioritize the allocated resources for the production of consumers goods instead capital goods (Kemal, 2006). The consumer good will improve living standard of the people immediately at the cost of lower rate of development. It is concluded that the inefficient growth of manufacturing sector is attributed by low quality manufactured goods, lack of infrastructure, lack of diversification and unskilled labor.

3.2. Economy of Punjab Province

Punjab is the largest province of Pakistan in terms of population and economic activities. It constitutes 110 million population (52.95% of national population), out of which 40.4 million are residing in urban areas where 69.6 million are in rural areas (Punjab Growth Strategy, 2018). The estimated share of the economy of this province was 54.2% in the national GDP during 2017-18 (Pasha, 2018). The agriculture sector of Punjab contributes 62.3% to the national economy of Pakistan. It contributes 39.8% and 55.7% in the industry and services sector respectively. The share of agriculture and services sector of Punjab is higher in national economy of Pakistan as compared to other provinces. Within the economy of Punjab, agriculture contributes 22.8%, Industry contributes 15.50% and services sector contributes 61.69%. In terms of total employment of Punjab, agriculture sector employs 40% of the labor force, Industrial sector employs 17.7% and services sector employs 43.3%. Administratively, the province is divided into 36 districts (prefectures) and 142 sub-districts (tehsil). Total geographic area of the province is 17,512 thousand hectares.

Table 3.6: Share of Sectors in GDP & Employment in Punjab

Sectors	**Share in GDP**	**Share of Employment**
Agriculture	22.8	40.01
Industry	15.50	25.32
Services	61.69	34.61

Source: Pasha, H.A. (2018), Growth and Inequality in Pakistan

The Table 3.7 shows the contribution of manufacturing firms in GDP of Punjab at factor cost in 2015-16. The manufacturing firms of the textile industry are the leading contributor to the GDP of the province followed by Food Products, Other Non-Metallic Minerals and Chemical and Chemical Products.

Table 3.7: Contribution to GDP at Factor Cost in '000' PKR in Punjab

Manufacturing firms		**Contribution to GDP**
PSIC Code	**Description**	

10	Food Products	265,203,933
11	Beverages	44,252,459
12	Tobacco products	20,013,129
13	Textile	327,725,491
14	Wearing Apparel	106,496,781
15	Leather and Related Products	35,136,041
16	Wood and Products	49,38,458
17	Paper and Paper products	16,234,271
18	Printing & reproduction of recorded media	5,751,143
19	Coke & Refined petroleum products	6,124,073
20	Chemical and Chemical products	128,251,288
21	Basic Pharmaceutical	29,899,165
22	Rubber and Plastic Products	24,437,433
23	Other Non-metallic mineral	202,342,803
24	Basic Metal	41,994,926
25	Fabricated Metal Products	26,344,809
26	Computer, Electronic and Optical Product	2,108,549
27	Electrical Equipment	22,583,166
28	Machinery and Equipment	20,700,148
29	Motor Vehicles and Trailers	18,762,221
30	Other Transport Equipment	14,982,710
31	Furniture	16,741,425
32	Other Manufacturing (surgical, sports)	36,248,711

The Table 3.8 shows average daily employment in manufacturing firms of Punjab. The textile industry leads the average daily employment followed by Other Non-Metallic Minerals and Paper and Paper Products.

Table 3.8: Average Daily Employment in Manufacturing Firms of Punjab

Manufacturing firms		Average daily employment
PSIC Code	**Description**	
10	Food Products	140,346
11	Beverages	16,777
12	Tobacco products	1,281
13	Textile	457,123
14	Wearing Apparel	160,722
15	Leather and Related Products	46,934
16	Wood and Products	7,655
17	Paper and Paper products	20,025
18	Printing & reproduction of recorded media	6,848
19	Coke & Refined petroleum products	1,952
20	Chemical and Chemical products	42,730
21	Basic Pharmaceutical	25,791
22	Rubber and Plastic Products	25,255
23	Other Non-metallic mineral	358,187
24	Basic Metal	18,697
25	Fabricated Metal Products	26,094
26	Computer, Electronic and Optical Product	2,460
27	Electrical Equipment	26,583
28	Machinery and Equipment	18,412
29	Motor Vehicles and Trailers	17,049
30	Other Transport Equipment	7,651
31	Furniture	14,059
32	Other Manufacturing (surgical, sports)	52,997

3.3. Manufacturing Firms in Punjab

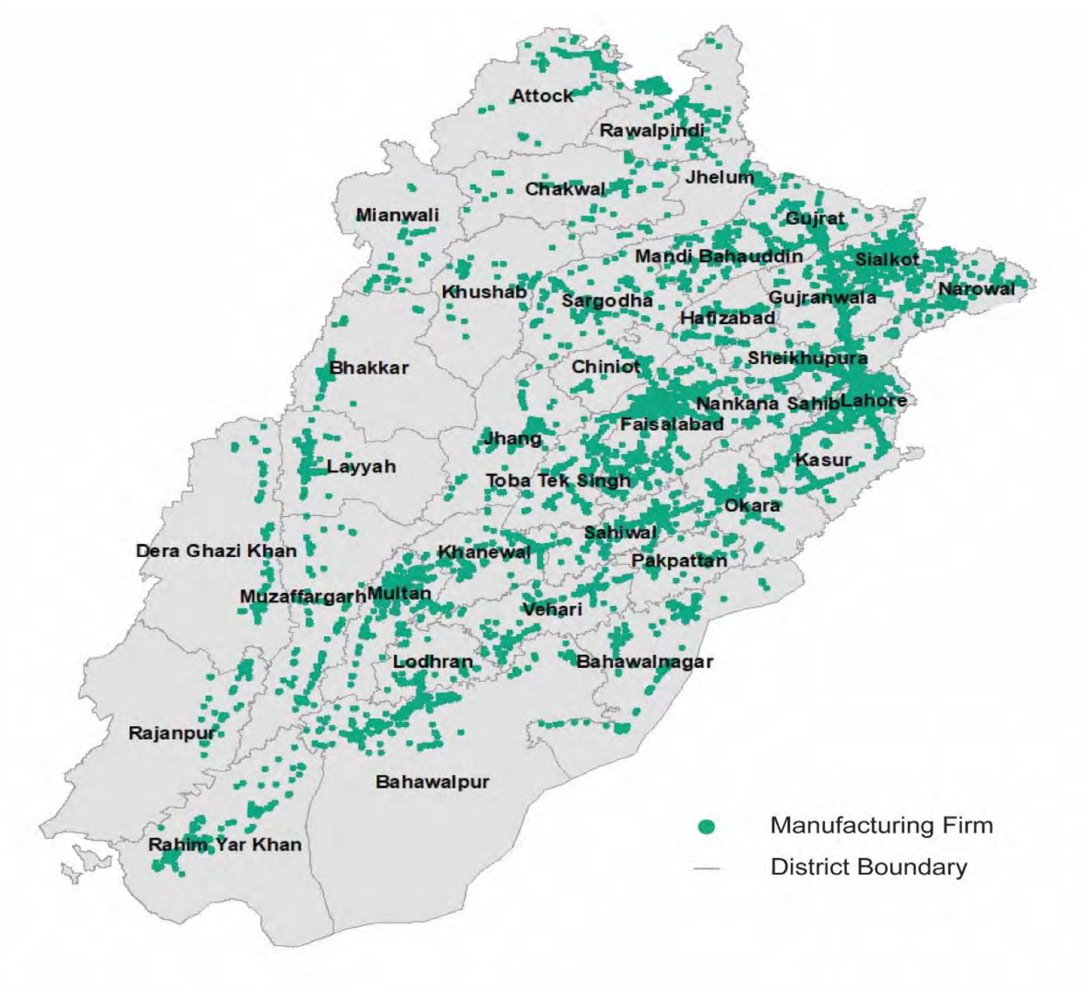
Pakistan Standard Industrial Classification (PSIC) has classified manufacturing industry according to the sectors at 2-digit, 3 digit and 4-digit levels keeping in view the International Standard Industrial Classifications (ISIC). There are 23 manufacturing related sectors at 2-digit level. Given the major economic indicators of Punjab, a further peek into the manufacturing sector reveals that it is highly concentrated. There are 32,256 manufacturing sector units across Punjab. The Table 3.9 shows the district with highest number of manufacturing units. It can be observed that Faisalabad (6,695) followed by Gujranwala (3,494), Lahore (3,421) and Sialkot (2,904) have the highest number of manufacturing firms whereas the districts with minimum number of manufacturing units are Pakpattan (59) and Mianwali (63).

Table 3.9: Number of Manufacturing Firms in Punjab

Districts	**No. Of manufacturing firms**	**Districts**	**No. Of manufacturing firms**
Faisalabad	6,695	Okara	212
Gujranwala	3,494	Hafizabad	211
Lahore	3,421	Vehari	185
Sialkot	2,904	Narowal	178
Sheikhupura	1,275	Rahim Yar Khan	159
Multan	927	Muzaffargarh	138
Gujrat	647	Chakwal	129
Kasur	619	Layyah	125
Sahiwal	503	Nankana Sahib	123
Rawalpindi	411	Chiniot	122
Mandi Bahauddin	300	Dera Ghazi Khan	105

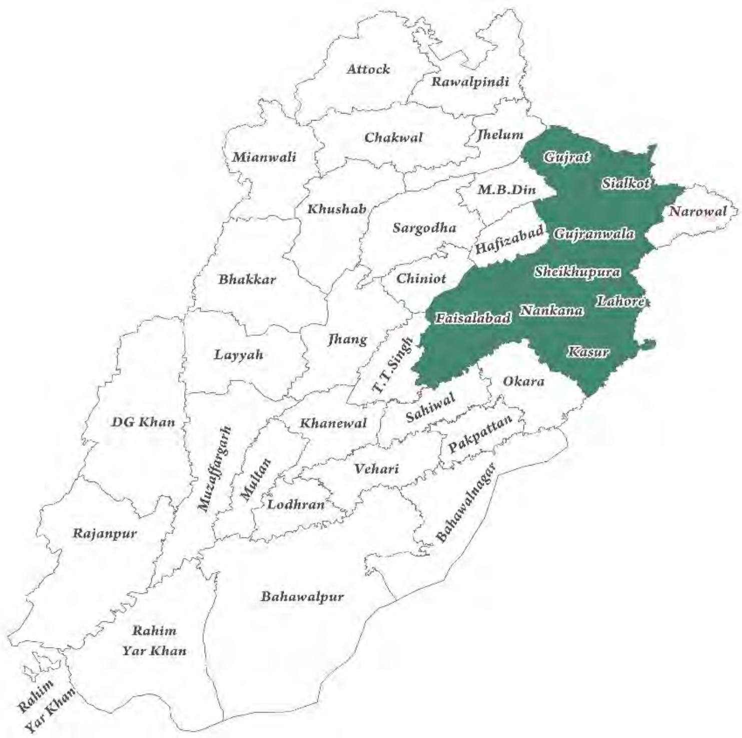
Bahawalpur	289	Jhelum	96
Toba Tek Singh	288	Bhakar	90
Jhang	261	Attock	84
Khanewal	257	Khushab	68
Sargodha	254	Lodhran	65
Rajanpur	232	Mianwali	63
Bahawalnagar	213	Pakpattan	59

Figure 3.1: Spatial Concentration of Manufacturing Firms in Punjab



A quick spatial review of the manufacturing sector in Punjab reveals that, manufacturing activity in Punjab is highly concentrated or clustered both in terms of districts and sectors. The spatial concentration of these industries is much skewed as 80% of these manufacturing units are concentrated in 8 districts only (as shown in figure 3.2). These 8 districts (Lahore, Gujranwala, Sheikhupura, Kasur, Sialkot, Gujrat, Faisalabad, Nankana Sahib) comprise of only 20% of the geographical area of Punjab. According to the population census 2017, the percentage of population in these 8 districts are 10%, 5%, 3%, 3%, 4%, 3%, 7%, and 1% respectively of the total population. Moreover, the percentage of labor force in these 8 districts according to Labor Force Survey 2017-18 are 11%, 3%, 2%, 2%, 3%, 2%, 8%, and 1% respectively of the total labor force. Similarly, manufacturing activity is concentrated in terms of sectors as well, as only 6 out of 23 (at 2-digit PSIC) sectors dominate the total manufacturing units. Those six sectors constitute 69% of total manufacturing units. Two major examples of industrial agglomeration in Punjab are the agglomeration of textile industry (PSIC-14) in Faisalabad district and surgical industry (PSIC-32) in Sialkot district. District Gujranwala presents the agglomeration of different sectors in one place. Such skewed concentration of manufacturing intrigues many questions for research.

Figure 3.2: Districts with Dominant Concentration of Manufacturing Firms in Punjab



Based on the above information it is observed that Punjab's Industrial activity is concentrated in certain regions. They are mainly clustered in the districts with more economic activities and better resources like large markets, employment opportunities, high road density and a large pool of educated and professional labor force. Moreover, due to the skewed industrial clusters, new business finds more market potential in these regions where these industrial clusters are located as compared to the other regions of Punjab, as new firms make their location choices based on availability of resource like labor, employment opportunities etc. Therefore, it is crucial to study where market as well as business potential concentrate in Punjab.

Punjab has a large number of industries concentrated in sectors like light engineering good, textiles and leather. Geographically, it has different industrial clusters, and the most famous industrial cluster is the industrial triangle constituting Sialkot, Gujrat and Gujranwala. In Punjab there are total seven industrial clusters which include Sialkot, Gujranwala, Lahore, Faisalabad, Sheikhupura, Rawalpindi and Wazirabad.

The Lahore district has most diversified industries such as carpets, food, furniture, textiles, automobile parts, machinery and equipment and printing. Faisalabad is the main hub of textiles industry of the country and also has concentration of light engineering goods. The specialization of Gujranwala is the textiles and electronics and Wazirabad is famous for the cutlery manufacturing firms. In Punjab, Sialkot is the most competitive and diversified in industrial clusters. It is also famous for the exports of the country and specializes in surgical, leather and sports goods. Lastly, the major industries of Sheikhupura are food, textiles and machinery and equipment.

In an economy, clusters are the major drivers of competitiveness and necessary economic unit. The significance of clusters derives from their fundamental role which they play in innovation, knowledge creation, skill development and accumulation, pooling of labors with special expertise. Clusters contribute significantly to the industrial development of country.

3.4. Sector Wise Concentration of Manufacturing Firms in Punjab

Figure 3.2: Textile Manufacturing Firms in Punjab

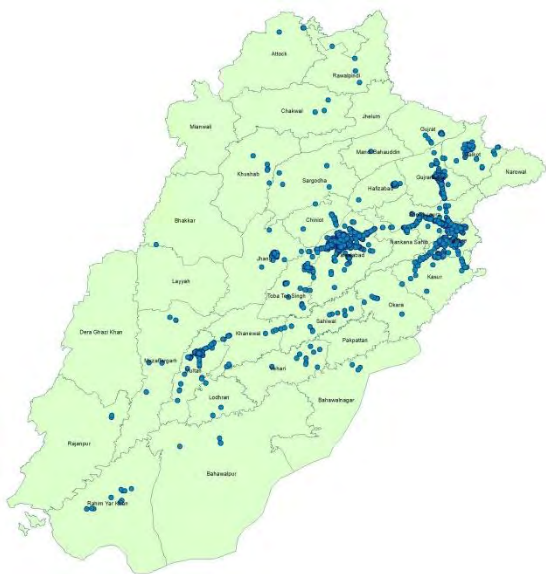
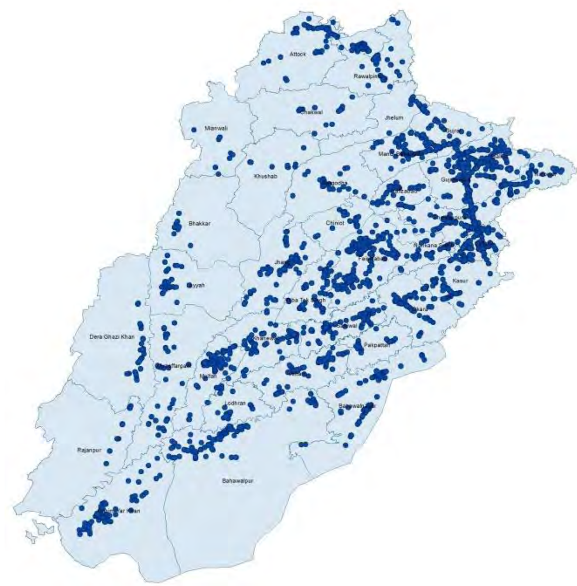


Figure 3.3: Food Product Firms



In Punjab, a total of 6747 textile industries are working which represents almost 70% of the whole textile industry. In the textile sector Punjab has the lions share and Faisalabad is the main hub of this industry. The district with highest number of textile industry is Faisalabad (4509) followed by Gujranwala (642), Lahore (586), Multan (226) and Sheikhupura (208). The textile sector is considered as the backbone of Pakistan's economy, as 46% of the manufacturing sector is textile, contributes 60% to the total exports and employs about 30.2% of the labor force. Textile industry in Faisalabad is the best example of localization economies.

There are 3493 food products manufacturing firms in Punjab. The district with the highest number of food products manufacturing is Faisalabad (476) followed by Gujranwala (248), Multan (202), Lahore (200) and Sialkot (199). It employs about 10.8% of total labor force. In Punjab, after the textile the food products are rich in numbers.

Figure 3.4: Other Non-Metallic Mineral Firms

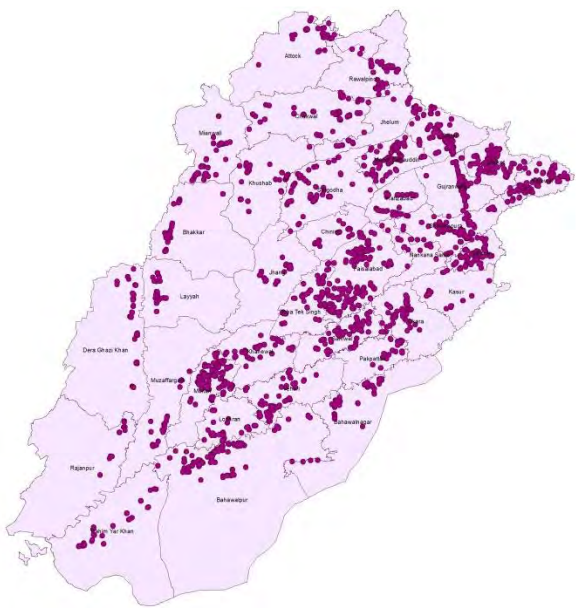
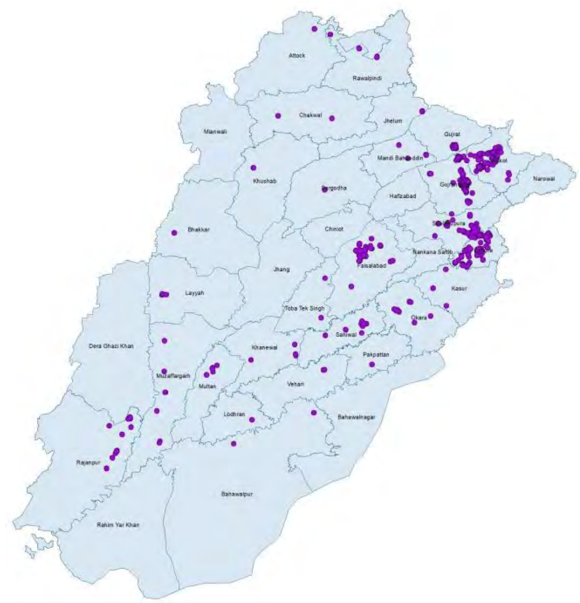


Figure 3.5: Fabricated Metal Firms



The classification of non-metallic mineral products includes manufacturing of tiles, ceramic products, glass and glass products, cement and plasters etc. A total of 2750 non-metallic mineral firms are located in Punjab. In terms of numbers of firms located in the district Gujrat leads with a total of 225 firms followed by Faisalabad (207), Gujranwala (195), Sialkot (137), Mandi Bahauddin (129) and Sheikhupura (125). These manufacturing firms are concentrated mainly in these seven districts.

The fabricated metal manufacturing firm includes the production of pure metal products like structures, parts and containers. A total of 1612 fabricated metal manufacturing firms are located in Punjab. The Gujranwala is the main hub of fabricated manufacturing firms, and 936 firms are located in this single district where, 218 firms are working in Sialkot and 171 are working in Lahore.

Figure 3.6: Wearing Apparel Firms

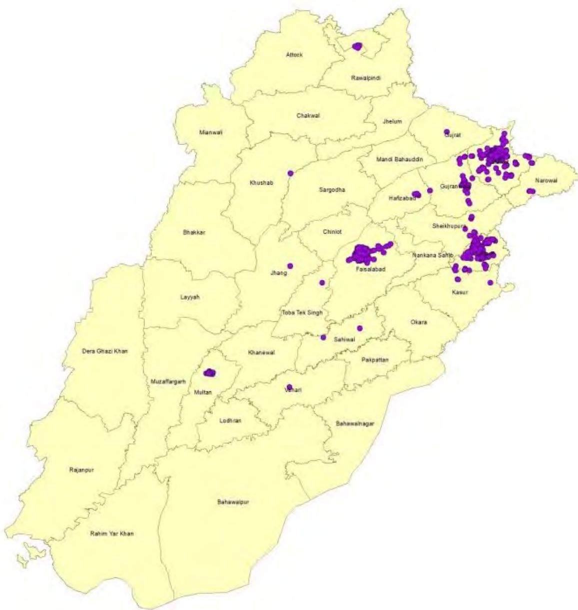
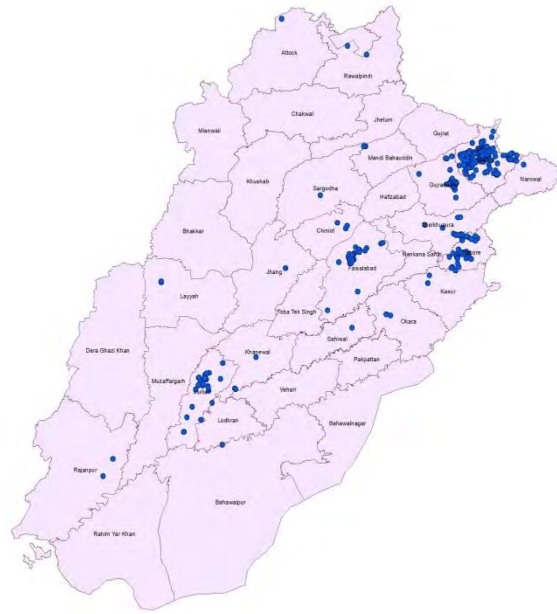


Figure 3.7: Other Manufacturing Firms



The wearing apparels manufacturing firms include all items and materials of cloths such as fabric, leather, crocheted and knitted etc. there is no difference between the clothing of men and women, children and adults. A total of 1653 number of wearing apparels firms are located in Punjab. The district with highest number of wearing apparels manufacturing firms is Sialkot (748) followed by Faisalabad (366), Lahore (344), Multan (84) and Gujranwala (70).

The manufacturing of sports, jewelry, toys and surgical equipment has been classified into other manufacturing group and criteria for grouping of this division has not been applied. This category includes a total of 1339 manufacturing firms in Punjab. From the total number of firms, 1134 are located in a single district i.e., Sialkot followed by 61 firms related to these groups are working in Lahore while 35 are working in Multan. The spatial portray of remaining sectors is given in Annexure I.

The phenomenon of clustering in the manufacturing firms can be a source of productive advantages from being able to work in close proximity to others. The word "cluster or concentration" is more of a generic term that is used to explain a broader concept, famously known as "agglomeration economies" in literature. Agglomeration economies allows the firms to expand physically and economically and is responsible for boosting their performances. The productivity advantage as the literature suggests may be because of one of the two types of agglomeration economies i.e., urbanization economies and localization economies.

Spatial review of Punjab province has revealed that the manufacturing firms are highly agglomerated in terms of location and sectors. From geographical perspective, 80% of the manufacturing firms are agglomerated in only 8 districts of Punjab. These are Lahore, Gujranwala, Sheikhupura, Kasur, Sialkot, Gujrat, Faisalabad and Nankana Sahib. Whereas these districts comprise of only 20% of the geographical area of Punjab. Similarly, manufacturing activity is also concentrated in terms of sectors as well, i.e., 6 out of 23 (2-digit PSIC) sectors dominates the manufacturing activity, which constitutes 69% of the total manufacturing firms. The revealed facts provide us with an overview of highly agglomerated manufacturing firms in the province. For so long the researchers have been intrigued by the question of how the agglomeration of manufacturing firms are affected by either of the two types of agglomeration economies i.e., urbanization economies and localization economies. Hence the identification of the nature of agglomeration and its effects, in a particular region, can provide valuable insights in formulating the industrial policies that can improve the city or a region.

CHAPTER 4

4. Theoretical Framework and Methodology

4.1. Theoretical Framework

The rise of industrial clusters all over the world has intrigued the researchers to find out the nature, causes and the effects of such agglomeration of industries. Many researchers have tried to explain this phenomenon of agglomeration through various lens keeping in view the indigenous country experiences, e.g., the nature and causes of industrial agglomeration in Canada may vary from China or Japan. Hence a consensual approach or theory may be difficult to find out since industrial agglomeration is an ongoing phenomenon. However, few theories are really relevant and can be regarded as universal approach on industrial agglomeration. These theories are somehow substantiated by the evolution of industrial agglomeration across the world.

4.1.1. Theory of Localization

The classical economist Alfred Marshall was the first that came forward with the theory of agglomeration. He explained that when firms agglomerate in adjacent geographic proximity, they experience increasing returns and external economies (Potter & Watts, 2014). He also identified three mechanisms that later came to be known as Marshallian externalities to explain the increasing returns and clusters of firms within the same industry. The first one is the presence of input providers which allow for productivity gains due to the specialization and vertical disintegration. The second one is labor pooling which facilitates the workers matching process because agglomeration enhances labor productivity. The last one is the knowledge spill over which shows the ability of industry to capture the specific knowledge when agglomerated at one industrial district [(Figueiredo, Guimaraes & Woodward, (2009) & (Potter & Watts, 2014)]. According to

Marshall, the firms operating at small scale are able to increase their productivity and specialization due to positive externalities. Marshall stated in his passages of *Industry and trade*:

> "For long ages industrial leadership depended mainly on the number and extent of centres of specialized skill in which [...] external economies abounded [...]. Each single business was on a small scale; and though it had access to many of the economies of production on a large scale, these were external to it, and common to the whole district. [...]. Thus, each firm, though of moderate size, might reasonably hope to obtain most of the advantages in production, which would be accessible only to vast businesses, if each had been mainly dependent on its own resources. Under these conditions, a very large capital in the aggregate was distributed over many firms of moderate size, each with its own individual life, its own power of initiative, and its own personal relations with its employees [...] Its own (Internal) economies were not great: but it took its part in affording a large market for firms in branches of manufacture, which supplied it with made or half-made materials: and in developing (External) economies of general organization, which gradually became common property". (Marshall, 1920, 114–115, 206)

This idea is reflected in literature of industrial district, where clusters are usually associated with smaller firms. Moreover, in his book "The Principles of Economics", He presented the idea that the productivity of the firms can be increased due to agglomeration economies, such that

> "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 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 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. (1920, pp.271).

Thus, Marshall presented the idea of localized agglomeration of the industries explaining the concentration of same type of industries in a specific geography. The Marshalls theory of localization agglomeration has experienced a substantial revival in theories of regional development and has remained most cited author. Now, ideas of Marshall are playing key role in different policies and theories.

After twenty years, Florence (1939) introduced measure of localization which is known as employment location quotient. It focused on calculating the agglomeration of industries at a given location. Location quotient can be calculated for every region and industry through the ratio between industrial total share of employment at regional and national level if the spatial scale of analysis is recognized. This measure lacks the theoretical foundation and is not accurate measure of localization agglomeration from Marshallian perspective. One major issue with location quotient is that it is not able to differentiate between the internal and external scale economies. The location quotient gives the same results whether the employment share of an industry at regional level results from a cluster of establishments or from a single establishment. Another issue of location quotient arises from potential inability to capture the randomness of the plant's location decisions which can explain the local concentration. Due to the isolated nature of phenomenon, there is a possibility to observe bogus concentration which occur by chance.

In 1997, Ellison & Glaeser developed measure of industrial localization agglomeration which overcame the problems of employment location quotient. They proposed an index based on the firm location model of Carlton (1983). The index was based on a probability model, which naturally accounts for hereditary randomness, that will be observed if location decisions are made by chance. The authors also claim that their approach eliminates the impact of internal economies of scale on the scale of localization of industry. Guimaraes et al. (2007) offered another statistic based on the plant counts which remove the impact of internal economies of scale, but it is consistent with the statistics of Glaeser and Ellison theoretically.

Table 4.1: Various Statistics to Measure Localization Agglomeration

Author	Year	Statistics
Florence	1939	Employment location quotient (This statistic calculates agglomeration of industries at a given location)
Glaeser and Ellison	1997	EG index (The index is based on a probability model, which naturally accounts for hereditary

		randomness, which will be observed if location decisions are made by chance)
Mori et al.	2013	D-index of localization (This test can be employed to test the relative degrees of localization among industries)
Guimaraes et al.	2007	This statistic is based on the plant counts which remove the impact of internal economies of scale, but it is consistent with the statistics of Glaeser and Ellison theoretically.

4.1.2. Theory of Urbanization

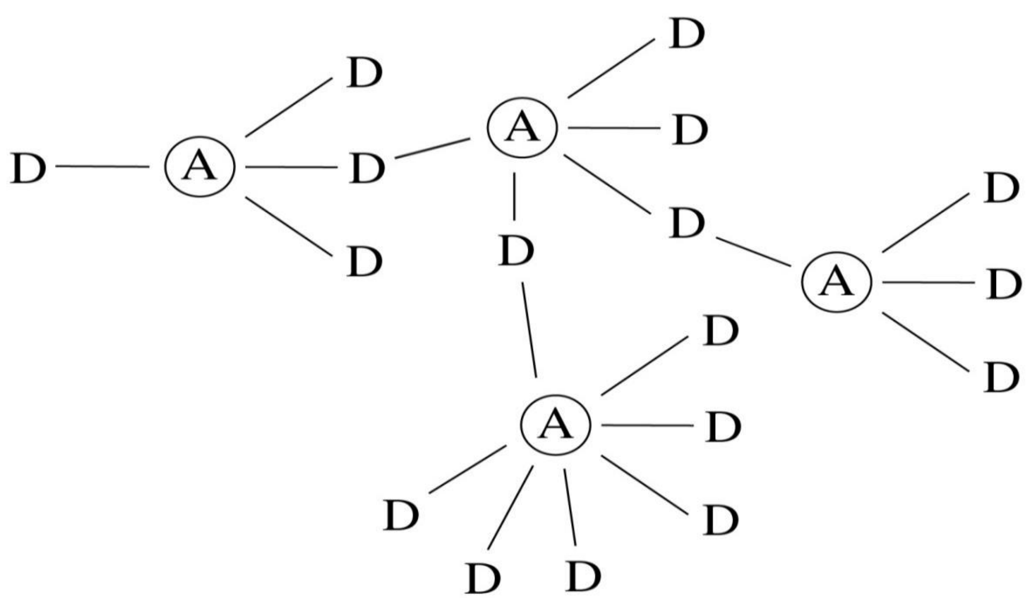
Many researchers have argued that the agglomeration of economic activities may be because of obtaining the benefits of environment and big cities. Hence, the city size or its density may also contribute towards increased productivity of the firms (Beaudry, & Schiffauerova, 2009). The urbanization agglomeration describes the benefits accrued because of the presence of other economic activities such as services sector, diversified labor skill set, consumption and population (Jacobs, 1969). The Jane Jacob stated:

> "The greater the sheer number of and variety of division of labour, the greater the economy's inherent capacity for adding still more kinds of goods and services" (Jacobs, 1969, p. 63)

Jacobs referred to urbanization economy as the total number and distribution of labor that is involved. It is considered different in some studies and is often measured by urban size and density.

Urbanization economies is based on the idea that old work adds to new work, and it multiplies the division of labor which is shown in the diagram below. The new work adds more divisions of labor in existing division of labor with the increased activities due to new yield up. The diagram below indicates four additions: the first D shows the initial activity for new manufacturing for example, dress fitting. The resulting nD shows different number of divisions of labor. The same process goes for the next Obsidians. The D + A…. nD formula is a handy way to express the process.

Figure 4.1: Idea of Urbanization Economies as described by (Jacobs, 1969)



According to Marshallian theory, localization contributes to knowledge spillover because of migration of workers within same industry whereas, Jacobs claimed that sources external to industries are imperative sources of knowledge spillover because firms operating in different industries gain from diversity of industries and new ideas cannot come from the similar kind of industries. In cities, the sources of diverse knowledge are greater and are the major source of innovation. The theory of Jacobs stressed that diversity of industries within a region play an important role in knowledge promotion, innovative activities and ultimately in economic growth and it is known as urbanization economies (Beaudry, & Schiffauerova, 2009).

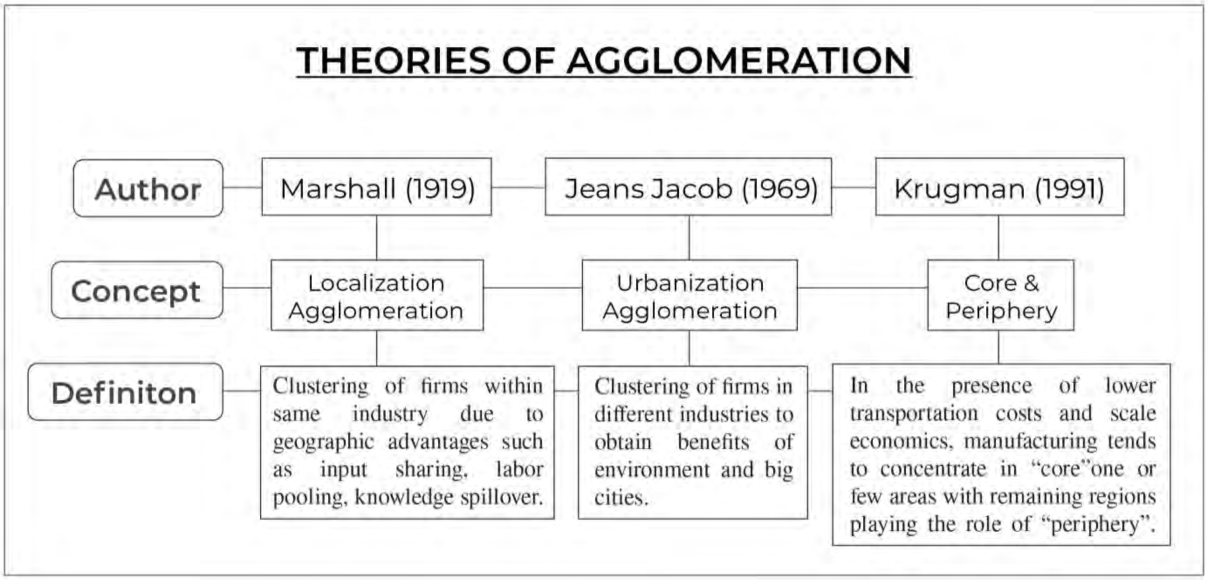
There are certain ways to measure the impact of urbanization agglomeration on industries. Mitra (1999) in his paper discussed two methods. First the direct impact of agglomeration on manufacturing industries. The researcher took labor, capital and indicator of agglomeration i.e., population to perform econometric regression. The author took the proxy of population and city work force to measure the impact of agglomeration. In second method the researcher used stochastic frontier model to measure the impact of agglomeration on efficiency of the manufacturing industries. This paper has provided very interesting and pertinent examples. This study has used two industrial sectors of Indian industries i.e., 1) electrical machinery and 2) cotton and cotton textiles to check for the agglomeration. Econometric model of regression was used, to estimate the impact of agglomeration. Results showed that the city population has a significant impact on cotton and cotton textiles. However, there is no empirical indication of any significant impact of city population or agglomeration on electrical machinery sector.

Transportation investment through agglomeration economies can encourage positive productive benefits with the increase in the efficiency and scale of local economic interactions. In order to estimate the gross benefits of transportation investment, there is need to understand the local level at which they are distributed externally. The study deals with the impact of urbanization agglomeration on productivity and how with the increase in distance the agglomeration externalities decline. It is based on data from a wide-ranging panel to estimate the impact of large-scale economic access on Total Factor Productivity (TFP) for large sectors of the economy. The econometric specification is based on a control function approach that addresses possible sources of integration associated with the production function and agglomeration-productivity relationship, and also allows for firm level heterogeneity. A nonlinear least square regression distance is used to provide a direct estimate of the decay. The results show an overall impact of 0.04 on all sectors of the economy. The study estimate elasticity of 0.02 for manufacturing and consumer services, 0.03 for construction and 0.08 for business services. The distance-closing parameter is approximately 1.0 for manufacturing, but approximately 1.8 for the consumer and business service sectors and 1.6 for construction. This shows that the effects of aggregation diminish more rapidly with distance from the sources of service industries than manufacturing (Graham, Gibbons & Martin, 2009).

4.1.3. Theory of Core and Periphery

After the theories of Marshall and Jacobs, Krugman (1991) presented the theory of “core” and “periphery” and provide economic rationale for regional divergence. Based on the geographic concentration of manufacturing activities, author argues that, in the presence of lower transportation costs and scale economics, manufacturing tends to concentrate in “core” one or few areas with remaining regions playing the role of “periphery” or agricultural suppliers to manufacturing “core”. The concentration of manufacturing occurs near the areas where there is a large demand market since it minimizes the transportation costs, and the markets will be large where the manufacturing is concentrated. Thus, Krugman advocated in favor of both localization and urbanization of agglomeration.

Figure 4.2: Theories of Agglomeration



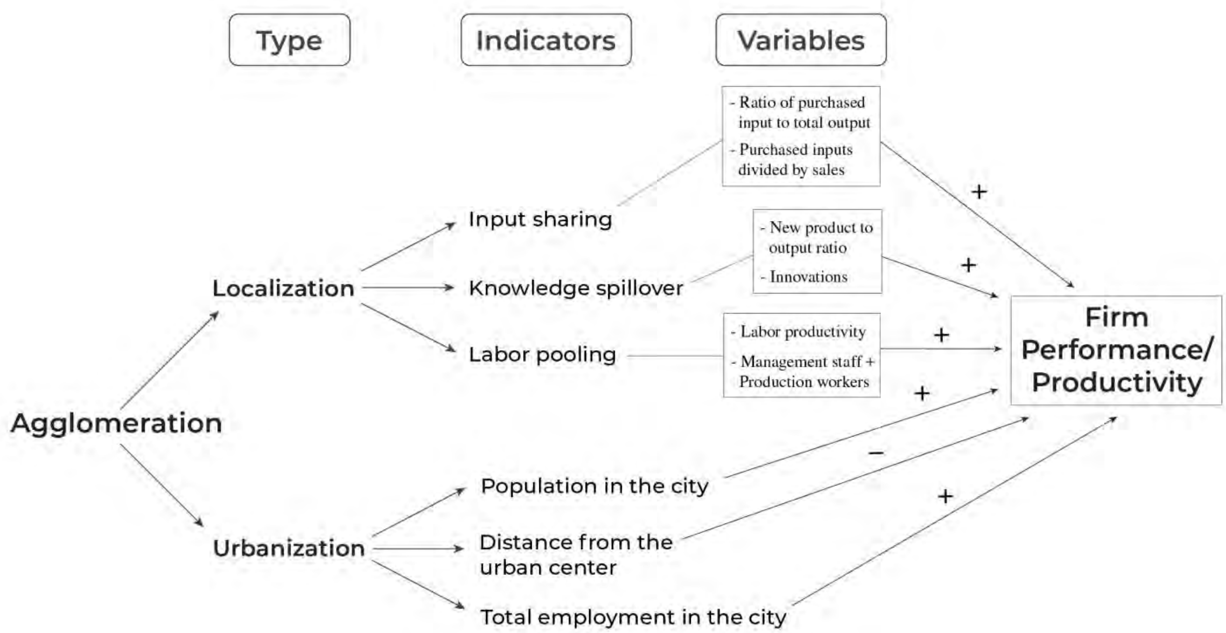
4.2. Application of the Theories

The theories regarding agglomeration economies have been highlighted in the previous chapter. Researchers have used different indicators and tools to identify localization agglomeration and urbanization agglomeration in different areas. In order to substantiate the localization agglomeration, the data on knowledge spillover has been used. Knowledge spillover has been calculated with different proxies such as portion of research and development in total sales and innovation. This proxy has been used by Feldman et al (2002); Audretsch and Feldman (1996); Rosenthal and Strange (2001). Another direct proxy of knowledge spillover, the new product to output ratio has been used. In literature, new products to output ratio have a positive and significant impact on the industrial agglomeration. The information sharing also contributed to spatial concentration of industries. By employing differencing method to data of information contribution the spatial concentration of activities can be calculated. The spatial concentration of industry on the basis of similar activities is linked to localization agglomeration.

A study assesses the concept of Marshall related to clustering of firms at one place. The authors used the data of intermediate input producers and examined the relationship between vertical disintegration and localization of industries. The results of the study showed a positive association between the input sharing and localization of industries. In this study the author employed panel data and dynamic estimation approach to calculate the input sharing as a determinant of agglomeration. The clustering of firms due to input sharing is directly linked to the concept of localization agglomeration (Holmes, 1999).

Different studies have been done to calculate the impact of urbanization agglomeration. Shefer (1973) considered cross section of group of industries and results of the study explained that if the size of city is doubled, it will increase the productivity from 14 to 27%. Another study in the same areas has calculated 6 to 7% in productivity due to urbanization economies (Sveikauskas, 1975). Segal (1976) has found that productivity of cities having two million or more population was 8% higher. If the population of the city is doubled it would increase the productivity by 10% and Moomaw (1998) claimed that it will increase by 2.7%. These studies showed that urbanization agglomeration have positive and significant impact on productivity.

Figure 4.3: Summarizing the link between Agglomeration and Firm's Performance



Some studies have calculated impact of both urbanization and localization on productivity. The Nakamura (1985) and Henderson considered both localization and urbanization and calculated the impact on productivity. Nakamura estimated production function for two-digit manufacturing industry in Japan and Henderson calculated production function for two-digit manufacturing industries in Brazil and US. Total employment in the city has been used as proxy for urbanization and total employment in industry for the estimations of localization. The results of the study have provided more evidence on the localization economies whereas, the evidence for urbanization economies was visible in some of the industries. The study of Nakamura also explained that if the population is doubled, it will increase the productivity by 4.5% whereas, if the population of city is doubled it will increase the productivity by 3.4%. The study of Henderson has shown substantial evidence of localization and no evidence for the urbanization.

Moreover, Moomaw (1998) found the evidence for both urbanization and localization agglomeration. The results were also similar for the 3-digit data and 2-digit data of industries. The study concluded that data of 2-digit industries does not enlarge the prominence of agglomeration economies. Further, Lambert et al. (2006) finds that the urban areas attract the manufacturers to select plant locations and the main factors affecting the location decision are population, skilled labor, and infrastructure. Martin et al (2008) used the micro data of firms and provided very insightful results that agglomeration economies in France benefit due to urbanization in short-run, however agglomeration economies are more beneficial due to localization in long-term.

4.3. Theoretical Framework of the Research

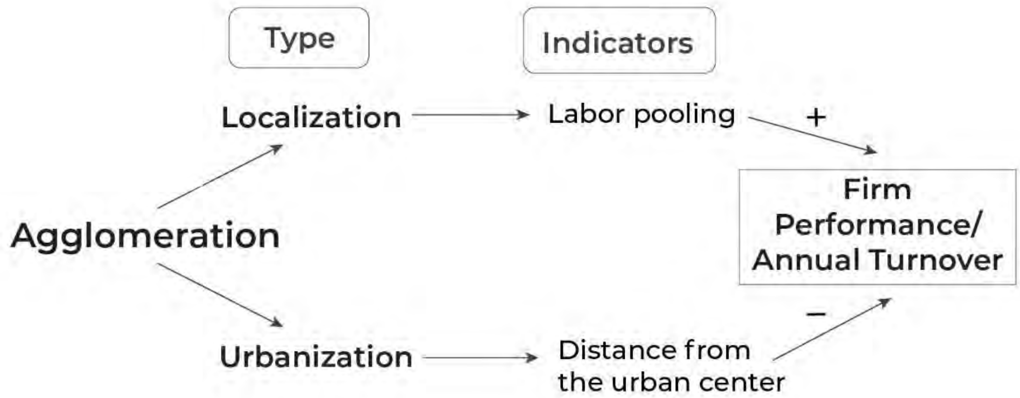
The theoretical framework of this study is based on the above-mentioned studies, linkages, variables and the nature of relationships. At first stage this research has constructed variable for agglomeration indicators. Variables for urbanization and localization have been constructed based on the nature of agglomeration. From theories, it has been observed that in urbanization agglomeration the firms take benefits from different types of activities from different industries. The studies have calculated urbanization agglomeration by using different proxies such as, population density, distance from the urban centre and total employment level in the city. For this study, a proxy has been used for urbanization agglomeration i.e., distance of the firm from the urban centre.

Localization agglomeration has measured by studies based on Marshallian externalities such as, input sharing, knowledge sharing and labor pooling. This study has constructed localization variable based on the labor pooling. The variable is required which can show the relationship between the firms based on the labor pooling. The literature suggests that Moran's *I* test of spatial autocorrelation can provide information whether firms are spatially autocorrelated or not, based on some indicator. So this study has employed spatial autocorrelation.

On second stage, this study will calculate the impact of localization and urbanization on annual turnover of the manufacturing firms of Punjab province. For this purpose, this study has developed production function. From the review of literature, it has expected that both urbanization and localization have positive impact on the annual turnover of the manufacturing firms. This study will calculate the overall impact of urbanization and localization on annual turnover of the firm and further, it will calculate the sector wise impact of both urbanization and localization. This will give a detail insight that whether the impact of urbanization is stronger or the impact of localization is stronger.

Moreover, this study will further explore the relationship of urbanization agglomeration with annual turnover of the firms through distance of the firms from urban centre. We have divided the distance of the firm from urban centre (proxy of urbanization) into three categories. The urbanization has measured by the proxy of distance of firm from the urban centre, so the study will divide distance into 3 categories 0-5km, 5-15km and above 15km. According to literature, the effect of urbanization attenuates with the increase in distance from the urban centre Rosenthal & Strange, (2003).

