THE GROWTH AND WELFARE EFFECTS OF INFRASTRUCTURE IN PAKISTAN

Pakistan Institute of Development Economics

By
Sabila Khatoon
PIDE2015FPHDECO06

Supervised By
Dr. Hafsa Hina
Assistant Professor PIDE

PhD Economics
PIDE School of Economics
Pakistan Institute of Development Economics, Islamabad
2023

Author's Declaration

I Ms. Sabila Khatoon hereby state that my PhD thesis titled **"The Growth and Welfare Effects of Infrastructure in Pakistan"** is my own work and has not been submitted previously by me for taking any degree from **Pakistan Institute of Development Economics, Islamabad'** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my PhD degree.

Ms. Sabila Khatoon
PIDE2015FPHDECO06

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"The Growth and Welfare Effect of Infrastructure in Pakistan"** is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the **HEC** and **Pakistan Institute of Development Economics, Islamabad** towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Students/Author Signature:________________

Ms. Sabila Khatoon

PIDE2015FPHDECO06

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"The Growth and Welfare Effects of Infrastructure in Pakistan"** was conducted by **Ms. Sabila Khatoon** under the supervision of **Dr. Hafsa Hina** No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics from **Pakistan Institute of Development Economics, Islamabad.**

Student Name: Ms. Sabila Khatoon
PIDE2015FPHDECO06

Signature: ______________

Examination Committee:

a) **External Examiner: Dr. Ather Maqsood Ahmed**
Professor
NUST
Islamabad

Signature: ______________

b) **Internal Examiner: Dr. Ghulam Mustafa**
Research Fellow,
PIDE, Islamabad

Signature: ______________

Supervisor: Dr. Hafsa Hina
Assistant Professor
PIDE, Islamabad

Signature: ______________

Dr. Shujaat Farooq
Head, PIDE School of Economics (PSE)
PIDE, Islamabad

Signature: ______________

Dedication

To my lovely children

Daneen, Ameed

&

Emaad

ACKNOWLEDGEMENTS

I am very delighted to pay homage and word of contentment to all my teachers for their guidance and mentorship that remained extremely helpful in this journey of PhD. I would like to special thanks to Dr. Atiya Yasmin and Dr. Javed for the academic support in the topic selection and framing the research question. Also, I am very thankful to Dr. Hafsa Hina for being encouraging and guiding during this journey. I am very thankful to my colleague Ms. Nazia for helping me logistically and liaison.

On personal front, I am very thankful to all batchmates of 2016 for their bonding and warm attitude that had made this academic journey livelier and more interactive. Also, I am very thankful to my parents, my siblings, my children and my husband Zaib Alam for their unconditional support, prayers, and love.

ABSTRACT

This dissertation explores the impact of economic, social, and financial infrastructure on Pakistan's economic growth, regional development, and social welfare. This dissertation employs multiple methodologies to analyze the research questions. The Principal component analysis (PCA) is applied to construct indexes for the dataset ranging from FY1980-81 to 2022-23 for time series regression analysis. An Autoregressive Distributive Lag (ARDL) estimation technique measures the long-run and short-run effects while Fully Modified Ordinary Least Square (FMOLS) estimation time series analysis for the robustness check. Dynamic Spatial Durbin panel data model estimated for provincial panel data analysis for dataset ranges from 1990 to 2020. The model incorporates spatial lags of both dependent and explanatory variables, capturing short-term and long-term direct and indirect effects, including spatial spillover effects on economic growth. Spatial panel data and cross-sectional spatial data analysis for the welfare analysis has been conducted at the district level using Pakistan Social and Living Standard Measurement (PSLM) survey data. This dissertation provides a comprehensive analysis of the research questions, ensuring robustness and capturing the various dimensions of the phenomena under investigation by employing multiple sets of methodologies and approaches.

This study found the significant importance of core, social infrastructure of education and health and financial infrastructure in generating long-term and short-term productivity growth effects. The core infrastructure increased the productivity growth in Pakistan, non-linearly. An inverted U-shape- nonlinear relationship of the physical stock of core infrastructure and education indicates the fundamental role of core and social infrastructure at the early stages of economic development. Based on the insignificant coefficient for the physical stock of education, it is inferred that Pakistan has not yet made sufficient investments in building educational infrastructure given to its growing population to reap the long-term benefits fully. However, despite the poor addition in physical infrastructure for education in per-worker terms, the utilization of educational services has been increased over time and contributed positively to Pakistan's long-run productivity growth. The physical infrastructure for health services does augment LR productivity growth linearly, while physical infrastructure for financial services augments LR productivity growth nonlinearly. It showed that investing in

the financial infrastructure had increased productivity growth at an accelerated rate over time. While, in the short run, investing in core, health, and financial infrastructure drives up the productivity growth signficantly.

In addition, this study also analyzes the asymmetric impact of each type of infrastructure by conducting country-specific analysis. Study found a significant asymmetric effect in the LR. Particularly the coefficient for positive-one is highly significant, representing all types of infrastructure critical for LR economic development of Pakistan. In the short run (SR), the asymmetric effect is significant for the core and educational infrastructure. Therefore, we inferred that core, social and financial infrastructure are the critical investment drivers of the productivity growth of Pakistan. The net reduction in the stock of infrastructure thus hampers the productivity growth.

The provincial-level analysis finds evidence for a positive spatial spillover effects. The findings support the proposition of the critical importance of the investment in the human capital stock, private capital stock, and core infrastructure for the economic development in provonices . The study found highly significant and positive spatial parameters (spatial lambda and spatial rho), indicating the presence of spatial dependence and positive regional spillover effects. The analysis reveals a reverse-spatial-temporal diffusion, as the negative and highly significant spatial diffusion parameter suggests a detrimental effect on economic development due to relative attractiveness resulting from lagged economic growth in neighboring areas. In the short run, both direct and indirect effects of core infrastructure are negative and highly significant. However, in the long run, direct spillover effects become positive and highly significant, with the total net effects (direct and indirect effects) being positive. These findings support the argument that Pakistan's current infrastructure investment strategy aims to maximize long-term gains. Therefore, it is highly recommended to allocate resources prudently towards infrastructure projects, taking into account the short-term tradeoffs associated with such investments. The existence of these tradeoffs can be primarily attributed to factors such as extensive implementation timelines, initial feasibility assessments, design vetting processes, contract tendering procedures, and the need for significant fiscal space. Hence, securing funding and technical assistance from international agencies becomes imperative for the successful implementation of infrastructure projects in Pakistan.

Additionally, this study examines the impact of transportation infrastructure investment on economic geography of social welfare in districts. The findings revealed investing in road infrastrcuture reduces social depreviation (reduction in multidimensional poverty). Using Spatial Error Model (SEM) and Spatial Lag Model (SLM) at the district level supported the evidence for positive network externalities (global effect). The observed disparities in the state of development at the district level are attributed to variations in infrastructure stock, population density, urbanization, and geographic coverage across districts. Hence, the degree of regional connectivity and economic integration significantly influences regional economic and social development in Pakistan. Based on these findings, it is highly recommended to prioritize investment in the development and maintenance of road infrastructure in the most impoverished and geographically large districts of Pakistan. The relative geographic size emerges as an important driver of social welfare in the country. Furthermore, effective public-administration management at the district level is crucial in addressing social deprivation in Pakistan. To alleviate social deprivation and economic marginalization, the study emphasizes on devolving policy and decision-making authority to the district level given the critical importance of enhance regional connectivity.