Figure 4.4: Theoretical Framework



4.4. Methodology

An enormous amount of literature has been generated on measuring the impact of localization and urbanization. The choice of methodologies used in the literature differs according to the availability of data. Due to the non-availability of the data on this topic, researchers have been using various proxies to measure the impact of urbanization and localization. For urbanization, the most widely used proxy is the population of the city and for localization the researcher have used the labor employed in the industry (Baldwin 2010, Nakamura 1985).

Second wave of literature on urbanization and localization have employed more advanced data and methodologies. With the availability of data about the location of the industry and the little introduction of geography into economic analysis has changed the thinking and analytical domain of agglomeration. Recently the researchers are using location-based indicators to measure the agglomeration factors such as distance between industries and regions (Duranton 2006).

In 1985, Nakamura employed cross-section data of manufacturing firms in Japan and explored the impact of agglomeration economies on the manufacturing industries. The study has employed translog production function in order to extensively analyze the behavior of the variables under consideration. The impact of agglomeration on sectors was also measured by using translog production function by Graham and Kim (2008). This paper constructed the translog inverse demand function. The paper has taken a standard translog production function which has the elements of labor, capital and 'U' as a measure of agglomeration of industries and services sectors in UK. Further, Garaham (2009) extended his previous work to measure the urbanization and localization externalities in manufacturing and services sectors of UK based on the translog production function with input demand system. The author has used the distance-based measure for localization and urbanization for the measurement of agglomeration.

Moreover, the researchers have also employed multi-stage estimation to identify the industrial agglomeration. The cluster level spatial autocorrelation was employed to show the concentration of industries. The study has also employed Cobb-Douglas production function to measure the productivity using value added, labor and capital (Tomoya & Smith, 2013). Numerous studies have employed translog production function to measure the effect of agglomeration on productivity and firm performance [(Fernandes et al., 2017), (Burki and Khan, 2010)].

Most of the studies that used panel data have employed GMM and 2SLS econometric models to identify the Marshallian characteristics and agglomeration indicators with respect to productivity of the firms [(Lu & Tao, 2006), Baldwin et al 2010)]. The OLS estimation technique has been used by different authors to analyze the concentration of manufacturing industries as well as to assess the determinants of agglomeration of industries. The studies have depicted those Marshallian externalities affect the agglomeration of manufacturing industries (Lu & Tao, 2009), (Burki and Khan, 2010), (Nasir, 2017).

In literature some studies have used the Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) to explore the technical efficiency of industries due to agglomeration [Fukao et al., 2014), (Sinain, Jones & Mygind, 2007), (Driffield & Munday, 2001), (Kim et al., 2009)]. A positive impact of agglomeration externalities was found on production frontier and technical efficiency when calculated with SFA. (Tveteras and Battese, 2006).

Some studies have employed spatial autocorrelation to depict the association between the firms based on some characteristics. The literature provides an insight that Moran I test of spatial autocorrelation gives information about the correlation of firms based on some indicators. López-Bazo et al., (1999) used Moran I statistics to observe the regional dynamics and convergence in EU. The study showed the spatial autocorrelation among neighboring regions in terms of productivity and GDP per capita. Another study has employed Moran I test to observe the spatial pattern of tuberculosis in the Si Sa Ket province of Thiland. The study has found the high-rate clusters in the northwest part of Si Sa Ket and low rate clusters in the Southeast area of the province. According to literature different studies has applied techniques of spatial autocorrelation such as (Rey and Montouri, 1999) & (Wang, Chang and Wang, 2019) & (Stankov et al., 2017).

This research has employed the tools of spatial econometrics to understand the agglomeration phenomenon in Punjab. The study has employed the geo-spatial point data on the industries of Punjab. This data includes the location of the firms, output level, annual turnover, labor employed, land utilization, export status of the firms and the PSIC classification at 2-digit and 3-digit. First, the research will provide the spatial portray of manufacturing firms in Punjab. Secondly, a spatial and econometric analysis will be conducted for research.

4.4.1. Spatial Portray

Under the spatial portray, the geographic data turned into useful information through ArcGIS. This study has given the maps of manufacturing firms in Punjab which are showing highly concentrated and clustered firms both in terms of districts and sectors (See chapter 3.3, 3.4 for detailed spatial portray of manufacturing firms).

4.4.2. Spatial Autocorrelation (Moran's *I* Test)

Spatial econometrics is used because we are working with firm level data which possess the information about the location of the firms and this location is an additional source of variation. Therefore, it is necessary to use quantitative tools that consider the characteristics unique to each of the observation as well as their location.

The spatial autocorrelation generally defined it as the measurement of the spatial association between a given variable. "It measures the trend of linear relationship between the variables and the degree of intensity of the spatial direction of a given variable with the same variable, but for a defined neighborhood". The first proposed statistical measurement for the spatial autocorrelation is the Moran's *I Test,* and it is the most widely used test of spatial autocorrelation. This test grounded on the measurement of covariance. However, this test only provides the idea of intensity of the average spatial autocorrelation in each sample for a given variable. It measures the relationship between value of suggested variable and value of this suggested variable in its surrounding (Legros & Dube, 2014). This test is robust to detect the presence of spatial autocorrelation between the variables and it provides calculations similar to the correlation coefficient. Due to these reasons, it is the most widely used measure of spatial autocorrelation.

The patterns of given variables can be assessed locally as well as globally and statistics mainly come from geography which measures the spatial dependence. The autocorrelation is constructed on the basis of relationship between the values of given variable and values of observations in its neighborhood. The neighborhood is shown in the form of spatial weights. The advantage of global spatial autocorrelation is real when the spatial observations (firms) are homogeneous, which is rarely the case. So, it is pertinent to consider whether there are local clusters of low and high values. In our case, we try to find the local clusters of firms based on labor pooling. Hence we used number of employees of firms as a variable for labor pooling. It will show the autocorrelation of those firms which have similar labor pooling indicators.

Local spatial autocorrelation allows us to recognize the individual contribution to global spatial autocorrelation. These measurements are used to study the significance of the spatial clusters around individual locations. The main advantage of this method is that it is a descriptive method which gives clear pattern of concentration of low and high values. The indicators of spatial autocorrelation named by LISA are used to test the random distribution of variables and to verify the contribution of couples of points.

Local Moran's *I* indices are written as:

$$I_i = (y_i - \bar{y}) \sum_{j=1}^{N} W_{ij}(y_j - \bar{y}) \qquad \text{for } i \neq j$$

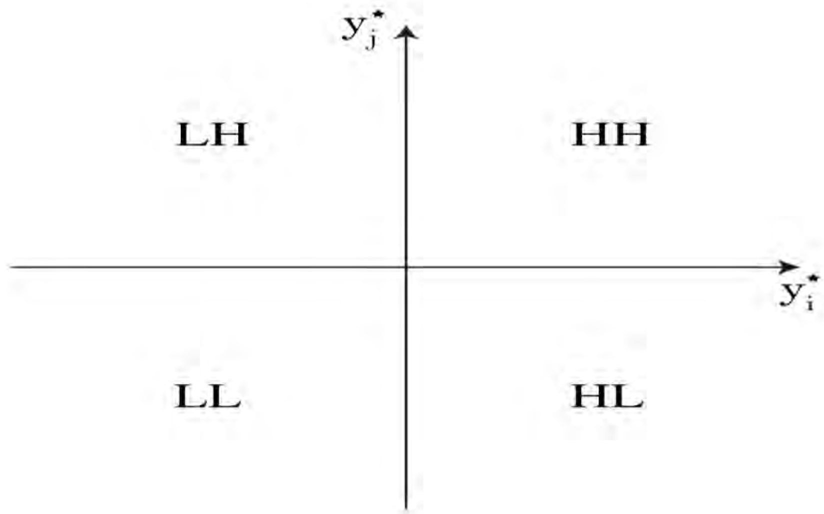
Where N represents the total number of observations, $\bar{y}$ is the arithmetic average of the values taken by the variable y over all the observations, and W_{ij} spatial weights matrix allowing us to link observation i with other observation j.

The significance test can be obtained by calculating the proportion p of the result of the permutations that provide the values of I_i that are greater than, less than, or equal to the observed value of I_i.

The Local Indicator of Spatial Autocorrelation (LISA) Moran's I test will be interpreted as if value of $p<0.10$ then it shows that given variable y_i is related with relatively high value of variable y_j in the neighborhood. An excessively high value of p means $p<0.90$ shows that y_i is related with relatively low value of variable y_j in the neighborhood. Consequently, local Moran's I test classifies the significant zones of spatial clustering for the variable y. It is a cluster of dissimilar values (LISA<0) or the similar values (LISA>0). In this way, the study will construct the indicator of localization where firms were low-low and high-high spatially correlated. It is spontaneous to understand that, if the firms are spatially autocorrelated based on efficiency, they are benefiting from the localization of agglomeration (Legros & Dube, 2014).

The local and global spatial autocorrelation can also be analyzed graphically. This graphical exploration of local indices depends on the Moran scatter plot. It represents original variable on x-axis and spatially lagged variable on y-axis as a function of original variable. The scatter plot is divided into four quadrants and the values of variables lies in one of these quadrants.

Figure 4.5: Moran Scatterplot



The High-High (HH) quadrant shows the case where the values of Moran *I* test for the given variables y* is high and it is also surrounded by the observations that also have high values. The Low-Low (LL) quadrant shows the opposite case of HH quadrant where the low values of given variables are surrounded by the low values of neighboring observations. Our interest is to find those firms which have values in these two quadrants. The values in these two quadrants show that the spatially autocorrelated firms have similar neighborhood (HH, LL) hence labor pooling is occurring in such types of firms. So, we construct a dummy [1] if the firms are in these two quadrants and [0] if the firms are in other quadrants or are insignificant.

The High-Low (HL) quadrant shows the situation where the value of given variable y* is high but it is surrounded by the observations that have low values. Moreover, the Low-High (LH) quadrant shows the inverse situation of HL where the low values of given variables are surrounded by the high values of neighboring observations. If the large number of the values passes through these two quadrants, then it means there is negative global spatial autocorrelation.

The positive spatial autocorrelation shows a typical localization. The observations which are located in HL quadrant are generally known as *diamonds in the rough* and in the opposite case where observations are located the LH quadrant are generally known as *black sheep*.

This spatial analysis (Moran I) is conducted on the labor pooling indicator of localization agglomeration. If the labor pooling is detected in the firms, it will show that the firms are located in a specific neighborhood due to specialization. Labor pooling is analyzed in this case using number of employees taken as an indicator. Once, we have the information for each firm about their nature of spatial dependence based on labor pooling, we construct the dummy variable for firms which are in first two quadrants and are spatially dependent on other firms. Such that:

$$X_2 = \begin{bmatrix} 1 & If\ the\ firms\ are\ in\ \ positive\ autocorrelation\ quadrant \\ 0 & otherwise \end{bmatrix}$$

4.4.3. Measures for Urbanization

For the analysis of urbanization agglomeration, there are two types of analyses performed, using two separate variables of urbanization. First, we constructed a dummy variable which shows whether the firm is located within the city centre or not. Secondly, for those firms which are outside the city centre, we will measure the impact of its distance on the annual turnover of the firm. For this purpose, we have conducted the spline regression to measure the impact of their distance from city centre.

4.3.3.1. Dummy Variable for overall Urban Distance

ArcGIS measures the distance once we input the point data of firm location and city sprawl. The software automatically measures the distance of firms from the polygon boundary of the city centre (see table 6.7 in chapter 6). Once we have distance for each firm from urban centre, dummy variable is created for urban distance representing whether a firm is located within or outside the city.

$$X_1 = \begin{bmatrix} 1 & firm\ located\ within\ city \\ 0 & firm\ located\ outside\ city \end{bmatrix}$$

4.3.3.2. *Spline Regression Analysis*

For those firms who are outside city centre, we employed spline regression to measure the impact of that distance on firm's annual turnover. Based on the linear or non-linear relationship between the dependent and independent variables, the regression analysis can vary. When the relationship between the dependent and independent variable is non-linear, we use number of models one of them, important yet neglected is spline regression model. The spline regression models by the look of it might sound a really complicated model but in reality, it is just a dummy variable model, an unrestricted dummy variable model to be precise that has been adjusted against few restrictions of continuity. Spline regression model can be considered one of the best if not the best alternative to the dummy variable model. Unlike the unrestricted dummy variable models, a spline regression model avoids any sudden jumps when joining the two regression lines. As indicated by (Marsh L. C., 1986) the lines generated by dummy variables are joined together by splines in order to remove sudden jumps in the regression line.

In a spline model, sample is divided into few segments in order to ensure the uniformity in the relationship between the segments created. The regression lines in the spline model are joined by what is known as a "spline knot" or just a "knot". A special and the simplest form of spline model is known as a piecewise linear regression model where the function is continuous, but the slope is not specifically at the point where the two regression lines meet i.e. at the knots. The major use of spline regression model is when a regression line is segregated into pieces and are separated by a point what we are calling a knot. In their book (Pindyck & Rubinfeld, 1998), indicated that the first step in setting up any spline model is to create a special dummy variable that takes the value 0 before we've reached the knot and after that it takes the value 1, as we move on it turns on the remaining dummy variables step by step. The number of dummy variables depends on the knots that we have taken for our model. A detailed discussion on setting up the spline regression model has been provided by (Smith, 1979) in his pioneer work, who developed the simple adjustment approach to spline models. It should be noted that the piece wise regression model in the spline models does account for the continuity restrictions but is not subject to the smoothness i.e., there will be no jumps in the regression lines, but the slopes will be subject to breaks meaning that there will be abrupt changes in slope.

In the estimation of spline models one thing that can change the outcome or the extent to which the data fits your model depends upon the number and location of the spline knots. The estimations of spline model are quite easy and simple if the spline knots are known in advance, however if this is not the case, the estimations become a little bit complicated. If the knots are known in advance, we could directly estimate for the continuity restrictions but if we are dealing with unknown number and location of the knots the traditional approach would be to present it as a maximum likelihood estimation problem. The alternative approach, however, will be to use the stepwise regression method (Marsh L. C., 1986) using the (Smith, 1979) adjustment approach. Using the adjustment approach as proposed by Smith in his pioneer work we start with the case where number, location and degree of knots are known then the unknown number, location and degree is addressed using the step wise regression method for dealing with the problem.

(Suits, Mason, & Chan, 1978) And (Smith, 1979) have in their articles presented a really good approaches in estimating spline regression functions when the spline knot locations are known. Their work was somewhat based on the development of this method by (Fuller, 1969), (Poirier, 1973), (Poirier, 1975), (Poirier, 1976) and (Buse & Lim, 1977).

(Bechard, 2020) Used the spline regression approach to show that a marginal increase in the number of days can have a negative impact on the monthly taxable sales in tourism-related sectors. The study gathered data from the Southwest Florida for estimation of their hypothesis. A 2- kink linear spline model (a model with only one spline knot) was used to see the further variation after the estimation from the linear OLS and quadratic OLS models. After the estimation of the model, it was found that their hypothesis that the marginal increase in the number of days can have a negative impact on the monthly taxable sales in tourism-related sectors was indeed true.

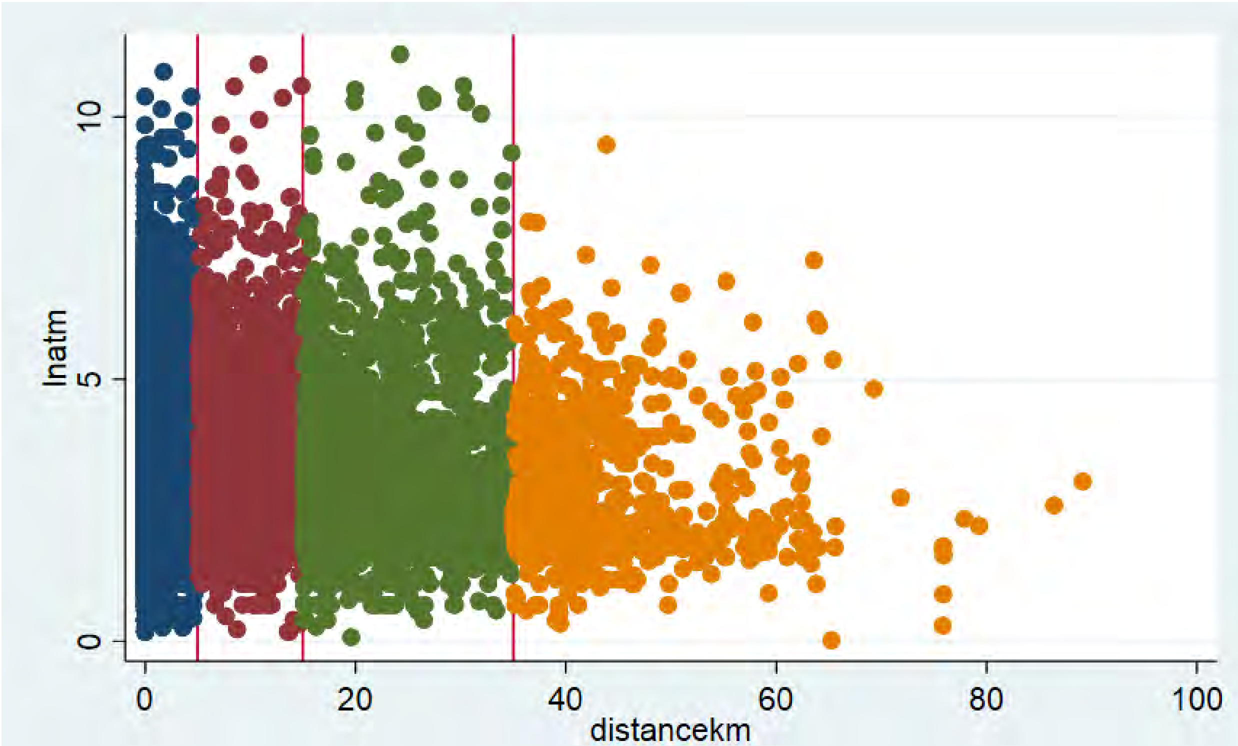
Though the spline regression models are rare in practice since its application is really limited but it does not mean that it hasn't been used entirely. For example, using the smoothing spline regressions confirmed the Bohn's conjecture, (Bechard, 2020) used the spline regression approach to show that a marginal increase in the number of days can have a negative impact on the monthly taxable sales in tourism-related sectors. They used spline regression model in their estimation of finding the relation between public debt and economic growth, (Zareipour, Bhattacharya, & Canizares, 2006) used multivariate adaptive regression splines approach for forecasting the short term behavior of hourly Ontario energy price. they used the spline linear regressions for evaluating the financial assets which proved to better follow the trend from the empirical evidence as compared to the other regression models. They examined the use of spline functions in linear, squared and cubic spline regression models for the estimation of the export-import ratio distribution in Turkey from the year 1923 till 2010, the researchers used the linear spline regression models to study the demand of tourism in Turkey from 2000 till 2010. They used the spline regression models in order to describe that the spline regressions are simple solutions when explaining the historical values and when we are forecasting the new ones.

4.3.3.3. Location and Number of Knots

The study followed a spline regression models to estimate whether the distance has an impact on the annual turnover of the industry by considering different distance points. We took three spline knots on the basis of the distance of the firms from urban centres. The number and location of the knots are presented through a scatter plot in the figure below.

The determined knots represent the distance at 5 kilometers, 15 kilometers and 35 kilometers. The location of the knots is represented by the vertical lines represented within the scatter plot for our data. After estimating the location and number of the knots the impact of the two variables was checked through linear, quadratic and cubic spline regression models overall and also for the individual industrial sectors.

Figure 4.6: Number and Location of Knots



4.3.3.4. Linear Spline Regression Model

A general linear spline regression model is given below.

$$Y = \alpha + \beta_0 x_i + \beta_1 D_{1i}(x_i - x^*) + e$$

Where,

Y= Dependent Variable

X= Independent Variable

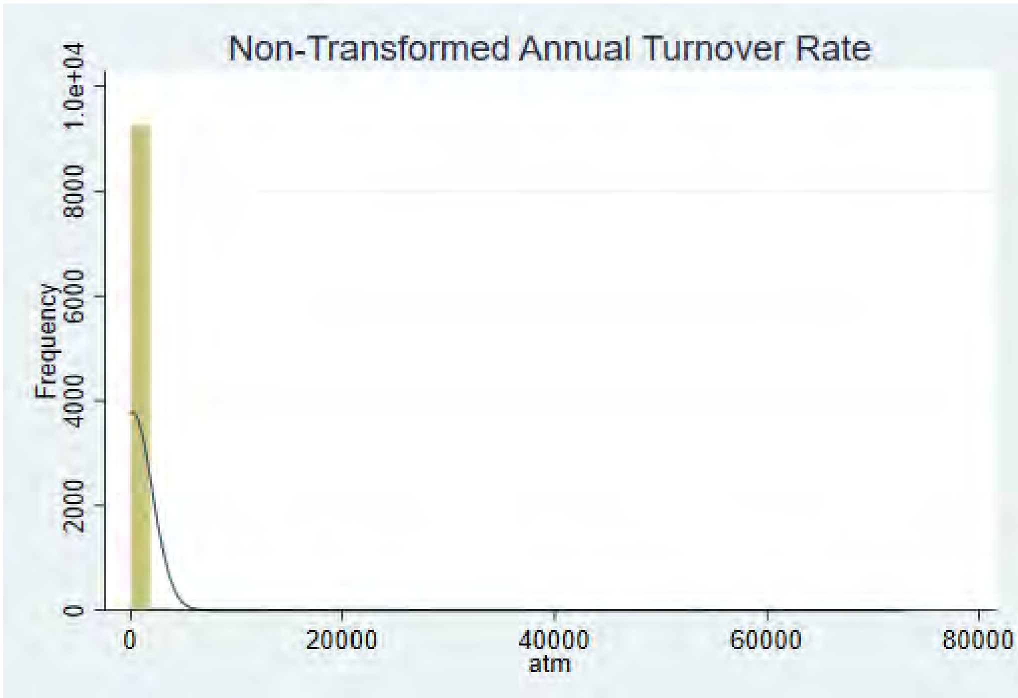
X^*= Spline Knot

D_{1i} = Dummy Variable which is equal to 0 when $\mathbf{X} \leq \mathbf{X}^*$ and 1 when $\mathbf{X} \geq \mathbf{X}^*$

e = Error Term

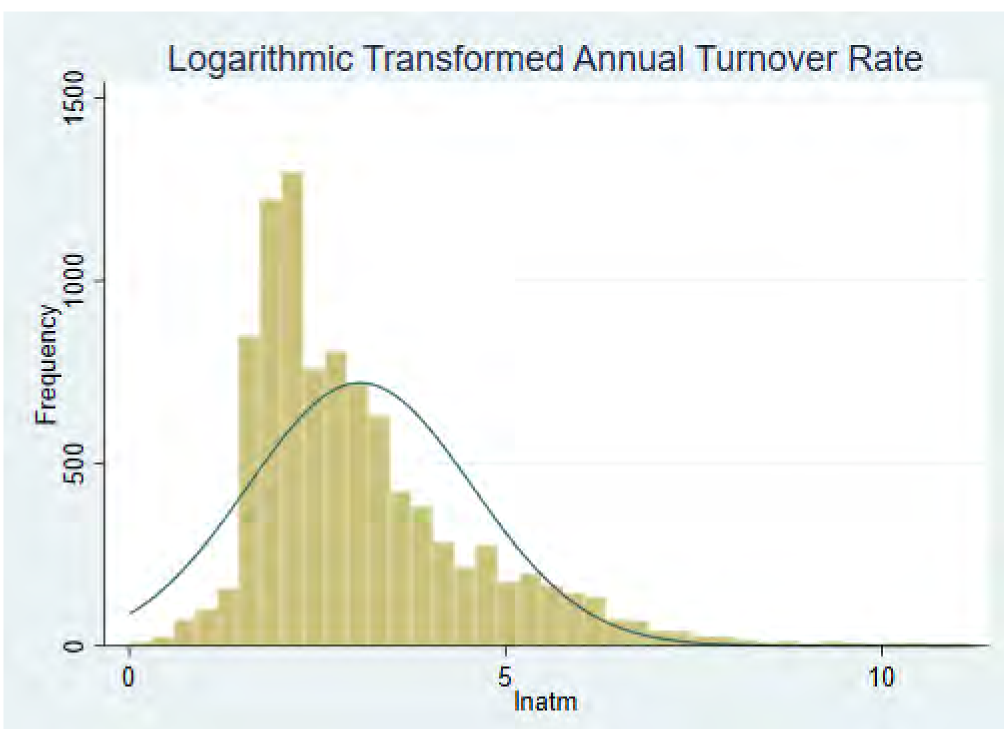
As indicated by the pioneer work in the domain of spline regression models the dummy variable creation is the first step in the estimation of spline model. In the general model given above the dummy variable is turned off when the value of X is less than that of the spline knot X^* and takes the value 1 when the value of X is greater than that of the spline knot X^*. The knots maybe determined through step wise regression or by visualizing the data. We have used a log-linear model for estimation in our study. The main reason behind the logarithmic transformation is to transform a highly skewed variable (annual turnover) into a distribution that is more approximately normal.

Figure 4.8: Non-Transformed Annual Turnover Rate



It can be seen, by plotting the histogram of annual turnover rate we see a significant right skew in the data which means that the data is clustered at lower values. However, when the histogram of

Figure 4.7: Logarithmic Transformation of Annual Turnover Rate



annual turnover rate was plotted, we get a distribution that looks more like a normal distribution as can be seen in the figure below.

The knots that are used for our model were selected subjectively, it was known from our observations that 0 km indicated that the industry is located within the city and as we start to move beyond the 0 mark, we the distance from the main centre increases based on those observations three knots were defined 5, 15 and 35. The following linear spline regression model will be used for our estimations.

$$\boldsymbol{lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5)}$$
$$\boldsymbol{+ \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e}$$

Where,

lnatm is the transformed form of the annual turnover of the industries by taking natural logarithm.

distancekm is the distance from the industry to the nearest centre i.e., city in kilometers

D_{15} is a dummy variable which is 0 when the distancekm is ≤ 5 and 1 when the distancekm is ≥ 5

D_{215} is a dummy variable which is 0 when the distancekm is ≤ 15 and 1 when the distancekm is ≥ 15

D_{335} is a dummy variable which is 0 when the distancekm is ≤ 35 and 1 when the distancekm is ≥ 35

e is the error term

The regression lines for each spline knot are estimated from the following models and are joined at points 5, 15 and 35.

For distancekm ≤ 5

$$\boldsymbol{lnatm = \alpha_0 + \beta_0(distancekm) + e}$$

For distancekm ≤ 15

$$\boldsymbol{lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e}$$

For distancekm ≤ 35

$$\boldsymbol{lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e}$$

For distancekm > 35

$$\boldsymbol{lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e}$$

4.3.3.5. Quadratic & Cubic Spline Regression Models

In order to a get a smoother regression line and to be sure that which model best fits the data we use the quadratic and cubic spline regression models. The choice about the more efficient and the smoothest spline regression model will be made based on the goodness of fit criterion. The quadratic spline regression model that is used for our estimations is given below (the results of these models are presented in annexure III).

$$\boldsymbol{lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2}$$
$$\boldsymbol{+ \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e}$$

The cubic spline regression model that will be used for our estimation

$$\boldsymbol{lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3}$$
$$\boldsymbol{+ \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3}$$
$$\boldsymbol{+ \beta_5 D_{335}(distancekm - 35)^3 + e}$$

4.4.4. Trans-log Production Function

The study has developed the measures of agglomeration factors i.e., localization and urbanization in the previous chapters. In this part, the study will develop a production function to measure the impact of those agglomeration factors on the production process.

Production Function

The trans-log production function has been discovered as a form of production function and it is also the approximation of CES production function. The primary form of trans-log function linked to the proposal presented by J.Kmenta made in 1967. He gave approximation of the CES function along with the Taylor second order conditions. In this condition the elasticity of substitution is very close to the unitary value, and it is also the case in Cobb-Douglas production function.

The Ringstad and Grilichs in 1972 suggested new forms of production function. They imposed the condition that $\alpha+\beta=1$ and in this way production function became the labor productivity function.

$$ln(Y/L) = lnA_2 + \alpha_2 . ln(K/L) + x_2 . ln^2(K/L)$$

The above function is considered single input and in logarithm it is second order polynomial.

The second form of production function was distinct from the previous one as it came with the relaxation of constraints which were imposed on the parameters of the Kmenta function to test the homotheticity assumptions. The function was written as:

$$lnY = lnA_{KL} + \alpha_K . ln\mathrm{K} + \alpha_L . ln\mathrm{L} + \beta_{\mathrm{K}^2} . ln^2\mathrm{K} + \beta_{\mathrm{L}^2} . ln^2\mathrm{L} + \beta_{KL} . ln\mathrm{K} . ln\mathrm{L}$$

This function is also called log-quadratic function and was used by Sargan in 1971. The translog production function was proposed in 1971 and 1973 by Christensen, Jorgenson and Lau and this term has abridged from "transcendental logarithmic production function". The main advantage of this production function is that it does not assume the rigid premises like perfect competition of production factors and perfect substitution between production factors.

In order to understand the production process, there are few assumptions which are required to be made to fully capture the nature of relationship between inputs and outputs. We assume that the manufacturing firms are competitive and have homogenous production function and the firms in the same industry/sector have identical production technologies. We further assume that the localization and urbanization are external factors for the manufacturing firms.

Assuming the separability between intermediate inputs (localization and urbanization) and the primary factors of land, labor, capital and output, the production function of a typical firm j in the sector i is written as:

$$y_{ij} = g_j(X_i) f_j(L_{ij}, E_{ij}, Q_{ij}) \qquad \text{(1A)}$$

Where y_{ij} is the annual turnover of the firm, X_i is the information on the localization and urbanization of the firm, L_{ij} is the labor employed, E_{ij} is the land utilized by the firm, Q_{ij} is the output of the firm. The function $gi(X_i)$ is firm specific function assumed to be independent of production technology of the firm j.

The explicit form the $gi(A_i)$ must be specified to measure the agglomeration economies. We use two variables of agglomeration, as derived in previous section, for urbanization and localization of the firms.

In this study, following Nakamura (1985), the transcendental logarithmic (Translog) production function is used. The translog specification of the equation (1) is given as under:

$$y = A \prod_j X_j^{\alpha_j} e^{\sum_j \alpha_j X_j}$$

$$lny \ = \ \alpha_0 + \sum_j \alpha_j lnX_j + \sum_j \alpha_j X_j$$

The translog full equation specification of the model becomes as flows:

$$lny_j = \alpha_0 + \alpha_1 X_{1j} + \alpha_2 X_{2j} + \alpha_l lnL_j + \alpha_e lnE + \frac{1}{2}\alpha_{ll}(lnL)^2 + \frac{1}{2}\alpha_{ee}(lnE)^2 + \alpha_{le}(lnL)(lnE) \tag{2A}$$

Where X_1 and X_2 are the agglomeration factors of urbanization and localization respectively. The data for these agglomeration variables are taken table 5.2. Equation (2A) is the final translog specification to be estimated in order to assess the impact of agglomeration factors on turnover of the firms.

CHAPTER 5

5. Data and Variable Construction

This chapter is organized as follows: section 5.1 presents the data and sources of data collection; section 5.2 presents construction of variables.

5.1. Sources of Data

Pakistan Bureau of Statistics (PBS) is the custodian of the classification of manufacturing industries in Pakistan. It is called Pakistan Standard Industrial Classification (PSIC) of all economic activities in which PBS has classified all economic activities according to the classification of International Standard Industrial Classification of All Economic Activities (ISIC) by United Nations. Currently, ISIC revision 4 is the latest version available according to which PSIC revision 4 is available from PBS. This classification has divided all the economic activities of economy into three basic sectors i.e., agriculture, industries, and services. Industrial sector is also called manufacturing sector, which has been further categorized into 23 sub-sectors at 2-digit PSIC codes (see table 5.1).

PBS has classified manufacturing industries into two major sectors i.e., Large Scale manufacturing and Small-Scale manufacturing Industry. Large scale manufacturing Industry is defined as those manufacturing establishments which are registered under factoring ordinance 1934 of the government, or qualify for registration, having at least ten or more employees.

This study has used the census data on manufacturing firms of Punjab, who fulfills the criteria of large-scale manufacturing industries, collected by The Urban Unit, Lahore in 2017. They conducted the manufacturing census of Punjab in 2017 in order to ascertain the latest status of manufacturing sector of the province.

5.2. Sample Size

This data set includes the information about manufacturing firms' output level, employees, land use, annual turnover, PSIC classification of the firms, and geographic location. There are a total of 32,500 manufacturing firms in Punjab as identified in the census data. However, after cleaning the data for missing observations, we were left with 25,191 observations. So this study is basically using the census data itself.

Two types of observations were dropped while cleaning the dataset. First, the manufacturing firms which had less than 10 employees were dropped from the data since those do not qualify the large scale manufacturing criteria. Secondly, those values were dropped in which data on output, number of employees, annual turnover was not available.

Table 5.1: Manufacturing Firms in Punjab at 2-Digit PSIC Code

PSIC Code	Sectors	No. of Observations
10	Food Products	3,493
11	Beverages	79
12	Tobacco Products	8
13	Textile	6,747
14	Wearing Apparel	1,653
15	Leather and Related Products	943
16	Wood and Products	469
17	Paper and Paper products	531
18	Printing & reproduction of recorded media	334

19	Coke & Refined petroleum products	32
20	Chemical and Chemical products	473
21	Basic Pharmaceutical	268
22	Rubber and Plastic Products	991
23	Other Non-metallic mineral	2,750
24	Basic Metal	727
25	Fabricated Metal Products	1,612
26	Computer, Electronic and Optical Product	35
27	Electrical Equipment	634
28	Machinery and Equipment	770
29	Motor Vehicles and Trailers	337
30	Other Transport Equipment	173
31	Furniture	793
32	Other Manufacturing (surgical, sports, toys, jewelry)	1,339
	Total	25,202

5.3. Limitations of Data

Although this is the latest dataset available, so far, on the manufacturing firms in Punjab, the questionnaire of the census was very brief and ascertained the information of manufacturing firms on few variables. This dataset contains information only on geographic location, output, land used, number of employees, and annual turnover. The questionnaire should have been comprehensive and include the other dimensions of manufacturing firms such as, status of R&D, wages, sales, costs, carbon emissions etc.

Nonetheless, this is the most recent and advanced dataset available on the manufacturing firms of Punjab, that we have used in this research.

5.4. Construction of Variables

At first stage, this study has constructed the variables for agglomeration indicators. The point data of firm location is used for this purpose. The detailed construction of variables is given below.

5.4.1. Variable for Urbanization

The variables of urbanization and localization were constructed based on the nature of agglomeration. Under urbanization agglomeration, firms take advantage of various types of activities and diversity in area such as cities. As the city expands, more economic benefits can be accrued by the firms. Thus, any firm which lies within the boundary of urban area, are assumed to be benefiting from urban economy.

There are 36 major cities in Punjab, and each represent their district headquarters as well. This study has also measured the distance of the firm from the city sprawl area. A city sprawl is defined as the spatial urban sprawl of the city. Hence, the distance of each firm is measured, in kilometers, from the nearest city. This process is completed in ArcGIS software. This gives us our first variable of urbanization agglomeration which shows distance of firms from urban centres by road (in Km).

Secondly a dummy variable is also constructed for urbanization agglomeration. If the firm lies within the city sprawl, the variable distance will assume the value zero. ArcGIS measures the distance once we input the point data of firm location and city sprawl. Once we have the distance for each firm, a dummy variable is created for urban distance representing whether a firm is located within city or outside city.

$$X_1 = \begin{bmatrix} 1 & firm\ located\ within\ city \\ 0 & firm\ located\ outside\ city \end{bmatrix}$$

5.4.2. Variable for Localization

Since the firms who are benefitting from agglomeration localization are expected to exhibit the Marshallian externalities i.e., input sharing, labor pooling and knowledge sharing. For analyzing the association between firms based on these characteristics, a variable need to be constructed. For that purpose, the variable of localization has been constructed on the basis of labor pooling. To identify the labor pooling, this research uses number of employees of the firm as a variable of it. The literature of economic geography suggests that Moran's *I* test of spatial autocorrelation provide information on whether firms are spatially autocorrelated or not, based on some indicator. This study has employed Moran's *I* test of spatial autocorrelation which explains the similarity of one object with the others surrounding it. We have conducted this test on the number of employees of each firm as a representative of labor pooling. The value of its coefficients lies between -1 and +1. This test gives four types of results HH (High high), HL (high low), LH (low high) and LL (low low). A dummy variable able isn constructed based on the test results where all the firms, which show HH (High high), LL (low low) concentration of labor, are considered spatially autocorrelated. The Moran's I test will provide us with an insight on whether the firms are benefitting from localization agglomeration based on labor pooling (see chapter 5.3.2 for details).

5.5. Other Independent Variables

The study has used land utilized by firm, output of the firm and labor employed as others independent variables which are denoted by E, Q and L respectively.

Table 5.2: Variables and Sources of Data

Variables	Symbols	Unit of Analysis	Variable description

Urbanization	X	In kilometers	Distance of manufacturing firms from the urban centres
	X1	1 = Firm located within the city 0 = Firm located outside the city	A dummy variable constructed for urban distance representing whether a firm is located within city or outside the city.
Localization	X2	1 = Firms are spatially autocorrelated 0 = Firms are not spatially autocorrelated	A dummy variable constructed for localization representing whether firms are spatially correlated with other firms or not.
Annual turn over	Y	PKR Million	It is a quantitative variable representing the annual sale of the firm.
Output	Q	Quantity (in numbers)	It is quantitative variable depicting output produced in quantity by the firm.
Labor employed	L	Number of employees	A quantitative variable showing amount of labor employed by the firm.
Land utilization	E	Marla	A quantitative variable showing land utilized by the firms.

CHAPTER 6

6. Results and Discussion

In this chapter the impact of localization and urbanization on annual turnover has been discussed. Literature shows that both localization and urbanization are playing crucial role in annual turnover. This chapter has been further divided into sub chapters. Chapter 6.1 gives descriptive statistics; chapter 6.2 presents results of urbanization indicators; chapter 6.3 provides results of Moran *I* test; chapter 6.4 presents results of estimation; chapter 6.5 gives detailed results of urbanization agglomeration

6.1. Descriptive Statistics

Descriptive statistics are basically used for describing the basic features of data, for example the measures of central tendency and measures of dispersion. It helps the researcher to manage the data and present the summary statistics for the variables. This section provides a detail insight on distance wise concentration of manufacturing firms. Further, the distance has been divided into 5 categories which includes zero kilometer (km) meaning that firms are located within the cities, 0-4km, 5-14km, 15-35km and above 35km.

6.1.1. Distance Wise Concentration of Manufacturing Firms

Table 6.1 shows the distance of manufacturing firms from their respective urban centres. The manufacturing firms of food products and other non-metallic minerals are located farthest from their respective urban centres with is 89 and 86 kilometers respectively. The fabricated metal, wood and wood products and furniture manufacturing firms are located at 63.49, 59.63 and 59.19 kilometers from their urban centres respectively. The manufacturing firms of computer, electronic and optical products are located within their urban cores. The maximum distance of these firms from their urban areas is 1.23 kilometers which shows that these manufacturing firms are located at closer distances from the urban centres.

Table 6.1: Distance-wise Concentration of Manufacturing Firms

Sectors	Distance in Kilometers from the Urban centre							
	No. of manufacturing firms	Average Distance	S.D	Min	.25	Median	.75	Max
Food Products	3493	13.18	14.85	0	0	7.65	23.16	89.14
Beverages	79	6.12	11.28	0	0	0	8.33	39.39
Textile	6747	2.09	5.83	0	0	0	0	48.2
Wearing Apparel	1653	1	3.64	0	0	0	0	29.81
Leather and Related Products	943	2.62	8.09	0	0	0	0	41.41
Wood and Products	469	19.76	17.45	0	0	17.42	39.38	59.63
Paper and Paper products	531	1.45	4.62	0	0	0	0.01	39.59
Printing & reproduction of recorded media	334	0.35	2.4	0	0	0	0	37.96
Coke & Refined petroleum products	32	6.54	7.45	0	0	4.54	9.43	26.49
Chemical and Chemical products	473	3.12	7.33	0	0	0	1.49	45.14
Basic Pharmaceutical	268	3.52	7.5	0	0	0	4.8	49.72
Rubber and Plastic Products	991	1.15	4.04	0	0	0	0	40.69
Other Non-metallic mineral	2750	13.93	14.49	0	0.31	9.79	24.3	86.42
Basic Metal	727	1.83	5.64	0	0	0	0	39.55
Fabricated Metal Products	1612	4.31	8.76	0	0	0	6.73	63.49
Computer, Electronic and Optical Product	35	0.08	0.28	0	0	0	0	1.23
Electrical Equipment	634	0.81	4.77	0	0	0	0	46.7
Machinery and Equipment	770	2.56	8.04	0	0	0	0	53.84
Motor Vehicles and Trailers	337	2.3	5.95	0	0	0	0	39.39
Other Transport Equipment	173	1.14	3.98	0	0	0	0	24.04
Furniture	793	8.14	13.39	0	0	0	13.02	59.19
Other Manufacturing (surgical, sports)	1339	2.88	6.4	0	0	0	0.69	37.4

6.1.2. Distance of Manufacturing Firms from Urban Centre

The below table summarizes the distance of manufacturing firms from their urban centres. It indicates that, more than half of the manufacturing firms are located within their respective urban centres, 10.88% are located within the 4km range, 10.41% are located within the 5-14 km range, 12.01% are located within the 15-35km range and 3.95% of the manufacturing firms are located beyond 35 kilometers from their urban centres.

Table 6.2: Overall Urban Distance of Manufacturing Firms

Distance in Kilometers (Km)	Frequency	Percent
0	15,811	62.74
0-4	2,743	10.88
5-14	2,624	10.41
15-35	3,028	12.01
Above 35	996	3.95

6.1.3. Sector-Wise Distance of Manufacturing Firms from Urban Centre

Table 6.3 summarizes the sector wise distance of manufacturing firms from the urban centres. It can be observed that a significant number of manufacturing firms among all the sectors are located within the cities with 91% of the manufacturing firms of electronic equipment and printing products and media located within the urban cores. As compared to the other sectors, the spread of food products manufacturing firms is huge. From the table we can establish a fact that as we move farther from the urban cores, the number of manufacturing firms decline, and the major chunk of manufacturing firms are located within the urban centres.

Table 6.3: Sector-wise Location of Manufacturing Firms

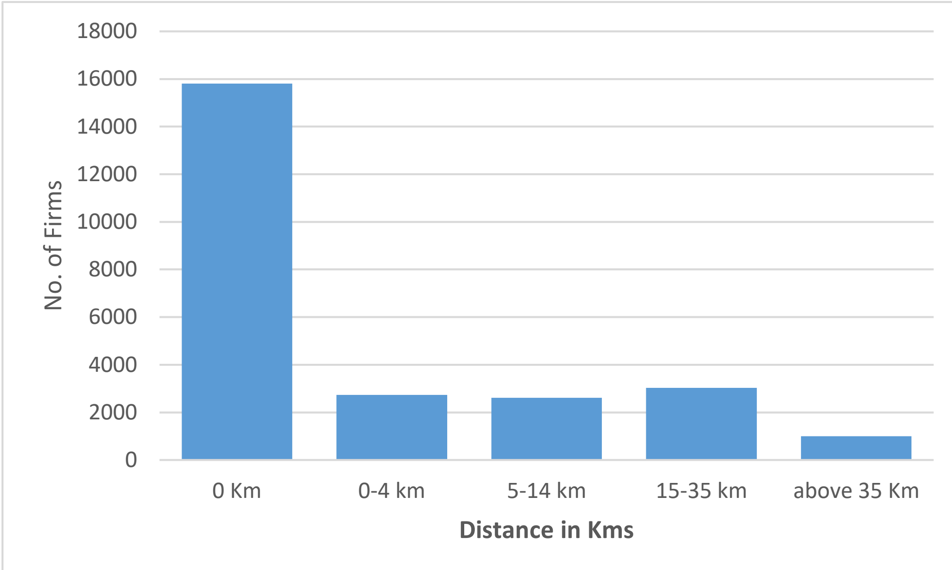
PSIC Code	Sectors	0 Km	0-4 km	5-14 km	15-35 km	above 35 Km
10	Food Products	1,002 (29%)	504 (14.428%)	597 (17.091%)	1,013 (29%)	377 (10.793%)
11	Beverages	45 (56.962%)	9 (11.392%)	11 (13.924%)	8 (10.126%)	6 (7.594%)
13	Textile	5,075 (75%)	805 (12%)	516 (8%)	326 (5%)	25 (0%)
14	Wearing Apparel	1,362 (82%)	182 (11%)	66 (4%)	43 (3%)	0 (0%)
15	Leather and Related Products	780 (82.714%)	48 (5.0901%)	44 (4.665%)	48 (5.090%)	23 (2.439%)
16	Wood and Products	124 (26%)	31 (7%)	51 (11%)	120 (26%)	143 (30%)
17	Paper and Paper products	395 (74%)	78 (15%)	41 (8%)	15 (3%)	2 (0%)
18	Printing & reproduction of recorded media	305 (91.317%)	17 (5.089%)	11 (3.293%)	0 (0%)	1 (0.299%)
19	Coke & Refined	10 (31.25%)	6 (18.75%)	9 (28.125%)	7 (21.875%)	0 (0%)

	petroleum products					
20	Chemical and Chemical products	306 (64.693%)	78 (16.490%)	47 (9.936%)	37 (7.822%)	5 (1.057%)
21	Basic Pharmaceutical	168 (62.686%)	32 (11.940%)	48 (17.910%)	17 (6.3432%)	3 (1.1194%)
22	Rubber and Plastic Products	837 (84.460%)	67 (6.7608%)	59 (5.953%)	27 (2.724%)	1 (0.1009%)
23	Other Non-metallic mineral	584 (21.236%)	473 (17.2%)	541 (19.672%)	880 (32%)	272 (9.890%)
24	Basic Metal	578 (79.504%)	63 (8.665%)	57 (7.840%)	22 (3.0261%)	7 (0.962%)
25	Fabricated Metal Products	1,123 (70%)	68 (4%)	268 (17%)	111 (7%)	42 (3%)
26	Computer, Electronic and Optical Product	31 (88.571%)	4 (11.428%)	0 (0%)	0 (0%)	0 (0%)
27	Electrical Equipment	583 (91.955%)	28 (4.416%)	9 (1.419%)	9 (1.419%)	5 (0.788%)

28	Machinery and Equipment	656 (85%)	30 (4%)	17 (2%)	54 (7%)	13 (2%)
29	Motor Vehicles and Trailers	262 (77.744%)	25 (7.418%)	21 (6.231%)	28 (8.308%)	1 (0.2967%)
30	Other Transport Equipment	149 (86%)	11 (6%)	6 (3%)	7 (4%)	0 (0%)
31	Furniture	472 (59.520%)	45 (5.674%)	82 (10.340%)	128 (16.141%)	66 (8.322%)
32	Other manufacturing (Surgical, sports etc.)	955 (71%)	136 (10%)	117 (9%)	127 (9%)	4 (0%)

A graphical summary of the above table 6.3 is given below:

Figure 6.1 Total Number of Firms against Distance in Kilometers



6.1.4. Sector-Wise Annual turnover of Manufacturing Firms

Table 6.4 provides the sector-wise annual turnover of the manufacturing firms. The beverages, chemical and chemical products and food products are the sectors with higher average annual turnover where, the wood products have the lowest reported average annual turnover as compared to the other manufacturing firms.

Table 6.4: Annual Turn-over of Manufacturing Firms (PKR Million)

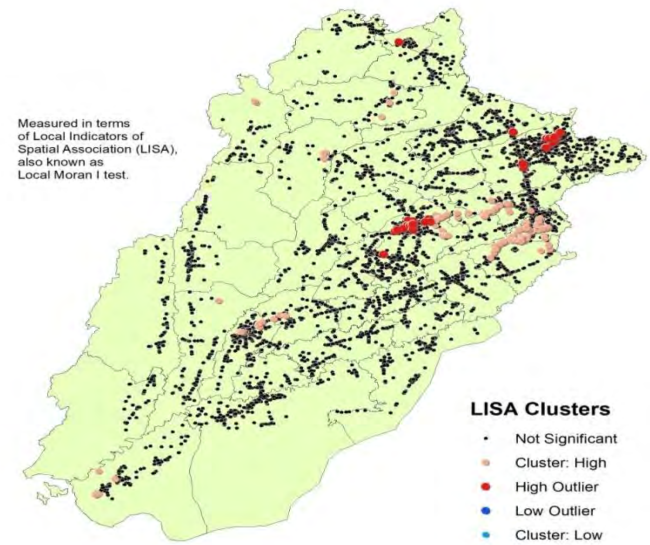
Sectors	Sector- Wise Annual turnover of Manufacturing Firms				
	No. of manufacturing firms	Average Annual Turn Over	S. D	Min	Max
Food Products	3493	242.7373	1895.492	1.02	60,000
Beverages	79	501.6908	1624.83	5	10,000
Textile	6747	115.1104	1240.391	1.2	52,348
Wearing Apparel	1653	81.93682	300.4024	1.2	6,024

Leather and Related Products	943	55.57481	294.8664	1.4	7,000
Wood and Products	469	12.55825	36.57026	1.5	500
Paper and Paper products	531	64.37271	219.2712	1.3	3,000
Printing & reproduction of recorded media	334	52.49187	157.785	1.5	2,000
Coke & Refined petroleum products	32	114.9338	212.0969	2	990
Chemical and Chemical products	473	471.5314	3877.532	1.5	72,877
Basic Pharmaceutical	268	140.8459	243.4574	2.38	2,880
Rubber and Plastic Products	991	81.87603	715.3452	1.5	18,900
Other Non-metallic mineral	2750	56.29089	853.1841	1.2	29,703
Basic Metal	727	211.1236	1369.408	1.5	30,694
Fabricated Metal Products	1612	26.26219	181.8085	1.32	7,000
Computer, Electronic and Optical Product	35	121.87	202.2801	4	1,000
Electrical Equipment	634	67.66746	247.6688	1.5	3,000
Machinery and Equipment	770	34.18072	126.0422	2	2,500
Motor Vehicles and Trailers	337	148.247	781.2342	3	12,000
Other Transport Equipment	173	120.7292	252.0317	3.5	2,160
Furniture	793	20.63983	73.05296	1.5	1,035
Other Manufacturing (surgical, sports)	1339	64.13352	415.4083	1.2	12,000

6.2. Results of Localization Indicator (Moran *I* test)

For this study, spatial autocorrelation among manufacturing firms in Punjab has been observed on the basis of labor pooling. The following map provides insight of overall spatial autocorrelation of manufacturing firms in Punjab. It shows how the firms are agglomerated based on localization indicator. This analysis provides a detailed information about the spatial autocorrelation of manufacturing firms in every sector.

Figure 6.2: Spatial Autocorrelation among Manufacturing Firms in Punjab



The above map is showing the spatial autocorrelation of manufacturing firms in Punjab. The legend provides results where all significant manufacturing clusters are, and which firms are insignificant. The map shows spatially autocorrelated firms and are differentiated on the basis of high or low neighboring values. The black dots on the map are showing insignificant results of Moran I test. The pink dots are depicting that the value of given variable (labor pooling) is high, and it is also surrounded by the high value of neighboring observations where, the red dots are showing the high value of given variable is surrounded by the observations that have low values. Hence the firms depicting cluster high (HH) and cluster low (LL) are showing the pooling of labor in a specific geography and show the significant clustering of firms based on localization indicator (labor pooling). The results and spatial portray of Moran I are provided below for few sectors, remaining sectors are provided in annexure II.

6.2.1. Results of Sector-Wise Spatial Autocorrelation

Figure 6.3: Spatial Autocorrelation in Textile Sector

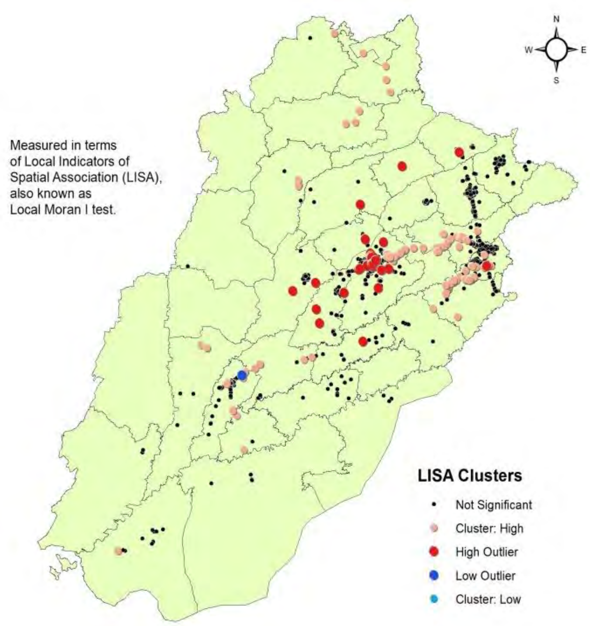
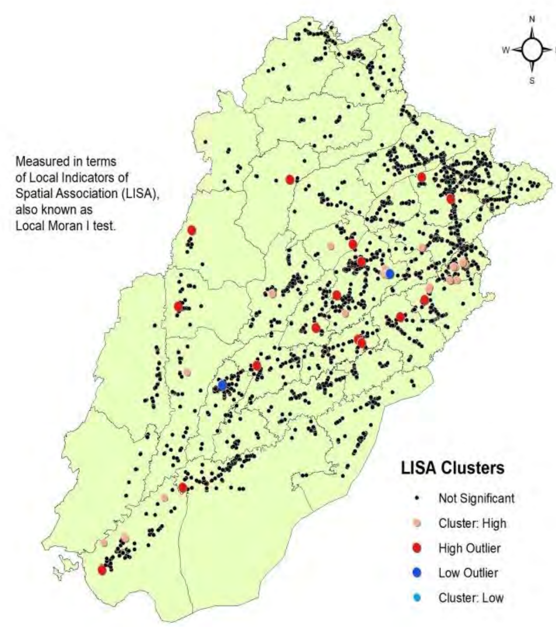


Figure 6.4: Spatial Autocorrelation in Food Products Sector



The left map is showing spatial autocorrelation of firms in the textile sector of Punjab. The pink dots are depicting High-High clusters of the firms in the textile sector of Punjab which means that high values of given variable are also surrounded by the high values of neighboring observations. The red dots are showing that high values of given variable are surrounded by low values of neighboring observations. The blue dots depict the low outliers which means that low values of given variable are surrounded by the observations which have high value. The black dots are the insignificant results of the Moran I test.

The right map is showing spatial autocorrelation in food sector of Punjab district. The black dots are the insignificant results of the Moran I test. The pink dots are explaining that high values of food sector are surrounded by high values of neighboring observations. The red dots are depicting the high value of given variable, but it is surrounded by the observations that have low values. Similarly, the blue dots which are also known as low outliers because the low value of given variable is surrounded by the high value of neighboring observations.

Figure 6.5: Spatial Autocorrelation in Other Non-Metallic Mineral Sector

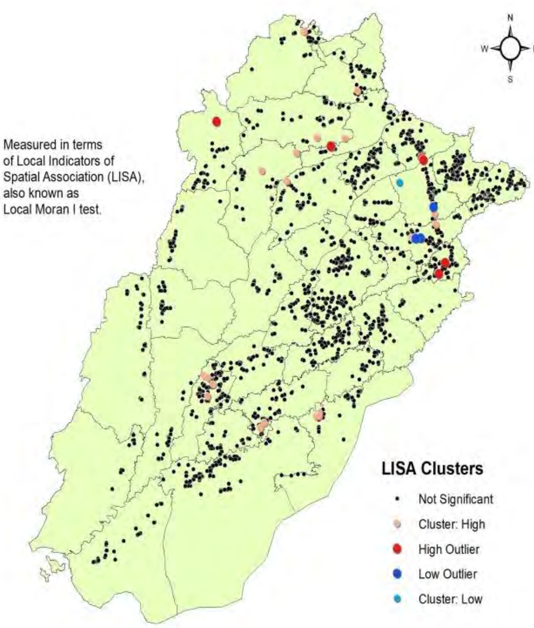
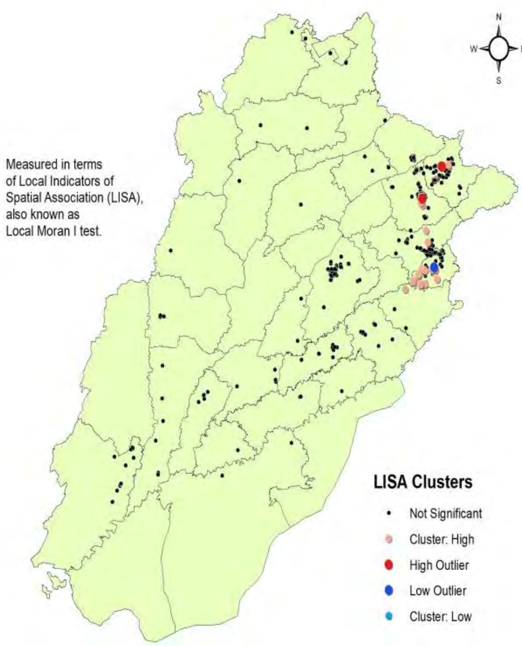


Figure 6.6: Spatial Autocorrelation in Fabricated Metal Products Sector



The left map is showing spatial autocorrelation in other non-metallic mineral sector of Punjab district. The black dots are the insignificant results of Moran I test. The pink dots are explaining the high clusters of other non-metallic mineral sector which means the high value of non-metallic mineral sector are surrounded by high values of neighboring observations. The red dots depict the high value of given variable, but it is surrounded by the observations that have low values also known as the high outlier. Similarly, the blue dots are known as low outliers because the low value of given variable is surrounded by the high value of neighboring observations.

The right map is showing spatial autocorrelation in fabricated metal product sectors in Punjab. The black dots are depicting insignificant results of the Moran I test. The pink dots in map are showing high clusters of fabricated metal products which means that high values of given variable are surrounded by the high value of neighboring observations. The red dots are depicting the high value of given variable, but it is surrounded by the observations that have low value also known as the high outlier. Similarly, the blue dots are known as the low outliers because the low values of given variable are surrounded by the high values of neighboring observations.

Figure 6.7: Spatial Autocorrelation in Wearing Apparel Sector

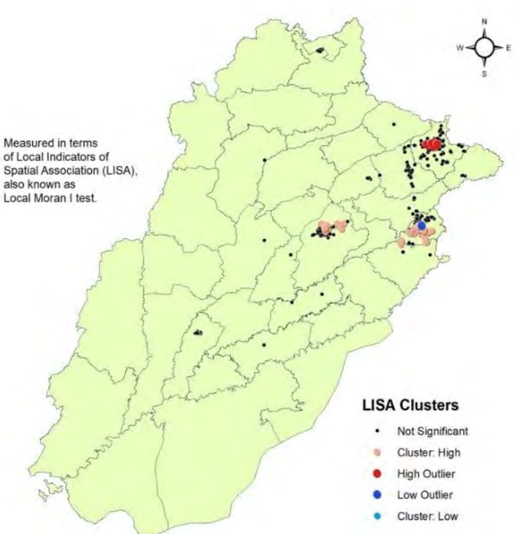
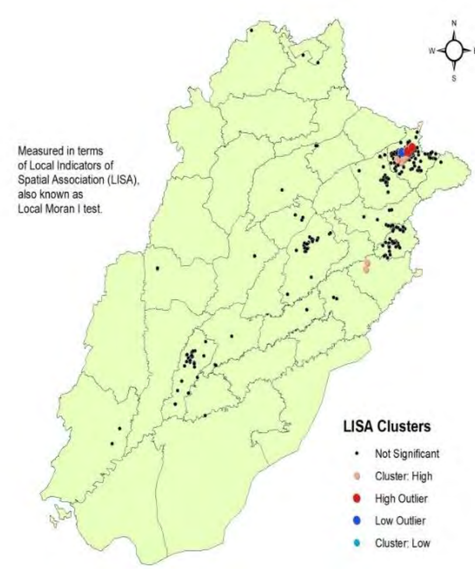


Figure 6.8: Spatial Autocorrelation in Other Manufacturing Sector



The left map is showing spatial autocorrelation in wearing apparel sector of Punjab. The black dots are the insignificant results of the Moran I test. The pink dots are explaining the high clusters of firms of wearing apparel sector which means the high values of the given variable are surrounded by high values of neighboring observations. The red dots are depicting the high value of given variable, but it is surrounded by the observations that have low value also known as the high outliers. Similarly, the blue dots are known as low outliers because the low value of given variable are surrounded by the high value of neighboring observations.

The right map is showing spatial autocorrelation in other manufacturing sectors of Punjab. The black dots are depicting insignificant results of the Moran I test. The pink dots in map are showing high clusters of firms in other manufacturing sector which means that the high values of the given variable are also surrounded by the high value of neighboring observations. The red dots are depicting the high value of given variable, but it is surrounded by the observations that have low value also known as the high outliers. Similarly, the blue dots are also showing negative spatial autocorrelation, and these are called low outliers because the low value of given variable is surrounded by the high value of neighboring observations.

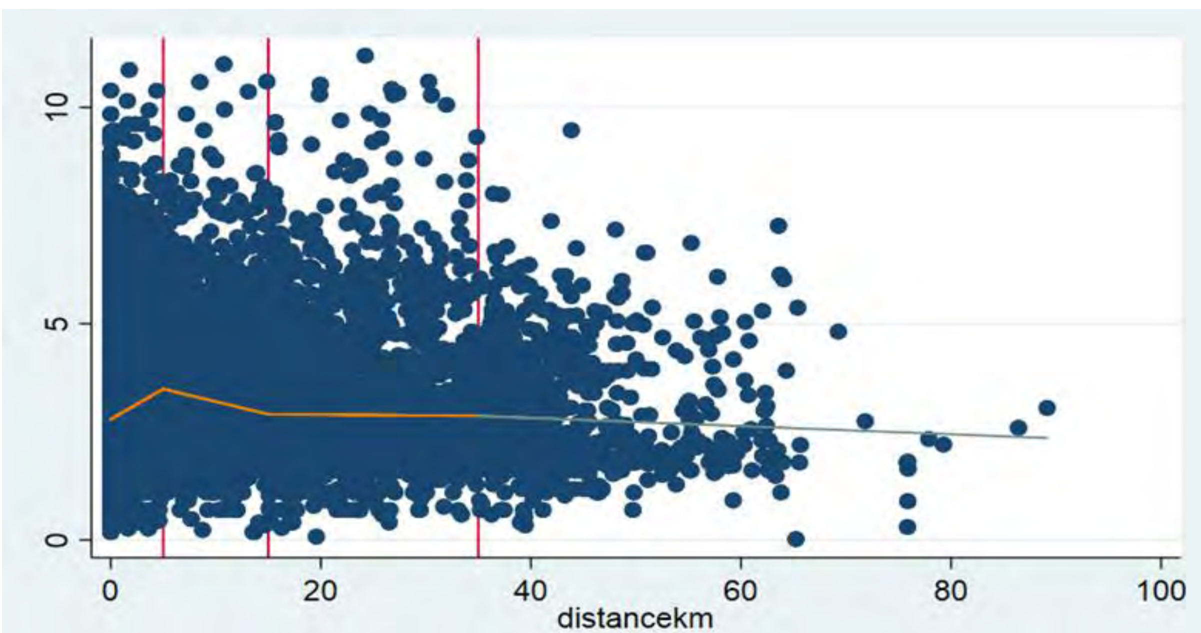
6.3. Results of Urbanization Indicators

The study has estimated the impact of distance on the annual turnover of the industries and for this purpose, this study has used a linear spline regression analysis (also known as piece-wise regression analysis). Different goodness of fit measures was utilized in order to estimate how well our data fits the models proposed.

6.3.1. Spline Regression for all Firms

From the estimation results the annual turn-over of the all the manufacturing firms has a positive relation with distance within the 5km. When the distance is less than or equal to 5km, 1 km increase in distance increases the annual turn-over of the industries by almost 15% which is equal PKR 421 million.

However, the annual turn-over decreases if the firm is located within the 5-15km range. If the distance is less than or equal to 15 km, 1 km increase in distance decreases the annual turn-over of the industries by almost 5.7% which in monetary terms is approximately equal to PKR 160 million. Similarly, the annual turn-over decreases if the firm is located within 15-35km range. When the distance is less than or equal to the 35-kilometer 1 km increase in distance decreases the annual turn-over by almost 0.19% which amounts to PKR 5.33 million. Similarly, the annual turn-over of



the firm decreases if the firm is located beyond the 35km range. When the distance exceeds the 35 kilometers mark it was observed that 1 km increase in distance decreases the annual turn-over of all the industries by almost 0.94% that amounts to PKR 26.4 million.

The results are in accordance with the theory of urbanization as proposed by (Jacobs, The Economies of Cities, 1969). The manufacturing firms are indeed benefiting when they are located closer to their urban centres i.e., when the manufacturing firms are closer to the urban centres a higher annual turn-over is reported as compared to those manufacturing firms that are located far from the urban centres. However, it cannot be generalized for all sectors as we will see later that the nature of manufacturing firms and other factors also has a role to play in determining the impact of distance on the annual turn-over of firms.

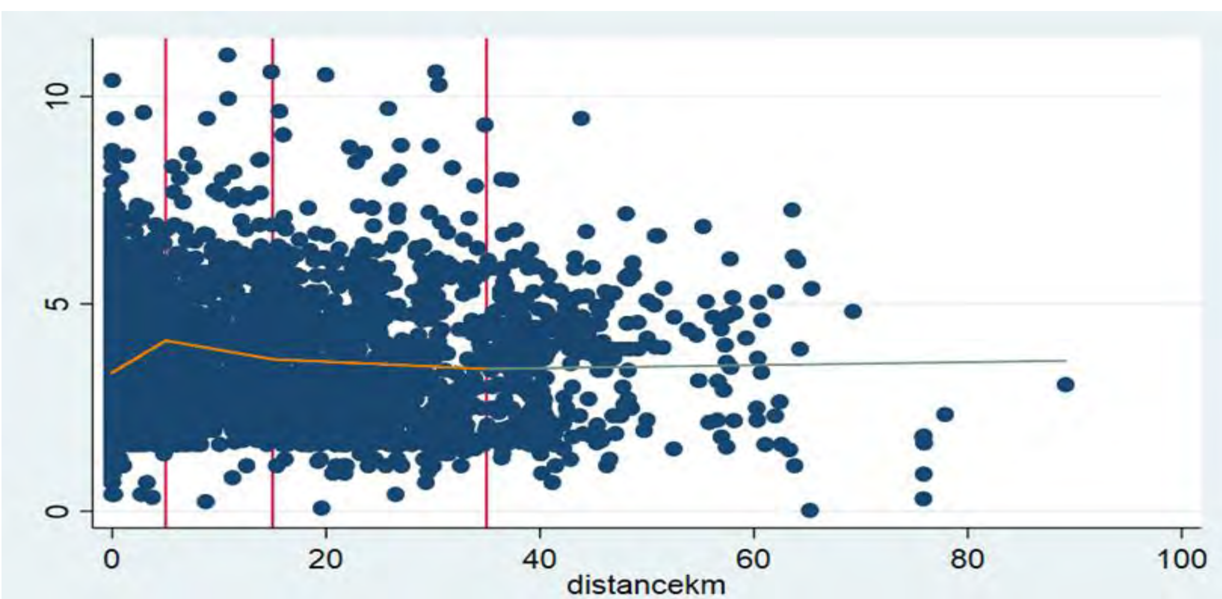
The estimated regression lines are graphically summarized below for few sectors whereas remaining all sectors' details are provided in annexure III.

6.3.2. Spline Regressions for Different Manufacturing Sectors

Food Products

According to the estimation results, the annual turn-over of the manufacturing firm in food sector has a positive relation with the distance from urban centre within 5km range. When the distance is

Figure 6.9: Spline Regression for Food Products



less than or equal to 5 km, 1 km increase in distance increases the annual turn-over of the industries by almost 16.7% in manufacturing firms of food sector.

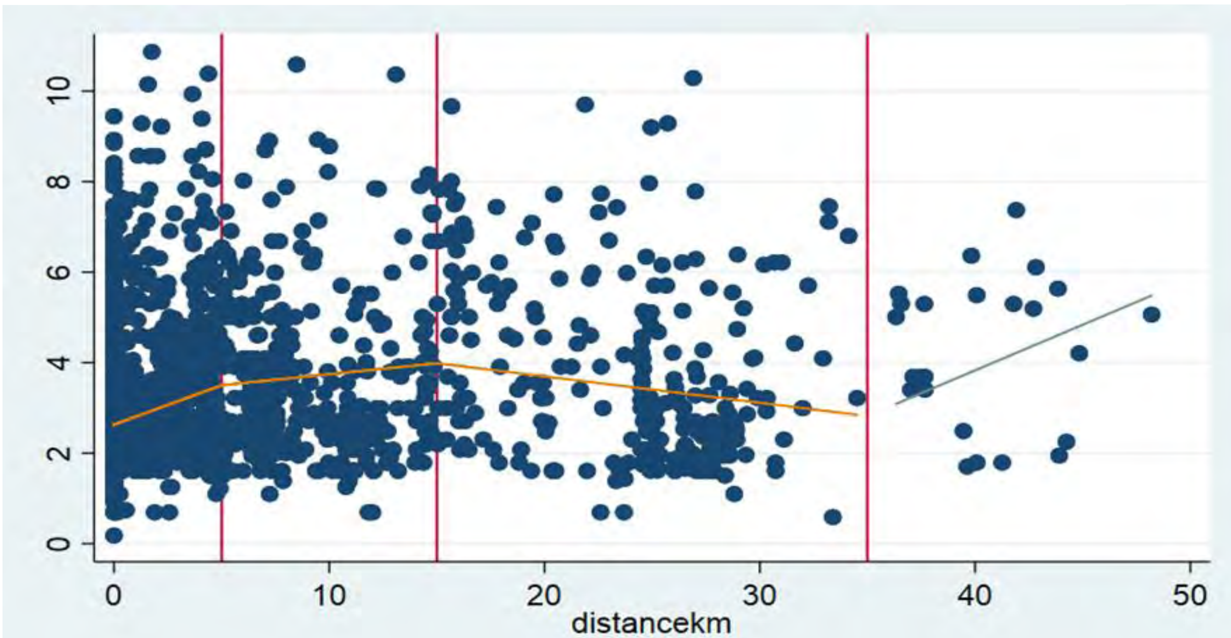
However, the annual turn-over decreases if the firm is located within 5-15km range. If the distance is less than or equal to 15 km, 1 km increase in the distance decreases the annual turn-over of the industries by almost 4.42%. Similarly, the annual turn-over decreases if the firm is located within 15-35km range. If the distance is less than or equal to the 35-kilometer 1 km increase in distance decreases the annual turn-over by almost 1.2%. Whereas the annual turn-over increases if the firm is located beyond the 35km range. When the distance exceeds the 35 kilometers mark it was observed that 1 km increase in distance increases the annual turn-over of food industries by almost 0.36% which amounts to PKR 0.9 million.

Textile

According to the estimation results, the annual turn-over of the manufacturing firms in textile sector has a positive relation with distance from urban centre within the 5km range. When the distance is less than or equal to 5 km, 1 km increase in distance increases the annual turn-over of the industries by almost 19%.

Similarly, the annual turn-over increases if the firm is located within the 5-15km range. If the distance is less than or equal to 15 km, 1 km increase in the distance increases the annual turn-over of the industries by almost 5%. However, the annual turn-over decreases if the firm is located within the 15-35km range. If the distance is less than or equal to the 35 kilometer 1 km increase in distance decreases the annual turn-over by almost 5.7% . Whereas the annual turn-over increases if the firm is located beyond the 35km range. When the distance exceeds the 35 kilometers mark it was observed that 1 km increase in distance increases the annual turn-over of textile industries by almost 22.3%. The estimated regression lines are graphically summarized below.

Figure 6.10: Spline Regression for Textile



Other Non-Metallic Minerals

According to the estimation results, the annual turn-over of the manufacturing firms in Other Non-Metallic Minerals has a positive relation with relation with distance within the 5km range. If the distance is less than or equal to 5 km, 1 km increase in distance increases the annual turn-over of the industries by almost 0.8%.

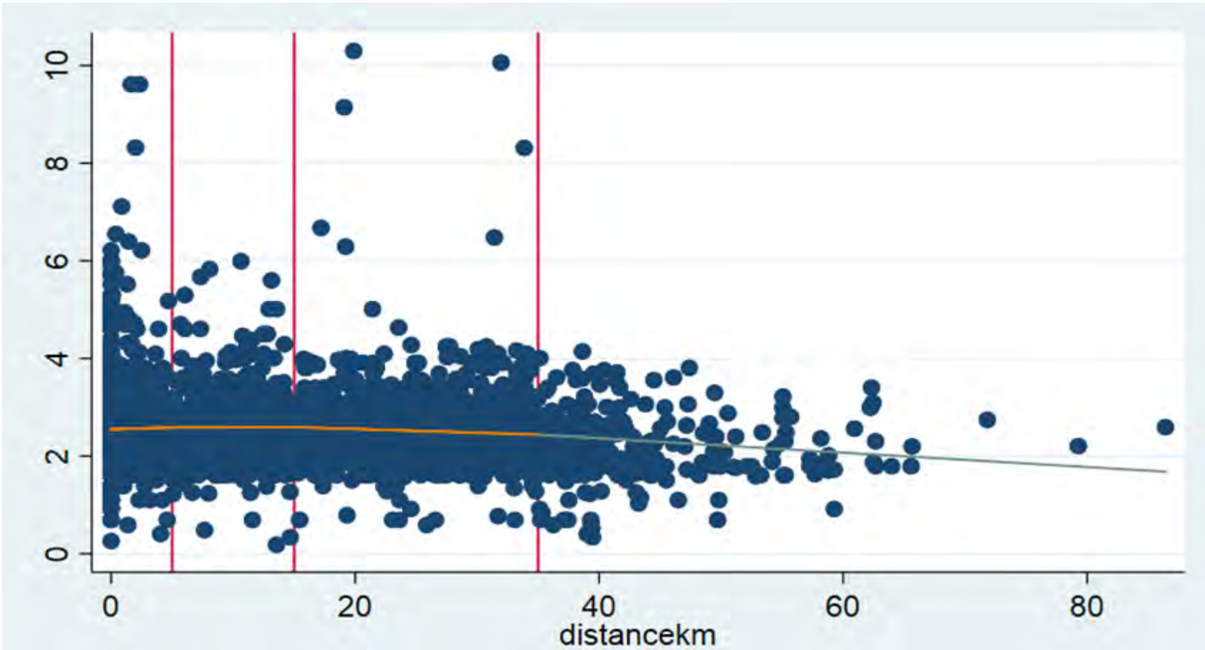


Figure 6.11: Spline Regression for Other Non-Metallic Minerals

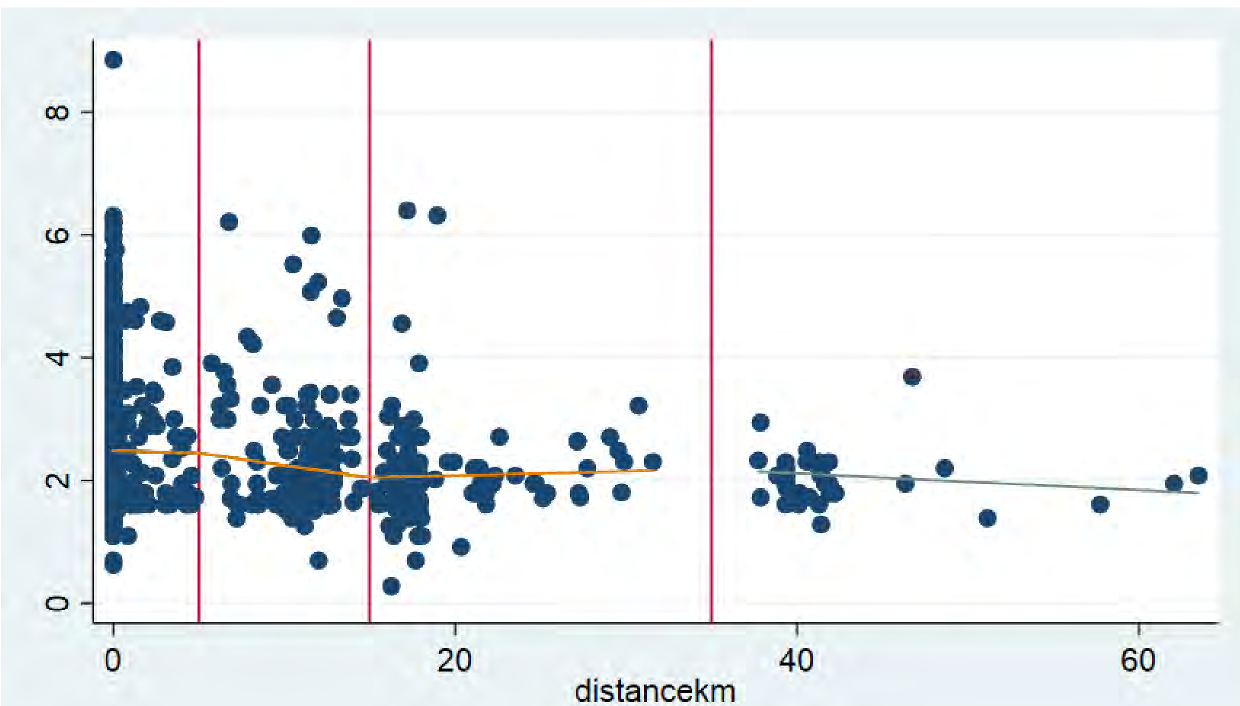
Similarly, the annual turn-over increases if the firm is located within the 5-15km range. When the distance is less than or equal to 15 km, 1 km increase in the distance increases the annual turn-over of the industries by almost 0.05%. However, the annual turn-over decreases if the firm is located within the 15-35km range. If distance is less than or equal to the 35-kilometer 1 km increase in distance decreases the annual turn-over by almost 0.8%. Similarly, the annual turn-over decreases if the firm is located beyond the 35km range. If the distance exceeds the 35 kilometers mark it was observed that 1 km increase in distance decreases the annual turn-over by almost 1.5%.

Fabricated Metal Products

According to the estimation results, the annual turn-over of manufacturing firms in fabricated metal products has a negative relation with distance within the 5km range. If the distance is less than or equal to 5 km, 1 km increase in distance decreases the annual turn-over of the industries by almost 0.72%.

Similarly, the annual turn-over decreases if the firm is located within 5-15km range. If the distance is less than or equal to 15 km, 1 km increase in the distance decreases the annual turn-over of the industries by almost 4%. However, the annual turn-over increases if the firm is located within 15-35km range. If the distance is less than or equal to the 35 kilometer 1 km increase in distance increases the annual turn-over by almost 0.7%. Whereas the annual turn-over decreases if the firm is located beyond 35km range. When the distance exceeds the 35 kilometers mark it was observed that 1 km increase in distance decreases the annual turn-over by almost 1.35%. The estimated regression lines are graphically summarized below.

Figure 6.12: Fabricated Metal Products

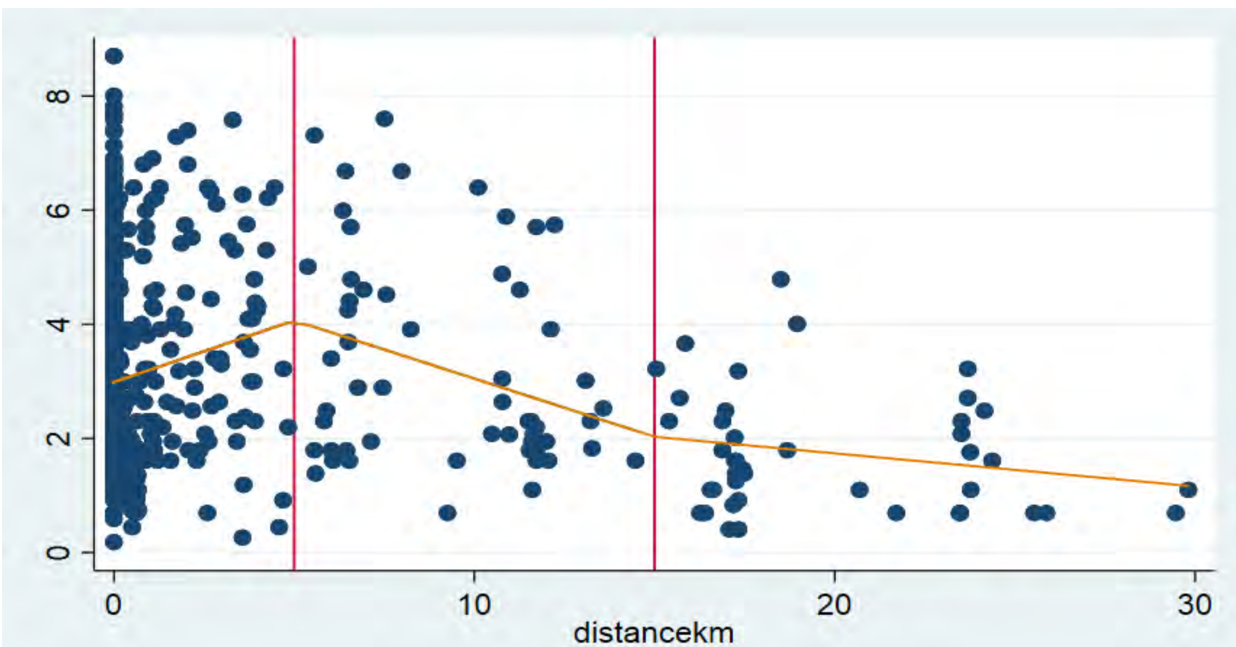


Wearing Apparel

According to the estimation results, the annual turn-over of the manufacturing firms in the wearing apparel sector have a positive relation with distance within the 5km range. When the distance is less than or equal to 5 km, 1 km increase in distance increases the annual turn-over of the industries by almost 24.2%.

However, the annual turn-over of firm decreases if the firm is located within the 5-15km range. If the distance is less than or equal to 15 km, 1 km increase in the distance decreases the annual turn-over of the industries by almost 18.5%. Similarly, the annual turn-over of firm decreases if the firm is located within 15-35km range. If the distance is less than or equal to the 35 kilometer, 1 km increase in distance decreases the annual turn-over by almost 5.63%. The estimated regression lines are graphically summarized below.

Figure 6.13: Spline Regression for Wearing Apparel

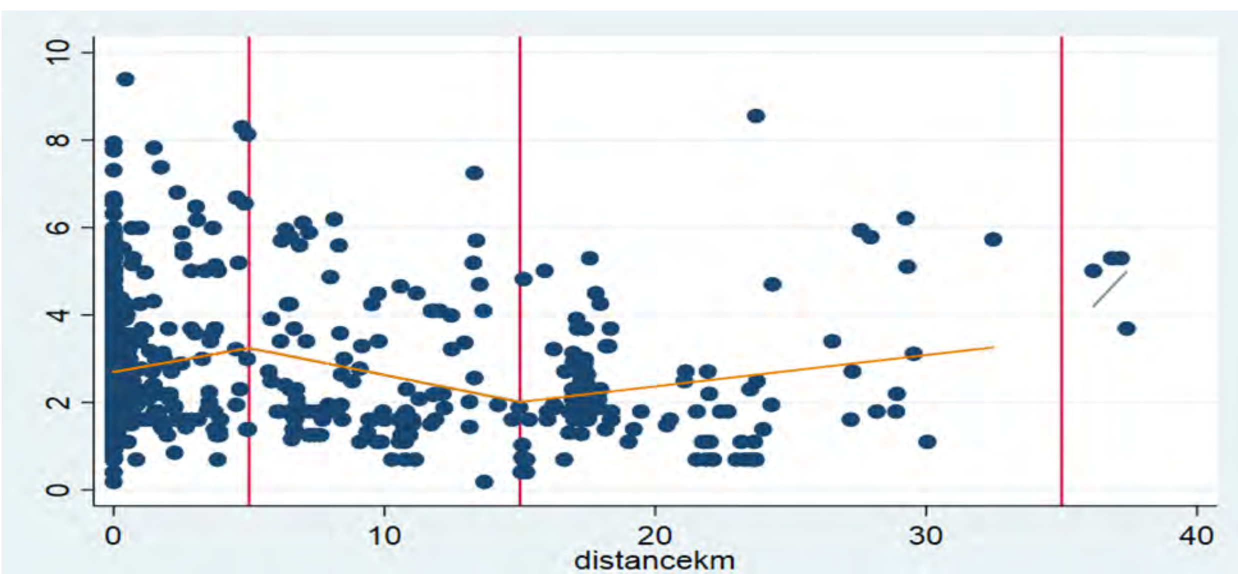


Other Manufacturing

According to the estimation results, the annual turn-over of the manufacturing firms in Other Manufacturing sector has a positive relation with distance within the 5km range. If the distance is less than or equal to 5 km, 1 km increase in distance increases the annual turn-over of the industries by almost 12%.

However, the annual turn-over decreases if the firm is located within 5-15km range. When the distance is less than or equal to 15 km, 1 km increase in the distance decreases the annual turn-over of the industries by almost 12%. Whereas the annual turn-over increases if the firm is located within 15-35km range. When the distance is less than or equal to the 35 kilometer 1km increase in distance increases the annual turn-over by almost 7%. Similarly, the annual turn-over increases if the firm is located beyond 35km range. When the distance exceeds the 35 kilometers mark it was observed that a unit increase in distance increases the annual turn-over by almost 91%. The estimated regression lines are graphically summarized below.

Figure 6.14: Spline Regression for Other Manufacturing Firms



.

Table 6.5: Summary of Spline Regression Results for 2-digit PSIC code Manufacturing Firms

PSIC Code	5Km Distance (β0)		5-15 Km			15-35 Km			35km and above		
	β0		β	Coefficient	Significance	β	Coefficient	Significance	β	Coefficient	Significance
Overall	**Coefficient**	**Significance**	**β0**	0.1407368*	0.000	**β0**	0.1407368*	0.000	**β0**	0.1407368*	0.000
	0.1407368*	0.000	**β1**	-0.1993568*	0.000	**β1**	-0.1993568*	0.000	**β1**	-0.1993568*	0.000
	-	-	-	-	-	**β2**	0.0566923*	0.000	**β2**	0.0566923*	0.000
	-	-	-	-	-	-	-	-	**β3**	-0.0074683	0.267
10	0.1547987*	0.000	**β0**	0.1547987*	0.000	**β0**	0.1547987*	0.000	**β0**	0.1547987*	0.000
	-	-	**β1**	-0.1999989*	0.000	**β1**	-0.1999989*	0.000	**β1**	-0.1999989*	0.000
	-	-	-	-	-	**β2**	0.0336216*	0.045	**β2**	0.0336216*	0.045
	-	-	-	-	-	-	-	-	**β3**	0.0152217	0.215
13	0.1739067*	0.000	**β0**	0.1739067*	0.000	**β0**	0.1739067*	0.000	**β0**	0.1739067*	0.000
	-	-	**β1**	-0.125115*	0.000	**β1**	-0.125115*	0.000	**β1**	-0.125115*	0.000
	-	-	-	-	-	**β2**	-0.1070284*	0.000	**β2**	-0.1070284	0.000
	-	-	-	-	-	-	-	-	**β3**	0.2597874*	0.000
14	0.2165951*	0.000	**β0**	0.2165951*	0.000	**β0**	0.2165951*	0.000	**β0**	0.2165951*	0.000
	-	-	**β1**	-0.4205678*	0.000	**β1**	-0.4205678*	0.000	**β1**	-0.4205678*	0.000
	-	-	-	-	-	**β2**	0.1459762*	0.000	**β2**	0.1459762*	0.000

PSIC Code	5Km Distance (β0)		5-15 Km			15-35 Km			35km and above		
	β0		β	Coefficient	Significance	β	Coefficient	Significance	β	Coefficient	Significance
	-	-	-	-	-	-	-	-	**β3**	-	-
15	0.223448*	0.000	**β0**	0.223448*	0.000	**β0**	0.223448*	0.000	**β0**	0.223448*	0.000
	-	-	**β1**	-0.35055*	0.000	**β1**	-0.35055*	0.000	**β1**	-0.35055*	0.000
	-	-	-	-	-	**β2**	0.1027615***	0.119	**β2**	0.1027615***	0.119
	-	-	-	-	-	-	-	-	**β3**	0.0125682	0.906
16	-0.0088163	0.775	**β0**	-0.0088163	0.775	**β0**	-0.0088163	0.775	**β0**	-0.0088163	0.775
	-	-	**β1**	-0.0178058	0.703	**β1**	-0.0178058	0.703	**β1**	-0.0178058	0.703
	-	-	-	-	-	**β2**	0.0365272***	0.106	**β2**	0.0365272***	0.106
	-	-	-	-	-	-	-	-	**β3**	-0.0259343***	0.139
17	0.2478089*	0.000	**β0**	0.2478089*	0.000	**β0**	0.2478089*	0.000	**β0**	0.2478089*	0.000
	-	-	**β1**	-0.4178352*	0.000	**β1**	-0.4178352*	0.000	**β1**	-0.4178352*	0.000
	-	-	-	-	-	**β2**	0.2714406*	0.000	**β2**	0.2714406*	0.000
	-	-	-	-	-	-	-	-	**β3**	-0.612765**	0.070
20	0.0463427	0.500	**β0**	0.0463427	0.500	**β0**	0.0463427	0.500	**β0**	0.0463427	0.500
	-	-	**β1**	-0.0128715	0.910	**β1**	-0.0128715	0.910	**β1**	-0.0128715	0.910
	-	-	-	-	-	**β2**	0.048717	0.558	**β2**	0.048717	0.558

PSIC Code	5Km Distance (β0)		5-15 Km			15-35 Km			35km and above		
	β0		β	Coefficient	Significance	β	Coefficient	Significance	β	Coefficient	Significance
	-	-	-	-	-	-	-	-	**β3**	-0.3681118**	0.056
22	0.1401376*	0.001	**β0**	0.1401376*	0.001	**β0**	0.1401376*	0.001	**β0**	0.1401376*	0.001
	-	-	**β1**	-0.2196331*	0.003	**β1**	-0.2196331*	0.003	**β1**	-0.2196331*	0.003
	-	-	-	-	-	**β2**	0.0757215	0.236	**β2**	0.0757215	0.236
	-	-	-	-	-	-	-	-	**β3**	0.1535433	0.565
23	0.0077258	0.547	**β0**	0.0077258	0.547	**β0**	0.0077258	0.547	**β0**	0.0077258	0.547
	-	-	**β1**	-0.0071813	0.708	**β1**	-0.0071813	0.708	**β1**	-0.0071813	0.708
	-	-	-	-	-	**β2**	-0.0081816	0.43	**β2**	-0.0081816	0.430
	-	-	-	-	-	-	-	-	**β3**	-0.0070808	0.356
25	-0.0072712	0.800	**β0**	-0.0072712	0.800	**β0**	-0.0072712	0.800	**β0**	-0.0072712	0.800
	-	-	**β1**	-0.0332707	0.481	**β1**	-0.0332707	0.481	**β1**	-0.0332707	0.481
	-	-	-	-	-	**β2**	0.0475015**	0.073	**β2**	0.0475015**	0.073
	-	-	-	-	-	-	-	-	**β3**	-0.0205513	0.498
27	0.4339841*	0.000	**β0**	0.4339841*	0.000	**β0**	0.4339841*	0.000	**β0**	0.4339841*	0.000
	-	-	**β1**	-0.4697406*	0.004	**β1**	-0.4697406*	0.004	**β1**	-0.4697406*	0.004
	-	-	-	-	-	**β2**	-0.0150158	0.898	**β2**	-0.0150158	0.898

PSIC Code	5Km Distance (β0)		5-15 Km			15-35 Km			35km and above		
	β0		**β**	**Coefficient**	**Significance**	**β**	**Coefficient**	**Significance**	**β**	**Coefficient**	**Significance**
	-	-	-	-	-	-	-	-	**β3**	-0.095204	0.505
28	0.0554517	0.390	**β0**	0.0554517	0.390	**β0**	0.0554517	0.390	**β0**	0.0554517	0.390
	-	-	**β1**	-0.0999338	0.327	**β1**	-0.0999338	0.327	**β1**	-0.0999338	0.327
	-	-	-	-	-	**β2**	0.0655954	0.222	**β2**	0.0655954	0.222
	-	-	-	-	-	-	-	-	**β3**	-0.1067025*	0.042
29	0.2368278*	0.008	**β0**	0.2368278*	0.008	**β0**	0.2368278*	0.008	**β0**	0.2368278*	0.008
	-	-	**β1**	-0.3889143*	0.008	**β1**	-0.3889143*	0.008	**β1**	-0.3889143*	0.008
	-	-	-	-	-	**β2**	0.0470802	0.701	**β2**	0.0470802	0.701
	-	-	-	-	-	-	-	-	**β3**	0.4424907	0.359
31	-0.0473442**	0.096	**β0**	-0.0473442**	0.096	**β0**	-0.0473442**	0.096	**β0**	-0.0473442**	0.096
	-	-	**β1**	0.024087	0.596	**β1**	0.024087	0.596	**β1**	0.024087	0.596
	-	-	-	-	-	**β2**	0.0298731	0.242	**β2**	0.0298731	0.242
	-	-	-	-	-	-	-	-	**β3**	-0.0162301	0.468
32	0.1100678*	0.001	**β0**	0.1100678*	0.001	**β0**	0.1100678*	0.001	**β0**	0.1100678*	0.001
	-	-	**β1**	-0.2336024*	0.000	**β1**	-0.2336024*	0.000	**β1**	-0.2336024*	0.000

PSIC Code	5Km Distance (β0)		5-15 Km			15-35 Km			35km and above		
	β0		β	Coefficient	Significance	β	Coefficient	Significance	β	Coefficient	Significance
	-	-	-	-	-	**β2**	0.1950903*	0.000	**β2**	0.1950903*	0.000
	-	-	-	-	-	-	-	-	**β3**	0.5751708	0.148

***Represents 5% significance level, **Represents 10% significance level, ***Represents 15% significance level**

6.3.3. Impact of Distance on the Annual Turnover of the Industrial Sectors

In order to understand the impact of distance on the annual turnover of the sectors, a spline regression model was developed in the previous section.