Keywords: Regional Connectivity; Core Infrastructure; Social Infrastructure, Health Infrastructure, Financial Infrastructure, Exploratory Spatial Data Analysis, Spillover effects, Spatial Regression Analysis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A2FS	Access to Finance Survey
ADB	Asian Development Bank
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ARCH	Autoregressive Conditional Heteroskedasticity
ARDL	Autoregressive Distributed Lag
ATM	Automated Teller Machines
BOT	Build-Operate-Transfer
BRI	Belt and Road Initiative (BRI)
CAB	Current Account Balance (CAB)
CPI	Consumer Price Index
CSD	Cross-sectional dependency
DFDI	Foreign Direct Investment
DPT	Diphtheria, Pertussis (Whooping Cough), and Tetanus
ECT	Error Correction Term
ESDA	Exploratory Spatial Data Analysis
FMOLS	Fully Modified Ordinary Least Squares
GDL	Global Data Lab (GDL)
GNIPC	Gross National Income Per Capita
HQIC	Hannan-Quinn Information Criterion
IDWM	Inverse-distance weight matric
IMF	International Monetary Fund IMF
IDW	Inverse-Distance Weight-Matrix
IPS	Im–Pesaran–Shin
KEBA	Korean Exim Bank Assistance
KPH	Kilometers per hour
LISA	Local Indicators of Spatial Association (LISA)
LPI	Logistics Performance Indicator
MHCI	Multidimensional Headcount

MPI	Multidimensional Poverty Index
NHA	National Highway Authority
NTP	National Transport Policy
NTRC	National Transport Research Centre
OCHA	Office for coordinating Humanitarian Affairs
OLS	Ordinary Linear Least-Squares
PCA	Principal Components Analysis
PNSC	Pakistan National Shipping Corporation
PP	Phillips-Perron
PPP	Public-Private Partnership (PPP)
PTA	Pakistan Telecommunication Authority
QML	Quasi-Maximum Likelihood
QCW	Queen Contiguity Weight-Matrix
RESET	Ramsey Regression Equation Specification Error Test
SAR	Spatial Auto-Regressive
SEM	Spatial Error Model
SLM	Spatial Lag Model
SDG	Sustainable Development Goals
SDM	Spatial Durbin Model
SIC	Schwarz Information Criterion
SOC	Social Overhead Capital
T&T	Telephone and Telegraph department
UNDP	United Nations Development Programme
VECM	Vector Error Correction Model
WB	World Bank
WDI	World Development Indicators

CHAPTER 1

INTRODUCTION

1.1 Background, and Rationale for the study

The role of infrastructure in regional development is fundamental, serving as a crucial foundation for economic activities and social mobility. Infrastructure plays a primary role in alleviating development constraints and bottlenecks, directly fostering economic growth (Nijkamp, 1986). In the literature of long run economic growth Arrow & Kurz (1970) and Weitzman (1970) developed the theoretical construction of the infrastructure-growth relationship. Later on the pioneering work of Solow(1956) defined process of the capital accumulation as a central investment driver of long run productivity growth and this ground-breaking work initiated had spurred a set of empirical literature calibrating cross-country convergence and differences in income level across countries (Barro & Sala-i-martin, 1992; Knight et al., 1993). At the same time, keeping the critical role of human-skills and knowledge in economic development, Mankiw et al., (1992) incorporated investing in human capital as critical driver of long run productivity growth. Similarly Aschauer (1989) empirical work elucidated founding role core infrastructure in the productivity growth, and yet it a widely explored topic in economic growth literature (Boarnet, 1998, Canning & Pedroni, 2008, Chatterjee & Turnovsky, 2012, Ferreira, 1995; Haughwout, 2005; Ottaviano 2008; Snieska & Simkunaite, 2009; Straub 2008; Straub et al. 2006; Nijkamp & Capello, 2009). Infrastructure affects economic growth through various channels. For instance, it enhances the productivity of factor inputs, creates a synergistic effect with private investment, reduces investment costs, increases the durability of private capital, and improves health and education outcomes (Agénor & Moreno-Dodson, 2012). Moreover, infrastructure stimulates economic activities, generates employment opportunities, enhances production facilities, and reduces transaction and trade costs, thereby enhancing competitiveness (Sahoo & Dash, 2012).

Infrastructure plays a crucial role in determining the distribution of wealth, income, and welfare, with its effects influenced by factors such as infrastructure spillovers, time of consideration, and source of infrastructure financing (Chatterjee & Turnovsky, 2012). By inducing economic growth, Infrastructure contributes in reducing inequality. For example, it improves access to productive opportunities, lower production and transaction costs facilitates industrial and agro-industrial development, increases the value of assets of the poor (Bajar & Rajeev, 2016). Furthermore, well-developed transport infrastructure enhance labor mobility by the geographical access, improves transportation service, facilitates information flow and enables disadvantaged individuals and marginalized communities to access productive opportunities through connectivity and networking. Many developing countries have witnessed a reduction in poverty as result of investment in the road infrastructure, as demonstrated in China and India (Fan et al., 1999; Fan & Chan-Kang, 2005).

Defining infrastructure is a complex task, as there is no universally agreed upon definition (Snieska & Simkunaite, 2009). Fourie (2006) presents two ways of defining infrastructure. First, infrastructure can be seen as 'social overhead capital' characterized by its public nature and role in public service. Alternatively, infrastructure can be defined by compiling a comprehensive list of its various types such as transport infrastructure, communications infrastructure, energy supply infrastructure, water infrastructure, environmental infrastructure, education infrastructure, etc. Gianpiero (2010) further categorizes infrastructure into three broad categories, based on its nature.. The initial category pertains to economic or core infrastructures, encompassing transport networks, communication systems, and energy networks. The second category involves social infrastructures, which include health, education, culture, and environmental facilities. Lastly, territory infrastructures encompass tourist attractions, trade-related facilities, and monetary intermediation structures. In essence, infrastructure is an overarching term encompassing various types of infrastructures that exhibit characteristics of public-capital goods, capable of generating externalities and spillover effects.

Improvements in infrastructure have a significant impact on the geographical distribution of economic activities (Ottaviano, 2008). Infrastructure enhances the local economic activity in the areas where it is located and generates benefits in the neighbouring areas through spatial diffusion processes and spillover effects. Positive spillover effects of transport infrastructure in particular, have extensively studied at the national and regional levels in various countries, including China (Shi et al., 2017; Yu et al., 2013; Zhang & Ji, 2019) and Spain (Cohen & Paul, 2004: Cantos et al., 2005; Cohen, 2010). Connectivity characteristics of infrastructure contributes to positive spillovers, while negative spillover effects occurs when factor migrate from the region with inadequate infrastructure to higher due to relative economic attractiveness. Boarnet (1998) estimated such negative spillovers in the counties of California, USA. Hence, efficient, and reliable infrastructure services are crucial for the socio-economic well-being of developing countries and determines geospatial patterns of economic development. Increased productivity, employment opportunities, regional growth, and improved welfare are the direct and indirect channels through which infrastructure positively impacts the well-being of the poor (Ali & Pernia, 2003; Bajar, 2013; Bajar & Rajeev, 2016; Chatterjee & Turnovsky, 2012).

The economic geography of development exhibits significant spatial disparities in Pakistan. It is elucidated from the trends in Gross National Income (GNI) per capita at the sub-national level of provinces in Pakistan, highlighting significant differences in economic development. Over the past three decades, there has been an overall increase in GNI per capita; however, the rate of change varies across provinces. Notably, the province of Punjab has experienced the highest increase in GNI per capita compared to other provinces.

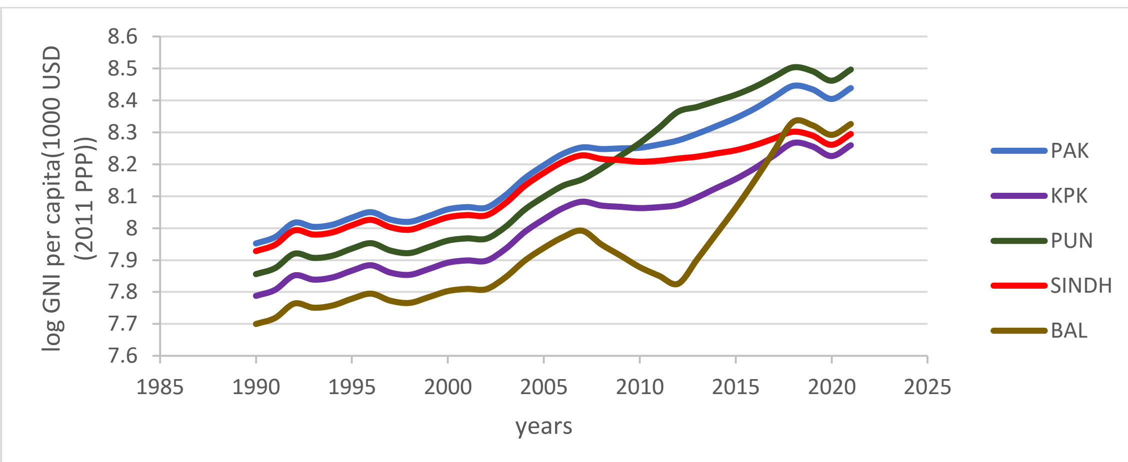


Figure 1. 1 **Province-wise trends: log GNI per capita (constant 2011, USD PPP)**[1]

Source: Global Data Lab, *Institute for Management Research, Radboud University, Netherland*

The trends indicate GNI per capita has increased during the last 3 decades, but the rate of change in GNI per capita varies across provinces, particularly GNI per capita increased at most in the province of Punjab, comparing other provinces of Pakistan. These differences and prominent disparities in terms of socio-economic development not only exist in provincial level but also visible at the district level. For instance, unequal district-wise socio-economic development was reported in the 1960s, 1970s, and 1980's (Hussain, 1996), and these trends are persisting (UNDP,

1 *The data has been obtained from GDL, UNDP expressed in terms of natural log*

2016). Over time, the statistics for poverty show a reduction at both national and provincial level. However, the poverty reduction trends is not equivocal across the provinces and districts (UNDP, 2016). For instance, a negative change has been observed in the Multidimensional Poverty Index (MPI) in Harnai, Killa Abdullah, Ziarat, Umerkot, Sherani, Kashmore, Panjgur, Chagai, Pishin, Tando Muhammad Khan, and Badin (UNDP, 2016). The negative change in poverty incidence has been reported for Chagai, Tharparker, Pishin, Ziarat, Tando Muhammad Khan, Kashmore, Killa Abdullah, Panjgour, Harnai, and Umerkot. In contrast to these negative trends, the districts Islamabad, Attock, Jhelum, Lahore, Karachi, Rawalpindi, and Sialkot outperformed all other districts in poverty reduction.

This appears that Pakistan's infrastructure is lacking in terms of supporting economic development. According to the 2018 Logistics Performance Index report, the country's economic infrastructures for trade facilitation and logistical service delivery are not on par with neighbouring countries like China and India, which are experiencing rapid development. Therefore, it is crucial to examine how infrastructural services contribute to the creation and maintenance of economic activities. Although there has been a sustained increase in infrastructure investment at the national level over the past three decades, however this investment is not evenly distributed among the provinces in Pakistan. Figures 1.2 and 1.3 present data trends on the quality of road infrastructure and electricity distribution networks across different provinces.

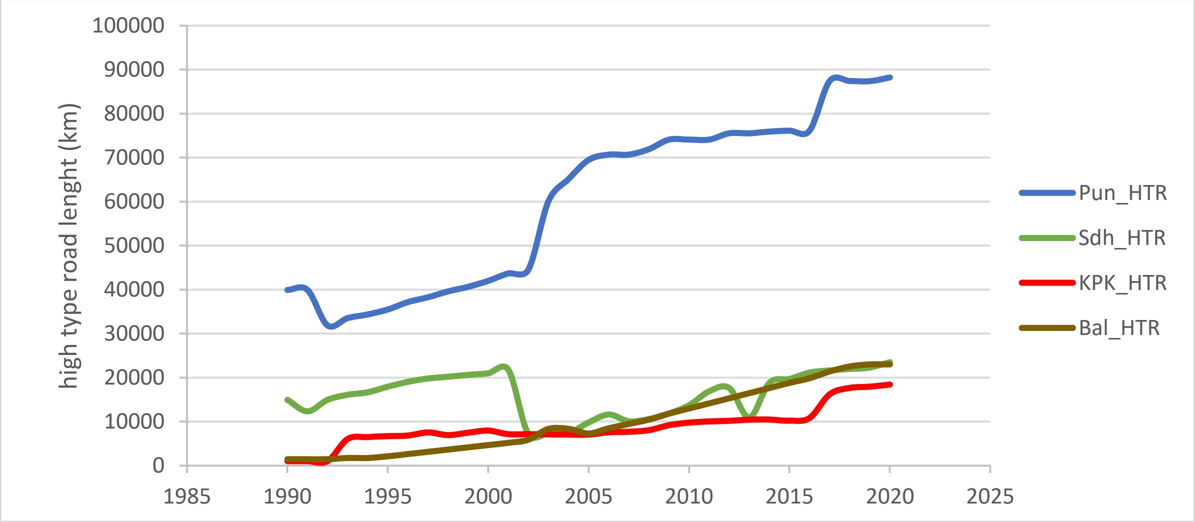


Figure 1. 2 **Province wise (log) of high type road length in Pakistan (FY 1990-2022)**