The manufacturing firms of different sectors behave differently across distance ranges. Within the 5km range, the annual turn-over of the manufacturing firms in Food, Textile, Other Non-Metallic Minerals, Wearing Apparel, Leather, Paper & Paper Products, Chemical & Chemical Products, Rubber & Plastic Products, Electrical Equipment, Machinery & Equipment, Motor Vehicles & Trailers and Other Manufacturing sector has a positive relation with the distance from urban centre. However, the annual turn-over of manufacturing firms in Fabricated Metal Products, Wood & Products and Furniture sector has a negative relation with distance.

Within the 5-15 km range, the annual turn-over of the manufacturing firms in Food, Fabricated Metal Products, Wearing Apparel, Leather, Wood & Products, Paper & Paper Products, Rubber & Plastic Products, Electrical Equipment, Machinery Equipment, Motor Vehicles & Trailer, Furniture and Other Manufacturing sectors has a negative relation with the distance from the urban centre. However, the annual turn-over of the manufacturing firms in Textile, Chemical & Chemical Products and Other Non-Metallic Minerals sectors has a positive relation with the distance from the urban centre.

Within the 15-35km range, the annual turn-over of the manufacturing firms in Food, Textile, Other Non-Metallic Minerals, Leather, Rubber & Plastic Product, Electrical Equipment, Motor Vehicles & Trailers and Wearing Apparel sectors has a negative relation with the distance from the urban centre. However, the annual turn-over of the manufacturing firms in Fabricated Metal Products, Wood & Products, Paper & Paper Products, Chemical & Chemical Products, Machinery

Equipment, Furniture and Other Manufacturing sectors has a positive relation with the distance from the urban centre. Beyond the 35km range, the annual turn-over of the manufacturing firms in Food, Textile and Other Manufacturing sectors has a positive relation with the distance from the urban centre. However, the annual turn-over of the manufacturing firms in Other Non-Metallic Minerals, Wood & Products, Paper & Paper Products, Chemical & Chemical Products, Electrical Equipment, Machinery & Equipment, Furniture and Fabricated Metal Products has a negative relation with the distance from the urban centre.

This study doesn't provide the understanding behind the unusual behavior of some of the manufacturing firms i.e., why in some cases the distance has an opposite impact on the annual turn-over of the manufacturing firms. For that purpose, detailed analysis of each sector is required which will provide a deeper understanding on why some sectors at different distance ranges behave differently.

Table 6.6: Summary of Impact of Distance on Firm's Annual Turnover

PSIC Code	Description	5Km Distance	5-15 Km Distance	15-35 Km Distance	35Km & above Distance
10	Food Products	Positive	Negative	Negative	Positive
13	Textile	Positive	Positive	Negative	Positive
14	Wearing Apparel	Positive	Negative	Negative	-
15	Leather & Related Products	Positive	Negative	Negative	Negative
16	Wood & Wood Products	Negative	Negative	Negative	Negative
17	Paper & Paper Products	Positive	Negative	Positive	Negative

PSIC Code	Description	5Km Distance	5-15 Km Distance	15-35 Km Distance	35Km & above Distance
20	Chemical & Chemical Products	Positive	Positive	Positive	Negative
22	Rubber & Plastic Products	Positive	Negative	Negative	-
23	Other Non-Metallic Minerals	Positive	Positive	Negative	Negative
25	Fabricated Metal Products	Negative	Negative	Positive	Negative
27	Electrical Equipment	Positive	Negative	Negative	Negative
28	Machinery & Equipment	Positive	Negative	Positive	Negative
29	Motor Vehicles & Trailers	Positive	Negative	Negative	-
31	Furniture	Negative	Negative	Positive	Negative
32	Other Manufacturing	Positive	Negative	Positive	Positive
Overall		Positive	Negative	Negative	Negative

6.4. Estimation Results of Trans-Log Regression

6.4.1. Nature of Industrial Agglomeration

Following the methodology developed in previous section, at first stage, we constructed the variables for agglomeration indicators of urbanization and localization. The dummy variable for urbanization has been constructed by measuring the location of the firm whether it was in city sprawl area or outside of city area. Table 6.6 shows the status of the firms' location.

Table 6.7: Sector-wise Urban Distance of the Firms

PSIC Code	Sector	Firms' Location		Total
		Outside City	Within City	
10	Food Products	2,491	1,002	3,493
11	Beverages	34	45	79
12	Tobacco Products	3	5	8
13	Textile	1,672	5,075	6,747
14	Wearing Apparel	291	1,362	1,653
15	Leather and Related Products	163	780	943
16	Wood and Products	345	124	469
17	Paper and Paper products	136	395	531
18	Printing & reproduction of recorded media	29	305	334
19	Coke & Refined petroleum products	22	10	32
20	Chemical and Chemical products	167	306	473
21	Basic Pharmaceutical	100	168	268
22	Rubber and Plastic Products	154	837	991
23	Other Non-metallic mineral	2,166	584	2,750
24	Basic Metal	149	578	727
25	Fabricated Metal Products	489	1,123	1,612
26	Computer, Electronic and Optical Product	4	31	35
27	Electrical Equipment	51	583	634
28	Machinery and Equipment	114	656	770
29	Motor Vehicles and Trailers	75	262	337
30	Other Transport Equipment	24	149	173
31	Furniture	321	472	793
32	Other Manufacturing (surgical, sports)	384	955	1,339

	Total	9,384	15,807	25,191

Based on the above table, a dummy variable is constructed for urbanization indicator. We performed the Local Moran's I (LISA) to find the spatial autocorrelation between the firms. Table 6.2 shows the number of firms in each sector which are spatially correlated and spatially uncorrelated.

Table 6.8: Spatially Autocorrelated Firms (Results of Moran I Test)

PSIC Code	**Sector**	**Spatially uncorrelated firms**	**Spatially auto-correlated firms**	**Total**
10	Food Products	3,446	47	3,493
11	Beverages	73	6	79
12	Tobacco Products	3	5	8
13	Textile	6,543	204	6,747
14	Wearing Apparel	1,595	58	1,653
15	Leather and Related Products	922	21	943
16	Wood and Products	468	1	469
17	Paper and Paper products	522	9	531
18	Printing & reproduction of recorded media	325	9	334
19	Coke & Refined petroleum products	31	1	32
20	Chemical and Chemical products	459	14	473
21	Basic Pharmaceutical	254	14	268
22	Rubber and Plastic Products	969	22	991
23	Other Non-metallic mineral	2,720	30	2,750

24	Basic Metal	717	10	727
25	Fabricated Metal Products	1,574	38	1,612
26	Computer, Electronic and Optical Product	33	2	35
27	Electrical Equipment	608	26	634
28	Machinery and Equipment	768	2	770
29	Motor Vehicles and Trailers	326	11	337
30	Other Transport Equipment	163	10	173
31	Furniture	769	24	793
32	Other Manufacturing (surgical, sports)	1,317	22	1,339
	Total	24,605	586	25,191

Based on the results of Moran's I test, we constructed the indicator of localization. Firms which are spatially correlated are assumed to be benefitting from each other. Thus, those firms are agglomerated based on localization, i.e., firms are benefitting from the agglomeration of other firms in the same vicinity.

6.4.2. Impact of Agglomeration on Firms' Annual Turnover

In order to understand the impact of agglomeration on annual turnover of the firms, we regress the equation (2A) constructed in chapter 4. Table 6.3 shows the parameter estimated, where α_1 and α_2 represent the agglomeration economies of urbanization and localization respectively. The overall regression analysis shows that both urbanization and localization impact positively annual turnover of the firms, showing the firms are benefiting from both city diversity and the clustering of firms. However, given the values of the parameters, it can be deducted that the impact of urbanization is greater than the impact of localization in Punjab. These results are consistent with the existing status of manufacturing sector development in the province.

When we regress separate regression equation for two-digit level sectors, the results are different. We found that out of 23 sectors, 13 showed positive impact of urbanization on annual turnover of the firms. These sectors are food products, textile, wearing apparel, wood and wood products, chemical and chemical products, other manufacturing (surgical, sports), printing & recorded media, basic pharmaceuticals, other non-metallic minerals, electrical equipment, furniture, fabricated metal products and motors and trailer.

The impact of localization economies is observed positive in 10 sectors, i.e., textile, leather and products wood and wood products paper and paper products, chemical and chemical products basic pharmaceuticals, electrical equipment, machinery and equipment and other manufacturing (surgical, sports and toys).

Table 6.9: Impact of Agglomeration Economies on Firms' Annual Turnover

PSIC Code	α_0	α_1	α_2	α_L	α_E	α_{LL}	α_{EE}	α_{LE}	R^2
10	-0.721 (-2.41)	0.213 (4.03) *	0.438 (1.83) **	0.923 (5.47) *	-0.665 (-10.24)*	0.007 (0.93)	-0.026 (-0.92)	-0.012 (-0.53)	0.4329
11	-4.303 (-3.50)	0.066 (0.26)	0.988 (1.39)	3.054 (5.13) *	-0.940 (-2.59) *	0.104 (1.49)	-0.050 (-0.64)	-0.278 (-1.82)**	0.557
12	no regression due to small sample size								
13	-.280 (-2.57)	-0.151 (-7.08) *	0.121 (1.34)	1.034 (15.98) *	-0.934 (-25.92) *	0.006 (1.16)	-0.035 (-2.70)*	0.013 (0.93)	0.528
14	-1.262 (-5.75)	0.318 (5.91) *	-0.065 (-0.39)	1.115 (8.65) *	-0.712 (-10.00) *	-0.051 (-4.38) *	-0.082 (-3.84) *	0.094 (3.89) *	0.475
15	-0.125 (-0.34)	0.093 (1.37)	0.550 (2.37) *	0.859 (3.76) *	-1.043 (-9.53)*	-0.001 (-0.10)	-0.068 (-1.50)	0.102 (2.46)*	0.516
16	0.617 (0.47)	0.275 (5.11) *	6.394 (0.67)	0.938 (0.87)	-1.574 (-7.15) *	0.011 (0.65)	-0.229 (-0.94)	0.263 (2.57)*	0.736
17	-1.489 (-2.69)	0.151 (1.61)	0.619 (1.30)	1.540 (4.87) *	-0.970 (-6.09)*	-0.000 (-0.03)	-0.160 (-2.47)*	0.090 (1.23)	0.463
18	-1.504 (-2.77)	0.476 (2.99) *	-0.638 (-1.70)**	1.134 (3.23) *	-0.921 (-5.85)*	0.037 (1.12)	0.012 (0.16)	-0.041 (-0.52)	0.546
19	-8.238 (-2.64)	0.407 (0.60)	6.024 (1.61)	7.001 (2.86) *	-1.775 (-1.65)	0.108 (0.59)	-1.056 (-1.42)	0.111 (0.18)	0.537
20	-2.432 (-4.53)	0.631 (5.85) *	0.360 (0.82)	1.421 (4.09) *	-0.641 (-4.35) *	0.030 (1.17)	-0.013 (-0.21)	-0.077 (-1.22)	0.413
21	-4.007 (-2.98)	0.211 (1.47)	0.033 (0.09)	2.254 (3.75)*	-0.229 (-0.57)	-0.050 (-0.99)	-0.199	0.009 (0.09)	0.339

PSIC Code	α_0	α_1	α_2	α_L	α_E	α_{LL}	α_{EE}	α_{LE}	R^2
							(-2.33) *		
22	-2.138 (-5.53)	0.217 (2.95)*	-0.137 (-0.58)	1.930 (8.11)*	-1.013 (-10.52)*	-0.018 (-1.24)	-0.234 (-6.81)*	0.153 (4.69)*	0.441
23	2.671 (9.33)	0.296 (7.20)*	0.509 (3.08)*	-1.024 (-6.00)*	-0.796 (-12.72)*	-0.032 (-3.82)*	0.156 (5.91)*	0.071 (2.90)*	0.755
24	-1.883 (-2.15)	0.024 (0.19)	-0.791 (-1.65)**	1.305 (2.65)*	-0.209 (-1.13)	-0.075 (-2.60)*	-0.164 (-1.99)*	0.109 (1.14)	0.164
25	-1.166 (-4.05)	0.163 (4.45)*	0.162 (1.10)	1.207 (6.33)*	-0.759 (-8.32)*	-0.034 (-3.27) *	-0.095 (-2.30) *	0.074 (1.890**	0.544
26	1.139 (0.22)	-0.679 (-0.94)	-3.036 (-1.04)	-1.053 (-0.33)	0.087 (0.10)	0.145 (1.04)	0.672 (1.16)	-0.573 (-1.03)	0.366
27	-1.444 (-2.82)	-0.282 (-2.10)*	0.679 (2.65)*	1.678 (5.50)*	-0.886 (-6.90)*	0.027 (1.08)	-0.096 (-1.67)**	-0.032 (-0.49)	0.387
28	-0.548 (-1.01)	0.012 (0.15)	1.183 (1.56)	0.806 (3.13)*	-0.611 (-4.58)*	-0.040 (-2.17)*	-0.058 (-3.77)*	0.077 (1.69)**	0.413
29	-0.095 (-0.14)	0.245 (1.99)*	-0.587 (-1.52)	0.793 (1.96)**	-1.008 (-4.70)*	0.026 (0.58)	0.011 (0.12)	0.022 (0.19)	0.360
30	-3.377 (-2.92)	0.104 (0.50)	-0.522 (-0.96)	2.521 (3.66)*	-0.732 (-2.18)*	-0.103 (-2.69)*	-0.343 (-3.06)*	0.254 (2.47)*	0.459
31	-0.014 (-0.03)	0.247 (5.09)*	0.101 (0.56)	0.945 (2.71)*	-1.184 (-8.23)*	-0.015 (-1.57)	-0.146 (-2.08)*	0.184 (3.30)*	0.546
32	-1.378 (-4.23)	0.237 (4.35)*	0.755 (2.91)*	1.279 (6.81)*	-0.807 (-8.97)*	-0.034 (-3.28)*	-0.137 (-4.28)*	0.121 (3.89)*	0.499
Overall	-0.557 (-7.24)	0.208* (14.41)	0.440 (7.82) *	0.669 (15.48)*	-0.586 (-30.46)*	-0.017 (-6.64)*	0.020 (3.13)*	-0.013 (-2.13)*	0.579
*Shows significant values at 5%, **shows significance at 10%, Values in parenthesis show t-values									

The following table shows the summary of sectors which show positive impact of agglomeration economies. Given the overall economic development of the province, the below table summarize the agglomeration effect which are close to reality. Pakistan in general and the province of Punjab in particular, have more labor-intensive industries as compared to capital intensive. Since industries in Punjab are not very advanced sectors oriented, the localization effect is less in industrial sectors as compared to urbanization effect as literature suggests that localization effect strengthens when sector specialize in advance sector.

Table 6.10: Summary of Effects of Agglomeration Economies

2-Digit Sectors	Positive Effects of Agglomeration Economies	
	Urbanization	Localization
PSIC-10: Food Products	✓	✓
PSIC-11: Beverages	✗	✗
PSIC-13: Textile	✓	✗
PSIC-14: Wearing Apparel	✓	✗
PSIC-15: Leather and Related Products	✗	✓
PSIC-16: Wood and Wood Products	✓	✗
PSIC-17: Paper and Paper products	✗	✗
PSIC-18: Printing & Reproduction of Recorded Media	✓	✓
PSIC-19: Coke and Refined Petroleum Products	✗	✗
PSIC-20: Chemical and Chemical Products	✓	✗
PSIC-21: Basic Pharmaceutical	✗	✗
PSIC-22: Rubber and Plastic Products	✓	✗
PSIC-23: Other Non-metallic Minerals	✓	✓
PSIC-24: Basic Metal	✗	✓
PSIC-25: Fabricated Metal Products	✓	✗
PSIC-26: Computer, Electronic and Optical Products	✗	✗
PSIC-27 Electrical Equipment	✓	✓
PSIC-28 Machinery and Equipment	✗	✗
PSIC-29: Motor Vehicles and Trailers	✓	✗
PSIC-30: Other Transport Equipment	✗	✗
PSIC-31: Furniture	✓	✗
PSIC-32: Other Manufacturing (surgical, sports, toys)	✓	✓

6.5. Discussion of the Results

We estimated the impact of urbanization agglomeration on firms' annual turnover for overall manufacturing sector and disaggregated at sectoral level. For this purpose, we used trans-log production function for those firms who are operating within the urban centre and spline regression analysis for the firms that are located outside the urban centre.

For the overall manufacturing sector, the estimation results of trans-log regression analysis shows that impact of urbanization agglomeration is positive on annual turnover of the firms which are operating within their respective urban centres. Whereas the spline regression analysis shows that the annual turnover of the firms increases when they are operating within 0-5 km distance. The annual turnover decreases if those manufacturing firms are located beyond the 5 km range. This result is consistent yet a little different from (Verstraten, Verweij & Zwaneveld, 2019) who argued that the benefits of urbanization increase within 5-10 km and decreases beyond that distance.

However, the results are a little different when estimated at disaggregated sectoral level. It is observed that urbanization agglomeration shows positive effect on the annual turnover of 13 out of 23 manufacturing sectors, which includes food products, textile, wearing apparel, wood & wood products, printing & reproduction of recorded media, chemical & chemical products, rubber & plastic products, other non-metallic minerals, fabricated metal products, electrical equipment, motor vehicles & trailers, furniture and other manufacturing (surgical, sports, toys). This implies that, if the manufacturing firms of the above-mentioned sectors are operating within city, they are taking the benefit from urbanization agglomeration i.e., large consumer markets, huge population and low transportation cost. These results are consistent with the literature on the benefits of urbanization agglomeration (Lambert et al., 2006), (Rosenthal & Strange, 2004), (Mitra, 1999), (Capello, 1999).

However, if the firms of these same sectors are operating from outside the city, the impact of urbanization agglomeration on annual turnover varies according to their distance from urban centre. For five out of thirteen sectors (wearing apparel, leather & related products, rubber & plastic products, electrical equipment, motor vehicles & trailers), the spline regression analysis shows that the annual turnover increases if the firms are operating from within the 0-5 km distance.

The annual turnover decreases if those manufacturing firms are located beyond 5 km distance. (Rosenthal and Strange, 2003) and (Graham and Melo, 2009) also concluded the similar results.

For three out of thirteen sectors (wood & wood products, fabricated metal products, furniture), the analysis shows that the annual turnover decreases if the firms are operating from within the 0-5 km distance. However, for two of these three sectors (fabricated metal products, furniture), the analysis shows positive annual turnover at 15-35 km distance while for the other sector (wood & wood products) the annual turnover decreases if the manufacturing firms are located beyond 5km range.

We estimated the impact of localization agglomeration on manufacturing firms' annual turnover for overall manufacturing sector and disaggregated at sectoral level as well. For this purpose, we used spatial autocorrelation and trans-log production function.

For the overall manufacturing sector, the estimation results of trans-log regression analysis shows that impact of localization agglomeration is positive on annual turnover of the firms. On the other hand, at the disaggregated sectoral level it is observed that localization agglomeration shows positive effect on the annual turnover of 07 out of 23 manufacturing sectors, which includes food products, leather & leather related products, electrical equipment, basic metal, printing & reproduction of recorded media, other non-metallic minerals and other manufacturing (surgical, sports, toys). This implies that, if the manufacturing firms of the above-mentioned are taking benefit from localization agglomeration i.e., labor pooling, input sharing and knowledge spillover (also known as Marshallian externalities). The positive impact of these externalities has been extensively advocated by (Hill & Brennan, 2000), (Henderson et al, 1995), and (Duranton and Overman, 2005).

While some sectors showed positive effect of urbanization, others showed positive effect of localization on the annual turnover of the manufacturing firms. However, many sectors showed positive effect of both on the annual turnover of the manufacturing firms. It is observed that both urbanization and localization agglomeration show positive effect on the annual turnover for 5 out of 23 manufacturing sectors, which includes food products, electrical equipment, printing & reproduction of recorded media, other non-metallic minerals and other manufacturing (surgical, sports, toys). All the studies (Ascani, et al., 2012), (Lu & Tao, 2009), (Moomaw, 1998), (Martin et al, 2008), (Garaham, 2009), (Nakamura, 1985) provide the similar results of the presence of both agglomeration indicators i.e., mutual existence of urbanization and localization.

Table 6.11: Summary Results

Impact of Agglomeration Economies	Sectors
Urbanization only	1. PSIC-13: Textile 2. PSIC-14: Wearing Apparel 3. PSIC-16: Wood and Wood Products 4. PSIC-20: Chemical and Chemical Products 5. PSIC-22: Rubber and Plastic Products 6. PSIC-25: Fabricated Metal Products 7. PSIC-29: Motor Vehicles and Trailers 8. PSIC-31: Furniture
Localization only	1. PSIC-15: Leather and Related Products 2. PSIC-24: Basic Metal
Both Urbanization & Localization	1. PSIC-10: Food Products 2. PSIC-18: Printing & Reproduction of Recorded Media 3. PSIC-23: Other Non-metallic Minerals 4. PSIC-27 Electrical Equipment 5. PSIC-32: Other Manufacturing (surgical, sports, toys)
No Impact of agglomeration economies	1. PSIC-11: Beverages 2. PSIC-17: Paper and Paper products 3. PSIC-19: Coke and Refined Petroleum Products 4. PSIC-21: Basic Pharmaceutical 5. PSIC-26: Computer, Electronic and Optical Products 6. PSIC-30: Other Transport Equipment

The objective of this analysis was to determine the performance (annual turnover) of any sector that might change with the increasing distance of firms from the urban centers. There were 23 sectors against which we have analyzed their performance on four categories of distance i.e., 0-5 km, 5-15 km, 15-35 km, and 35 & above. Our results showed as in Table 6.5 there are insignificant relationships between distance and performance of firms but overall most of the sectors showed a positive relationship between the distance (0-5 km) and performance and negative relationship as we move beyond city i.e., greater than 0-5 km.

The results are however, consistent with theory that with the increase in distance the benefits of agglomeration economies attenuate [Rosenthal & Strange, 2004 & Graham & Melo, 2009]. However, the sectoral variance also shows that some of the sectors show insignificant relationship between distance and performance of the firms. This might be due to peculiar behavior of some sectors in regard to agglomeration economies. Hence, insignificant results of some sectors are part of the analysis and do not negate/influence overall trend of any results.

The insignificant results can be attributed to limitations of data that is required in order to understand the comprehensive behavior of manufacturing firms. For that purpose, the analysis would require more data on variables like R&D, Knowledge spending, Labor, Output, Energy usage, GFCF, etc. If we have data on all these variables then perhaps we will be able to understand the behavior that led towards the insignificant results for urbanization and localization of firms. Since our data does not provide information on such variables we cannot comment on the exact nature of the manufacturing firms with respect to agglomeration economies. So that in turn is our limitation of our dataset as it does not provide information on such variables.

However, we may still be able to generally correlate our findings with capital intensity of different sectors. UNIDO classifies the manufacturing sectors by technological intensity i.e., technological classification which is a method widely applied for the purpose of policy relevant analysis. The classification of sectors on the basis of their technological usage is an important one. Recently, share of medium-high and high-technology has been incorporated as one of the SDG indicators related to industrialization.

Before going on to emphasizing our point that manufacturing sector of Pakistan is relatively less capital intensive, we will look on to the basis of technological classification of UNIDO. The basis of technological classification lies in the expenditure incurred on Research and Development (R&D) during the production of manufactured goods. Those manufacturing industries that are R&D intensive are high technology industries and vice versa. The intensity of R&D can be defined as the expenditure on R&D to the output that is usually the gross value added. On the basis of technological intensity, UNIDO has classified industries into three main groups i.e., Medium-High and High Technology, Medium Technology, and Low Technology. The classification of manufacturing industries on the basis of their technological intensity is given in the table below.

Table 6.12 Classification of Manufacturing Sectors by Technological Intensity (ISIC Revision 4)

Medium-High and High Technology	
Division 20	Chemicals and Chemical Products
Division 21	Pharmaceuticals
Division 26	Computer, Electronic, Optical Products
Division 27	Electrical Equipment
Division 28	Machinery and Equipment
Division 29	Motor Vehicles, Trailers, and Semi-Trailers
Division 30	Other Transport Equipment except Ships and Boats
Medium Technology	
Division 22	Rubber and Plastic Products
Division 23	Other Non-Metallic Mineral Products
Division 24	Basic Metals
Division 32	Other Manufacturing
Division 33	Repair and Installation of Machinery and Equipment
Low Technology	
Division 10	Food Products
Division 11	Beverages
Division 12	Tobacco Products
Division 13	Textiles
Division 14	Wearing Apparel
Division 15	Leather and Related Products
Division 16	Wood and Products of Wood
Division 17	Paper and Paper Products
Division 18	Printing and Reproduction of Recorded Media
Division 19	Coke and Refined Petroleum Products
Division 25	Fabricated Metal Products
Division 31	Furniture

Source: United Nations Industrial Development Organization (UNIDO)

As it can be established from the table above that the level of technology in most of the manufacturing firms of Pakistan is not high. From the theory of agglomeration economies, we know that the industries that are capital intensive experience the effects of localization as compared to the effects of urbanization and vice versa. In Pakistan, most of the manufacturing firms are less capital intensive the effect of urbanization is more dominant as compared to localization (see Table 6.11). In our results only one high tech manufacturing sector i.e., electrical equipment, shows positive effect of localization on annual turnover. All other significant sectors are either meium or low-tech industries. It is also specified in the UNIDO classification of the manufacturing groups that most of the manufacturing industries of Pakistan have been classified under low technology meaning that most of the manufacturing firms in Pakistan are labor rather than capital intensive.

CHAPTER 7

7. Conclusion

The study aimed to identify the nature and type of agglomeration in manufacturing sectors of Punjab and also to explore the effect of those agglomeration economies on the annual turnover of the manufacturing firms. For that purpose, this study has used the spatial econometric tools, trans-log production function and spline regression models. The study used the spatial point data of the manufacturing firms to estimate the nature of agglomeration in Punjab. Further, to estimate the effect of agglomeration on the annual turnover of the firms' trans-log production function and spline regression analysis were employed. The results show that more firms are enjoying the benefits from being locating inside the urban centres as compared to taking benefits from the Marshallian externalities.

First objective of this research was to assess the nature of agglomeration of the manufacturing firms in the Punjab province. From the urbanization point of view, it was found that from a total of 25,191 firms about 15,807 firms are located within the city, with the firms from Textiles having the most firms located within the city followed by Wearing Apparel, Fabricated Metal Product and Food Product. On the other hand, from the perspective of localization, the results show that, on the basis of labor pooling, only 586 out of 25,191 firms are spatially autocorrelated with the Textile sector having the most spatially autocorrelated firms. From the results, there has been a clear indication of firms benefiting from the urbanization economies of agglomeration.

After the nature of agglomeration has been identified, the second objective of the study was to find out the impact of agglomeration on the annual turnover of the firms. It has been estimated that both urbanization and localization positively impact the annual turnover of the firms showing that the firms are benefiting from both city diversity and the clustering of firms. Out of the 23 sectors, 13 showed positive impact of urbanization on annual turnover of the firms including food products, textile, wearing apparel, wood & wood products, printing & reproduction of recorded media, chemical & chemical products, rubber & plastic products, other non-metallic minerals, fabricated metal products, electrical equipment, motor vehicles & trailers, furniture and other manufacturing (surgical, sports, toys). On the other hand, 07 sectors i.e., food products, leather & leather related products, electrical equipment, basic metal, printing & reproduction of recorded media, other non-metallic minerals and other manufacturing (surgical, sports, toys) showed a positive impact of localization economies. However, the impact of urbanization was found to be greater than the impact of localization in our targeted area i.e., Punjab.

Further, the study attempted to find how the distance from the urban centres effects the annual turnover of the firms. The results show that as the firms move away from their urban centres, it negatively effects the annual turnover of the firms. With exception of few industries at different distances used in the analysis (like at 5 km distance, wood and wood products, Fabricated metal products and Furniture showed negative behavior) rest of the industries confirmed the hypothesis that moving away from the urban centres negatively effects the annual turnover of the firms which also seems logical from the urbanization economies perspective as being located inside the urban cores increases the productivity of the firms. This phenomenon is also substantiated by Ascani, et al. (2012).

From the estimations, it has been noted that urbanization economies are more dominant in Punjab as compared to the localization economies. The results are in fact closer to reality as Punjab specifically have more labor-intensive industries and are not very advanced and capital oriented. The impact of localization economies is less because as the literature suggests Jane Jacobs (1969), the effect of localization is dominant when each sector specialize and is more technologically oriented. Further, the study found out that overall agglomeration economies (urbanization and localization) positively impact the annual turnover of the firms (this result answers our second research question).

In a nutshell, the effect of urbanization economies influences more on the annual turnover of the firms as compared to localization. Most of the industrial sectors experience productivity from the phenomena of urbanization economies however, the localization economies cannot be completely ruled out as some of the industries in Punjab's manufacturing sector experience their productive advantages from localization economies.

7.1. Policy Implications

The study provides the evidence that the agglomeration of manufacturing sector in Punjab province is indeed benefiting from the agglomeration economies. The effects of agglomeration economies can be categorized into urbanization and localization. The results have indicated that some manufacturing sectors show positive effects of urbanization while some show positive effects of localization. The results have significant policy implications in reference to the regional development of the country.

Given the proof that the manufacturing sector of Pakistan are benefiting from the agglomeration economies, the regional development of the province should be in line with that, as it will help identify what will be beneficial for the manufacturing sector. The manufacturing sector that are benefited from the urbanization economies should be in close proximity to the urban centres for example, food products, textile, wearing apparel, wood & wood products, printing & reproduction of recorded media, chemical & chemical products, rubber & plastic products, other non-metallic minerals, fabricated metal products, electrical equipment, motor vehicles & trailers, furniture and other manufacturing (surgical, sports, toys) as indicated in the study are positively affected by urbanization economies and should be in close proximity to the urban centres. On the other hand, the manufacturing sector benefited from the localization economies including food products, leather & leather related products, electrical equipment, basic metal, printing & reproduction of recorded media, other non-metallic minerals and other manufacturing (surgical, sports, toys) should be placed in special economic zones. The placement of manufacturing sector is crucial from the regional development perspective. Misplacement of the manufacturing sector can significantly impact the annual turnover of the manufacturing sector.

The analysis in this study is limited to the 2-digit PSIC classification, which however can be further extended to 3-digit PSIC classification which will further classify each manufacturing firm. The further classification of the manufacturing sector can be used for identifying benefits from the agglomeration economies of each firm in the manufacturing sector. Its significance is of high importance as it will help in identification of manufacturing related problems at grass root level.

7.2. Limitations of the Study

The results of the study have proved that manufacturing sector of Punjab province is benefiting from the agglomeration economies. However, there are a few limitations that must be kept in mind while interpreting these results. There are two major limitations affiliated with my study that could be addressed in the future studies. First limitation concerns the dataset that is used in the study. Even though, the dataset used in the study is unique because it uses the geo-spatial location of the manufacturing sector, the detailed description and characteristics of manufacturing sector are missing. For the detailed description and characteristics of manufacturing sector, the Census of Manufacturing Industries (CMI) could prove fruitful if the data of manufacturing sector is collected on the basis of their geo-spatial location. The other limitation of the study involves the PSIC classification of the manufacturing sector. The study used the 2-digit code of the PSIC classification. Future studies could use the 3-digit code of the PSIC classification for further extension of the study.

8. References

Abel, J. R., & Gabe, T. M. (2009). *Labor Market Pooling and Occupational Agglomeration.* Federal Reserve Bank of New York Staff Report No. 393.

Åberg, Y. (1973). Regional Productivity Differences in Swedish Manufacturing. *Regional and Urban Economics, 3*(2), 131-155.

Addario, S. D., & Patacchini, E. (2008). Wages and the City: Evidence from Italy. *Labour Economics, 15*(5), 1040-1061.

Afraz, N., Hussain, S. T., & Khan, U. (2014). Barriers to the Growth of Small Firms in Pakistan: A Qualitative Assessment of Selected Light Engineering Industries. *Lahore Journal of Economics, 19*(Special E), 135-176.

Agovino, M., & Rapposelli, A. (2015). Agglomeration Externalities and Technical Efficiency in Italian Regions. *Quality & Quantity, 49*(5), 1803-1822.

Ahmed, G., Khan, M. A., Mahmood, T., & Afzal, M. (2017). Trade Liberalization and Industrial Productivity: Evidence from Manufacturing Industries in Pakistan. *The Pakistan Development Review, 56*(4), 319-348.

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics, 6*(1), 21-37.

Akhtar, S. M. (1955). Problems of Industrialization in Under-Developed Countries. *The Punjab University Economist, 3*(1), 3-17.

Alkan, Ö., Oktay, E., Genç, A., & Çelik, A. K. (2014). An investigation of export–import ratios in Turkey using spline regression models. *Economic Research-Ekonomska Istraživanja, 30*(1), 223-237.

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science, 45*(7), 905-917.

Amend, E., & Herbst, P. (2008). Labor Market Pooling and Human Capital Investment Decisions. *IAB-Discussion Paper No. 4/2008*.

Andersson, M., & Karlsson, C. (2007). Knowledge in Regional Economic Growth - The Role of Knowledge Accessibility. *Industry & Innovation, 14*(2), 129-149.

Andersson, M., & Lööf, H. (2011). Agglomeration and Productivity: Evidence from firm-level Data. *The Annals of Regional Science, 46*(3), 601-620.

Andini, M., Blasio, G. d., Duranton, G., & Strange, W. C. (2013). Marshallian labour market pooling: Evidence from Italy. *Regional Science and Urban Economics, 43*(6), 1008-1022.

Arauzo-Carod, J.-M., Liviano-Solis, D., & Manjón-Antolín, M. (2010). Empirical Studies in Industrial Location: An Assessment of their Methods and Results. *Journal of Regional Science, 50*(3), 685-711.

Arif , B. W. (2012). Industrial Clusters, Schumpeterian Innovations and Entrepreneurs' Human and Social Capital: A Survey of Literature. *Pakistan Economic and Social Review, 50*(1), 71-95.

Arvanitis, R., & Haixiong, Q. (2009). Research for Policy Development: Industrial Clusters in South China. In M. Graham, & J. Woo, *Fueling Economic Growth: The Role of Public-Private Research in Development* (pp. 39-85). Practical Action Publishing.

Ascani, A., Crescenzi, R., & Iammarino, S. (2012). Regional Economic Development: A Review. *SEARCH Working Paper WP1/03*, 2-26.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, 630-640.

Autant-Bernard, C., & LeSage, J. P. (2010). Quantifying Knowledege Spillovers Using Spatial Econometric Models. *Journal of Regional Science, 51*(3), 471-496.

Azhar, A., & Adil, S. (2019). The Effects of Agglomeration on Socio-Economic Outcomes: A District Level Panel Study of Punjab. *The Pakistan Development Review, 58*(2), 159-176.

Baldwin , J. R., Brown, W. M., & Rigby, D. L. (2010). Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing. *Journal of Regional Science, 50*(5), 915-934.

Baldwin, J. R., Beckstead, D., Brown, W. M., & Rigby, D. L. (2008). Agglomeration and the Geography of Localization Economies in Canada. *Regional Studies, 42*(1), 117-132.

Baldwin, R. E., & Okubo, T. (2006). Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography, 6*(3), 323-346.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy, 38*(2), 318-337.

Bechard, A. (2020). The Economic Impacts of Harmful Algal Blooms on Tourism: An Examination of Southwest Florida Using a Spline Regression Approach. *Natural Hazards: Journal of International Society for the Prevention and Mitigation of Natural Hazards, 104*(1), 593-609.

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive Cities: Sorting, Selection and Agglomeration. *Journal of Political Economy, 122*(3), 507-553.

Bernard, A. B., Jensen , J. B., & Lawrence, R. Z. (1995). Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987. *Brookings Papers on Economic Activity. MicroEconomics, 1995*(1995), 67-119.

Bökemeier, B., & Greiner, A. (2015). On the relation between public debt and economic growth:. *Economics and Business Letters, 4*(4), 137-150.

Braunerhjelm, P., & Borgman, B. (2004). Geographical Concentration, Entrepreneurship and Regional Growth: Evidence from Regional Data in Sweden, 1975-99. *Regional Studies, 38*(8), 929-947.

Brunow, S., & Blien, U. (2015). Agglomeration effects on labor productivity: An assessment with microdata. *The Journal of ERSA, 2*(1), 33-53.

Burki, A. A., & Khan, M. A. (2010). Spatial inequality and Geographic Concentration of Manufacturing Industries. *27th Annual General Meeting and Conference of the Pakistan Society of Development Economics.* Islamabad: PIDE.

Burki, A. A., Munir, K. A., Khan, M. A., Khan, M. U., Faheem, A., Khalid, A., & Hussain, S. T. (2010). Industrial Policy, Its Spatial Aspects and Cluster Development in Pakistan. *Analysis Report to the Industrial Policy 2010*(1).

Buse, A., & Lim, L. (1977). Cubic Splines as a Special Case of Restricted Least Squares. *Journal of the American Statistical Association, 72*(357), 64-68.

Cainelli, G., & Ganau, R. (2018). Distance-based Agglomeration Externalities and Neighbouring Firms' characteristics. *Regional Studies, 52*(7), 922-933.

Capello, R. (1999). Agglomeration Economies and Urban Productivity: The case of the High-tech Industry in the Milan Metropolitan Area. *In 39th European Regional Science Association Congress in Dublin.*

Caragliu, A., De Dominicis, L., & De Groot, H. L. (2016). Both Marshall and Jacobs were Right! *Economic Geography, 92*(1), 87-111.

Carlton, D. W. (1983). The Location and Employment Choices of New Firms: An Econometric Model with. *The Review of Economics and Statistics, 65*(3), 440-449.

Changes in Agglomeration Economies and Linkage Externalities for Japanese Urban Manufacturing Industries: 1990 & 2000. (2008). *Discussion Papers 08040*.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1971). Conjugate Duality and the Transcendental Logarithmic Production Function. *Econometrica, 39*(4), 255-256.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics, 55*(1), 28-45.

Cingano, F., & Schivardi, F. (2004). Identifying the Sources of Local Productivity Growth. *Journal of European Economic Association, 2*(4), 720-742.

Combes, P. P., & Duranton, G. (2006). Labour Pooling, Labour Poaching, and Spatial Clustering. *Regional Science and Urban Economics, 36*(1), 1-28.

Combes, P. P., Duranton, G., & Overman, H. G. (2005). Agglomeration and the adjustment of the spatial economy. *Papers in Regional Science, 84*(3), 311-349.

Costa, D. L., & Kahn, M. E. (2003). Civic Engagement and Community Heterogeneity: An Economist's Perspective. *Perspectives on Politics, 1*(1), 103-111.

De Almeida, E. T., & Rocha, R. D. (2018). Labor pooling as an agglomeration factor: Evidence from the Brazilian Northeast in the 2002–2014 period. *EconomiA, 19*(2), 236-250.

Driffield, N., & Munday, M. (2001). Foreign Manufacturing, Regional Agglomeration and Technical Efficiency in UK Industries: A Stochastic Production Frontier Approach. *Regional Studies, 35*(5), 391-399.

Duranton, G., & Overman, H. G. (2005). Testing for Localization Using Micro-Geographic Data. *The Review of Economic Studies, 72*(4), 1077-1106.

Duranton, G., & Overman, H. G. (2008). Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data. *Journal of Regional Science, 48*(1), 213-243.

Duranton, G., & Puga, D. (2004). Microfoundations of Urban Agglomeration Economies. In J. V. Henderson, & J. F. Thisse, *Handbook of Regional and Urban Economics* (pp. 2063-2117). Elsevier.

Eberts, R. W., & McMillen, D. P. (1999). Agglomeration Economies and Urban Public Infrastructure. In P. Cheshire, & E. S. Mills, *Handbook of Regional and Urban Economics* (Vol. 3, pp. 1455-1495).

Ellison, G., & Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy, 105*(5), 889-927.

Faberman, R. J., & Freedman, M. (2016). The urban density premium across establishments. *Journal of Urban Economics, 93*, 71-84.

Faggio, G., Silva, O., & Strange, W. C. (2017). Heterogeneous Agglomeration. *Review of Economics and Statistics, 99*(1), 80-94.

Fang, C., & Yu, D. (2017). Urban Agglomeration: An Evolving Concept of an Emerging Phenomenon. *Landscape and Urban Planning, 162*, 126-136.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, 120*(3), 253-290.

Feldman, M. P., Feller, I., Bercovitz, J. L., & Burton, R. M. (2002). University Technology Transfer and the System of Innovation. In M. P. Feldman, & N. Massard, *Institutions and Systems in the Geography of Innovation. Economics of Science, Technology and Innovation* (Vol. 25, pp. 55-77). Boston: Springer.

Feldman, M., Feller, I., Bercovitz, J., & Burton, R. (2002). Equity and the Technology Transfer Strategies of American Research Universities. *Management Sciences, 48*(1), 105-121.

Fernandes, M., Santos, S., & Gouveia, A. F. (2017). The Empirics of Agglomeration Economies: The Link with Productivity. *Revue D'economie Industrielle, 123*(3), 87-109.

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 in Great Britian by Political and Economic Planning. *The Economic Journal, 49*(194), 331-335.

Forslid, R., & Okubo, T. (2014). Spatial Sorting with Heterogeneous Firms and Heterogeneous Sectors. *Regional Science and Urban Economies, 46*, 42-56.

Frenken, K., Oort, F. V., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies, 41*(5), 685-697.

Fukao, K., & Kwon, H. U. (2006). Why did Japan's TFP Growth Slowdown in the Last Decade? An Empirical Analysis Based on Firm-Level Data of Manuacturing Firms. *The Japanese Economic Review, 57*(2), 195-228.

Fukao, K., Kravtsova, V., & Nakajima, K. (2014). How Important is Geographical Agglomeration to Factory Efficiency in Japan's Manufacturing Sector? *The Annals of Regional Science, 52*(3), 659-696.

Fuller, W. A. (1969). Grafted Polynomials as Approximating Functions. *Australian Journal of Agricultral Economics, 13*(1), 35-46.

Gardezi, U. N. (2013). Labor Pooling as a Determinant of Industrial Agglomeration. *CREB Working papers 4-2013*.

Geambaşu, L., & Stancu, I. (2011). Spline Regression and Research of Capital Market Under Financial Crises Impact. *Theoretical and Applied Economics, 5*(5), 278-286.

GEAMBAŞU, L., JIANU, I., & GEAMBAŞU, C. (2010). Spline Linear Regression Used for Evaluating Financial Assets. *Annals-Economic Series, 4*, 310-321.

Glaeser, E. L., & Mare, D. C. (2001). Cities and Skills. *Journal of Labor Economics, 19*(2), 316-342.

Gornig, M., & Schiersch, A. (2019). Agglomeration Economies and the Firm TFP: Different Effects Across Industries. *DIW Berlin Discussion Paper No. 1788*.

Gottman, J. (1957). Megalopolis, or the urbanization of the North-Eastern seaboard. *Economic Geography, 33*(7), 189-200.

Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science, 88*(1), 63-84.

Graham, D. J., & Kim, H. Y. (2008). An Empirical Analytical Framework for Agglomeration Economies. *The Annals of Regional Science, 42*(2), 267-289.

Graham, D. J., & Melo, P. C. (2009). Agglomeration Economies and Labor Productivity: Evidence from Longitudinal Worker Data for GBs Travel-to-Work Areas. *SERC Discussion Papers 0031* .

Graham, D., Gibbons, S., & Martin, R. (2009). Transport Investment and the Decay of Agglomeration Benefits.

Greenaway, D., & Kneller, R. (2008). Exporting, Productivity and Agglomeration. *European Economic Review, 52*(5), 919-939.

Greiner, A., & Kauermann, G. (2007). Sustainability of US Public Debt: Estimating Smoothing Spline Regressions. *Economic Modelling, 24*(2), 350-364.

Guevara-Rosero, G. C., Riou, S., & Autant-Bernard, C. (2019). Agglomeration externalities in Ecuador: do urbanization and tertiarization matter? *Regional Studies, 53*(5), 706-719.

Guimarães, P., Figueiredo, O., & Woodward, D. (2007). Measuring the Localization of Economic Activity: A Parametric Approach. *Journal of Regional Science, 47*(4), 753-774.

Gutierrez, T. T., & Ordóñez, J. A. (2019). Agglomeration Economies and Urban Productivity. *REGION, 6*(1), 17-24.

Han, F., Xie, R., & Fang, J. (2018). Urban Agglomeration Economies and Industrial Energy Efficiency. *Energy, 1*, 45-59.

Harrison, S., & Dourish, P. (1996). Re-place-ing space: the roles of place and space in collaborative systems. *CSCW '96: Proceedings of the 1996 ACM conference on Computer supported cooperative work*, (pp. 67-76).

Helsely, R., & Strange, W. (1990). Matching and Agglomeration Economies in a System of CIties. *Regional Science and Urban Economics, 20*(2), 189-212.

Helsley, R. W., & Strange, W. C. (2014). Coagglomeration, Clusters, and the scale and composition of cities. *Journal of Political Economy, 122*(5), 1064-1093.

Henderson, J. V. (1986). Efficiency of Resource Usage and City Size. *Journal of Urban Economics, 19*(1), 47-70.

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

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy, 103*(5), 1067-1085.