Source: Provincial Development Statistics Reports, various issues

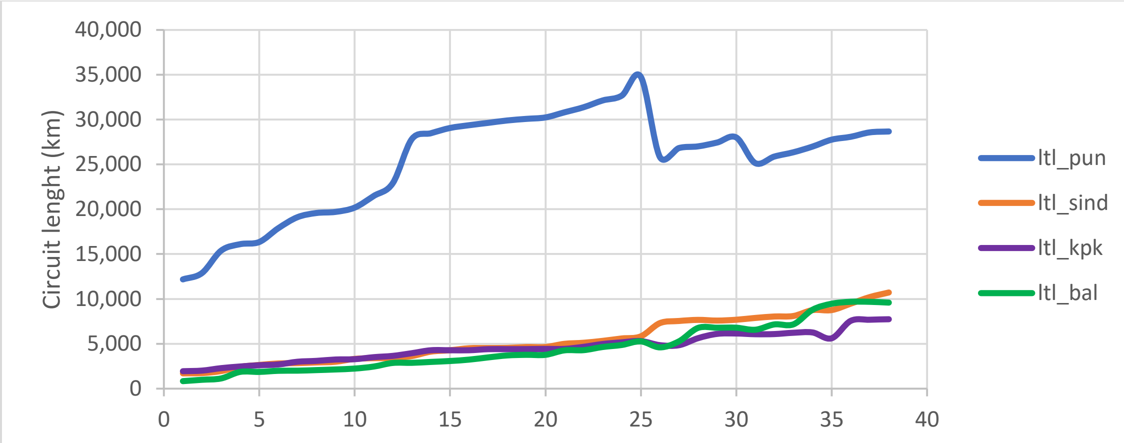


Figure 1. 3 **Province wise length of transmission lines (circuit) length[2] in Pakistan (FY1980-2022)**

Source: NTDC (2022) Power System Statistics, 47th edition.

The observed data trends indicate a growth in the scale of transportation and energy networks; however, this expansion is not consistent among the provinces. Table 1.1 presents statistics on the distribution of road and energy networks per 000 sq km of geographic area, categorized by province. The data reveals that Punjab has a greater absolute and proportional density of road networks compared to other provinces. On the other hand, in terms of road network density relative to the size of each province, KPK has a denser road network in Pakistan. Similarly, when considering the distribution network of energy, Punjab exhibits higher absolute and relative density in both aspects.

Table 1. 1 *Road and energy network per 000 sq km area*

	Punjab	Sindh	KPK	Balochistan
High type of Road (HTR) per area (000, sq km)				
1990-91	4.97	106.02	13.69	4.24
2000-01	38.75	148.77	106.77	13.49
2010-11	47.55	97.20	131.01	37.41
2020-21*	89.63	166.69	246.99	66.26
Length of Transmission lines (LTL) length per area (000, sq km)				
1990-91	95.95	21.33	43.63	6.13
2000-01	144.30	32.13	59.14	10.01
2010-11	128.64	52.90	64.47	18.10
2020-21	140.09	66.60	78.60	27.08

Data Source: Pakistan's Economic Survey (various issues).

[2] length of transmission lines (which is expressed in natural log), data has been obtained from (NTDC, 2021) power system statistics report

These statistical trends highlight the presence of a causal relationship between infrastructure and the development of states at the sub-national level in Pakistan, It appears that the varying levels of infrastructure within districts in Pakistan are one of the primary factors contributing to differences in development and people's well-being. Therefore, it is crucial to examine the role of infrastructure in Pakistan's economic development, both at the national and sub-national levels, considering the inherent characteristics of infrastructure services such as networking, cohesion, indirect productivity, and spillover effects. Given the nature of infrastructure and its wide-ranging impacts on the economy, it is essential to analyse the growth and welfare effects in order to recommend infrastructure investment as a policy tool for long-term economic development in Pakistan.

To the best of my knowledge, no similar study has been conducted in Pakistan that specifically examines the direct, indirect, and spatial spillover effects of infrastructure at the sub-national level, focusing on provinces and districts. Furthermore, this study will utilize the physical stock of infrastructure as a composite index to measure both the economic/core and social aspects of infrastructure, incorporating a comprehensive set of indicators. Therefore, this study aims to complement the existing literature on the significance of infrastructure development in Pakistan by adopting a holistic approach to analyze its impact. Specifically, it will assess the influence of infrastructure on aggregate productivity and welfare indicators, including the incidence and intensity of multidimensional poverty. By doing so, this study seeks to provide a comprehensive understanding of the role of infrastructure in Pakistan's development context.

1.2 Research Question and Objectives

The present study aims to address the following research questions:

1. What is the role of economic and social infrastructure in economic growth?
2. Does infrastructure development have an impact on welfare in Pakistan?
3. What is the significance of both aggregate infrastructure and its individual components?
4. How do the spillover effects of infrastructure development integrate geographical regions within Pakistan?

The study's objectives are organized into three main premises. The first premise focuses on economic growth, the second on the welfare effects of infrastructure in Pakistan, and the third on the infrastructure-policy nexus within the country. Therefore, this study will construct an aggregated index to analyze the relationship between infrastructure's components and output growth in Pakistan. These components encompass the economic, social, and financial infrastructure. The study specifically investigates the growth effects resulting from investments in core, social, and financial infrastructure, considering their physical presence and utilization. Additionally, it aims to identify the most influential type of infrastructure for driving productivity growth in Pakistan.

The sub-objectives are as follows:

- Examine the growth effects of physical stock and utilization of core infrastructure.
- Investigate the role of social infrastructure in fostering productivity growth in Pakistan.
- Explore the contribution of financial infrastructure to Pakistan's economic growth.
- Determine the critical types and components of infrastructure for boosting output growth in Pakistan.
- To examine the asymmetric and non-linear growth effects of each type of infrastructures.
- Examine how infrastructure spillovers promote regional economic development in Pakistan.
- Quantify infrastructure's dynamic direct, indirect, and spillover effects on Pakistan's economic performance.
- Assess the impact of infrastructure provisioning on regional poverty at the district level.
- Examine inter-district variations in infrastructure availability and its relationship with social welfare.
- Qualitatively evaluate the efficacy of infrastructure development policies, highlighting policy and implementation gaps, with respect to Sustainable Development Goal (SDG) number 9, which focuses on sustainable infrastructure development.

1.3 Contribution of the Study

This study investigates the impact of infrastructure and its component's impact on Pakistan's Gross Domestic Product (GDP) growth and social welfare. The findings of this study would augment informed future infrastructure policies in Pakistan, aligning with the country's long-term socio-economic development goals.

The contribution of the study is multifaceted:

1. Infrastructure development plays a significant role in Pakistan’s policy agenda. The country has been actively enhancing its infrastructure, particularly in transportation and energy, through international collaborations such as China-Pakistan Economic Corridor (CPEC) under the Belt and Road Initiative (BRI). This study is of critical importance as many infrastructure projects have been and or are expected to complete.
2. The national framework for sustainable development of Pakistan based on the Sustainable Development Goals (SDGs), emphasizes the importance of building resilient infrastructure, promoting inclusive and sustainable industrialization, and fostering innovation (MoPDR, 2018). Under this framework, Pakistan has prioritized objectives such as increasing manufacturing value added and employment opportunities.
3. The national priority SDG indicators 9.2.1 and 9.2.2 state that Pakistan's objective is to increase manufacturing value added by 16 % of GDP and to increase employment opportunities by 18 % of total employment in 2030. The critical policy question is whether building a mass of physical infrastructure of economic, social, and financial intermediation would be able to deliver more significant economic development in the country. Another question is finding which infrastructure types are essential for long-term economic development, i.e., for economic growth promotion and reduction of the incidence and intensity of poverty. This study is vital in addressing the global commitment to achieving sustainable development goals. Pakistan has prioritized the SDGs, intending to join the league of upper-middle-class countries by 2030. In addition, Pakistan aimed to achieve a sustainable development path by eliminating poverty and reducing inequalities. Providing efficient and resilient infrastructure is a crucial service delivery area to meet SDG objectives. This study will guide further in prioritizing national development objectives.
4. The previous empirical literature on the long-run impact of investing in infrastructure on Pakistan's economic growth has provided mixed results. These discrepancies can be attributed to methodological differences, datasets, and indicators. For example, Ghani & Din (2006) found that public investment did not generate a significant long-term growth effect in Pakistan. However, Javid (2019) analyzed the component-wise impact of public investment in infrastructure and found that it significantly contributed to long-term economic growth. (Javid, 2019) specifically examined public investment in sub-sectors

such as electricity generation and distribution, gas distribution, and transport and communication.

5. This study aims to fill that gap by analysing the growth effects of major components of core infrastructure, including economic, social, and financial aspects. It employs a comprehensive set of indicators to represent these different types of infrastructure. Additionally, new indicators, such as the total length of transmission lines for energy infrastructure and the number of hospital beds and child immunization for health infrastructure, have been introduced. These indicators have not been used in previous studies on Pakistan.
6. By answering the key policy question of which types of infrastructure are critical for generating a growth effect in Pakistan, this study provides important policy-related insights. It is particularly valuable considering Pakistan's limited fiscal space. The findings of this study can shape and refine Pakistan's economic and social policy objectives, making it a significant contribution to the country's development.

1.4 Organization of the study

This study is divided into 9 chapters. After the introduction, the second chapter focuses on the literature review. The third chapter provides an overview and situational analysis of Pakistan's current infrastructure state. The fourth chapter presents the theoretical framework, which includes a mathematical model developed to calibrate the growth and spillover effects of infrastructure. The remaining five to nine chapters empirically examine each objective outlined in the study. The fifth chapter provide provided the estimation techniques and methodology to analyses the aggregate as well as component wise effect of infrastructure. Moreover, the sixth chapter provides a comprehensive account for estimation of the infrastructural spillovers alongside other key drivers of economic development in Pakistan. The seventh and eight chapter presents an Exploratory Spatial Data Analysis (ESDA) for road infrastructure and social welfare and regression analysis to assess the impact of connectivity on the social welfare of Pakistan. Finally, the last chapter offers a comprehensive conclusion, summarizing the study's findings and providing key policy insights.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter thoroughly examines the existing scholarly literature, focusing on three key areas: the impact of infrastructure on economic growth, the influence of infrastructure on spatial spillovers and regional welfare, and the role of infrastructure in social welfare. The chapter is organized into three main sections for clarity.

The first section provides a concise overview of the theoretical and empirical connections between infrastructure and productivity, emphasizing their impact on economic growth. It explores the existing literature in this field. The second section delves into the literature that investigates how infrastructure contributes to spatial externalities in regional economic development. It examines the studies that explore the link between infrastructure and regional welfare, emphasizing the concept of spatial spillovers. Lastly, the chapter discusses the literature that underscores the significance of infrastructure for social welfare. It highlights the studies that shed light on the role of infrastructure in improving the overall well-being of society. Overall, this chapter serves as a comprehensive review of the scholarly works related to infrastructure and its implications for economic growth, spatial spillovers, and social welfare.

2.2 Productivity Growth and infrastructure

The relationship between infrastructure investment and factors driving productivity growth has been explored using various analytical approaches. The theoretical literature often interchangeably uses the terms "public capital" and "physical infrastructure," considering them crucial inputs in aggregate production functions within exogenous (Arrow & Kurz, 1970; Barro, 1990) and endogenous growth models.

Studies based on the exogenous growth framework have indicated that changes in the stock of public capital/infrastructure only have a temporary impact, with technological progress being the primary driver of long-term growth (Arrow & Kurz, 1970). In contrast, within the endogenous growth model framework, an increase in infrastructure stock leads to higher steady-state income per capita (Barro, 1990; Holtz-Eakin & Schwartz, 1995). Furthermore, Glomm & Ravikumar

(1994) examined the role of infrastructure as a growth-promoting input using a general equilibrium growth model. Their analysis treated infrastructure as an external augmentable input in the production functions under constant returns.

Within the literature, the term "infrastructure" is frequently used interchangeably with "public expenditure," "public investment," or "public capital stock." Typically, public investment or capital expenditure for infrastructure is measured in monetary terms, although this approach has raised concerns about its suitability as an infrastructure indicator. Scholars have suggested that measuring infrastructure in physical units, especially in developing countries, would be more appropriate. Pritchett (2000) emphasized that not all public sector spending labeled as "investment" generates economically valuable "capital." Additionally, Timilsina et al. (2020) highlighted several challenges in measuring infrastructure investment, including a lack of comprehensive data, difficulties in distinguishing between new infrastructure investment and maintenance costs, varying time lags for the impacts of social and physical infrastructure, and the challenge of quantifying infrastructure quantity and quality based solely on public expenditures.

The empirical literature extensively documents a positive and well-established relationship between public expenditure or public capital and regional economic growth in developing areas (Yu et al., 2012, 2013). The role of public expenditure in infrastructure as a catalyst for growth has been widely recognized, although its effectiveness is contingent upon policy implementation (Esfahani & Ramírez, 2003). The literature underscores the significant importance of infrastructure for economic growth, considering the absence of adequate infrastructure as a hindrance to development (Sanchez-Robles, 1998). Consequently, the empirical evidence overwhelmingly supports a positive association between public capital and output growth. In addition, (Bom & Ligthart, 2014)Bom & Ligthart (2014) (Bom & Ligthart, 2009)conducted a meta-analysis of 67 published research studies from 1983 to 2008, accounting for publication bias, and found the true unconditional output elasticity of public capital to be 0.146. Furthermore, Elburz et al., (2017) critically evaluated 42 studies using a meta-analysis technique and an ordered probit model. These studies, examining the impact of public investment in infrastructure from 1995 to 2004, reported varying effects—negative, positive, or insignificant. Elburz et al. found that the outcomes of the primary studies were influenced by factors such as infrastructure type, research methodology, time span, infrastructure measurement approach, and geographical scale. Notably, studies conducted in developed countries tended to report a lesser impact of infrastructure investment on growth. For example, studies focused on the United States demonstrated a reduced impact on productivity growth compared to other nations, likely due to differences in developmental stages. Consequently, Elburz et al., (2017) argued that the impact of public expenditure on infrastructure in driving economic growth tends to be more pronounced in less-developed regions compared to already-developed countries.

Infrastructure exerts both direct and indirect influences on economic growth through various channels. The direct impact arises from infrastructure acting as an input in production processes, thereby enhancing the productivity of other essential inputs. However, the effectiveness of these channels varies depending on the type of infrastructure and its role within the economy. Existing literature categorizes infrastructure into three broad classifications: core (economic or hard) infrastructure, social (soft) infrastructure, and territorial infrastructure (Gianpiero, 2010).