Hill, E. W., & Brennan, J. F. (2000). A Methodology for Identifying the Drivers of Industrial Clusters: The Foundation of regional competitive advantage. *Economic Development Quarterly, 14*(1), 65-96.

Holmes, T. J. (1999). Localization of Industry and Vertical Disintegration. *Review of Economics and Statistics, 81*(2), 314-325.

Hussain, S. T., Khan, U., Malik, K. Z., & Faheem, A. (2012). Constraints Faced by Industry in Punjab, Pakistan. *Lahore Journal of Economics, 17*(Special Edition), 135-189.

Iqbal , A., & Siddiqi, M. W. (2013). Spatial Agglomeration and Productivity of Textile and Leather Manufacturing in the Punjab Province of Pakistan. *Pakistan Journal of Commerce and Social Sciences, 7*(1), 27-42.

Jacobs, J. (1969). *The Economies of Cities.* New York: Random House.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics, 108*(3), 577-598.

Karlsson, C., & Pettersson, L. (2005). Regional Productivity and Accessibility to Knowledge and Dense Markets. *Working Paper Series in Economics and Institutions of Innovation 32*.

Kemal, A. R. (2006). Key Issues in Industrial Growth in Pakistan. *Lahore Journal of Economics, 11*(Special Edition), 49-74.

Kerr, W. R., & Kominers, S. D. (2015). Agglomerative Forces and Cluster Shapes. *Review of Economics and Statistics, 97*(4), 877-899.

Kim, M. K., Harris, T. R., & Vusovic, S. (2009). Efficiency Analysis of the US Biotechnology Industry: Clustering Enhances Productivity. *AgBioForum, 12*(3&4), 422-436.

Kim, S. (1990). Labor Heterogeneity, Wage Barganing and Agglomeration Economies. *Journal of Urban Economics, 28*(2), 160-177.

Klaesson, J., & Larsson, H. (2009). Wages, Productivity and Industrial Composition - Agglomeration Economies in Swedish Regions. *Working Paper Series in Economics and Institutions of Innovation 203*.

Kmenta, J. (1967). On Estimation of CES Production Function. *International Economic Review, 8*(2), 180-189.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy, 99*(3), 483-499.

Kukalis, S. (2010). Agglomeration Economies and Firms Performance: The Case of Industry Clusters. *Journal of Management, 36*(2), 453-481.

Lakner, S., Cramon-Taubadel, S. V., & Brummer, B. (2012). Technical Efficiency of Organic Pasture Farming in Germany: The Role of Location Economies and of Specific Knowledge. *Renewable Agriculture and Food Systems, 27*(3), 228-241.

Lambert, D. M., McNamara, K. T., & Garrett, M. I. (2006). An Application of Spatial Poisson Models to Manufacturing Investment Location Analysis. *Journal of Agriculture and Applied Economics, 38*(1), 105-121.

Legros, D., & Dube, J. (2014). *Spatial Econometrics Using Microdata.* John Wiley & Sons.

López-Bazo, E., Vayá, E., Mora, A. J., & Suriñach, J. (1999). Regional Economic Dynamics and Convergence in the European Union. *The Annals of Regional Science, 33*(3), 343-370.

Lu, J., & Tao, Z. (2009). Trends and determinants of China's industrial agglomeration. *Journal of Urban Economics, 65*(2), 167-180.

Mahmood , K., Ahmed, R., & Jafferi, N. B. (2016). Urban Development and Industrial Clustering in Pakistan: A Study Based on Geographical Perspective. *Journal of Basic and Applied Sciences, 12*, 32-40.

Malmberg, A., Malberg, B., & Lundequist, P. (2000). Agglomeration and Firm Performance: Economies of Scale, Localization, and Urbanisation among Swedish Export Firms. *Environment and Planning, 32*(2), 305-321.

Maré, D. C., & Graham, D. J. (2013). Agglomeration Elasticities and Firm Heterogeneity. *Journal of Urban Economics, 75*, 44-56.

Markusen, A. (1996). Sticky Places in Slippery Space: A Typology of Industrial Districts. *Economic Geography, 72*(3), 293-313.

Marsh, L. C. (1986). Estimating the Number and Location of Knots in Spline Regressions. *Journal of Applied Business Research, 2*(3), 60-70.

Marsh, L. C., & Cormier, D. R. (2002). *Spline Regression Models.* Thousand Oaks, California: SAGE Publications, Inc.

Marshall, A. (1920). *Principles of Economics.* London: Macmillan.

Martin, P., Mayer, T., & Mayneris, F. (2008). Spatial Concentration and Firm-Level Productivity in France. *Journal of Urban Economics, 69*(2), 182-195.

McCoy, K., & Moomaw, R. L. (1995). Determinants of Manufacturing Efficiency in Canadian Cities: A Stochastic Frontier Approach. *The Review of Regional Studies, 25*(3), 317-330.

Miao, C., & Wang, H. (2005). Analyzing China's Urban Agglomeration Development. *Journal of Urban Development Studies, 12*, 11-14.

Mitra, A. (1999). Agglomeration Economies as Manifested in Technical Efficiency at the Firm Level. *Journal of Urban Economics, 45*(3), 490-500.

Mitra, A. (2014). Agglomeration Economies and Wellbeing: Evidence from India. *Athens Journal of Health, 1*(1), 23-36.

Montana, J. P., & Nenide, B. (2008). The Evolution of regional Industry Clusters and their Implications for Sustainable Economic Development: Two Case Illustrations. *Economic Development Quareterly, 22*(4), 290-302.

Moomaw, R. L. (1998). Agglomeration Economies: Are They Exaggerated by Industrial Aggregation? *Regional Science and Urban Economics, 28*(2), 199-211.

Mori, T., & Smith, T. E. (2013). A Spatial Approach to Identifying Agglomeration Determinants. *Discussion Papers 13014*.

Myles, S. J., & Flyer, F. (2000). Agglomeration Economies, Firm Heterogeneity, and Foreign Direct Investment in the United States. *Strategic Management Journal, 21*(12), 1175-1193.

Nakamura, R. (1983). Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities. *Journal of Urban Economics*.

Nakamura, R. (1985). Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities. *Journal of Urban Economics, 17*(1), 108-124.

Nakamura, R. (2008). Agglomeration Effects on Regional Economic Disparities: A Comparison between UK and Japan. *Urban Studies, 45*(9), 1947-1971.

Nasir, M. (2017). Agglomeration and Firm Turnover in Punjab. *The Lahore Journal of Economics, 22*(1), 19-36.

Nathan, M., & Overman, H. (2013). Agglomeration, Clusters, and Industrial Policy. *Oxford Review of Economic Policy, 29*(2), 383-404.

Ning, Y. (2015). *Issues in China's Urban Agglomeration Studies and New Exploration for China's Urban Agglomeration selection and Nurturing.* Beijing: Science Press.

North, D. C. (1955). Location Theory and Regional Economic Growth. *Journal of Political Economy, 63*(3), 243-258.

Oktay, E., Talase, E., Alkan, Ö., & Genç, A. (2012). Modeling with Linear Spline Regression of Turkish Tourism Demand. *Journal of Selçuk University Natural and Applied Science, 1*, 10-22.

Otsuka, A., & Goto, M. (2015). Estimations and Determinants of Energy Efficiency in Japanese Regional Economies. *Regional Science Policy and Practice, 7*(2), 89-101.

Otsuka, A., Goto, M., & Sueyoshi, T. (2010). Industrial Agglomeration Effects in Japan: Productive Efficiency, Market Access, and Public Fiscal Transfer. *Papers in Regional Science, 89*(4), 819-840.

Overman, H. G., & Puga, D. (2010). Labor Pooling as a Source of Agglomeration: An Empirical Investigation. In E. L. Glaeser, *Agglomeration Economies* (pp. 133-150). University of Chichago Press.

Pakistan, G. o. (2015-16). *Census of Manufacturing Industries.* Ministry of Planning Development & Special Initiatives. Islamabad: Pakistan Bureau of Statistics.

Pakistan, G. o. (2018). *Punjab Growth Strategy.* Planning & Development Department, Government of Punjab.

Pakistan, G. o. (2018). *Punjab Growth Strategy, Accelerating Economic Growth and Improving Social Outcomes.*

Pakistan, G. o. (2022). *Pakistan Bureau of Statistics*. Retrieved from http://www.pbs.gov.pk

Partridge, M. D., & Rickman, D. S. (2008). Distance from Urban Agglomeration Economies and Rural Poverty. *Journal of Regional Science, 48*(2), 285-310.

Pasha, H. (2018). *Growth and Inequality in Pakistan.* Friedrich-Ebert-Stiftung.

Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric Models and Economic Forecasts.* New York.

Poirier, D. J. (1973). *Journal of the American Statistical Association, 68*(343), 515-524.

Poirier, D. J. (1975). On the Use of Bilinear Splines in Economics. *Journal of Econometrics, 3*(1), 23-24.

Poirier, D. J. (1976). *The Econometrics of Structural Change with Special Emphasis on Spline Functions.* Amsterdam: North Holland Publishing Company.

Potter, A., & Watts, H. D. (2014). Revisiting Marshall's Agglomeration Economies: Technological Relatedness and the Evolution of the Sheffield Metals Clusters. *Regional Studies, 48*(4), 603-623.

Rasiah, R., & Nazeer, N. (2015). The State of Manufacturing in Pakistan. *The Lahore Journal of Economics, 20*(Special Edition), 205-224.

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

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial Determinants of Productivity: Analysis for the regions of Great Britian. *Regional Science and Urban Economics, 36*, 727-752.

Ringstad, V., & Grilichs, Z. (1972). Economies of Scale and the Form of the Production Function. *Recherches Économiques de Louvain/ Louvain Economic Review, 38*(2), 234.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy, 98*(5), 71-102.

Rosenthal, S. S., & Strange, W. C. (2001). The Determinants of Agglomeration. *Journal of Urban Economics, 33*(2), 191-229.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In J. V. Henderson, & J. F. Thisse, *Handbook of Regional and Urban Economics* (pp. 2118-2147).

Rosenthal, S. S., & Strange, W. C. (2006). The micro-empirics of agglomeration economies. In R. J. Arnott, & D. P. Mcmillen, *A Companion to Urban Economics* (pp. 7-23).

Rosenthal, S., & Strange, W. (2003). Geography, Industrial Organization, and Agglomeration. *The Review of Economics and Statistics, 85*(1), 377-393.

Ryohei, N. (1985). Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities. *Journal of Urban Economics, 17*(1), 108-124.

Saito, H., & Matsuura, T. (n.d.). Agglomeration Economies, Productivity, and Quality Upgrading: Evidence from Japan. *Discussion Papers 16085*.

Sargan, D. (1971). Production Functions. In R. Layard, *Qualified Manpower and Economic Performance.* London: Allan Lane.

Segal, D. (1976). Are There Returns to Scale in City Size? *Review of Economics and Statistics, 58*, 339-350.

Shefer, D. (1973). Localization Economies in SMSAs: A Production Function Analysis. *Journal of Regional Science, 13*, 55-64.

Sinain, E., Jones, D. C., & Mygind, N. (2007). *Determinants of firm level technical efficiency: A Stochastic Frontier Approach.* Copenhagen Business School.

Smith, P. L. (1979). Splines as a Useful and Convenient Statistical Tool. *The American Statistician, 33*(2), 57-62.

Stankov, U., Armenski, T., Klauco, M., Pavlukovic, V., Cimbaljevic, M., & Drakulic-Lovacevic, N. (2017). Spatial Autocorrelation analysis of Tourist Arrivals Using Municipal Data: A Serbian Example. *Geographica Pannonica, 21*(2), 106-114.

Statistics, P. B. (2017-18). *Pakistan Labor Statistics*. Retrieved from PBS: http://www.pbs.gov.pk/sites/default/files/LabourForce/publications/lfs2017_18/AnnualReportofLFS2017-18.pdf

Suits, D. B., Mason, A., & Chan, L. (1978). Spline Functions Fitted by Standard Regression Method. *The Review of Economics and Statistics, 60*(1), 132-139.

Sveikauskas, L. (1975). The Productivity of Cities. *Quaterly Journal of Economics, 89'*, 393-413.

Tomota, M., & Smith, T. E. (2013). A Spatial Approach to Identifying Agglomeration Determinants. *Discussion Papers 13014*.

Tveteras, R., & Battese, G. E. (2006). Agglomeration Externalities, Productivity, and Technical Inefficiency. *Journal of Regional Science, 46*(4), 605-625.

Van Soest, D. P., Gerking, S., & Van Oort, F. G. (2006). Spatial Impacts of Agglomeration Externalities. *Journal of Regional Science, 46*(5), 881-899.

Verstraten, P., Verweij, G., & Zwaneveld, P. J. (2019). Complexities in the Spatial Scope of Agglomertion Economies. *Journal of Regional Science, 59*(1), 29-55.

Wagner, A. (1891). Marshall's Principles of Economics. *The Quarterly Journal of Economics, 5*(3), 319-338.

Wang, W. C., Chang, Y. J., & Wang, H. C. (2019). An Application of the Spatial Autocorrelation Method on the Change of Real Estate Prices in Taitung City. *International Journal of GeoInformation, 8*(6), 249.

Widodo, W., Salim, R., & Bloch, H. (2015). The Effects of Agglomeration Economies on Technical Efficiency of Manufacturing Firms: Evidence from Indonesia. *Applied Economics, 47*(31), 3258-3275.

Wu, W. J., Zhao, S. Q., Zhu, C., & Jiang, J. L. (2015). A Comparative Study of Urban Expansion in Beijing, Tainjin and Shijiazhuang over the Past Three Decades. *Landscape and Urban Planning, 134*, 93-106.

Yao, S., Zhu, Y., & Chen, Z. (2001). *China's Urban Agglomeration.* Hefei: University of Science and Technology of China Press.

Yezer, A., & Goldfarb, R. (1978). An Indirect Test of Efficient City Sizes. *Journal of Urban Economics, 5*(1), 46-65.

Zareipour, H., Bhattacharya, K., & Canizares, C. (2006). Forecasting The Hourly Ontario Energy Price by Multivariate Adaptive Regression Splines. *IEEE*.

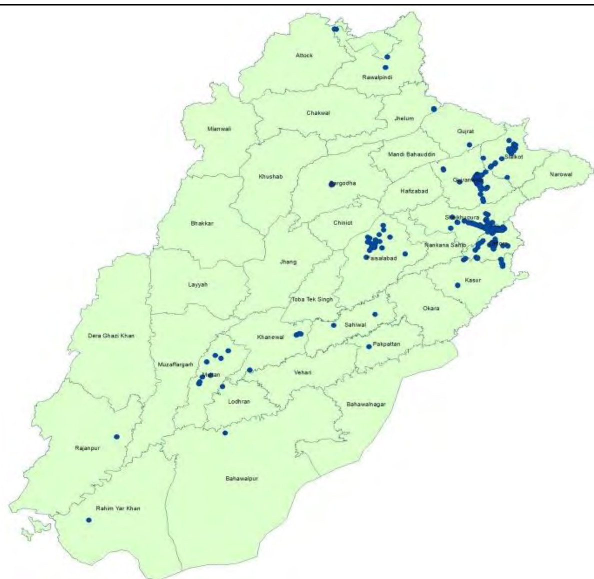
Zeng, D. Z. (2013). China's Special Economic Zones and Industrial CLusters: Success and Challenges. *Journal of International Commerce, Economics and Policy, 3*(3), 3-5.

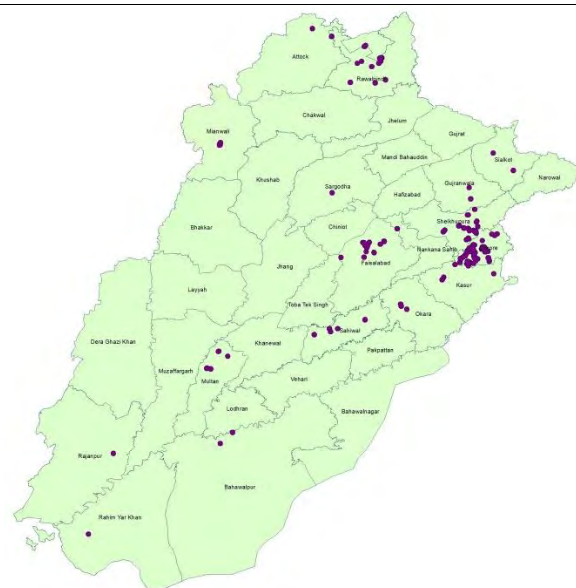
Zhou, Y., & Shi, Y. (1995). Towards Establishing the Concept of Physical Urban Areas in China. *Acta Geograaphica Sinica, 50*(5), 17-25.

9. Annexures

Annexure I: Sector-Wise Concentration of Manufacturing Firms

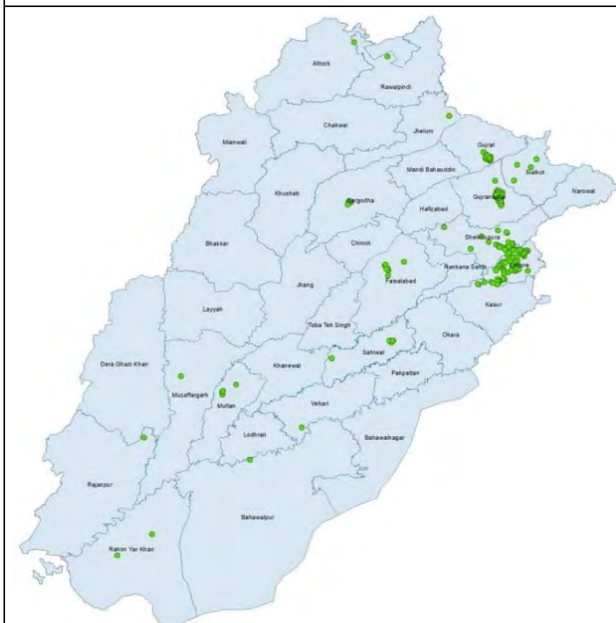
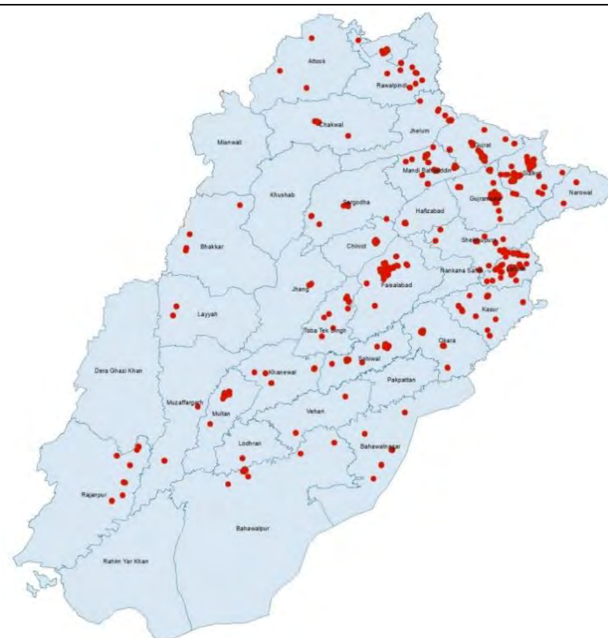
<table>
<tr><th>Basic Metal manufacturing firms in

Punjab</th><th>Basic pharmaceutical firms</th></tr>
<tr><td>


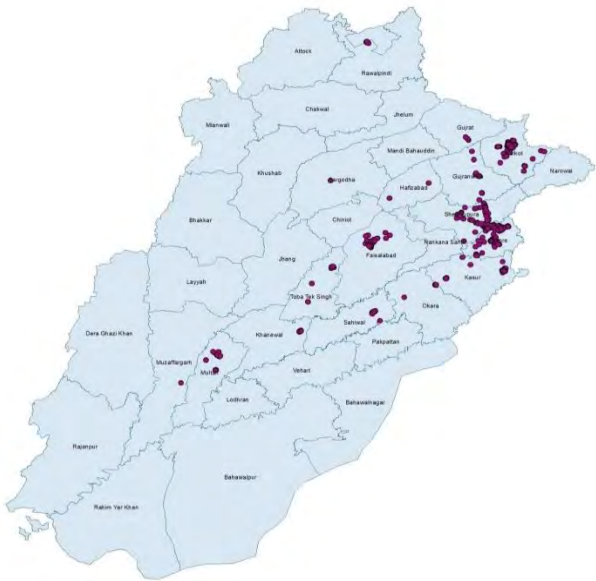
</td><td>


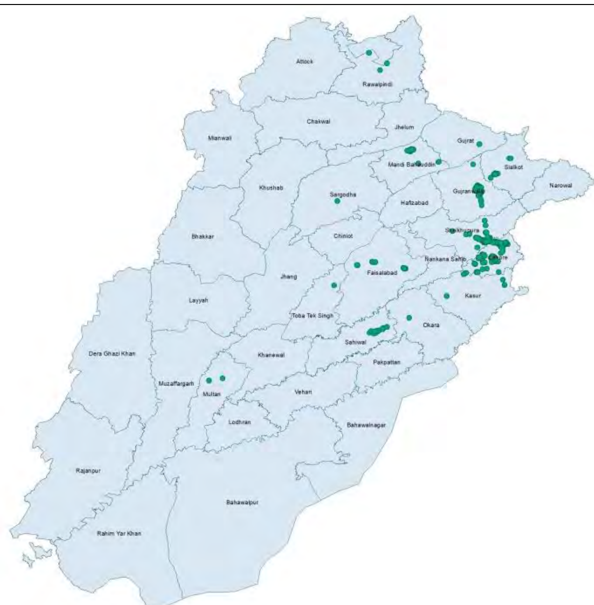
</td></tr>
</table>

Coke and Petroleum products	Computer and Electronic Products
	

Electrical Equipment	Furniture Products
	

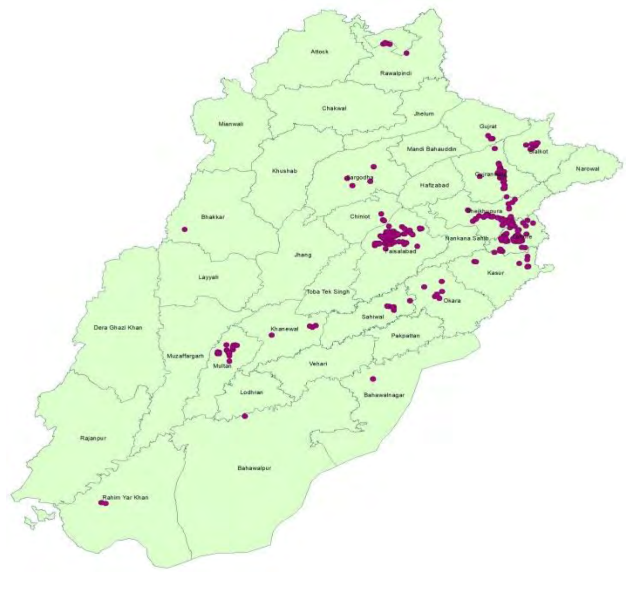
.

Leather and Leather Products	**Repairing and Installation of Machinery Equipment Sector**
	

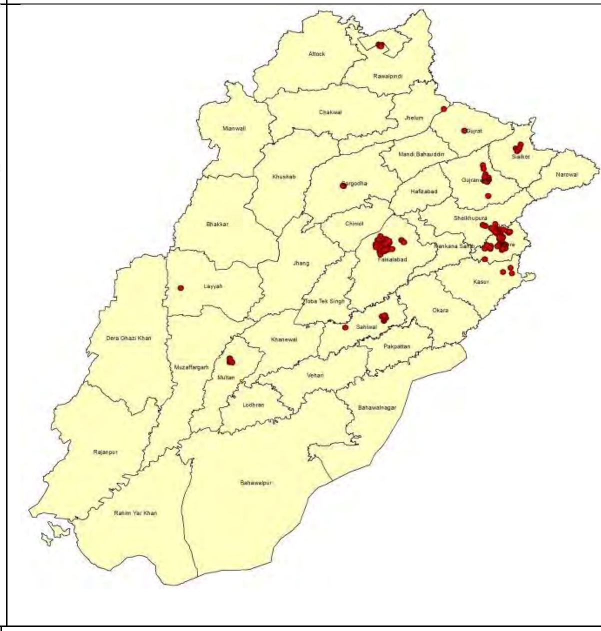
Motor Vehicles and Trailers Sector	**Other Transport Equipment Sector**
	

.

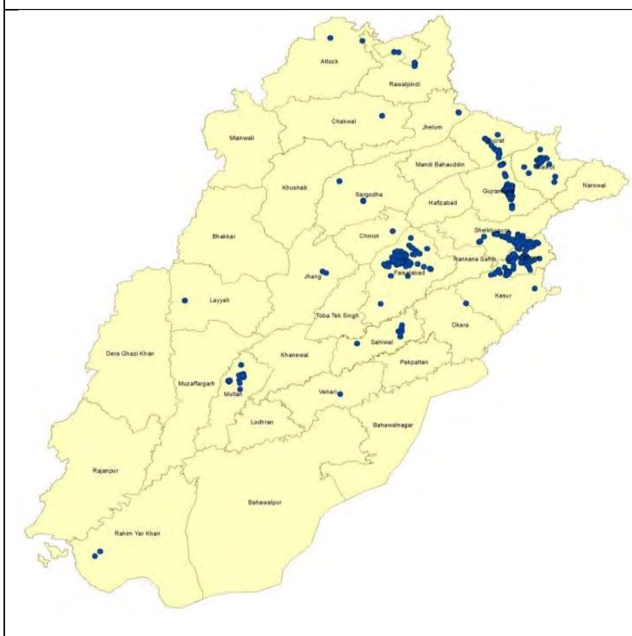
Paper and Paper Product Sector



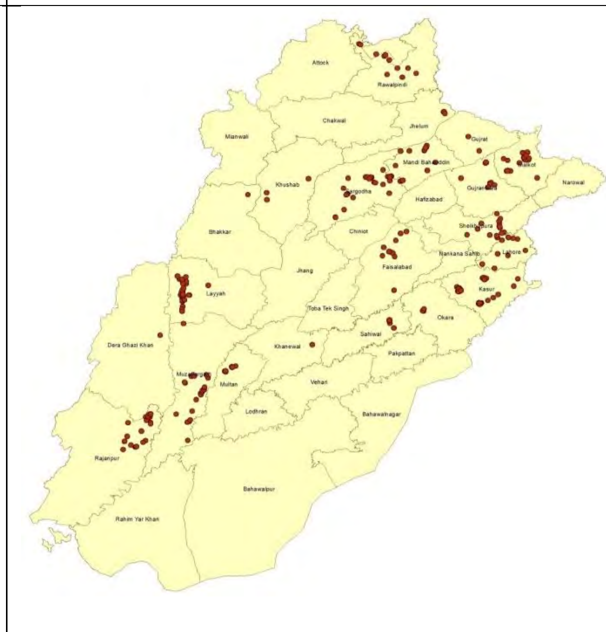
Printing and Reproduction of Recorded Media Sector



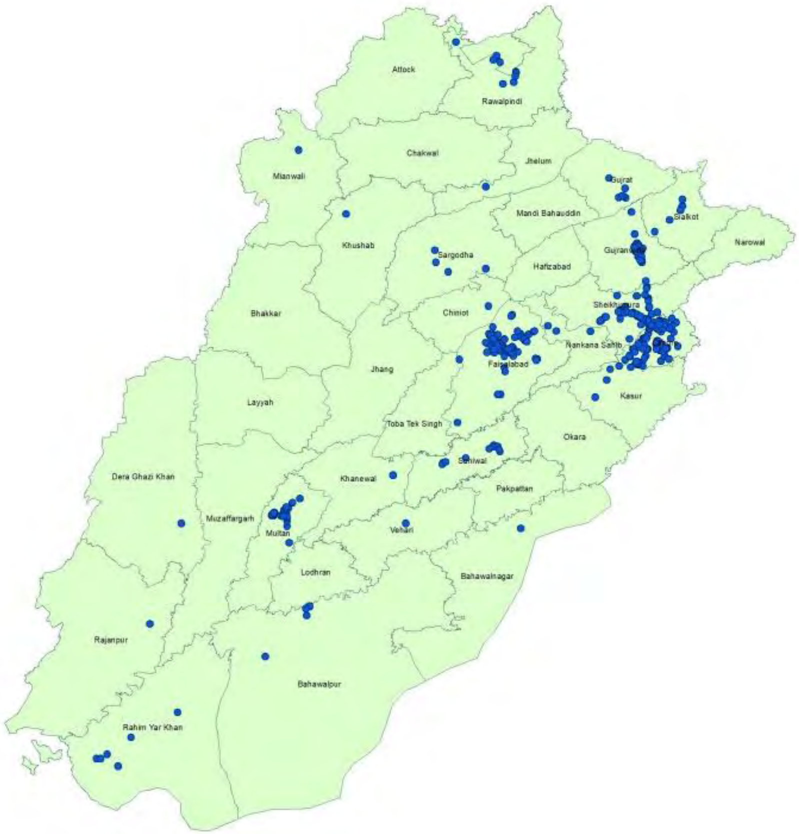
Rubber and Plastic Products Sector



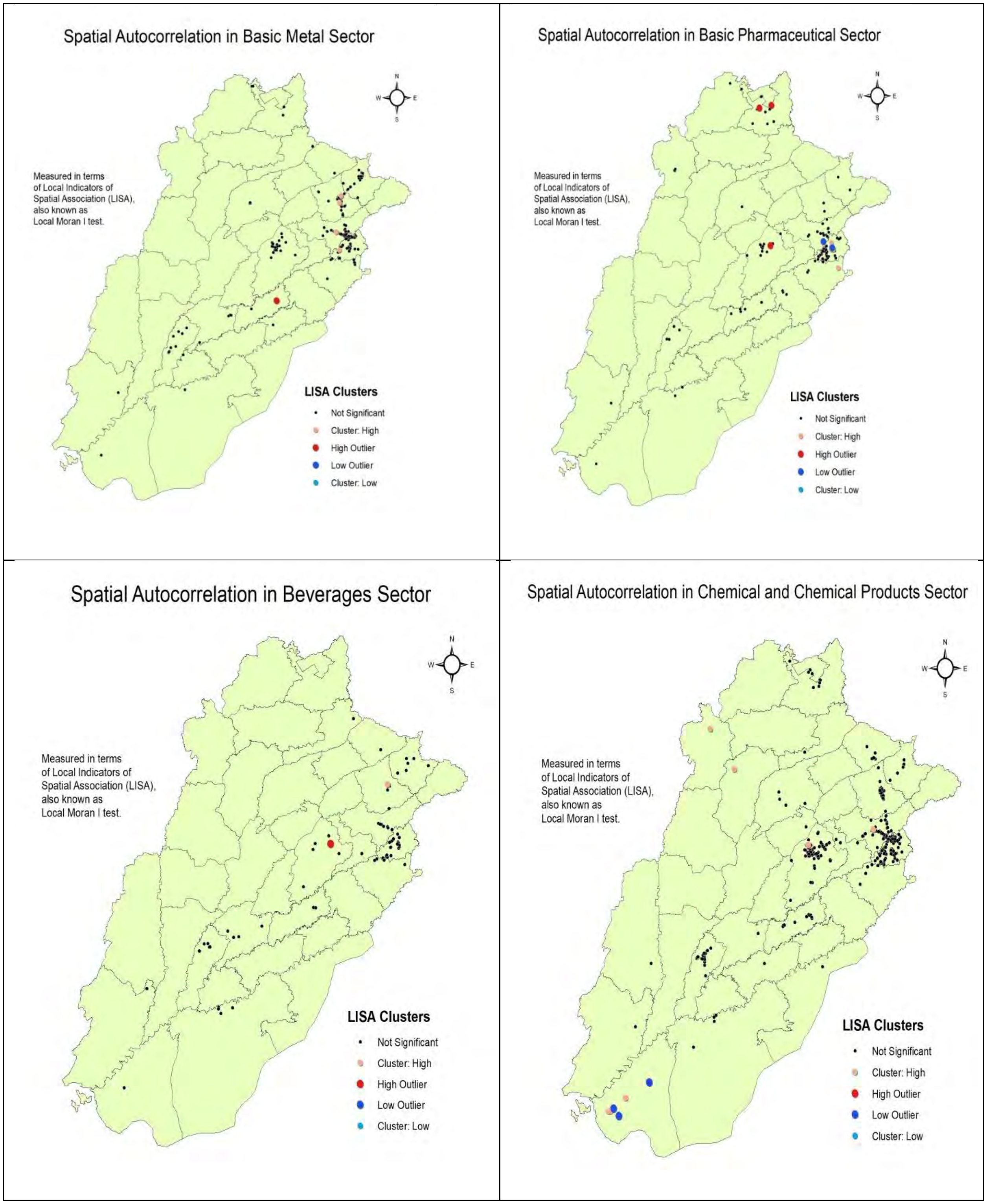
Wood and Wooden Products Sector



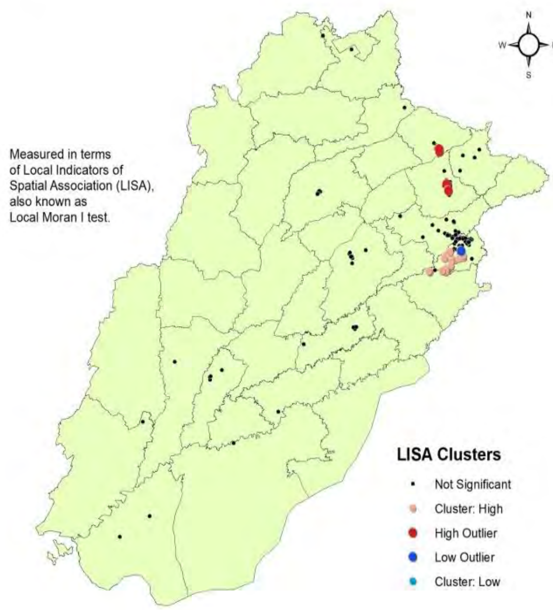
.

Chemical and Chemical products firms


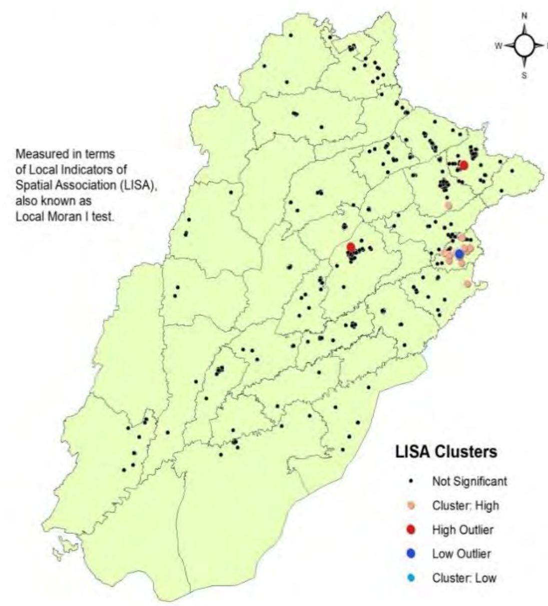
Annexure II: Results of Moran *I* test



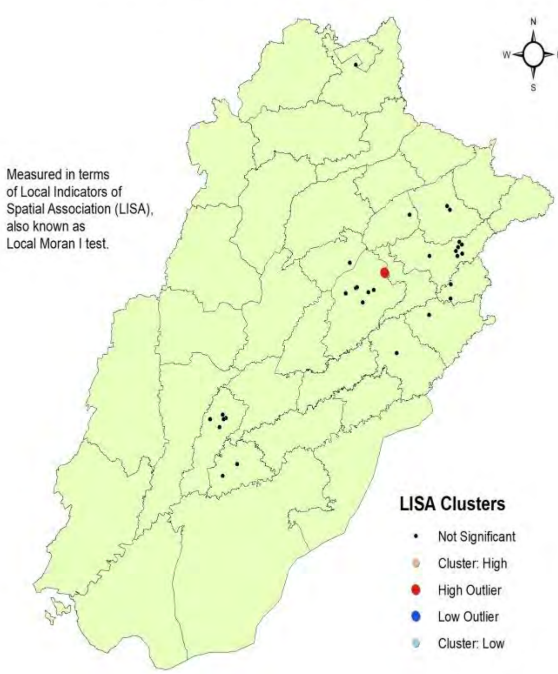
Spatial Autocorrelation in Electrical Equipment Sector



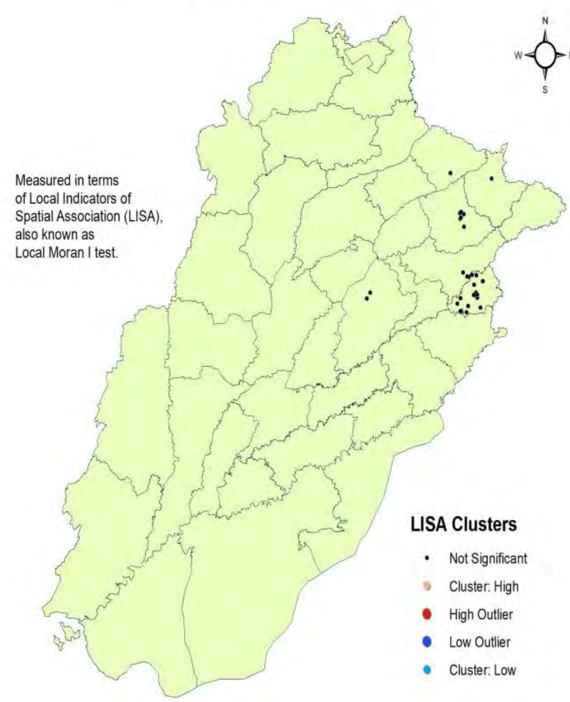
Spatial Autocorrelation in Furniture Sector



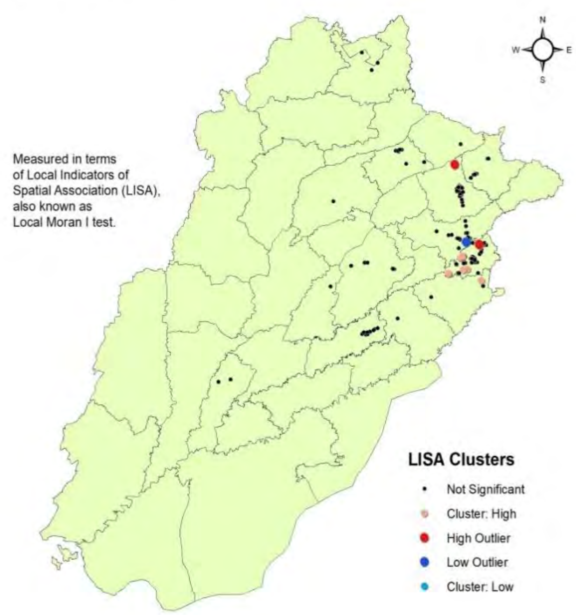
Spatial Autocorrelation in Coke and Refined Petroleum Products Sector



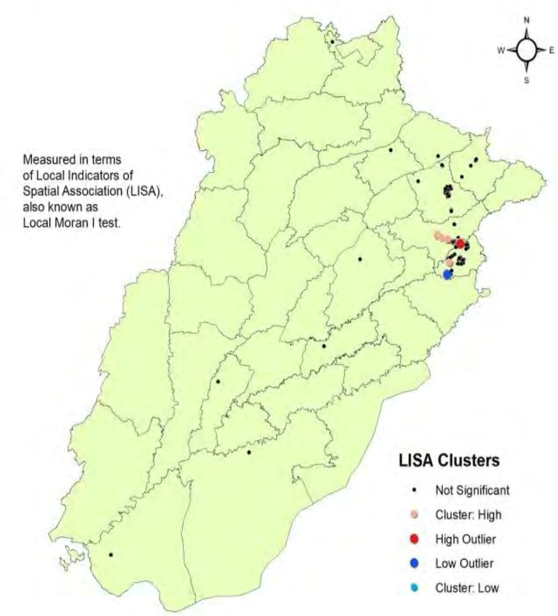
Spatial Autocorrelation in Computer, Electronic and Optical Products Sector



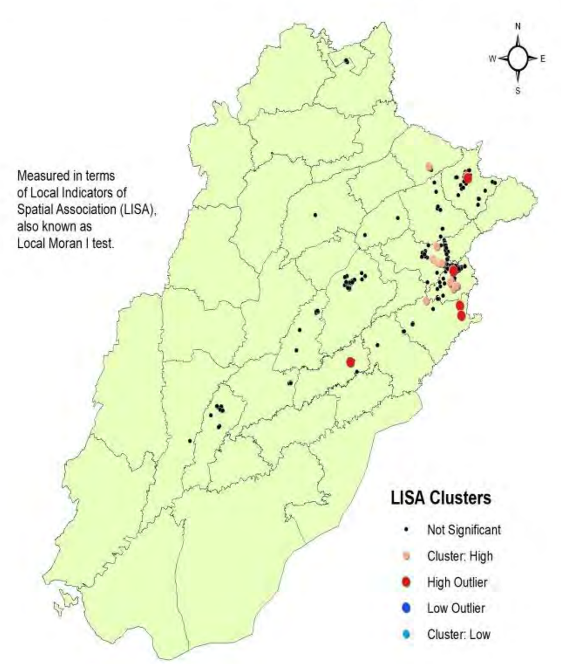
Spatial Autocorrelation in Motor Vehicles and Trailers Sector



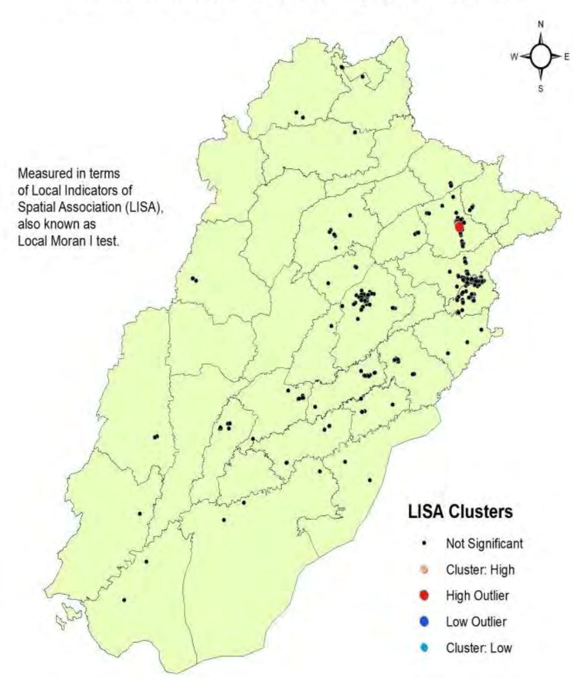
Spatial Autocorrelation in Other Transport Equipment Sector

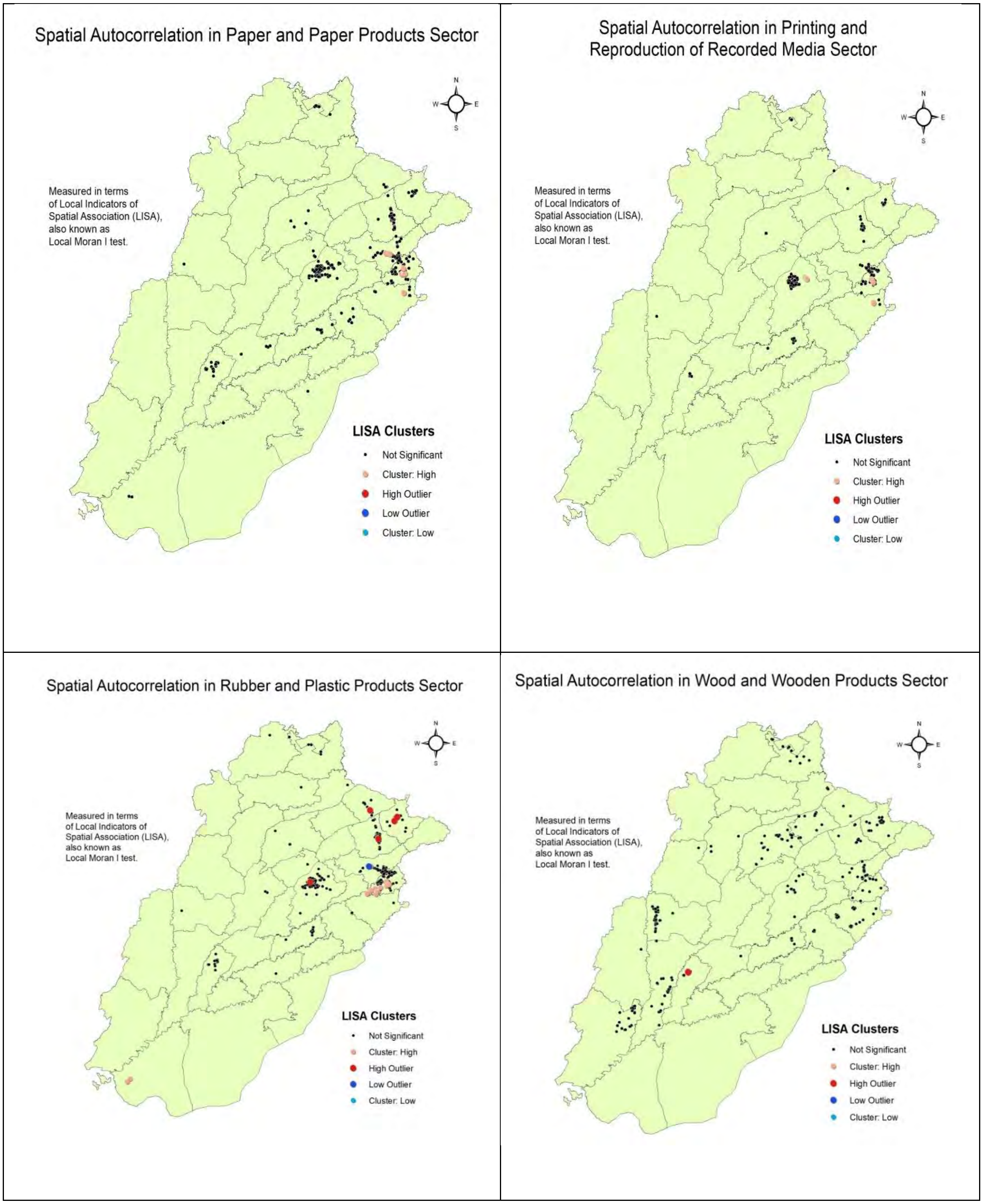


Spatial Autocorrelation in Leather and Related Products Sector



Spatial Autocorrelation in Repairing and
Installation of Machinery and Equipment Sector




Spatial Autocorrelation in Paper and Paper Products Sector
Measured in terms of Local Indicators of Spatial Association (LISA), also known as Local Moran I test.
LISA Clusters
Not Significant
Cluster: High
High Outlier
Low Outlier
Cluster: Low
Spatial Autocorrelation in Printing and Reproduction of Recorded Media Sector
Measured in terms of Local Indicators of Spatial Association (LISA), also known as Local Moran I test.
LISA Clusters
Not Significant
Cluster: High
High Outlier
Low Outlier
Cluster: Low
Spatial Autocorrelation in Rubber and Plastic Products Sector
Measured in terms of Local Indicators of Spatial Association (LISA), also known as Local Moran I test.
LISA Clusters
Not Significant
Cluster: High
High Outlier
Low Outlier
Cluster: Low
Spatial Autocorrelation in Wood and Wooden Products Sector
Measured in terms of Local Indicators of Spatial Association (LISA), also known as Local Moran I test.
LISA Clusters
Not Significant
Cluster: High
High Outlier
Low Outlier
Cluster: Low

Annexure III: Estimations of Spline Regression Models

Overall Manufacturing: All Industries

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

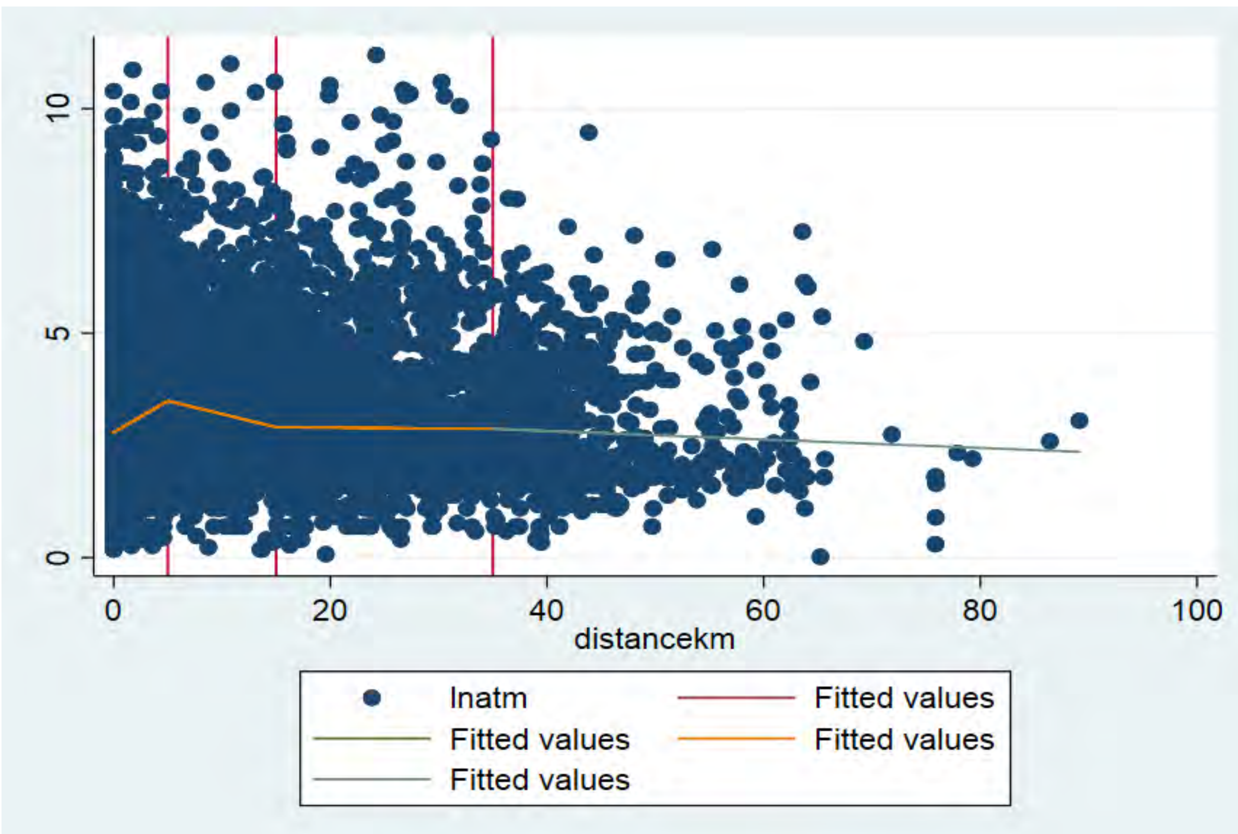
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**T**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.7872* (0.0100002)	278.71	Constant	2.772469*(0.0101082)	274.28	Constant	2.767335* (0.010167)	272.19
β0	Distancekm	0.1407368* (0.0071112)	19.79	distancekm	0.3491667* (0.0181426)	19.25	Distanced	0.5277953* (0.0334434)	15.78
β1	distancekm – 5	-0.1993568* (0.0115657)	-17.24	$distancekm^2$	-0.0441686* (0.0026814)	-16.47	$distancekm^2$	-0.1277633* (0.0104194)	-12.26
β2	distancekm – 15	0.0566923* (0.0073825)	7.68	$(distancekm - 5)^2$	0.0488028*(0.0032948)	14.81	$distancekm^3$	0.0089558* (0.0008119)	11.03
β3	distancekm – 35	-0.0074683 (0.0067252)	-1.11	$(distancekm - 15)^2$	-0.0049412* (0.0009064)	-5.45	$(distancekm - 5)^3$	-0.0091704* (0.0008843)	-10.37
β4	-	-	-	$(distancekm - 35)^2$	0.0005584 (0.0004322)	1.29	$(distancekm - 15)^3$	0.0002142* (0.0000958)	2.24
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	-2.36e -06 (0.0000253)	-0.09
No. of Observations	9391			9391			9391		
Significance	**F = 108.16 p = 0.0000**			**F = 103.88 p = 0.0000**			**F = 89.22 p = 0.0000**		

Distances	**Coefficients**	**Coefficient Value**	**Adjusted Coefficient[1]**
<5	α0	2.7872	-
	β0	0.1407368	0.151121633
<15	α0'	3.783984	-
	β0'	-0.05862	-0.056934934
<35	α0''	2.9335995	-

	β0''	-0.0019277	-0.001925843
>35	α0'''	3.19499	-
	β0'''	-0.009396	-0.009351996
[1] Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**84997.1**	**85037.77**
Quadratic Spline Regression	84913.92	84962.73
Cubic Spline Regression	84900.28	84957.22



.