Core infrastructure encompasses primary components such as transportation (e.g., roads, airports, ports, and rail networks), energy, and communication infrastructure. These components function in a complementary manner, providing vital support to the economic system and enhancing the productivity of other inputs. Transport infrastructure, as the first core component, plays a crucial role in facilitating factor mobility, trade, and spatial connectivity, thus serving as a key driver of economic development (Redding & Turner, 2015). Similarly, energy, as the second core component, is a fundamental ingredient in production processes. Efficient generation, transportation, and storage of energy are essential infrastructure services that meet the energy demands of an economy. A resilient energy infrastructure, along with fair pricing mechanisms, generates a growth effect. The theoretical and empirical literature extensively supports the proposition of a positive correlation between reliable energy infrastructure and economic development, and the nation's economic and political standing (Bridge et al., 2018). The third core component, communication infrastructure, plays a significant role in reducing geographical barriers and improving connectivity by facilitating the exchange of information. It reduces information asymmetry and links geographically distant markets, thereby reducing business costs and enabling offshore contracting. Empirical evidence underscores the profound impact of communication infrastructure on long-term economic growth (Czernich et al., 2011).

The empirical literature on specific countries provides robust support for the positive association between transport infrastructure and economic growth. For instance, Garcia-Milà & McGuire, (1992) employed a production function approach to demonstrate a significant and profound impact on output generation resulting from highway investments across the contiguous 48 states of the United States from 1969 to 1983. Similarly, Rives & Heaney (1995) conducted a study that linked economic development to both local and network infrastructure in the United States. Local amenities, such as local roads and sewerage systems, were considered community-centric point infrastructures, while network infrastructures like highways, railroads, and canals served not only specific local communities but also larger geographic areas. The study developed an index to measure infrastructure using physical units and indicators such as average sewer capacity, water plant capacity, and the number and length of US highways in different locations. The findings highlighted the strong role of these components of community infrastructure in locational development.

Furthermore, the Chinese experience of economic growth further reinforces the crucial role of core infrastructure in driving regional economic disparities. During the 1990s, infrastructure investments became a national priority in China (Fan & Chan-Kang, 2005). Road projects played a significant part in China's regional development strategy, with substantial investments made in road construction and the expansion of the road network within counties and towns, focusing on improving road quality and increasing expressway mileage. Initially, resources were primarily allocated to coastal regions but later shifted to western regions (Fan & Chan-Kang, 2005, p. 15). Additionally, Démurger (2001) examined the regional disparities using a dataset encompassing 24 provinces of China from 1985 to 1998 The study identified geographical location and infrastructure endowments as key factors contributing to regional growth gaps. Transport facilities were highlighted as a significant differentiating factor in explaining growth disparities, while the role of telecommunications was crucial in reducing the burden of isolation.

Instead of relying on individual variables such as roads or the number of telephones, Calderón & Servén (2004) employed principal components analysis (PCA) to construct synthetic indices representing the quantity and quality of infrastructure. Their approach involved creating an aggregate index using data from the telecommunication, power, and transportation sectors. Analyzing a panel dataset covering over 100 countries from 1960 to 2000, they used a single equation approach and GMM estimators to address endogeneity. The study yielded robust results, demonstrating a positive causal impact of infrastructure stock on economic growth.

Building on Calderón & Servén’s PCA methodology, Sahoo & Dash, (2012) empirically examined the relationship between core infrastructure and economic growth in South Asian countries.

Compared to Calderón & Servén (2004), they employed a broader range of heterogeneous indicators, including electricity power consumption, telephone lines, rail density, air transport, and paved roads Using panel data from four South Asian countries (India, Pakistan, Sri Lanka, and Bangladesh), they estimated a positive and significant output elasticity of infrastructure, ranging from 0.24 to 0.26, using FMOLS panel cointegration estimation approach.

In addition to core infrastructure, social infrastructure, also known as soft infrastructure, plays a complementary role in supporting economic productivity and fostering human capital development. Social infrastructure encompasses facilities for health, education, and cultural services, such as schools, libraries, universities, clinics, hospitals, courts, museums, theatres, playgrounds, parks, fountains, and statues. The theoretical literature highlights that differences in human capital across nations significantly contribute to variations in output per worker (Hall & Jones, 1999). They defined social infrastructure more broadly to include institutions and government policies that enhance productivity and output per worker, which plays a crucial role in this disparity. Moreover, the literature emphasizes the importance of the quality of social infrastructure services in driving long-term economic growth. Hanushek & Woessmann (2008) discovered that cognitive skills have a profound impact on economic development. They observed that the mere existence of social infrastructure, such as schools and school attendance, does not stimulate sustainable economic growth; rather, it depends critically on educational quality. They also revealed a significant skills deficit among individuals in developing countries compared to their counterparts in developed nations.

Despite the crucial importance of soft infrastructure in developing human skills, many developing countries, including Pakistan, struggle with inadequate access to and utilization of social infrastructure. In Pakistan, approximately 18% of the population lacks access to educational infrastructure, while about 33% lack basic health facilities (UNDP, 2016). Theoretical connections suggest that an extensive network of education and health infrastructure is essential for productivity, as it enables the acquisition of higher skills through education and promotes longer life expectancy in developing countries. However, there has been no study to date that measures the role of social infrastructure in the productivity growth of Pakistan at the national level using the growth accounting framework. Nonetheless, firm-specific evidence supports the critical role of social infrastructure in firm productivity. Ahmed (2016) conducted an empirical study examining the impact of social infrastructure indicators at the district level on firm productivity, using firm-level data from Pakistan. The findings reveal a positive and significant relationship between health and education indicators and firm-level productivity in urban settings within the manufacturing industries in Pakistan.

Furthermore, investing in territorial infrastructure is recognized as a significant driver of economic development. Territorial infrastructure encompasses private investments and activities that contribute to locational attractiveness, quality of life, and regional development dynamics (Gianpiero, 2010, p. 16). It includes resources for commerce, tourism, and monetary intermediation. The physical infrastructure of financial institutions, such as banks and credit cooperative societies, plays a crucial role in facilitating financial lubrication and intermediation services. This infrastructure enables the integration of financial markets for trade, commerce, and tourism activities. A reliable and well-connected infrastructure for monetary intermediation generates locational effects, while resilient and efficient financial intermediation is essential for modern economic operations. The financial infrastructure serves as the foundation for financial markets, enabling services such as financial risk management, liquidity provision, information dissemination, and achieving allocative efficiency. It also plays a vital role in providing basic access to financial services, with banks serving as essential providers. The literature extensively documents the positive impact of financial development on economic growth, primarily attributed to increased ease of trading between economic agents (Stulz, 2001). The process of financial development heavily relies on the quality of physical infrastructure, service delivery, and the proper functioning of financial markets.