PSIC 10: Food Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

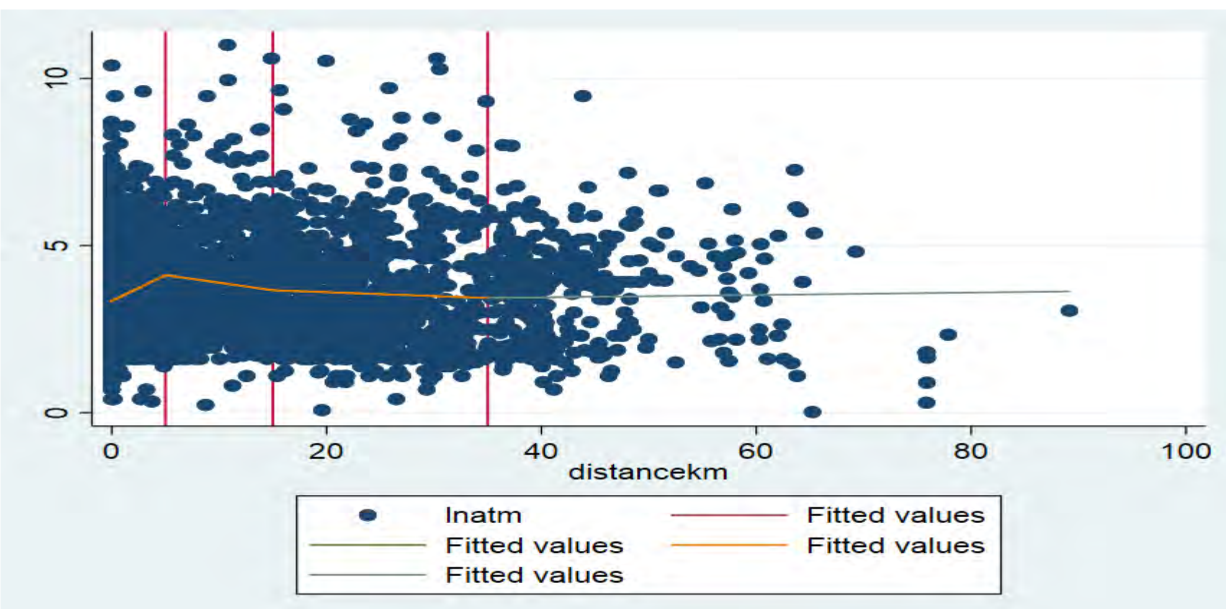
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	3.339951* (0.044808)	74.54	Constant	3.305247* (0.0464427)	71.17	Constant	3.288575* (0.0472673)	69.57
β0	Distanced	0.1547987* (0.0208342)	7.43	Distancekm	0.3026104* (0.0515056)	5.88	Distancekm	0.4329748* (0.0917256)	.72
β1	distancekm – 5	-0.1999989* (0.312668)	-6.4	distancekm^2	-0.0316761* (0.0071316)	-4.44	distancekm^2	-0.0840759* (0.0272523)	-3.09
β2	distancekm – 15	0.0336216* (0.016745)	2.01	(distancekm - 5)^2	0.0302535* (0.0084234)	3.59	distancekm^3	0.0053359* (0.002091)	2.55
β3	distancekm – 35	0.0152217 (0.0122868)	1.24	(distancekm - 15)^2	0.0027599*** (0.0017602)	1.57	(distancekm - 5)^3	-0.0051632* (0.0022569)	-2.29
β4	-	-	-	(distancekm - 35)^2	-0.0000491* (0.0000124)	-3.96	(distancekm - 15)^3	0.0001722 (0.0002182)	-0.79
β5	-	-	-	-	-	-	(distancekm - 35)^3	-0.0000453 (0.0000479)	-0.94

No. of Observations	3,493	3,493	3,493
Significance	F = 16.59 p = 0.0000	F = 18.00 p = 0.0000	F = 15.49 p = 0.0000

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	3.339951	-
	β0	0.1547987	0.167422935
<15	α0'	4.3399455	-
	β0'	-0.0452002	-0.04419389
<35	α0''	3.8356215	-
	β0''	-0.0115786	-0.011511826
>35	α0'''	3.302862	-
	β0'''	0.0036431	0.003649744
*Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**13041.27**	**13072.06**
Quadratic Spline Regression	13020.07	13057.02
Cubic Spline Regression	13019.16	13062.27



.

PSIC 13: Textile Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

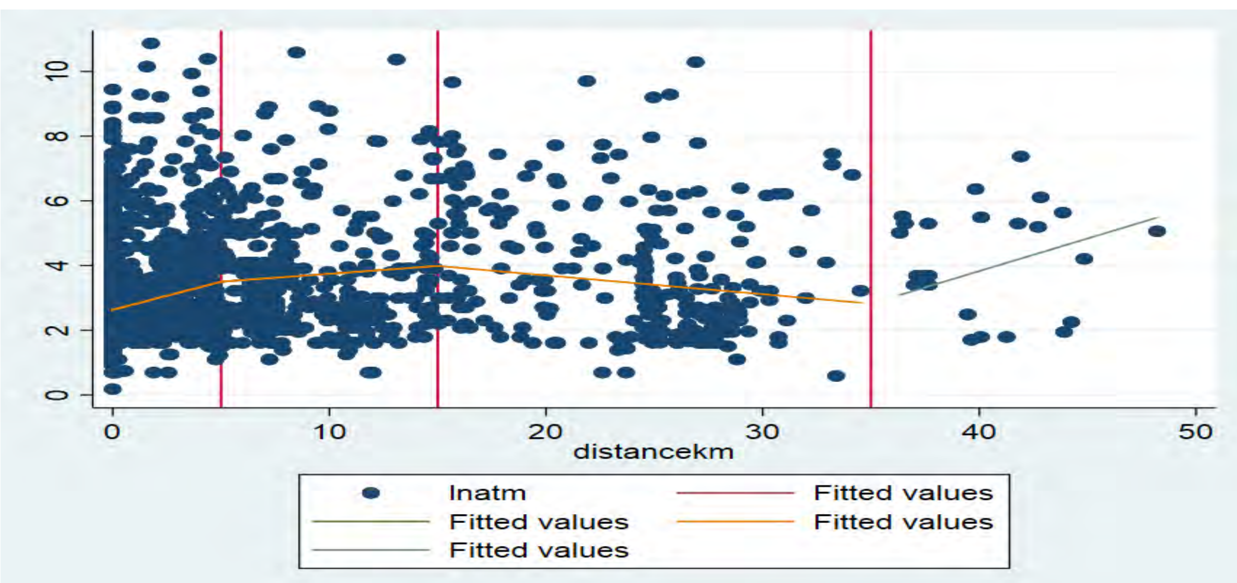
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.629229* (0.0155018)	169.61	Constant	2.622016* (0.0156421)	167.63	Constant	2.61545* (0.0156156)	167.49
β0	Distance km	0.1739067* (0.0108935)	15.96	distancekm	0.2776158* (0.0281707)	9.85	Distanced	0.6056339* (0.0592264)	0.23
β1	distancekm – 5	-0.125115* (0.0200492)	-6.24	distancekm^2	-0.0214102* (0.0046265)	-4.63	distancekm^2	-0.1535926* (0.0201952)	-7.61
β2	distancekm – 15	-0.1070284* (0.199968)	-5.35	(distancekm - 5)^2	0.016892* (0.0061106)	2.76	distancekm^3	0.01199* (0.0016436)	7.29
β3	distancekm – 35	0.2597874* (0.0472614)	5.5	(distancekm - 15)^2	0.0050834* (0.0025792)	1.97	(distancekm - 5)^3	-0.0135041* (0.001849)	-7.3
β4	-	-	-	(distancekm - 35)^2	0.0125111** (0.0074979)	1.67	(distancekm - 15)^3	0.0022706* (0.000309)	7.35
β5	-	-	-	-	-	-	(distancekm - 35)^3	-0.003698* (0.000787)	-4.7

No. of Observations	6,747	6,747	6,747
Significance	**F = 156.45 p = 0.0000**	**F = 121.65 p = 0.0000**	**F = 114.22 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.629929	-
	β0	0.1739067	0.189944539
<15	α0'	3.255504	-
	β0'	0.0487917	0.050001613
<35	α0''	4.86093	-
	β0''	-0.0582367	-0.056573388
>35	α0'''	-4.231629	-
	β0'''	0.2015507	0.223298257
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	20824.25	20858.34
Quadratic Spline Regression	**20842.26**	**20883.16**
Cubic Spline Regression	20773.74	20821.46



•

PSIC 14: Wearing Apparel Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

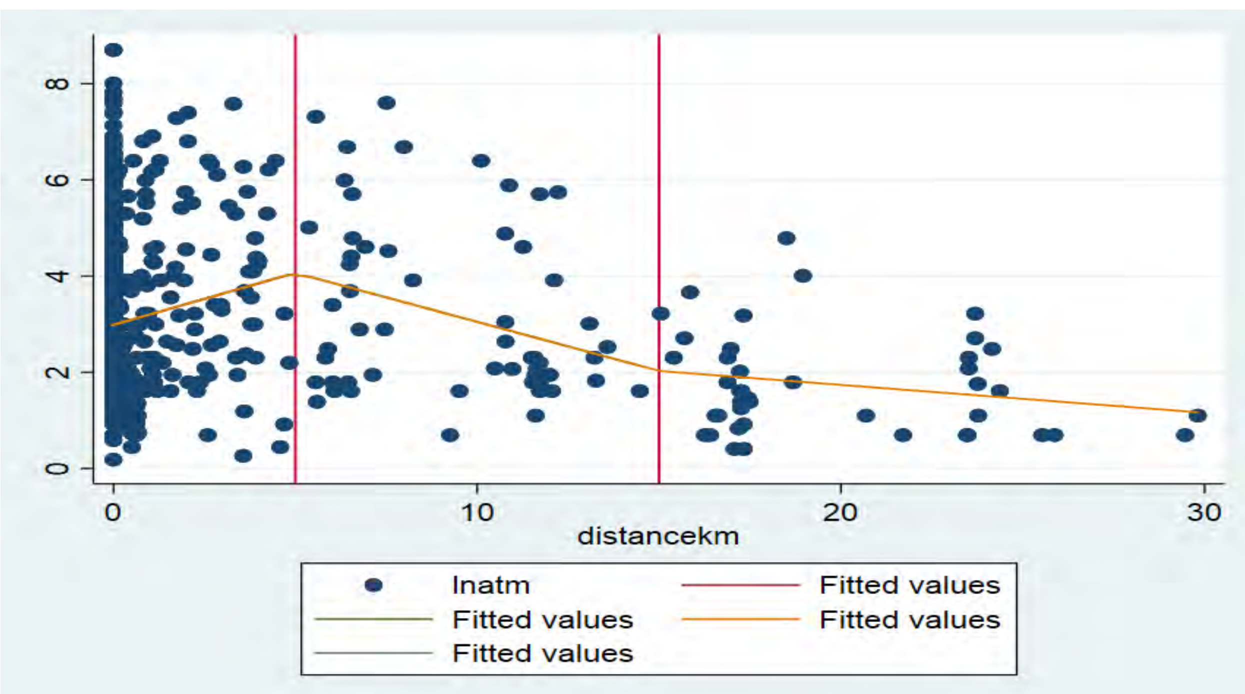
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	2.984237* (0.0358508)	83.24	Constant	2.977793* (0.0362716)	82.1	Constant	2.98292* (0.0364995)	81.72
β0	distancekm	0.2165951* (0.0419931)	5.16	distancekm	0.3937812* (0.1001787)	3.93	distancekm	0.2298499 (0.1768752)	.30
β1	distancekm – 5	-0.4205678* (0.0726567)	-5.79	distancekm^2	-0.0496209* (0.0159303)	-3.11	distancekm^2	0.0081372 (0.0618456)	0.13
β2	distancekm – 15	0.1459762* (0.0726027)	2.01	(distancekm - 5)^2	0.0455148* (0.0213025)	2.14	distancekm^3	-0.0036873 (0.0051594)	-0.71
β3	distancekm – 35	-	-	(distancekm - 15)^2	0.0130705 (0.0127179)	1.03	(distancekm - 5)^3	0.0061872 (0.0059834)	1.03
β4	-	-	-	(distancekm - 35)^2	-	-	(distancekm - 15)^3	-0.0040661* (0.0018487)	-2.2
β5	-	-	-	-	-	-	(distancekm - 35)^3	-	-
No. of Observations	1,653			1,653			1,653		

Significance	F = 20.72 p = 0.0000	F = 15.28 p = 0.0000	F = 12.93 p = 0.0000

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.984237	-
	β0	0.2165951	0.241841179
<15	α0'	5.087076	-
	β0'	-0.2039727	-0.184515366
<35	α0''	2.897433	-
	β0''	-0.0579965	-0.05634675
>35	α0'''	-	-
	β0'''	-	-
*Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	5739.727	5761.368
Quadratic Spline Regression	**5742.704**	**5769.756**
Cubic Spline Regression	5741.237	5773.7



PSIC 15: Leather Products Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

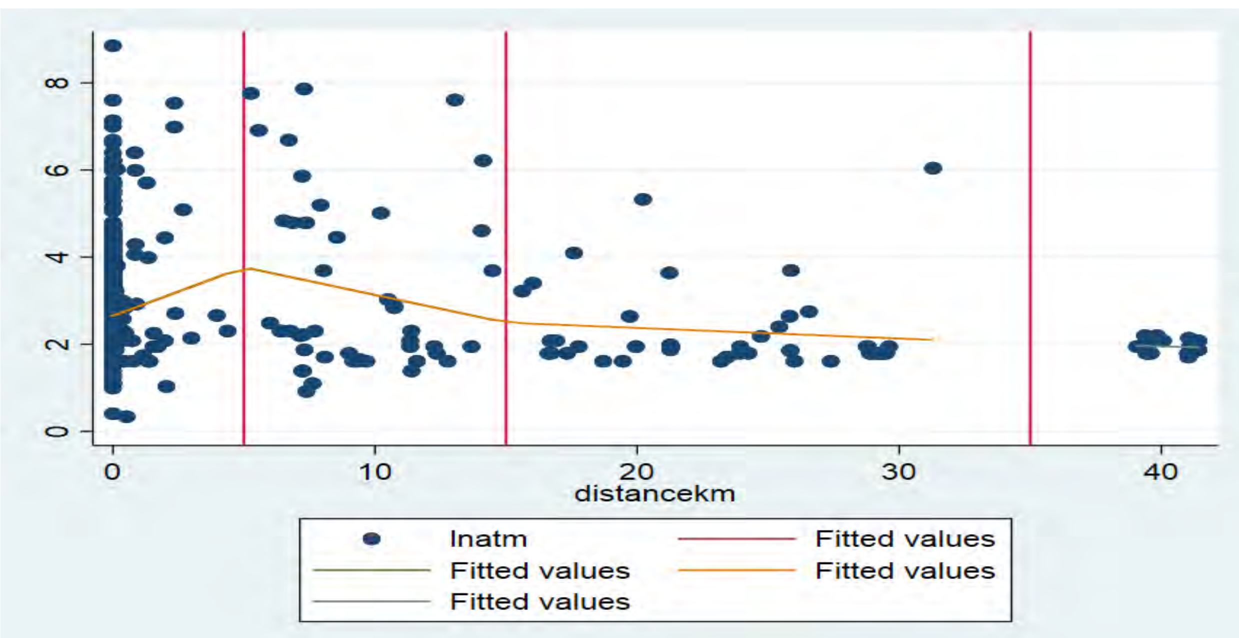
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.649369* (0.0399536)	66.31	Constant	2.640465* (0.040125)	65.81	Constant	2.639636* (0.0402778)	65.54
β0	distancekm	0.223448* (0.0526053)	4.25	distancekm	0.6361077* (0.1593156)	3.99	distancekm	0.8121862* (0.2576357)	3.15
β1	distancekm – 5	-0.35055* (0.0913764)	-3.84	distancekm^2	-0.0831014* (0.0236262)	-3.52	distancekm^2	-0.1928378* (0.0783847)	-2.46
β2	distancekm – 15	0.1027615*** (0.0657835)	1.56	(distancekm - 5)^2	0.0909023* (0.0288819)	3.15	distancekm^3	0.0132042* (0.006085)	2.17
β3	distancekm – 35	0.0125682 (0.1059632)	0.12	(distancekm - 15)^2	-0.0070692 (0.0081848)	-0.86	(distancekm - 5)^3	-0.0133297* (0.006651)	-2
β4	-	-	-	(distancekm - 35)^2	-0.0035689 (0.0315404)	-0.11	(distancekm - 15)^3	0.000126 (0.0008955)	0.14
β5	-	-	-	-	-	-	(distancekm - 35)^3	-0.0009322 (0.0049687)	-0.19

No. of Observations	943	943	943
Significance	F = 8.42 p = 0.0000	F = 7.57 p = 0.0000	F = 6.07 p = 0.0000

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.649369	-
	β0	0.223448	0.250380619
<15	α0'	4.402119	-
	β0'	-0.127102	-0.119356158
<35	α0''	2.8606965	-
	β0''	-0.0243405	-0.024046659
>35	α0'''	2.4208095	-
	β0'''	-0.0117723	-0.011703278
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the cubic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	2925.66	2949.905
Quadratic Spline Regression	2923.601	2952.695
Cubic Spline Regression	**2926.942**	**2960.885**



.

PSIC 16: Wood & Products Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

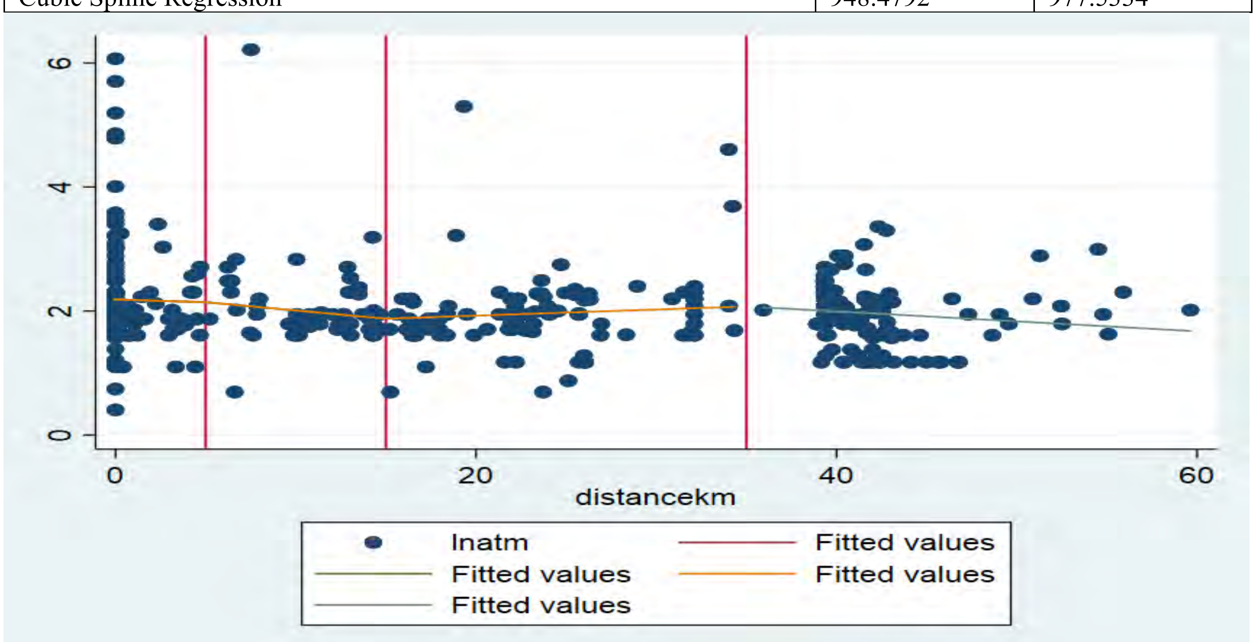
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	2.185898* (0.0564644)	38.71	Constant	2.195678* (0.0577229)	38.04	Constant	2.208265* (0.0578132)	38.2
β0	Distance km	-0.088163 (0.0308529)	-0.29	distancekm	-0.0374975 (0.0799273)	-0.47	distancekm	-0.2608818** (0.1508496)	1.72
β1	distancekm – 5	-0.0178058 (0.0466604)	-0.38	distancekm^2	0.0016114 (0.0112087)	0.14	distancekm^2	0.0804321** (0.0451493)	1.78
β2	distancekm – 15	0.0365272***(0.0225358)	1.62	(distancekm - 5)^2	-0.0003468 (0.0133128)	-0.03	distancekm^3	-0.0064068** (0.0034614)	-1.85
β3	distancekm – 35	-0.0259343*** (0.0175038)	-1.48	(distancekm - 15)^2	-0.0014326 (0.0029721)	-0.48	(distancekm - 5)^3	0.0071289** (0.0037349)	1.91
β4	-	-	-	(distancekm - 35)^2	0.0001926 (0.0013762)	0.14	(distancekm - 15)^3	-0.0009049* (0.0003631)	-2.49

β5	-	-	-	-	-	-	(distancekm - 35)^3	0.0004449* (0.0001488)	2.99
No. of Observations	469			469			469		
Significance	**F = 3.50 p = 0.0079**			**F = 2.30 p = 0.0438**			**F = 3.48 p = 0.0022**		

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.185898	-
	β0	-0.088163	-0.08438838
<15	α0'	2.274927	-
	β0'	-0.1059688	-0.100547289
<35	α0''	1.727019	-
	β0''	-0.0694416	-0.067085386
>35	α0'''	2.6347195	-
	β0'''	-0.0953759	-0.090968835
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	951.2795	972.0325
Quadratic Spline Regression	**955.7032**	**980.6068**
Cubic Spline Regression	948.4792	977.5334



.

PSIC 17: Paper & Paper Products Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

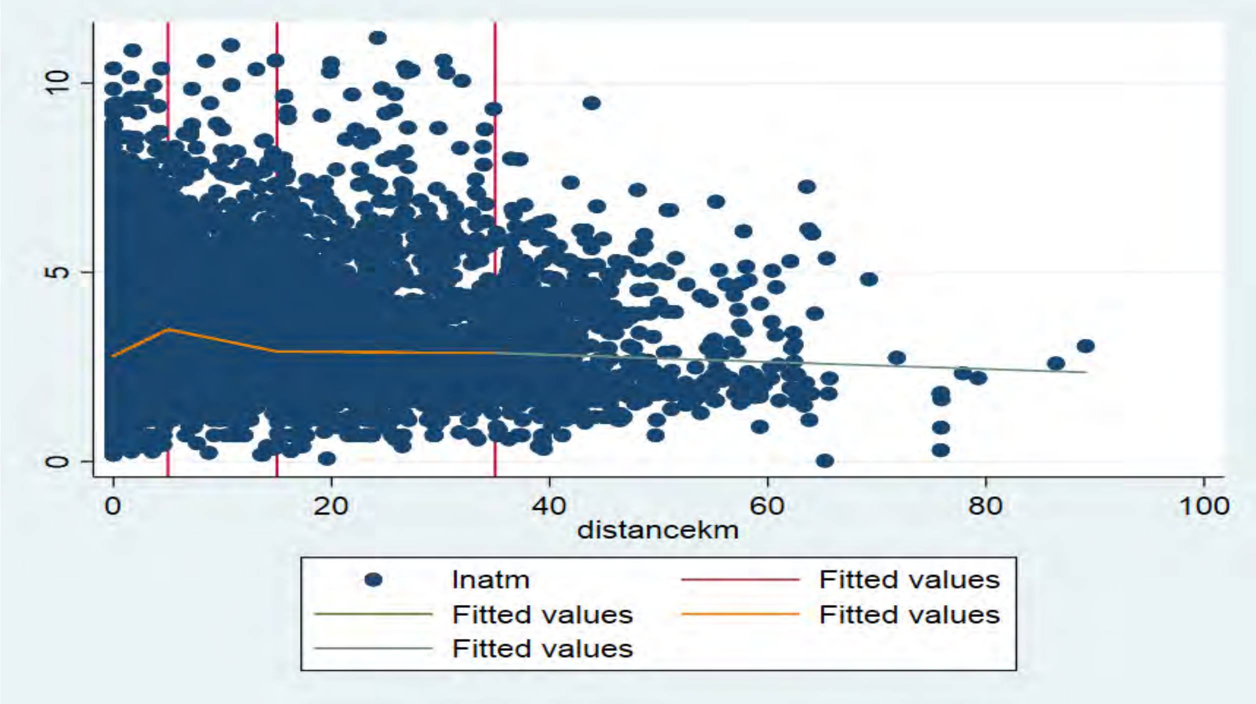
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	2.799476* (0.0590187)	47.43	Constant	2.781994* (0.600099)	46.36	Constant	2.786608* (0.0602943)	46.22
β0	distancekm	0.2478089* (0.0463391)	5.35	Distancekm	0.5514383* (0.1292123)	4.27	distancekm	0.4600761* (0.2268577)	2.03
β1	distancekm – 5	-0.4178352* (0.0835569)	-5	$distancekm^2$	-0.071328* (0.0210287)	-3.39	$distancekm^2$	-0.0572477* (0.0759837)	-0.75
β2	distancekm – 15	0.2714406* (0.0773873)	3.51	$(distancekm - 5)^2$	0.0780292* (0.0280789)	2.78	$distancekm^3$	0.0014117* (0.006211)	0.23
β3	distancekm – 35	-0.612765 (0.3376156)	-1.81	$(distancekm - 15)^2$	-0.00163437* (0.012457)	-0.13	$(distancekm - 5)^3$	0.0008252* (0.0070777)	0.12
β4	-	-	-	$(distancekm - 35)^2$	-0.1536703 (0.086899)	-1.77	$(distancekm - 15)^3$	-0.003446* (0.0016039)	-2.15
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	0.0017338 (0.0275332)	0.06

No. of Observations	531	531	531
Significance	**F = 8.10 p = 0.0000**	**F = 6.30 p = 0.0000**	**F = 5.80 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.799476	-
	β0	0.2478089	0.281215069
<15	α0'	4.888652	-
	β0'	-0.1700263	-0.156357371
<35	α0''	0.817043	-
	β0''	0.1014143	0.106735067
>35	α0'''	22.263818	-
	β0'''	-0.5113507	-0.400314963
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	1739.882	1761.225
Quadratic Spline Regression	**1742.671**	**1768.32**
Cubic Spline Regression	1741.464	1771.387



•

PSIC 20: Chemical & Chemical Products Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

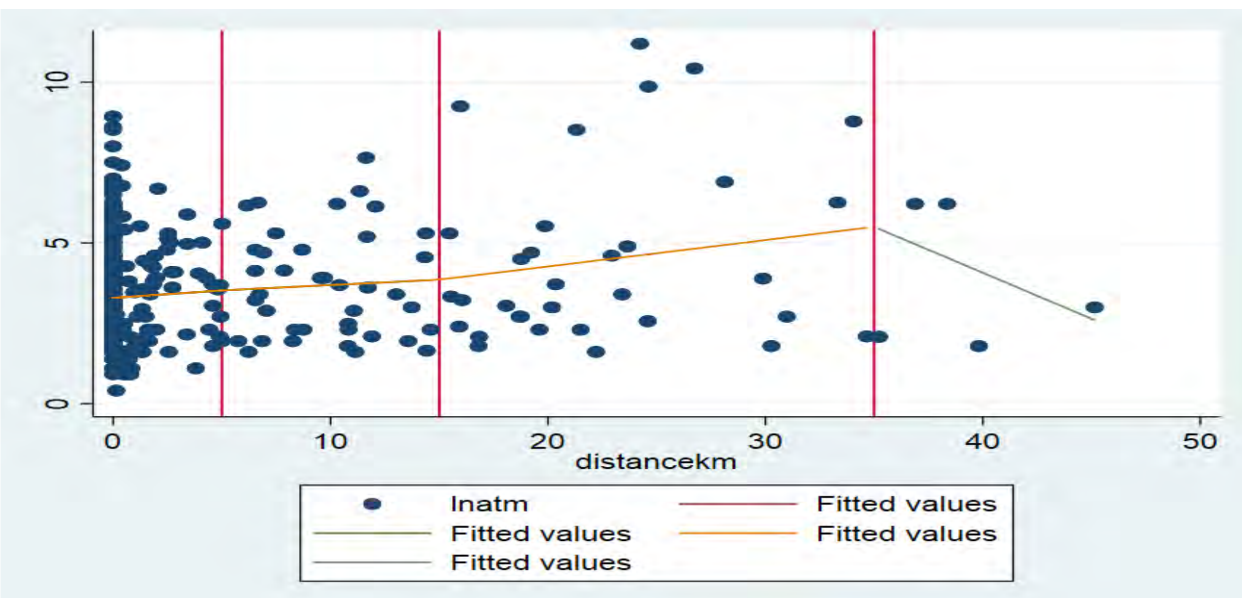
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	3.289924* (0.0930287)	35.36	Constant	3.292947* (0.0954111)	34.51	Constant	3.315771* (0.0962234)	34.46
β0	Distancekm	0.0463427 (0.06873)	0.67	distancekm	0.0597301 (0.17219)	0.35	distancekm	-0.2784446 (0.2912509)	-0.96
β1	distancekm – 5	-0.0128715 (0.1140579)	-0.11	$distancekm^2$	-0.0066729 (0.0260539)	-0.26	$distancekm^2$	0.1190191 (0.0944928)	1.26
β2	distancekm – 15	0.048717 (0.0831756)	0.59	$(distancekm - 5)^2$	0.0132564 (0.0329797)	0.4	$distancekm^3$	-0.0101668 (0.0075902)	-1.34
β3	distancekm – 35	-0.3681118** (0.1919312)	-1.92	$(distancekm - 15)^2$	-0.0106659 (0.0116924)	-0.91	$(distancekm - 5)^3$	0.0120309 (0.0084939)	1.42

β4	-	-	-	(distancekm - 35)^2	-0.0135516 (0.0268308)	-0.51	(distancekm - 15)^3	-0.0029132* (0.0014592)	-2
β5	-	-	-	-	-	-	(distancekm - 35)^3	0.0062255*** (0.0042177)	1.48
No. of Observations	473			473			473		
Significance	**F = 5.10 p = 0.0005**			**F = 4.12 p = 0.0011**			**F = 4.20 p = 0.0004**		

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	3.289924	-
	β0	0.0463427	0.047433305
<15	α0'	3.3542815	-
	β0'	0.0334712	0.034037663
<35	α0''	2.6235265	-
	β0''	0.0821882	0.085660112
>35	α0'''	15.5074395	-
	β0'''	-0.2859236	-0.248679985
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	1867.706	1888.501
Quadratic Spline Regression	**1869.446**	**1894.4**
Cubic Spline Regression	1868.08	1897.194



.

PSIC 22: Rubber & Plastic Product Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

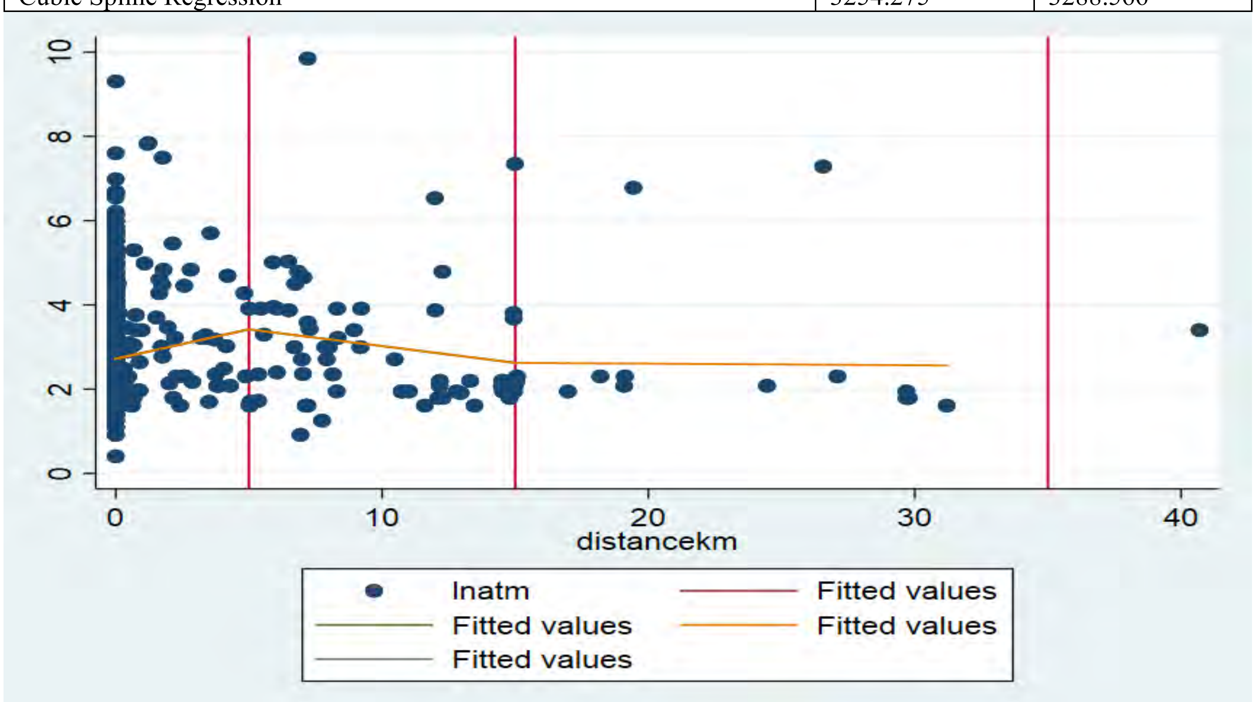
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.718691* (0.0423709)	64.16	Constant	2.706137* (0.0424985)	63.68	Constant	2.703023* (0.0425364)	63.55
β0	Distancekm	0.1401376* (0.0428352)	3.27	Distancekm	0.5079328* (0.1256519)	4.04	distancekm	0.6805168* (0.2291515)	2.97
β1	distancekm – 5	-0.2196331* (0.0744708)	-2.95	$distancekm^2$	-0.0739324* (0.0199989)	-3.7	$distancekm^2$	-0.1636825* (0.0754987)	-2.17
β2	distancekm – 15	0.0757215 (0.0639096)	1.18	$(distancekm - 5)^2$	0.0918237* (0.026863)	3.42	$distancekm^3$	0.0108367** (0.0061252)	1.77
β3	distancekm – 35	0.1535433 (0.2667668)	0.58	$(distancekm - 15)^2$	-0.0292526* (0.0143445)	-2.04	$(distancekm - 5)^3$	-0.009968 (0.0069817)	-1.43
β4	-	-	-	$(distancekm - 35)^2$	0.1475813** (0.0880905)	1.68	$(distancekm - 15)^3$	-0.0031525*** (0.0019875)	-1.59
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	0.1115555* (0.0474072)	2.35

No. of Observations	991	991	991
Significance	**F = 2.83 p = 0.0237**	**F = 3.86 p = 0.0018**	**F = 4.13 p = 0.0004**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.718691	-
	β0	0.1401376	0.150432087
<15	α0'	3.8168565	-
	β0'	-0.0794955	-0.076417824
<35	α0''	2.681034	-
	β0''	-0.003774	-0.003766887
>35	α0'''	-2.6929815	-
	β0'''	0.1497693	0.161566238
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**3263.591**	**3288.085**
Quadratic Spline Regression	3257.686	3287.078
Cubic Spline Regression	3254.275	3288.566



.

PSIC 23: Other Non-Metallic Minerals Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

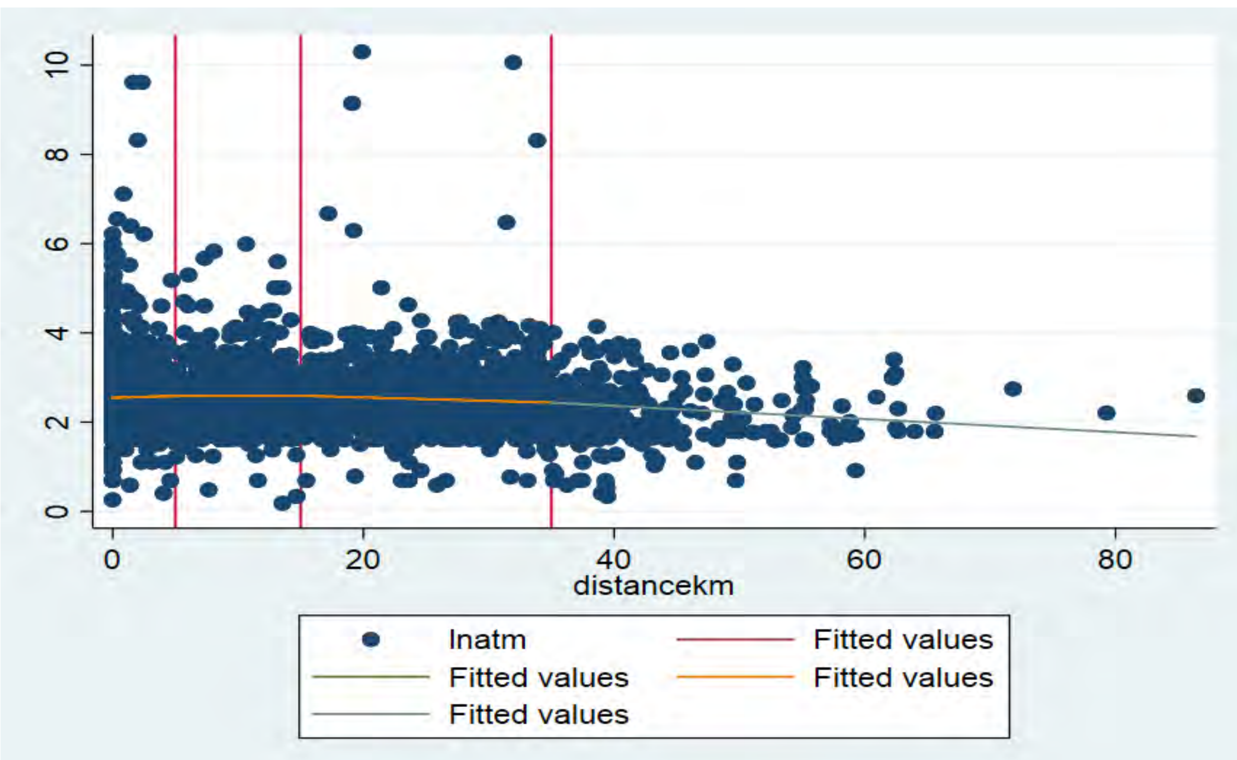
The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.546049* (0.0296309)	85.93	Constant	2.515287* (0.0311851)	80.66	Constant	2.493074* (0.0320485)	77.79
β0	distancekm	0.0077258 (0.0128166)	0.6	distancekm	0.0935919* (0.0334394)	2.8	distancekm	0.2389417* (0.0585556)	4.08
β1	distancekm - 5	-0.0071813 (0.0192013)	-0.37	$distancekm^2$	-0.0132674* (0.0046352)	-2.86	$distancekm^2$	-0.0689805* (0.0169885)	-4.06
β2	distancekm - 15	-0.0081816 (0.0103629)	-0.79	$(distancekm - 5)^2$	0.01601* (0.005505)	2.91	$distancekm^3$	0.0052101* (0.0012932)	4.03
β3	distancekm - 35	-0.0070808 (0.007674)	-0.92	$(distancekm - 15)^2$	-0.0038667* (0.0012588)	-3.07	$(distancekm - 5)^3$	-0.0055731* (0.0013902)	-4.01
No. of Observations	2,750			2,750			2,750		
Significance	**F = 6.91 p = 0.0000**			**F = 8.41 p = 0.0000**			**F = 8.10 p = 0.0000**		

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.546049	-
	β0	0.0077258	0.007755721
<15	α0'	2.5819555	-
	β0'	0.0005445	0.000544648
<35	α0''	2.7046795	-
	β0''	-0.0076371	-0.007608011
>35	α0'''	2.9525075	-
	β0'''	-0.0147179	-0.014610121
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**6950.916**	**6980.512**
Quadratic Spline Regression	6938.616	6974.133
Cubic Spline Regression	6934.173	6975.608



.

PSIC 25: Fabricated Metal Products Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

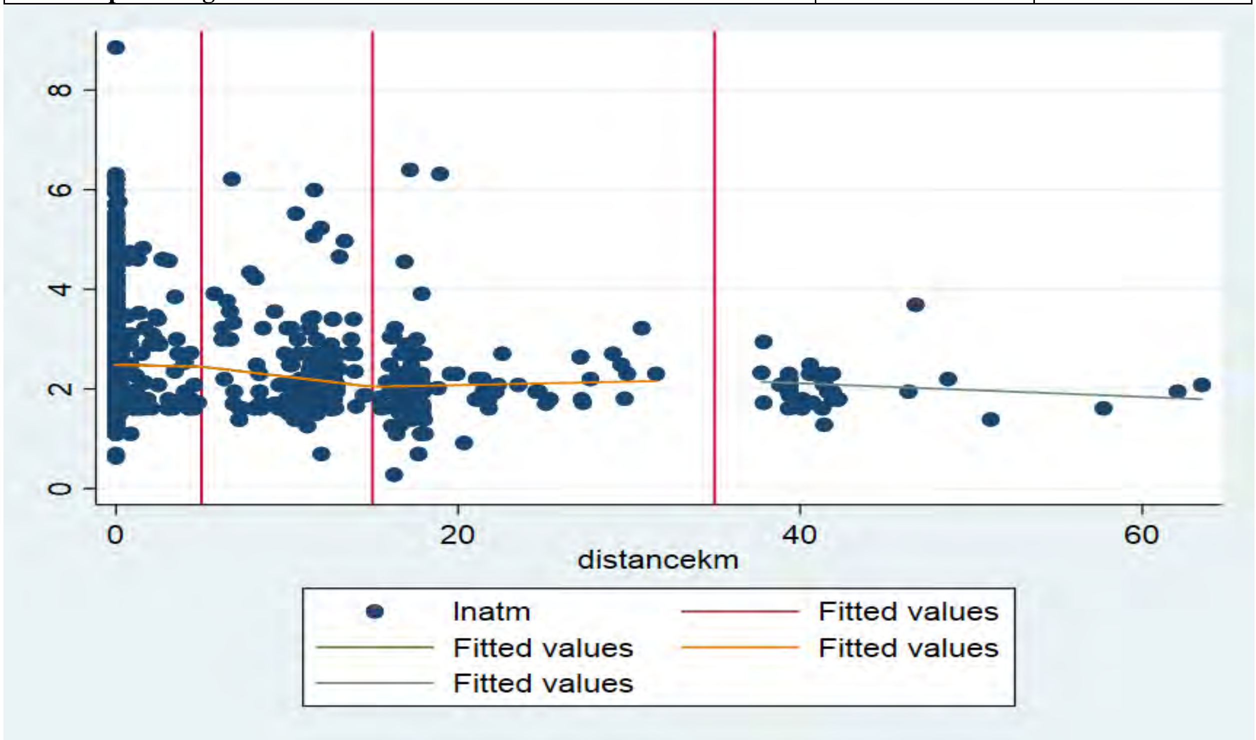
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.484099* (0.0267905)	92.72	Constant	2.482535* (0.0269365)	92.16	Constant	2.482597* (0.0270073)	91.92
β0	distancekm	-0.0072712 (0.0286541)	-0.25	distancekm	0.0632596 (0.0790896)	0.8	distancekm	0.0776743 (0.1366261)	0.57
β1	distancekm - 5	-0.0332707 (0.0472038)	-0.7	$distancekm^2$	-0.0134999 (0.0110765)	-1.22	$distancekm^2$	-0.0271086 (0.0403832)	-0.67
β2	distancekm - 15	0.0475015** (0.0265217)	1.79	$(distancekm - 5)^2$	0.0172048 (0.013112)	1.31	$distancekm^3$	0.0018992 (0.0030968)	0.61
β3	distancekm - 35	-0.0205513 (0.0302896)	-0.68	$(distancekm - 15)^2$	-0.0038619 (0.0030837)	-1.25	$(distancekm - 5)^3$	-0.001863 (0.0033567)	-0.56
β4	-	-	-	$(distancekm - 35)^2$	0.000049 (0.0018573_	0.03	$(distancekm - 15)^3$	-0.0001152 (0.0003903)	-0.3
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	0.0001628 (0.0002102)	0.77

No. of Observations	1,612	1,612	1,612
Significance	**F = 11.13 p = 0.0000**	**F = 8.99 p = 0.0000**	**F = 7.49 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.484099	-
	β0	-0.0072712	-0.007244829
<15	α0'	2.6504525	-
	β0'	-0.0405419	-0.039731072
<35	α0''	1.93793	-
	β0''	0.0069596	0.006983874
>35	α0'''	2.6572255	-
	β0'''	-0.0135917	-0.01349975
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the cubic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	4282.974	4309.9
Quadratic Spline Regression	4284.548	4316.859
Cubic Spline Regression	**4286.491**	**4324.187**



.

PSIC 27: Electrical Equipment Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

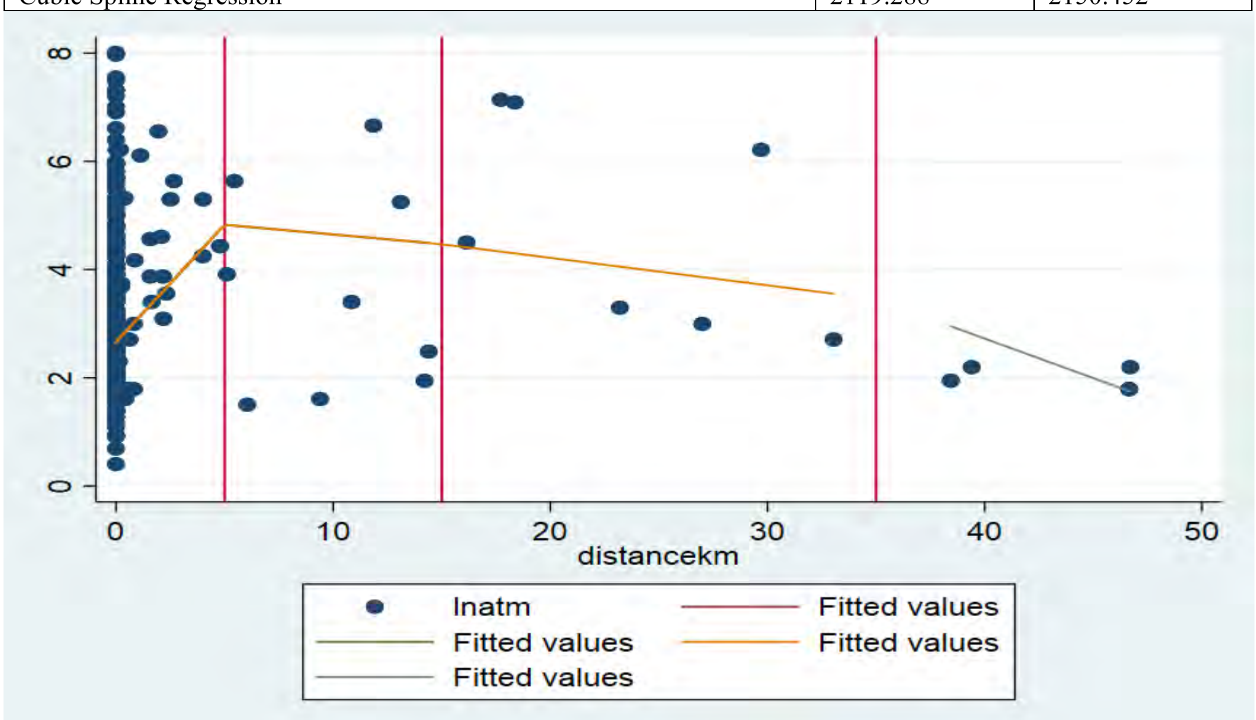
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.657698* (0.0528149)	50.32	Constant	2.643757* (0.0525336)	50.33	Constant	2.638547* (0.0526209)	50.14
β0	distancekm	0.4339841* (0.0972852)	4.46	Distancekm	1.095975* (0.215546)	5.08	distancekm	1.794908* (0.3829445)	4.69
β1	distancekm - 5	-0.4697406* (0.1632531)	-2.88	distancekm^2	-0.1432694* (0.0341027)	-4.2	distancekm^2	-0.5040061* (0.1363327)	-3.7
β2	distancekm - 15	-0.0150158 (0.1170498)	-0.13	(distancekm - 5)^2	0.1733807* (0.0445923)	3.89	distancekm^3	0.0386024* (0.0112551)	3.43
β3	distancekm - 35	-0.095204 (0.1426076)	-0.67	(distancekm - 15)^2	-0.0466624* (0.0176898)	-2.64	(distancekm - 5)^3	-0.0418078* (0.0127313)	-3.28
β4	-	-	-	(distancekm - 35)^2	0.0437817* (0.0212846)	2.06	(distancekm - 15)^3	0.0035691** (0.0021252)	1.68
β5	-	-	-	-	-	-	(distancekm - 35)^3	0.0005994(0.0031876)	0.19

No. of Observations	634	634	634
Significance	**F = 10.32 p = 0.0000**	**F = 10.96 p = 0.0000**	**F = 9.37 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.657698	-
	β0	0.4339841	0.543394328
<15	α0'	5.006401	-
	β0'	-0.0357565	-0.035124788
<35	α0''	5.231638	-
	β0''	-0.0507723	-0.049504926
>35	α0'''	8.563778	-
	β0'''	-0.1459763	-0.135821816
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**2129.441**	**2151.702**
Quadratic Spline Regression	2118.693	2145.405
Cubic Spline Regression	2119.288	2150.452



•

PSIC 28: Machinery & Equipment Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

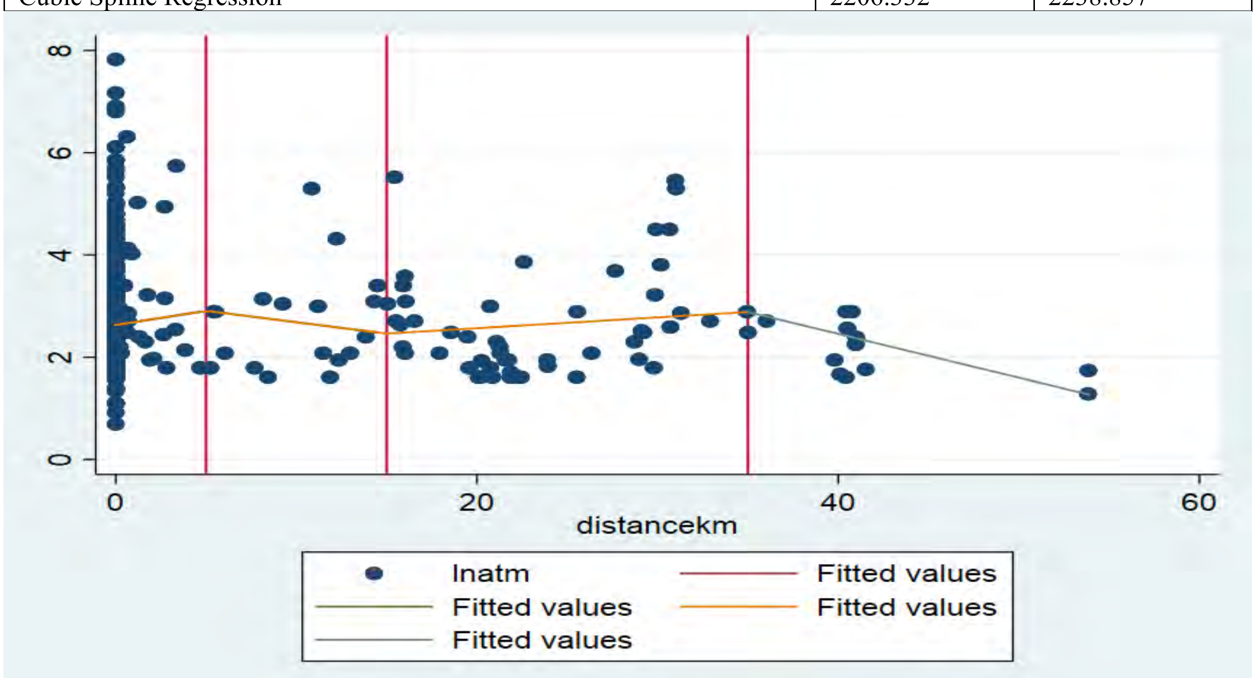
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	2.628959* (0.0389817)	67.44	Constant	2.62437* (0.0391539)	67.03	Constant	2.622226* (0.0391432)	66.99
β0	distancekm	0.0554517 (0.0644814)	0.86	Distancekm	0.2156691 (0.1546607)	1.39	distancekm	0.3265028 (0.2714161)	1.20
β1	distancekm – 5	-0.0999338 (0.1018514)	-0.98	$distancekm^2$	-0.0292106 (0.0226104)	-1.29	$distancekm^2$	-0.0651283 (0.0869302)	-0.75
β2	distancekm – 15	0.0655954 (0.0536522)	1.22	$(distancekm - 5)^2$	0.0327747 (0.02751)	1.19	$distancekm^3$	0.0037483 (0.0068492)	0.55
β3	distancekm – 35	-0.1067025* (0.0524713)	-2.03	$(distancekm - 15)^2$	-0.003004 (0.0073007)	-0.41	$(distancekm - 5)^3$	-0.003126 (0.0075167)	-0.42
β4	-	-	-	$(distancekm - 35)^2$	-0.0052454 (0.0049266)	-1.06	$(distancekm - 15)^3$	-0.0009625 (0.0009183)	-1.05
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	0.0008471*** (0.0005707)	1.48

No. of Observations	770	770	770
Significance	**F = 1.39 p = 0.2356**	**F = 1.10 p = 0.3564**	**F = 1.68 p = 0.1228**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.628959	-
	β0	0.0554517	0.057017962
<15	α0'	3.128628	-
	β0'	-0.0444821	-0.043507279
<35	α0''	2.144697	-
	β0''	0.0211133	0.021337763
>35	α0'''	5.8792845	-
	β0'''	-0.0855892	-0.082028744
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the quadratic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	2206.866	2230.098
Quadratic Spline Regression	**2208.896**	**2236.774**
Cubic Spline Regression	2206.332	2238.857



.

PSIC 29: Motor Vehicles & Trailers Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

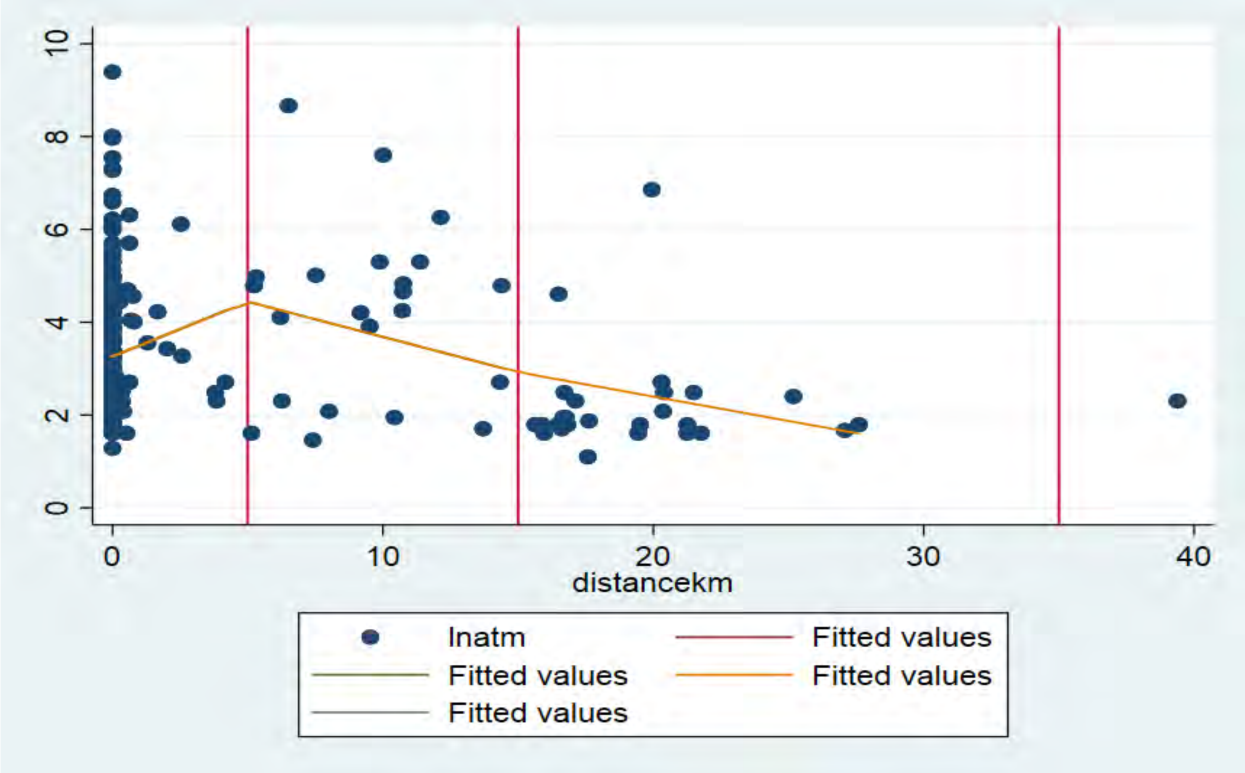
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**
α0	Constant	3.259047* (0.0846219)	38.51	Constant	3.262234* (0.0849155)	38.42	Constant	3.259644* (0.0857307)	38.02
β0	distancekm	0.2368278* (0.0880991)	2.69	Distancekm	0.1280432 (0.2651656)	0.48	distancekm	0.0908915 (0.4864622)	0.19
β1	distancekm - 5	-0.3889143* (0.1448518)	-2.68	distancekm^2	0.0037063 (0.0400317)	0.09	distancekm^2	0.0538691 (0.1557567)	0.35
β2	distancekm - 15	0.0470802 (0.1225132)	0.38	(distancekm - 5)^2	-0.0287198 (0.050444)	-0.57	distancekm^3	-0.0065933 (0.0124756)	-0.53
β3	distancekm - 35	0.4424907 (0.4816795)	0.92	(distancekm - 15)^2	0.046434* (0.0216899)	2.14	(distancekm - 5)^3	0.0084656 (0.0140996)	0.6
β4	-	-	-	(distancekm - 35)^2	-0.2810347 (0.2633067)	-1.07	(distancekm - 15)^3	-0.00197474 (0.0039332)	-0.5
β5	-	-	-	-	-	-	(distancekm - 35)^3	-0.0097459 (0.2098558)	-0.05

No. of Observations	337	337	337
Significance	**F = 4.32 p = 0.0020**	**F = 4.37 p = 0.0007**	**F = 3.39 p = 0.0029**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	3.259047	-
	β0	0.2368278	0.267222883
<15	α0'	5.2036185	-
	β0'	-0.1520865	-0.141086019
<35	α0''	4.4974155	-
	β0''	-0.1050063	-0.099681149
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the cubic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	1193.896	1212.997
Quadratic Spline Regression	1191.48	1214.4
Cubic Spline Regression	**1194.844**	**1221.585**



.

PSIC 31: Furniture Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

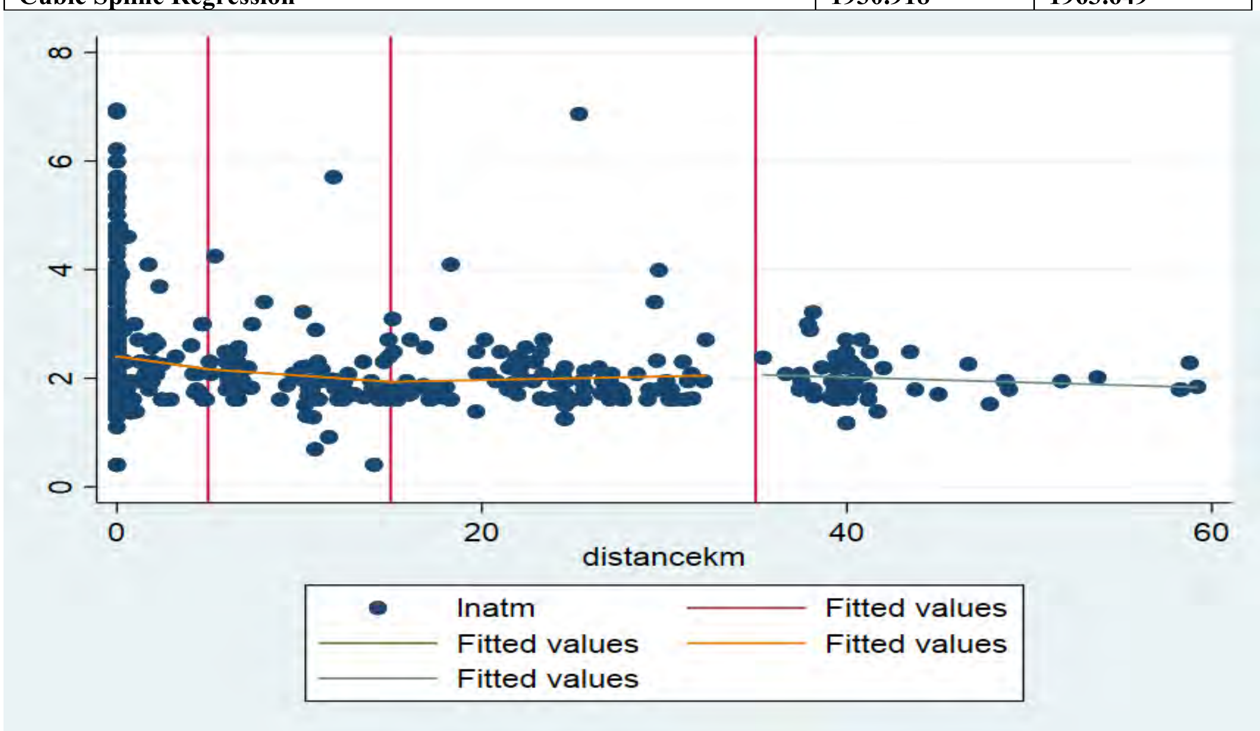
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	2.402285* (0.0364508)	65.9	Constant	2.402322* (0.0368335)	65.22	Constant	2.403396* (0.0370499)	64.87
β0	distancekm	-0.0473442* (0.0284242)	-1.67	distancekm	-0.030916 (0.0822754)	-0.38	distancekm	-0.0498371 (0.1508253)	-0.33
β1	distancekm – 5	0.024087 (0.0453946)	0.53	$distancekm^2$	-0.0025129 (0.0118586)	-0.21	$distancekm^2$	-0.0005631 (0.0455451)	-0.01
β2	distancekm – 15	0.0298731 (0.0255398)	1.17	$(distancekm - 5)^2$	0.0058417 (0.0142285)	0.41	$distancekm^3$	0.0001984 (0.0035033)	0.06
β3	distancekm – 35	-0.0162301 (0.022361)	-0.73	$(distancekm - 15)^2$	0.0037071 (0.0032869)	-1.13	$(distancekm - 5)^3$	-0.0002356 (0.0037916)	-0.06
β4	-	-	-	$(distancekm - 35)^2$	0.000367 (0.0016716)	0.22	$(distancekm - 15)^3$	-7.50e-06 (0.0003975)	-0.02
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	0.0001051 (0.0001861)	0.56

No. of Observations	793	793	793
Significance	**F = 10.45 p = 0.0000**	**F = 8.43 p = 0.0000**	**F = 6.98 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient*
<5	α0	2.402285	-
	β0	-0.0473442	-0.046240943
<15	α0'	2.28185	-
	β0'	-0.0232572	-0.022988836
<35	α0''	1.8337535	-
	β0''	0.0066159	0.006637833
>35	α0'''	2.401807	-
	β0'''	-0.0096142	-0.009568131
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the cubic spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	1927.112	1950.
Quadratic Spline Regression	1928.739	1956.794
Cubic Spline Regression	**1930.918**	**1963.649**



.

PSIC 32: Other Manufacturing Sector

Linear spline regression model that is being used for the estimation is given below:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1 D_{15}(distancekm - 5) + \beta_2 D_{215}(distancekm - 15) + \beta_3 D_{335}(distancekm - 35) + e$$

For distancekm ≤ 5

$$lnatm = \alpha_0 + \beta_0(distancekm) + e$$

For distancekm ≤ 15

$$lnatm = (\alpha_0 - 5\beta_1) + (\beta_0 + \beta_1)distancekm + e$$

For distancekm ≤ 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2) + (\beta_0 + \beta_1 + \beta_2)distancekm + e$$

For distancekm > 35

$$lnatm = (\alpha_0 - 5\beta_1 - 15\beta_2 - 35\beta_3) + (\beta_0 + \beta_1 + \beta_2 + \beta_3)distancekm + e$$

For quadratic and cubic spline regression models we estimate the following equations:

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2 D_{15}(distancekm - 5)^2 + \beta_3 D_{215}(distancekm - 15)^2 + \beta_4 D_{335}(distancekm - 35)^2 + e$$

The cubic spline regression model that will be used for our estimation

$$lnatm = \alpha_0 + \beta_0(distancekm) + \beta_1(distancekm)^2 + \beta_2(distancekm)^3 + \beta_3 D_{15}(distancekm - 5)^3 + \beta_4 D_{215}(distancekm - 15)^3 + \beta_5 D_{335}(distancekm - 35)^3 + e$$

The estimated linear, quadratic and cubic spline regressions along with the adjusted coefficients of linear spline regression are given below.

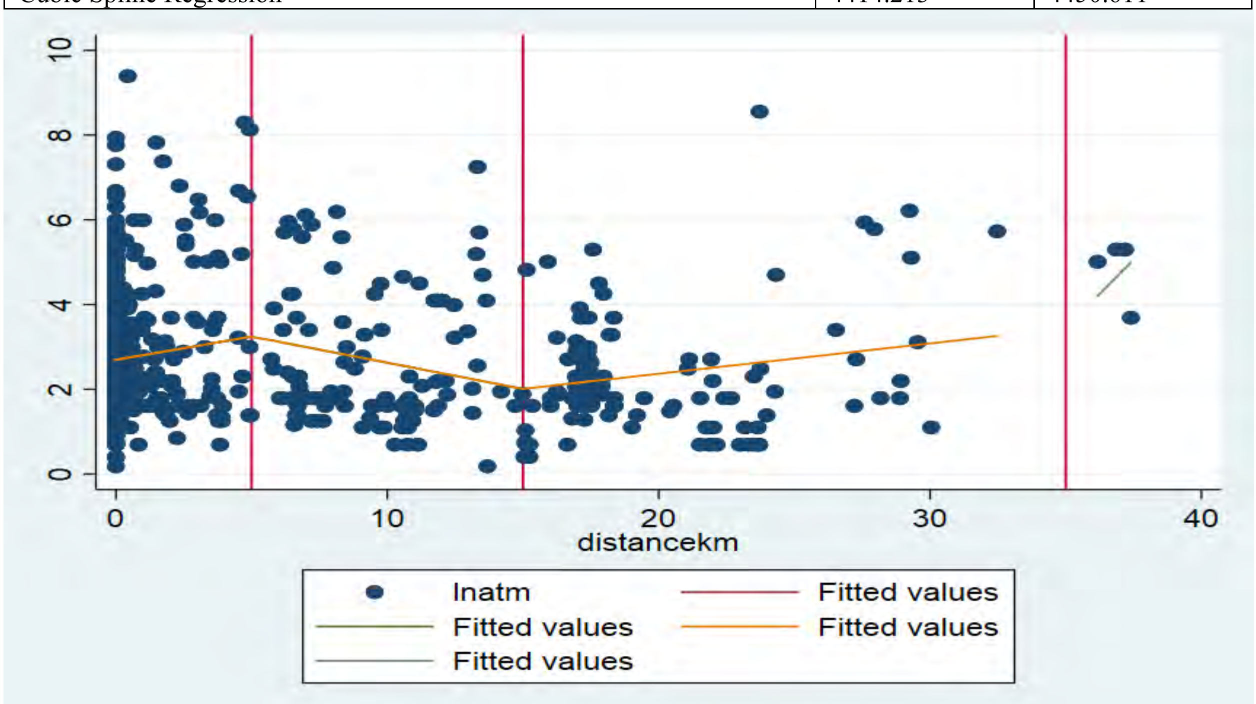
	Linear Spline Regression			**Quadratic Spline Regression**			**Cubic Spline Regression**		
Coefficients	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**t**	**Variables**	**Coefficients (Std. Error)**	**T**
α0	Constant	2.69434* (0.0393736)	68.43	Constant	2.686174* (0.0397426)	67.59	Constant	2.684032* (0.0399763)	67.14
β0	distancekm	0.1100678* (0.0325913)	3.38	distancekm	0.2763927* (0.0882937)	3.13	distancekm	0.4069699* (0.1532128)	2.66
β1	distancekm - 5	-0.2336024* (0.0533941)	-4.38	$distancekm^2$	-0.0383864* (0.0133645)	-2.87	$distancekm^2$	-0.1045545* (0.0486096)	-2.15
β2	distancekm - 15	0.1950903* (0.0433496)	4.5	$(distancekm - 5)^2$	0.0402934* (0.0169226)	2.38	$distancekm^3$	0.0072167** (0.003857)	1.87
β3	distancekm - 35	0.5751708*** (0.397776)	1.45	$(distancekm - 15)^2$	0.0089624 (0.006739)	1.33	$(distancekm - 5)^3$	-0.0073108** (0.0042943)	-1.7
β4	-	-	-	$(distancekm - 35)^2$	-0.2932623 (0.2514357)	-1.17	$(distancekm - 15)^3$	0.000627 (0.000815)	0.77
β5	-	-	-	-	-	-	$(distancekm - 35)^3$	-0.2628025** (0.1417025)	-1.85

No. of Observations	1,339	1,339	1,339
Significance	**F = 9.69 p = 0.0000**	**F = 9.66 p = 0.0000**	**F = 8.20 p = 0.0000**

Distances	Coefficients	Coefficient Value	Adjusted Coefficient
<5	α0	2.69434	-
	β0	0.1100678	0.116353757
<15	α0'	3.862352	-
	β0'	-0.1235346	-0.116208938
<35	α0''	0.9359975	-
	β0''	0.0715557	0.074177981
>35	α0'''	-19.1949805	-
	β0'''	0.6467265	0.909280558
***Coefficients have been adjusted ($e^{\beta 0}$-1) for the proper interpretation of the log-linear model**			

The most efficient of the three models is the linear spline regression model i.e., the model that best fits our data.

Models/Model Criteria	AIC	BIC
Linear Spline Regression	**4420.433**	**4446.432**
Quadratic Spline Regression	4413.144	4444.342
Cubic Spline Regression	4414.213	4450.611



.

Annexure IV: Knots Based on Maximum Likelihood Estimation

Introduction:

In this code, we've implemented a Spline Regression using Maximum Likelihood Estimation (MLE) to strategically position knots, optimizing the model's predictive accuracy. Spline Regression, a potent technique, segments the data and fits polynomial functions to each segment, empowering the model to discern intricate patterns.

Workflow:

Data Loading:

We initiate by loading economic data from an Excel file, where 'Distance from Urban Centre' serves as the independent variable (x), and 'Annual Turnover' as the dependent variable (y).

Knot Selection:

Defining the number of knots to test and a range for grid search, we iterate through knot combinations to minimize the Root Mean Squared Error (RMSE), signifying optimal model performance.

Model Evaluation:

Our “evaluate_model_performance” function utilizes Generalized Linear Models (GLM) with B-spline basis functions to calculate RMSE. Maximum Likelihood Estimation (MLE) ensures the identification of the knot configuration that maximizes the likelihood of our observed data.

Final Model Fitting:

Leveraging the best knot configuration, we fit the final Spline Regression model. MLE finetunes model parameters specifically for our dataset, enhancing the model's adaptability.

Plotting:

The code generates a visual representation of the data and the optimized Spline Regression line, with knots marked on the plot, illustrating their impact on the regression.

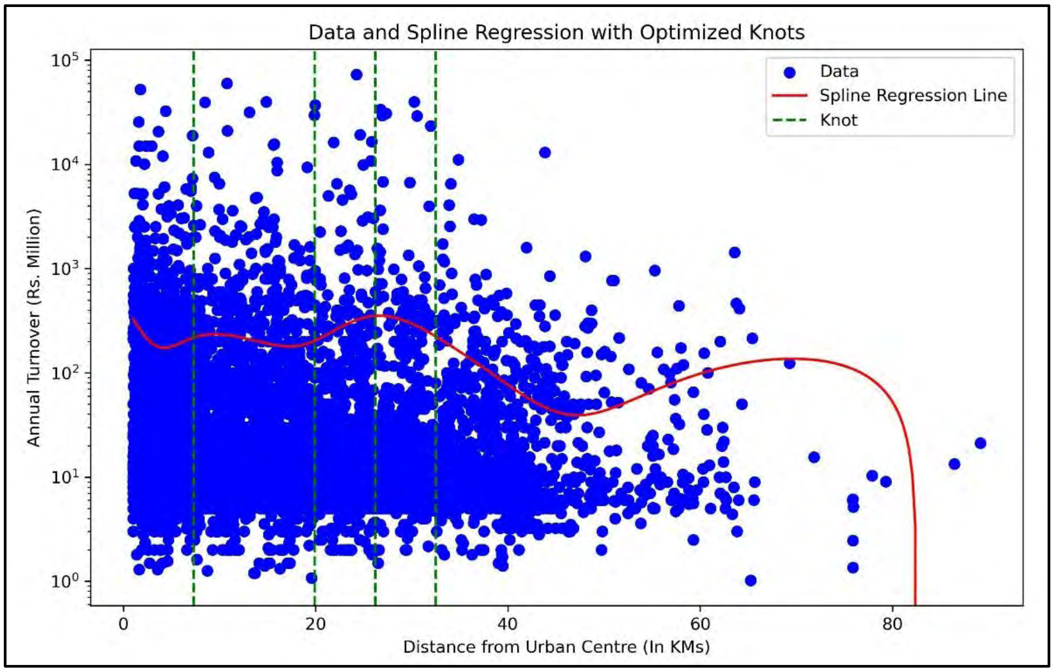
Best Knots:

The optimal knot configuration that minimizes the RMSE is as follows:

Best Knots: (7.297785009656634, 19.88961090360369, 26.18552385057722, 32.48143679755075)

Visual Representation:

Below is an image depicting the data, the Spline Regression line, and the marked knots. This visualization provides a clear understanding of how the knots are strategically placed along the regression line.



Significance:

Optimized Knots:

The code dynamically identifies the most effective knot configuration, adapting to the unique characteristics of our dataset.

Maximum Likelihood Estimation:

MLE maximizes the likelihood of our observed data under the model, ensuring that our model parameters are precisely tuned for the dataset. This enhances adaptability and generalization, making the model robust across different scenarios.

Benefits of MLE:

MLE not only optimizes the model for our data but also provides a statistical framework for estimating parameters. It maximizes the likelihood of observing the given data under our model assumptions, offering a principled approach to parameter estimation.

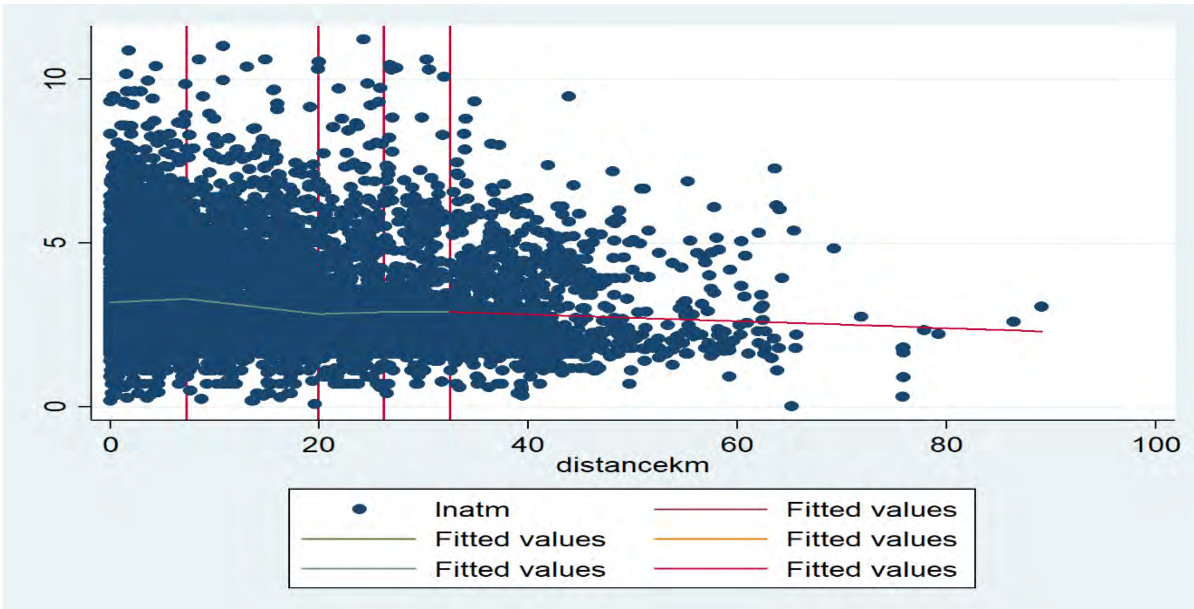
Visual Interpretation:

The resulting plot offers a clear interpretation of the Spline Regression, revealing the nuanced relationship between distance from urban centres and annual turnover.

This code represents a collective effort in deploying Spline Regression with MLE for knot optimization. MLE's benefits extend beyond mere optimization, providing a statistically sound foundation for parameter estimation. The methodology enhances our model's ability to capture subtle patterns in economic data, offering valuable insights for decision making. The visual representation not only aids in understanding complex relationships but also serves as a powerful tool for effective communication of our model's findings to stakeholders.

Results Based on New Knots

Overall Manufacturing: All Industries

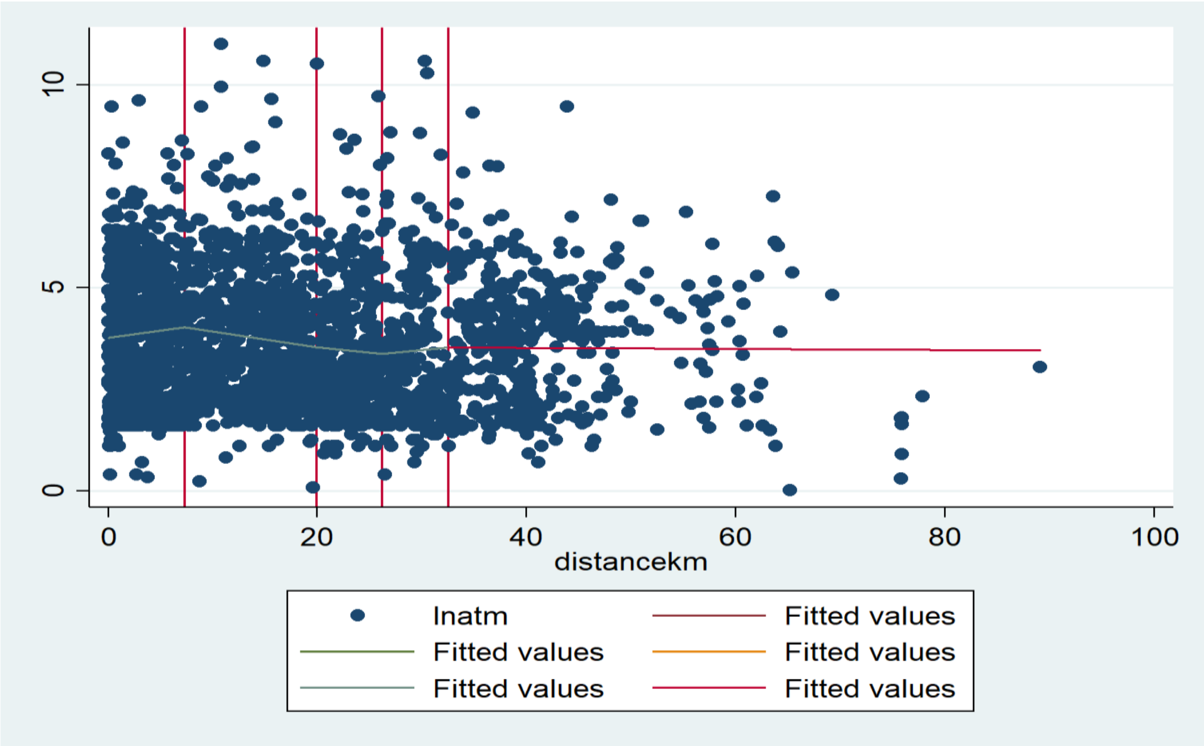



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4

      Source |       SS           df       MS      Number of obs   =     9,391
-------------+----------------------------------   F(5, 9385)      =     27.09
       Model |   297.40932         5  59.4818641   Prob > F        =    0.0000
    Residual |  20609.1202     9,385  2.19596379   R-squared       =    0.0142
-------------+----------------------------------   Adj R-squared   =    0.0137
       Total |  20906.5295     9,390  2.22646747   Root MSE        =    1.4819

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .0150771   .0085297     1.77   0.077    -.0016431    .0317972
      dist_1 |  -.0519733   .0130131    -3.99   0.000    -.0774819   -.0264648
      dist_2 |   .0470291   .0180305     2.61   0.009     .0116854    .0823729
      dist_3 |  -.0099169   .0259211    -0.38   0.702    -.0607279    .0408941
      dist_4 |  -.0106013   .0180179    -0.59   0.556    -.0459202    .0247176
       _cons |   3.184119   .0371783    85.64   0.000     3.111241    3.256996
------------------------------------------------------------------------------
```

PSIC 10: Food Sector

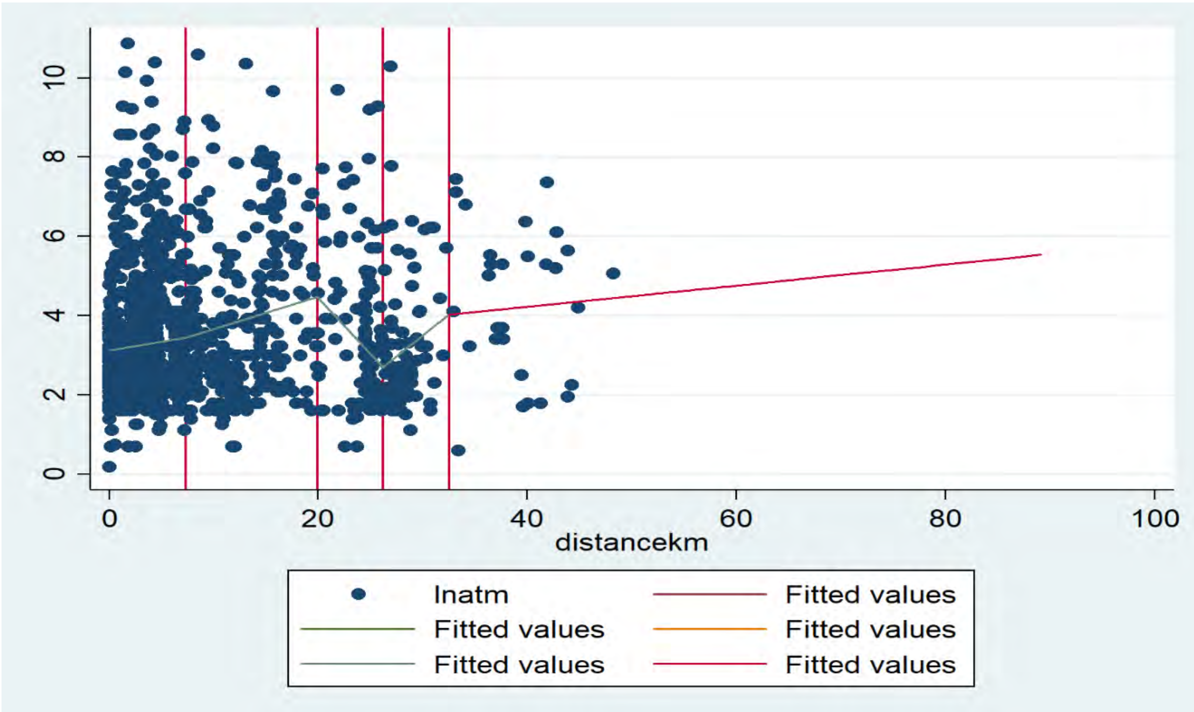



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==10

      Source |       SS           df       MS      Number of obs   =     2,491
-------------+----------------------------------   F(5, 2485)      =      7.67
       Model |  95.8019124         5  19.1603825   Prob > F        =    0.0000
    Residual |  6205.22912     2,485  2.49707409   R-squared       =    0.0152
-------------+----------------------------------   Adj R-squared   =    0.0132
       Total |  6301.03103     2,490  2.53053455   Root MSE        =    1.5802

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .0353871   .0204826     1.73   0.084    -.0047777    .0755518
      dist_1 |  -.0744046   .0297307    -2.50   0.012    -.1327042    -.016105
      dist_2 |    .013041    .034922     0.37   0.709    -.0554381    .0815201
      dist_3 |   .0508517   .0483989     1.05   0.294    -.0440546     .145758
      dist_4 |  -.0261335   .0312535    -0.84   0.403    -.0874191    .0351521
       _cons |   3.772373   .0932677    40.45   0.000     3.589483    3.955264
------------------------------------------------------------------------------
```