Particularly in the context of Pakistan, the performance of the financial sector is crucial for long-run economic growth. For example, Hina & Qayyum (2019) conducted a study on the episodes of financial crises and their impact on economic growth in Pakistan from 1972 to 2012. They used indicators such as domestic credit to the private sector as a percentage of GDP and money and quasi-money (M2) as a percentage of GDP to measure the size of the financial capital. The study assessed the effect of financial reforms by analyzing the weighted average lending rate to the weighted average deposit rate and developed a composite index based on banking, currency, and stock market crisis ratios to evaluate financial crises. The data supported the positive role of financial capital in promoting growth while emphasizing the detrimental impact of financial sector inefficiencies on Pakistan's economic development. Moreover, financial inclusion remains a significant policy issue in Pakistan. According to the recent 'Access to Finance Survey (A2FS)

2015', 53 percent of the population remains financially excluded from the formal financial system (SBP, 2018). Rizvi et al., (2017) demonstrated the potential beneficial role of modern communication infrastructure in increasing financial inclusion in Pakistan, particularly with the widespread penetration of mobile telephony and secure data streaming. Empirical studies have highlighted both demand-side and supply-side determinants of financial exclusion. Zulfiqar et al. (2016) examined micro-level reasons for low financial inclusion in Pakistan using data from the Global Findex database (2014) and employed probit estimation techniques. Their findings indicated that low education levels, low income, and a significant gender gap contribute to demand-side factors of financial exclusion. Additionally, Adil & Jalil (2020) found that the demographic outreach of banks is crucial for increasing financial inclusion in Pakistan. Therefore, financial infrastructure plays a significant role in economic growth by fostering financial inclusion and development in Pakistan.

2.3 The Spatial Spillovers and Infrastructure

The recent literature and theory of the New Economic Geography (NEG) justification for an unbalanced regional economic development based on the functional relationship of the economic interactions. NEG recognizes the existence of infrastructural spillovers, a dense labour market, and enhanced market access through inter-firm linkages, which lead to a change in productivity within regions and the size of spatial clusters (Venables, 2010). Empirical studies grounded in the theoretical framework of NEG also acknowledge the crucial role of public infrastructure in fostering economic, and social cohesion and interaction. Specifically, the transport infrastructure has been found to generate sectoral and regional 'spatial spillovers' as a result of regional connectivity and economic integration (Boarnet, 1998; Holtz-Eakin & Schwartz, 1995, Hu & Liu, 2010). Spatial spillovers refer to the effects that extend beyond the geographical boundaries and positive spatial spillovers arise as a result of investment in the transport infrastructure leading to increased market access, labor mobility, and economies of scale (Cohen & Paul, 2004, Cantos et al., 2005, Yu et al., 2013, Hu & Liu, 2010). Conversely, negative spatial spillovers occur when resources flow out from less economically attractive regions to economic centers including the outflow of financial capital, human resources, and other factors inputs (Boarnet, 1998).

The diffusion of positive spillovers takes place through several channels. Firstly, advancements in transport and communication technology facilitate the spread of technological innovations across regions. Secondly, the presence of infrastructural services in different regions enables better trade facilitation and capital inflow (Straub, 2008). Thirdly, neighboring regions upgrade their infrastructure and policies through imitation in response to positive spillovers. Fourthly, inward foreign investment in regions with better infrastructure generates spatial spillovers and encourages factors of production to relocate from regions with lower infrastructure capacity (Rietveld, 1994). A wide range of empirical literature has estimated spatial spillover effects as a result of investment in the transport infrastructure for many countries of the world. For instance, Cohen & Paul (2004) used the firm's level data (U.S. manufacturing firms) and by using the cost-function model for 1982–1996 found the productive spatial adaptations and the positive spatial spillovers that arise due to the cost savings as results of the intrastate public infrastructure. In addition to this, Cantos et al., (2005) estimated the positive transport infrastructure spillovers in Spain from 1965 to 1995 by using accounting and production approaches both. Their results are similar to both methodologies. Therefore, they confirm the networking effects of transport infrastructures on private-sector productivity within Spain. Moreover, Yu et al. (2013) measured the spatial spillover effects of transport infrastructure in the regions of China using spatial econometric techniques. They used Spatial Durbin Model for the time period 1978–2009 and found positive spillovers of infrastructure due to connectivity and networking. At the regional levels, these spillover effects vary across time. They found Eastern China enjoyed positive spillovers all the time. The Northeastern China region had no significant spillover effects in 1978–1990, negative spillovers in 1991–2000, and positive spillovers in 2001–2009. Central China had negative spillovers for the three periods. Western China experienced negative spillovers after the 1990s. It is found that the changes in spillovers among regions are due to the migration of production factors within China during the last decades.

In addition, literature found that transport infrastructure has an important role in achieving a balanced regional development growth objective of countries. For instance, Xueliang, (2013) found an unbalanced regional growth in China, despite her balanced regional growth policy. He used provincial panel data for FY 1993-2009 using four spatial weight matrixes of contiguity matrix, transport network matrix, population density matrix, and per capita GDP and found positive intra-regional infrastructure spillovers in China.

The empirical literature on the spillover effect of infrastructure had provided mixed results for many countries of the world. A list of recent studies that have estimated direct and indirect effects across the globe using spatial econometrics is provided in Table 2.1 The literature for Pakistan is silent so far regarding the role of core infrastructure at the sub-national level, and this study is contributing not only by attempting the direct effects but also by incorporating the geography into account by employing spatial econometric tools of estimation. And so far to the best of my knowledge, there is no study has yet been made attempted to calibrate intra-geographic spillovers of infrastructure at the sub-national level and hopefully generate a discussion in the body of literature in Pakistan.

Apart from this, the role of education infrastructure in building cognitive skills is critical for increasing productivity. In addition to education infrastructure, health infrastructure plays a significant role in economic growth, as health is a critical ingredient that is embedded in the labor force and human capital formation. Cross-country experiences provide ample evidence to support the argument that investing in the health sector promotes economic growth in the long run.

Table 2. 1 *Spatial spillovers of infrastructure and regional growth: empirical literature*

Authors	**Model**	**Key Findings**
Nawaz & Mangla, (2021)	2006–2016 35Asian economies SDPM	The positive and significant effect, both direct and spillover, on regional development. The quality and the complementarity of infrastructure with the institutions and regional integration, act as a stimulus to enlarge the spillover effects of the infrastructure.
Jiang et al., (2016)	1985-2012 Chinese provinces	The results confirm positive and significant non-homogeneous spillover effects are found. Highly positive spillovers are observed between economically similar provinces as results of the industrial reallocation and market expansion contributes to the positive spillovers. In the under-developed provinces of China, high network connectivity often results in low or negative spillovers due to the mobility and migration of production factors.
Álvarez et al., (2016)	1980–2007. Provinces of Spain SAR and SARAR model	Results documented the positive spillover effect of the capital stock of road infrastructure due to the increased trade flows, and access to markets and complementing the other forms of capital.
Isaev, (2015)	2000-2013 Russian-Regions SAR	Road infrastructure has a positive and direct influence on regional growth, but the impact of the railroad infrastructure on the regional economy depends on the impact of the congestion effect.