PSIC 13: Textile Sector

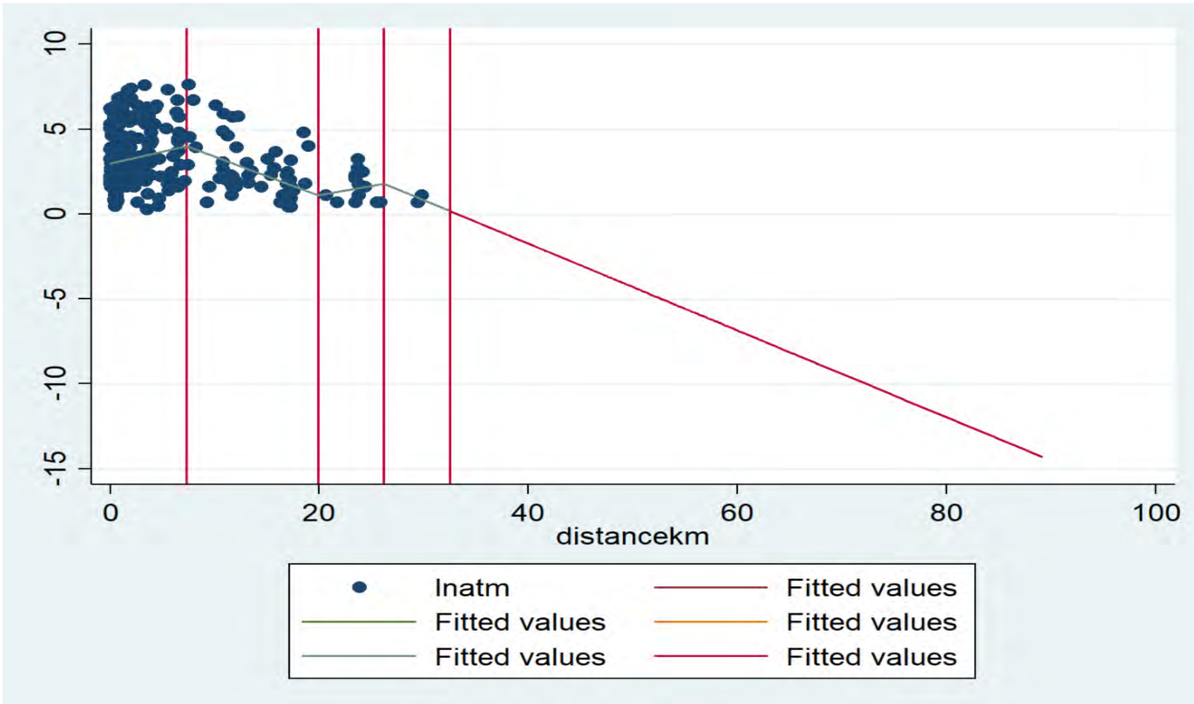



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==13

      Source |       SS           df       MS      Number of obs   =     1,672
-------------+----------------------------------   F(5, 1666)      =     15.20
       Model |   181.74127         5  36.348254   Prob > F        =    0.0000
    Residual |  3983.68898     1,666  2.39116986   R-squared       =    0.0436
-------------+----------------------------------   Adj R-squared   =    0.0408
       Total |  4165.43025     1,671  2.49277693   Root MSE        =    1.5463

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .0437131   .0199311     2.19   0.028     .0046205    .0828057
      dist_1 |     .03822   .0322337     1.19   0.236    -.0250028    .1014427
      dist_2 |  -.3637507   .0530039    -6.86   0.000    -.4677118   -.2597895
      dist_3 |   .4923434   .0929871     5.29   0.000     .3099596    .6747272
      dist_4 |   -.183879   .1120374    -1.64   0.101    -.4036279      .03587
       _cons |   3.122957   .0836147    37.35   0.000     2.958956    3.286958
------------------------------------------------------------------------------
```

PSIC 14: Wearing Apparel Sector

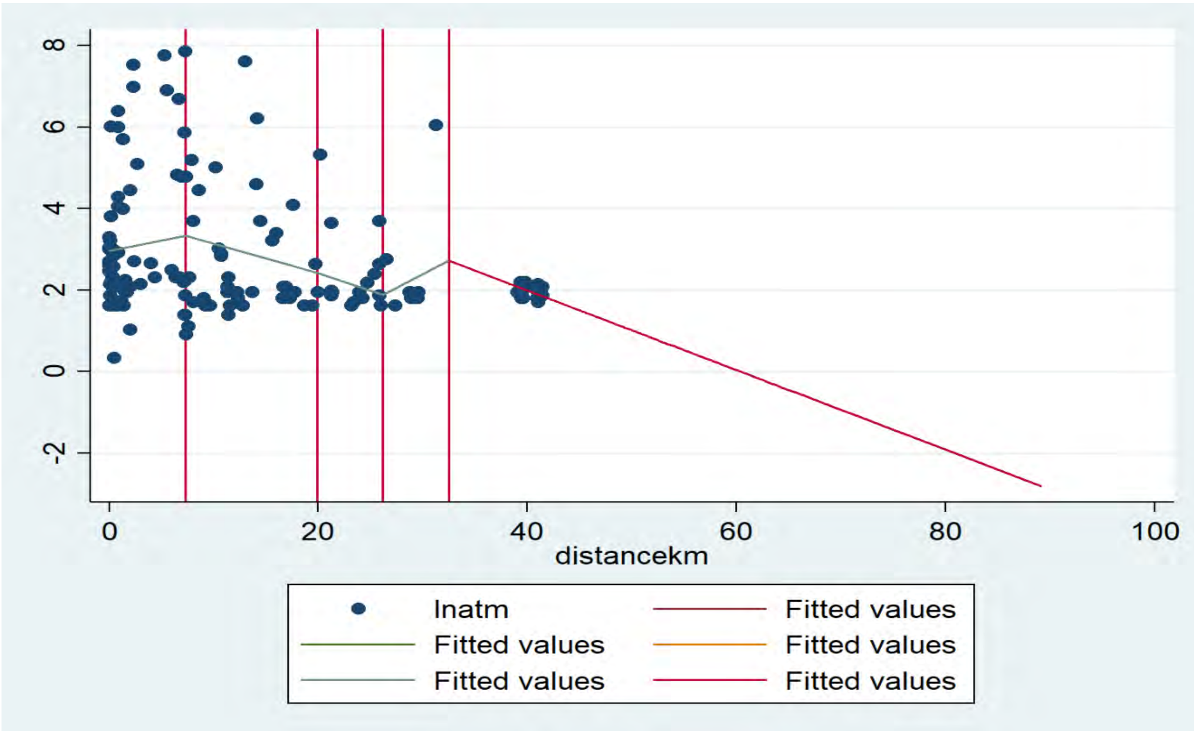



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==14
note: dist_4 omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =       291
-------------+----------------------------------   F(4, 286)       =     10.77
       Model |  113.190815         4  28.2977037   Prob > F        =    0.0000
    Residual |  751.493282       286  2.62759889   R-squared       =    0.1309
-------------+----------------------------------   Adj R-squared   =    0.1187
       Total |  864.684097       290   2.9816693   Root MSE        =     1.621

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .1401009    .047235     2.97   0.003     .0471285    .2330733
      dist_1 |  -.3701377   .0833949    -4.44   0.000    -.5342834   -.205992
      dist_2 |   .3391759   .1708724     1.98   0.048     .0028489    .6755029
      dist_3 |  -.3646815   .4755913    -0.77   0.444    -1.300785    .5714217
      dist_4 |          0  (omitted)
       _cons |   2.977396   .1490174    19.98   0.000     2.684086    3.270706
------------------------------------------------------------------------------
```

PSIC 15: Leather Products Sector



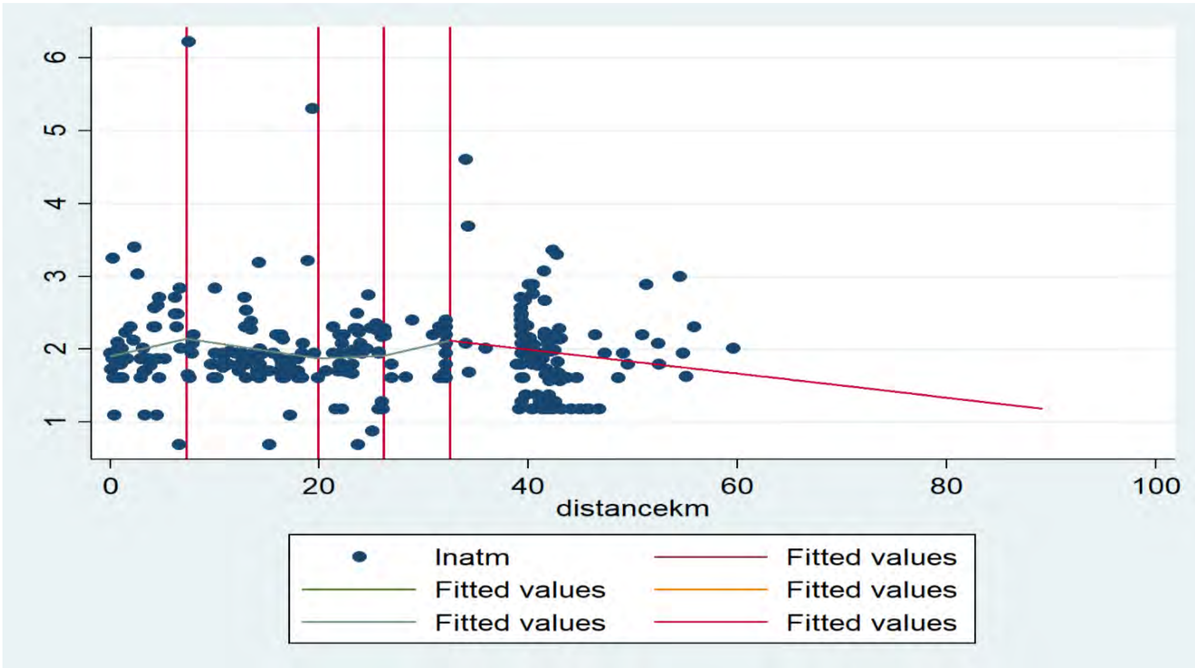

```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==15
```

Source	SS	df	MS
Model	39.5677029	5	7.91354057
Residual	337.859785	157	2.15197315
Total	377.427488	162	2.32979931

Number of obs	=	163
F(5, 157)	=	3.68
Prob > F	=	0.0036
R-squared	=	0.1048
Adj R-squared	=	0.0763
Root MSE	=	1.467

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	.0512956	.0535977	0.96	0.340	-.0545699	.1571611
dist_1	-.123785	.0853262	-1.45	0.149	-.2923204	.0447504
dist_2	-.0132601	.1385425	-0.10	0.924	-.2869079	.2603876
dist_3	.2193371	.2698291	0.81	0.418	-.3136264	.7523005
dist_4	-.2312311	.2970925	-0.78	0.438	-.8180451	.3555829
_cons	2.961446	.2422799	12.22	0.000	2.482897	3.439994

PSIC 16: Wood & Products Sector



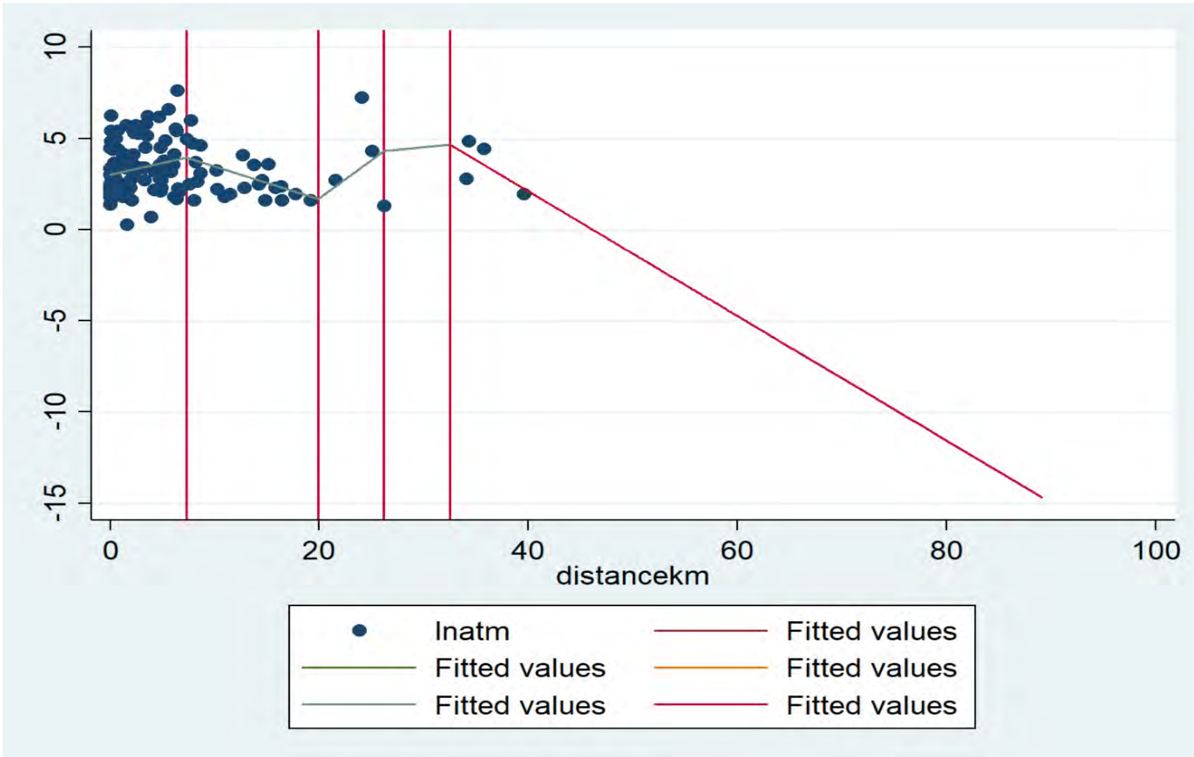
. **regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==16**

Source	SS	df	MS
Model	1.91719447	5	.383438893
Residual	106.033612	339	.312783517
Total	107.950807	344	.313810484

Number of obs	=	345
F(5, 339)	=	1.23
Prob > F	=	0.2965
R-squared	=	0.0178
Adj R-squared	=	0.0033
Root MSE	=	.55927

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	.0330311	.0267657	1.23	0.218	-.0196168	.0856789
dist_1	-.0544168	.0377527	-1.44	0.150	-.1286758	.0198422
dist_2	.0272493	.0379838	0.72	0.474	-.0474643	.101963
dist_3	.0270643	.0461336	0.59	0.558	-.0636798	.1178083
dist_4	-.049372	.0303488	-1.63	0.105	-.1090677	.0103238
_cons	1.903414	.1296652	14.68	0.000	1.648364	2.158464

PSIC 17: Paper & Paper Products Sector



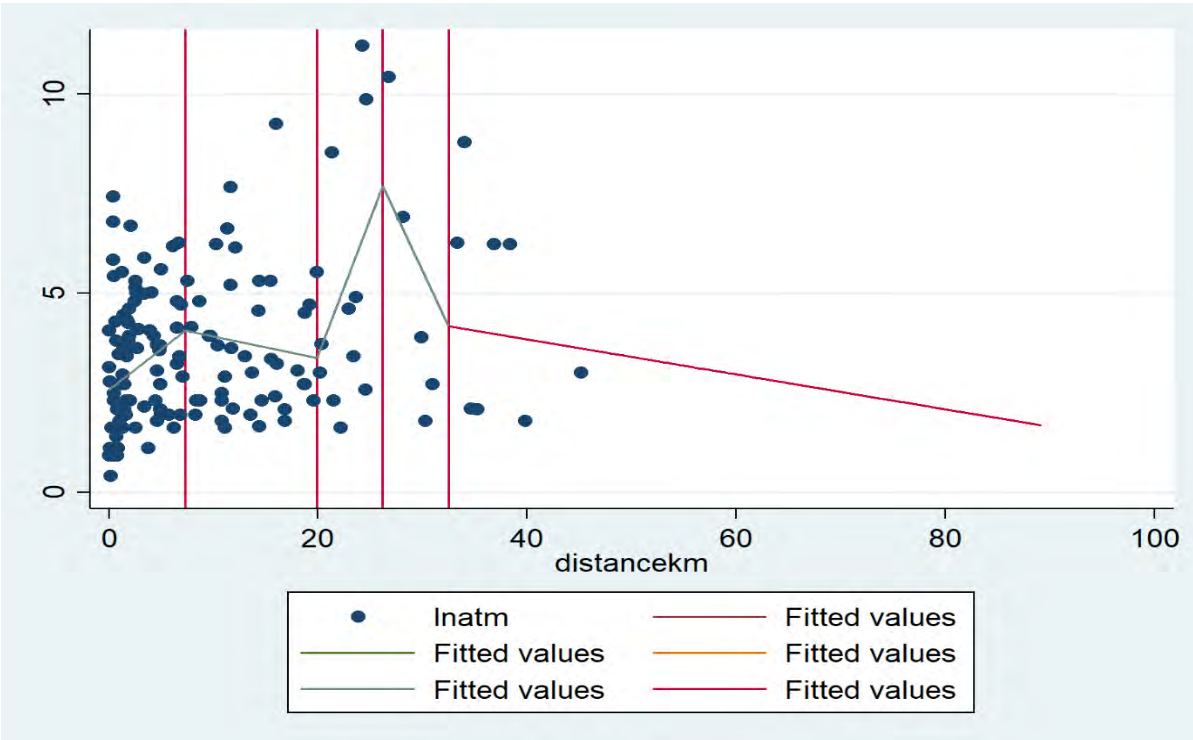
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==17

Source	SS	df	MS
Model	23.2844912	5	4.65689823
Residual	242.574842	130	1.86596032
Total	265.859333	135	1.96932839

Number of obs	=	136
F(5, 130)	=	2.50
Prob > F	=	0.0341
R-squared	=	0.0876
Adj R-squared	=	0.0525
Root MSE	=	1.366

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	.127196	.0521028	2.44	0.016	.0241168	.2302752
dist_1	-.3070963	.0960286	-3.20	0.002	-.4970774	-.1171152
dist_2	.5969619	.2253668	2.65	0.009	.1511006	1.042823
dist_3	-.3599129	.3844005	-0.94	0.351	-1.120403	.4005775
dist_4	-.3987579	.5178057	-0.77	0.443	-1.423175	.6256588
_cons	3.022759	.1887323	16.02	0.000	2.649374	3.396143

PSIC 20: Chemical & Chemical Products Sector

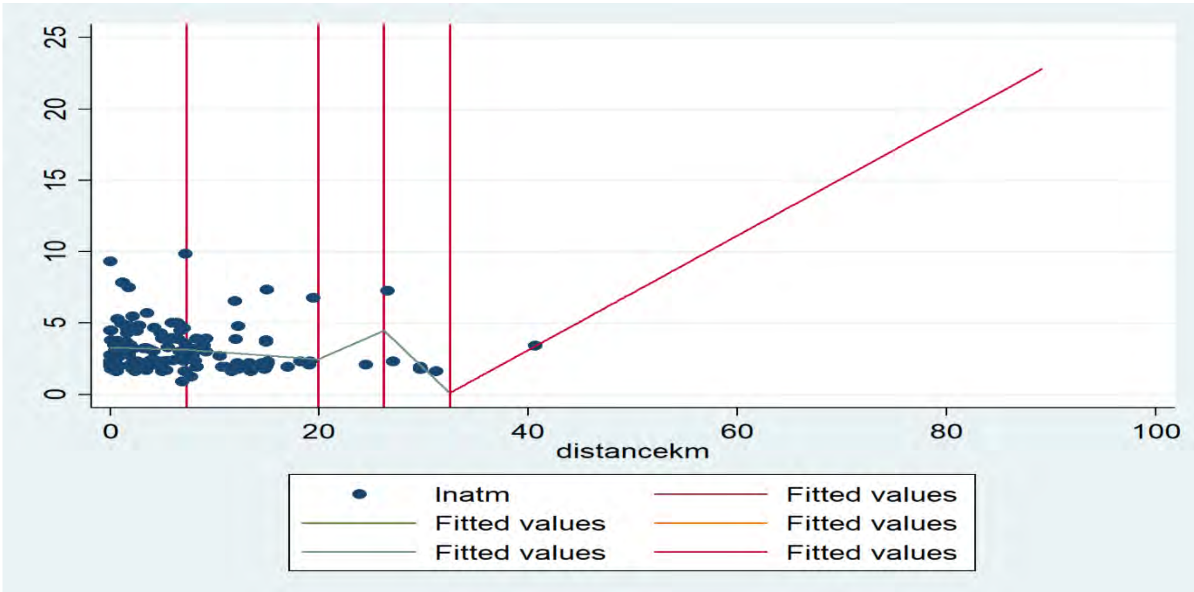



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==20

      Source |       SS       df       MS              Number of obs =     167
-------------+------------------------------           F(5, 161)     =    7.52
       Model |  126.559625     5  25.3119251           Prob > F      =  0.0000
    Residual |  541.877227   161  3.36569706           R-squared     =  0.1893
-------------+------------------------------           Adj R-squared =  0.1642
       Total |  668.436853   166  4.02672803           Root MSE      =  1.8346

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .2030835   .0696699     2.91   0.004     .0654988    .3406681
      dist_1 |  -.2576842   .1135954    -2.27   0.025    -.4820134   -.0333551
      dist_2 |    .740214   .2171163     3.41   0.001     .3114509    1.168977
      dist_3 |  -1.244561   .3779968    -3.29   0.001    -1.991033   -.4980904
      dist_4 |   .5150148   .3427118     1.50   0.135    -.1617753    1.191805
       _cons |   2.574458   .2593413     9.93   0.000     2.062309    3.086608
------------------------------------------------------------------------------
```

PSIC 22: Rubber & Plastic Product Sector



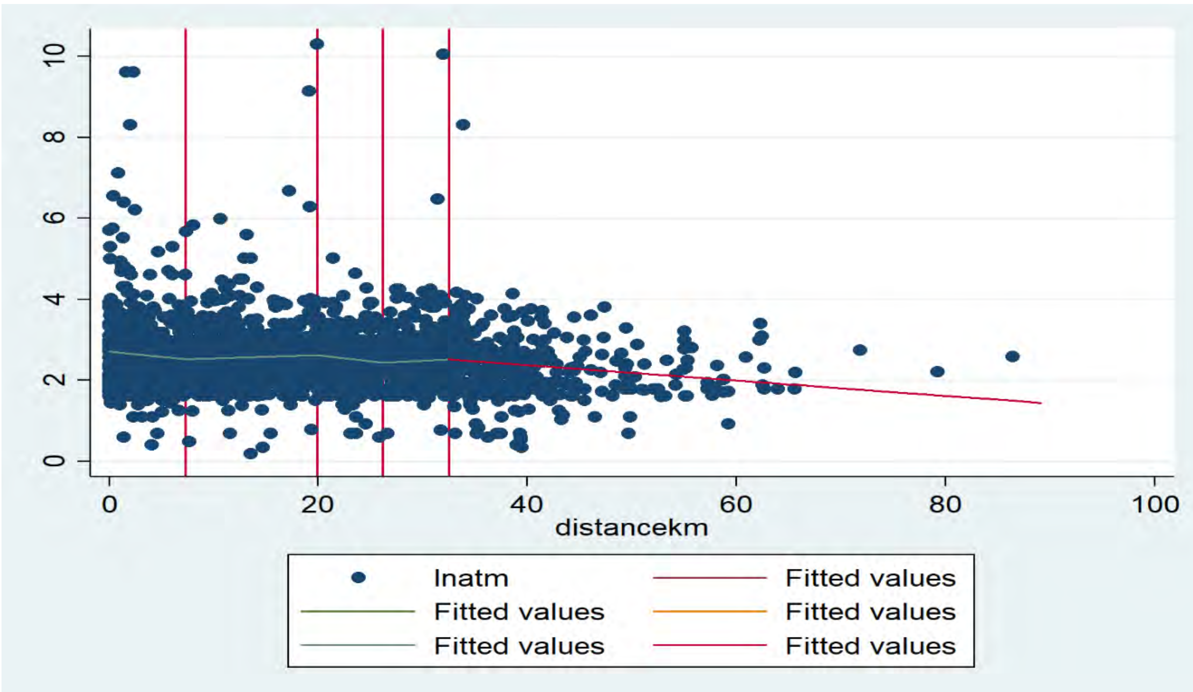
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==22

Source	SS	df	MS
Model	17.5147782	5	3.50295564
Residual	365.892113	148	2.47224401
Total	383.406891	153	2.5059274

Number of obs	=	154
F(5, 148)	=	1.42
Prob > F	=	0.2214
R-squared	=	0.0457
Adj R-squared	=	0.0134
Root MSE	=	1.5723

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	-.019366	.0553639	-0.35	0.727	-.1287719	.0900399
dist_1	-.0339359	.0953668	-0.36	0.722	-.2223923	.1545206
dist_2	.3723298	.2224637	1.67	0.096	-.0672857	.8119454
dist_3	-1.012638	.5142898	-1.97	0.051	-2.028938	.0036614
dist_4	1.094204	.5379937	2.03	0.044	.0310625	2.157345
_cons	3.283557	.237261	13.84	0.000	2.8147	3.752414

PSIC 23: Other Non-Metallic Minerals Sector



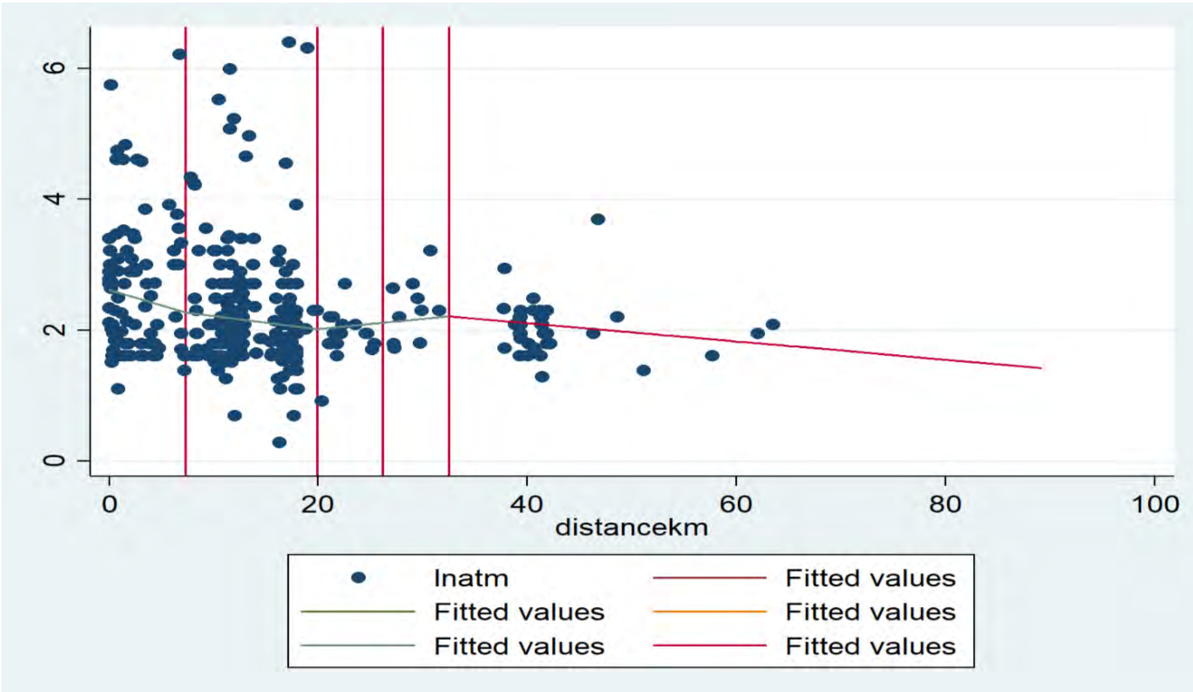
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==23

Source	SS	df	MS
Model	32.3158527	5	6.46317053
Residual	1537.80729	2,160	.71194782
Total	1570.12314	2,165	.725230089

Number of obs	=	2,166
F(5, 2160)	=	9.08
Prob > F	=	0.0000
R-squared	=	0.0206
Adj R-squared	=	0.0183
Root MSE	=	.84377

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	-.0252603	.0106663	-2.37	0.018	-.0461776	-.0043431
dist_1	.0332178	.015988	2.08	0.038	.0018643	.0645714
dist_2	-.0378074	.02065	-1.83	0.067	-.0783032	.0026885
dist_3	.0428136	.0276619	1.55	0.122	-.011433	.0970603
dist_4	-.0318894	.0176279	-1.81	0.071	-.0664587	.00268
_cons	2.707987	.0480838	56.32	0.000	2.613691	2.802282

PSIC 25: Fabricated Metal Products Sector

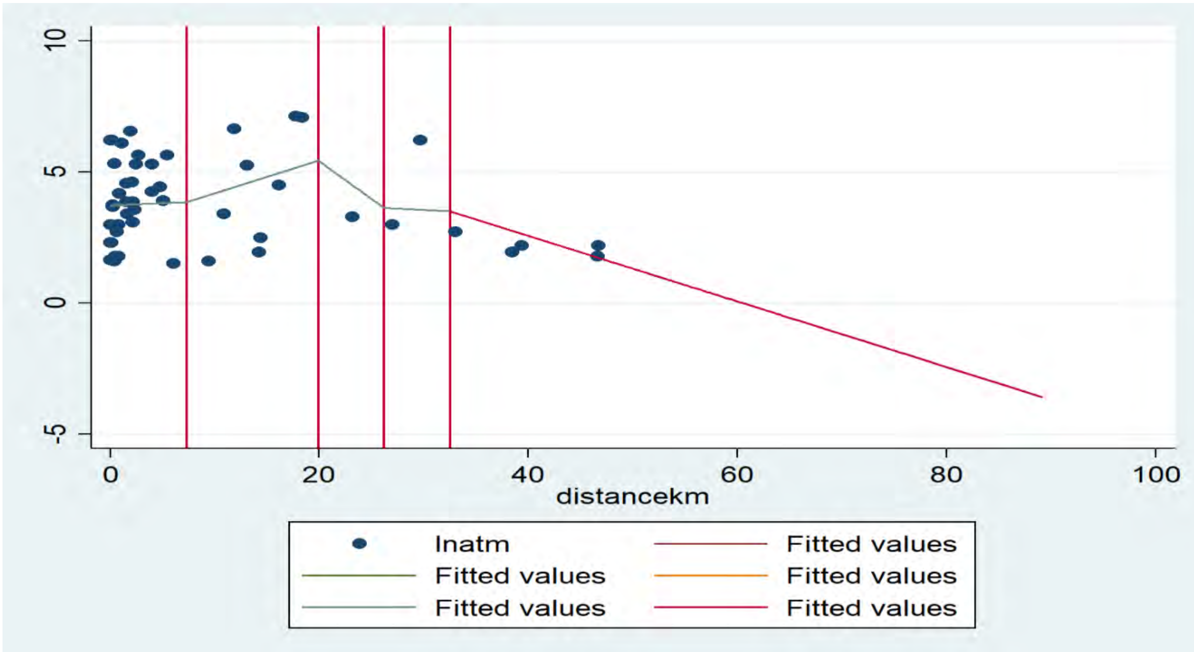



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==25

      Source |       SS       df       MS              Number of obs =     489
-------------+------------------------------           F(5, 483)     =    3.86
       Model |  12.0843152     5  2.41686303           Prob > F      =  0.0020
    Residual |  302.741601   483  .626794205           R-squared     =  0.0384
-------------+------------------------------           Adj R-squared =  0.0284
       Total |  314.825916   488  .645135074           Root MSE      =   .7917

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |  -.0474421   .0218709    -2.17   0.031    -.0904159   -.0044683
      dist_1 |   .0272712   .0325492     0.84   0.403    -.0366842    .0912267
      dist_2 |   .0358576   .0585139     0.61   0.540    -.0791157    .1508309
      dist_3 |  -.0003885   .1068127    -0.00   0.997    -.2102634    .2094864
      dist_4 |  -.0292032   .0758058    -0.39   0.700    -.1781532    .1197467
       _cons |   2.616128   .1163684    22.48   0.000     2.387477    2.844779
------------------------------------------------------------------------------
```

PSIC 27: Electrical Equipment Sector

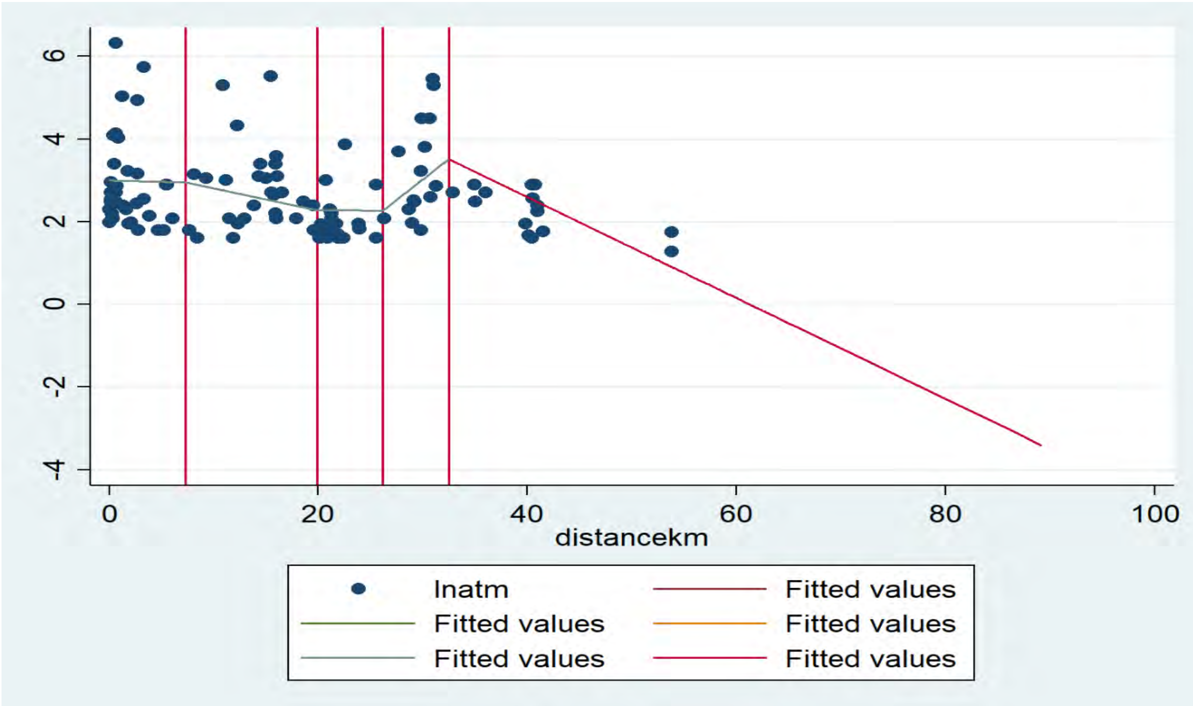



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==27

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(5, 45)        =      1.87
       Model |  24.4058341         5  4.88116681   Prob > F        =    0.1189
    Residual |  117.568047        45  2.61262327   R-squared       =    0.1719
-------------+----------------------------------   Adj R-squared   =    0.0799
       Total |  141.973881        50  2.83947762   Root MSE        =    1.6164

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .0166684   .1364008     0.12   0.903    -.2580568    .2913937
      dist_1 |   .1092012   .2397814     0.46   0.651    -.3737432    .5921457
      dist_2 |  -.4123005    .387188    -1.06   0.293    -1.192137    .3675362
      dist_3 |   .2660232    .607437     0.44   0.664    -.9574177    1.489464
      dist_4 |   -.104703   .4341852    -0.24   0.811    -.9791969     .769791
       _cons |   3.731746   .3843388     9.71   0.000     2.957648    4.505844
------------------------------------------------------------------------------
```

PSIC 28: Machinery & Equipment Sector

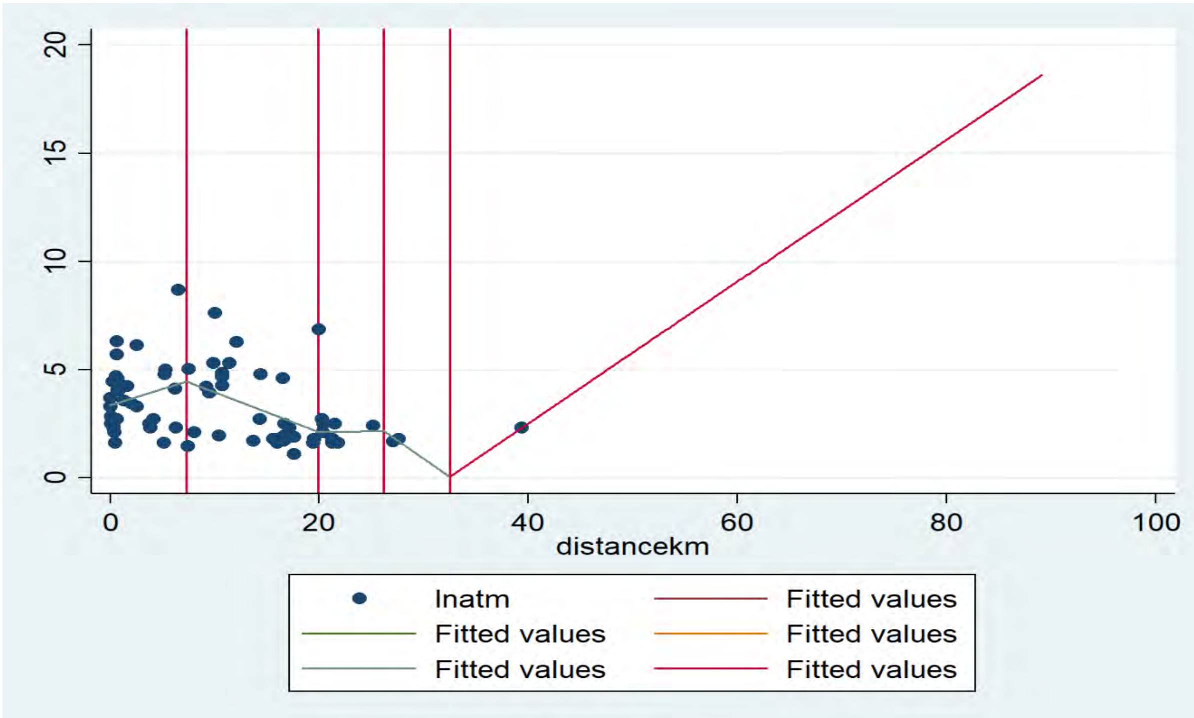


. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==28

Source	SS	df	MS			
Model	17.7811631	5	3.55623261	Number of obs	=	114
Residual	103.546158	108	.958760724	F(5, 108)	=	3.71
Total	121.327321	113	1.07369311	Prob > F	=	0.0039
				R-squared	=	0.1466
				Adj R-squared	=	0.1070
				Root MSE	=	.97916

lnatm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distancekm	-.0063951	.0589052	-0.11	0.914	-.1231554	.1103652
dist_1	-.0473475	.089111	-0.53	0.596	-.223981	.129286
dist_2	.0532279	.1074128	0.50	0.621	-.1596829	.2661387
dist_3	.1975121	.1728049	1.14	0.256	-.1450173	.5400414
dist_4	-.319004	.1237324	-2.58	0.011	-.564263	-.073745
_cons	2.998259	.2206689	13.59	0.000	2.560855	3.435663

Sector 29 **PSIC 28: Machinery & Equipment Sector**

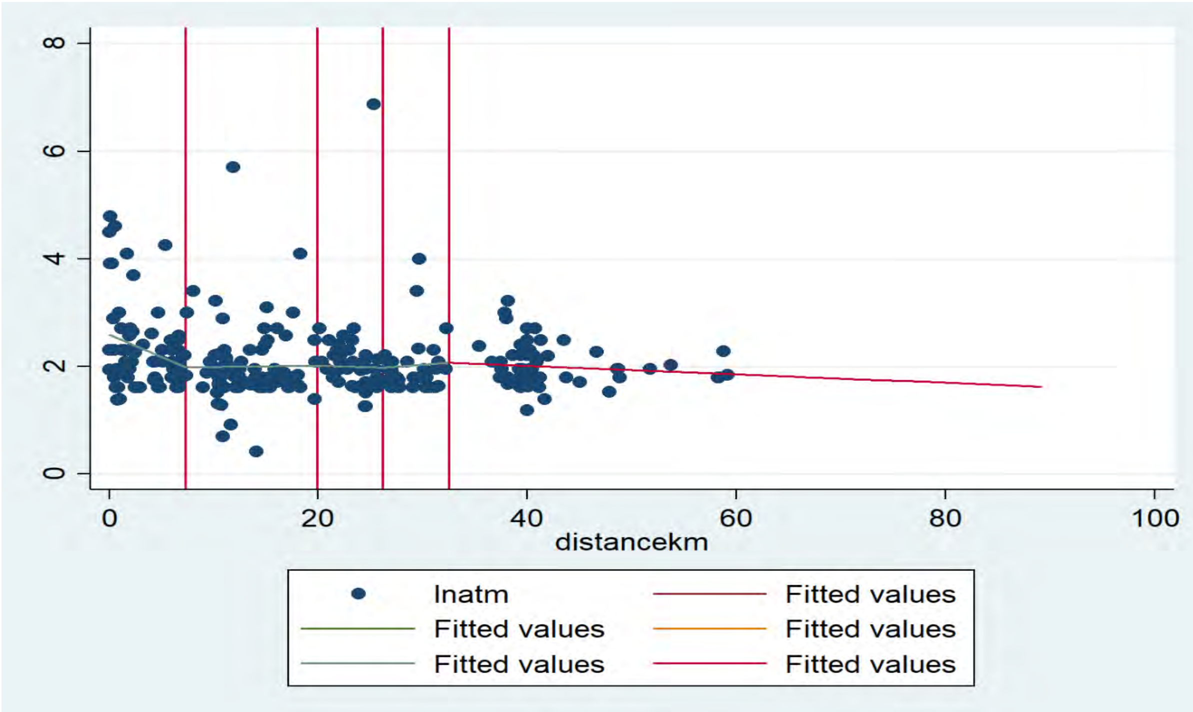



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==29

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(5,    69) =    3.38
       Model |  39.0140253     5  7.80280506           Prob > F      =  0.0087
    Residual |  159.392008    69  2.31002909           R-squared     =  0.1966
-------------+------------------------------           Adj R-squared =  0.1384
       Total |  198.406033    74  2.68116261           Root MSE      =  1.5199

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   .1499734   .0834536     1.80   0.077     -.016512    .3164589
      dist_1 |  -.3348461   .1272643    -2.63   0.010    -.5887315   -.0809607
      dist_2 |   .1902161   .2949971     0.64   0.521    -.3982871    .7787193
      dist_3 |  -.3412724   1.768641    -0.19   0.848    -3.869614     3.18707
      dist_4 |   .6639395   2.792777     0.24   0.813    -4.907499    6.235378
       _cons |   3.352385   .3513862     9.54   0.000     2.651389    4.053382
------------------------------------------------------------------------------
```

PSIC 31: Furniture Sector

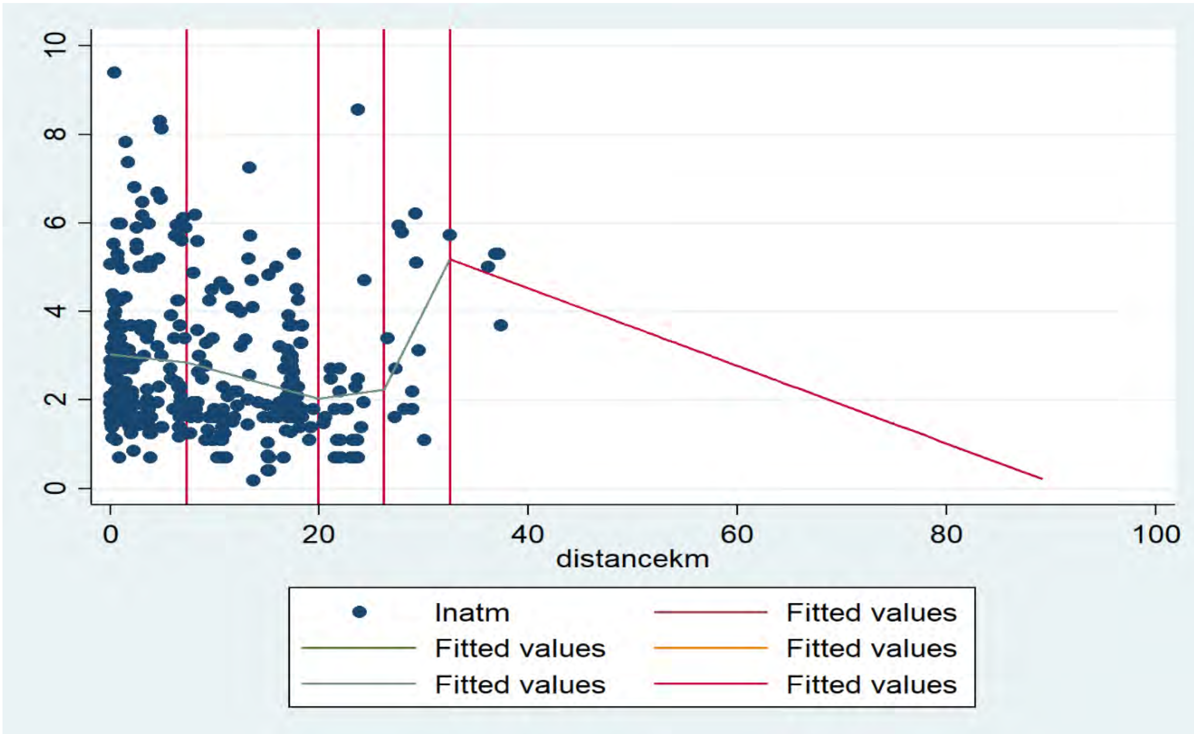



```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==31

      Source |       SS           df       MS      Number of obs   =       321
-------------+----------------------------------   F(5, 315)       =      4.62
       Model |  9.70290366         5  1.94058073   Prob > F        =    0.0004
    Residual |  132.249619       315  .419840061   R-squared       =    0.0684
-------------+----------------------------------   Adj R-squared   =    0.0536
       Total |  141.952523       320  .443601634   Root MSE        =    .64795

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |  -.0829726   .0239549    -3.46   0.001    -.1301045   -.0358407
      dist_1 |    .085369   .0346628     2.46   0.014     .0171691    .1535689
      dist_2 |  -.0084226   .0407657    -0.21   0.836      -.08863    .0717849
      dist_3 |   .0206836   .0555438     0.37   0.710       -.0886    .1299673
      dist_4 |  -.0224909   .0407257    -0.55   0.581    -.1026196    .0576378
       _cons |   2.586958   .1158822    22.32   0.000     2.358957    2.814959
------------------------------------------------------------------------------
```

PSIC 32: Other Manufacturing Sector




```
. regress lnatm distancekm dist_1 dist_2 dist_3 dist_4 if sector==32

      Source |       SS           df       MS      Number of obs   =       384
-------------+----------------------------------   F(5, 378)       =      6.29
       Model |  74.6659621         5  14.9331924   Prob > F        =    0.0000
    Residual |  896.884927       378  2.37271145   R-squared       =    0.0769
-------------+----------------------------------   Adj R-squared   =    0.0646
       Total |  971.550889       383  2.53668639   Root MSE        =    1.5404

------------------------------------------------------------------------------
       lnatm |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  distancekm |   -.024514   .0386833    -0.63   0.527    -.1005754    .0515475
      dist_1 |  -.0405102   .0587519    -0.69   0.491    -.1560316    .0750112
      dist_2 |   .0974253   .1105486     0.88   0.379     -.119942    .3147926
      dist_3 |    .435842   .2845903     1.53   0.126    -.1237363    .9954204
      dist_4 |  -.5557896    .463172    -1.20   0.231    -1.466506    .3549268
       _cons |   3.027944    .169141    17.90   0.000     2.695368    3.360519
------------------------------------------------------------------------------
```

.