		Road infrastructure generates a negative spillover effect representing that rapid road infrastructure development in some regions moves mobile factors of production away from adjacent regions retarding their economic development. The spillover effect of railroad infrastructure is significant and negative again only if the congestion effect is considered.
Chen & Haynes, (2015)	1997 to 2011 SDM 48 Contiguous US states and social accounting matrix (SAM)	Significant and positive indirect effects are found in sectors of manufacture, utility and construction, truck, transit, and pipeline, which indicate that wage-rental ratios from adjacent regions have positive impacts on the local region itself. The results further confirm the existence of spatial dependence among these sectors.
Alvarez-Áyuso & Delgado-Rodriguez, (2012)	1980 to 2008 Spain regions fixed-effect model; the neighborhood: average public capital of neighboring regions.	Results show a positive and significant effect of high-capacity roads (HCR) on the private sector through spillover effects derived from the network infrastructure.
Cohen, (2010)	1996 US state manufacturing Kelejian and Prucha's (1998) spatial model	Ignoring the broader effects of a spatially lagged dependent variable can lead to underestimating infrastructure elasticity and incorrect estimation of the overall productive impacts of public infrastructure.
Yu et al., (2013)	1978-2009 Chinese regions SDM	At the regional level, transport infrastructure spillover effects vary over time among China's four macro-regions: the eastern region enjoyed positive spillovers all the time; the northeastern region had no significant spillover effects in 1978-1990, negative spillovers in 1991-2000, and positive spillovers in 2001-2009; the central region had negative spillovers for the three sub-periods; for the western region, negative spillovers can be observed after the 1990s. The changes in spillover are mainly due to the migration of production factors in China.

2.4 Social Welfare and Infrastructure

The theoretical literature distinguishes the impact of infrastructure on income, economic growth, and the impact on wealth distribution. Ferreira (1995) introduces capital market imperfections in the production function that includes public capital as complementary to private capital with three different social classes — subsistence workers, middle-class entrepreneurs, and upper-class entrepreneurs. He used the general equilibrium model for wealth dynamics. He described if public capital provided free of the cost falls below a minimum level, the middle class disappears, and a decrease in the level of public investment results in higher levels of inequality. Lower-income households will have no access to infrastructure, whereas "private infrastructure-owning" upper-class entrepreneurs will benefit more. This study also theoretically proved that an increase in productive public investment has a negative impact on distribution (inequality) but that results in greater output (production) and advocates increasing government infrastructure provisioning. In this addition to this, Getachew (2010) developed a joint theory of public capital, inequality, and growth, in a two-sector growth model that yields complete analytical solutions. This model finds public capital plays an important role in long-run growth by enhancing productivity and complementing the accumulation of private inputs. This model predicts inequality is bad for growth in the presence of credit market imperfection. Certain public services and investments may benefit the poor more than proportionally and thus improve the distribution of income, and hence, improve economic growth through an indirect channel. The key mechanism linking the distribution of income to public capital is its disproportional effect on the economy that affects factor shares of capital.

In continuation to the previous models; Chatterjee & Turnovsky (2012) incorporated the time dimension, and also distinguish between the impact of public capital on the distribution of pre and post-tax income and distribution of wealth in a general equilibrium framework. Three main conclusions are drawn from their analysis. First, government spending on public capital leads to a persistent increase in wealth inequality over time, regardless of how it is financed. Second, the wealth inequality generated by government spending depends critically on externalities and allocation decisions, financing policies, and time of consideration (short run, transition path, or the long run). Third, government expenditure on infrastructure improves average welfare and increases its dispersion.

The literature evidences the spatial spillovers of transport infrastructure for productivity growth but few studies that take locational the specific effect of the infrastructure into the analysis of social welfare except Haughwout (2005). He acknowledges the role of infrastructure as the main contributor to the quality of life and highlighted the difficulty in the measurement of utility that people drive from the quality-of-life benefits from the core infrastructure.

The literature on social welfare signals the key importance of road infrastructure in reducing poverty and economic depreciation, particularly in developing countries. For instance, Ali, et al (2014) used multiple indicators of welfare such as crop and livestock revenue, non-agricultural income, MPI poverty reduction, and wealth index as the indicators of welfare in Nigeria. They found 10 percent reduction in transport costs increases welfare by increasing crop revenues by 9.7%, non-agriculture income by 4.6 %, wealth index by 2.1%, and MPI reduced by 2.4% in Nigeria. In addition, Fan et al. (2000, 2002, 2004) estimated the effect of infrastructure investments on economic growth and poverty in rural India, China, and Thailand. By estimating a system of equations, to explicitly account for the simultaneous effects of infrastructure investment in the factor and product markets. Results from these studies consistently show the importance of road investments in promoting productivity growth and poverty reduction in the countries of India, China, and Thailand.

In the context of Pakistan, literature signifies the role of investment in infrastructure is a key factor of welfare improvement and it is a macroeconomic determinant for inequality reduction in Pakistan (Jamal, 2006). So far, there is no literature available (as per best of knowledge) that has accounted for the spatial spillover effects of transport infrastructure in the analysis of the social welfare of Pakistan (district (meso-level)). Therefore, this essay is contributing to the body of literature by estimating through spatial econometric modelling to gauge the direct and indirect spatial spillover effect of transport infrastructure in determining the landscape of social welfare at the district level in Pakistan.

2.5 The research gap

A research gap exists, particularly in empirical research focusing on Pakistan. This study aims to address this gap by evaluating the long-term effects of different types of infrastructure on productivity growth. The infrastructure types considered in this study are classified as core (transport, energy, and communication), social (education and health), and territorial (financial) infrastructure. Previous literature on Pakistan primarily focuses on transport infrastructure, thus limiting our understanding of the broader impact of different infrastructure types. By examining the role and collective aspects of each infrastructure type, this study aims to provide policy-oriented insights and contribute to the existing body of literature. To the best of our knowledge, no similar study has been conducted for Pakistan thus far.

CHAPTER 3

AN OVERVIEW: A SITUATIONAL ANALYSIS

3.1 Introduction

Quality infrastructure is a fundamental requirement for building a modern economy. However, Pakistan faces challenges in meeting this requirement due to its high population density and comparatively lower infrastructure stock compared to developed and fast-growing developing countries. The Logistics Performance Indicator (LPI), which benchmarks trade logistics across 160 countries, ranked Pakistan poorly in infrastructure in 2018, placing it at 122nd position, while neighbouring countries China and India ranked much higher at 44th and 26th positions, respectively. As a result, Pakistan has suffered an annual GDP loss of approximately 4 to 6 percent (around $6 billion) and faced a 30 percent increase in production costs due to inadequate infrastructure services (SBP, 1996). Consequently, substantial fiscal space is required for infrastructure development, given the heavy financial investments required. This further exacerbates financing risks and debt burden, considering Pakistan's macroeconomic management challenges and limited fiscal space (IMF, 2018).

Since 2013, Pakistan has actively engaged in international collaboration with China through the China-Pakistan Economic Corridor (CPEC), a flagship project of the Belt and Road Initiative (BRI). CPEC aims to enhance connectivity, including the construction of an integrated transport system, the development of information network infrastructure, and the strengthening and building of energy infrastructure.

This chapter provides a brief overview of the current trends and existing policies for each component of infrastructure in Pakistan. The first subsection presents an overview of core infrastructure, the second subsection provides insights into social infrastructure, and the last subsection presents a trend analysis of financial infrastructure in Pakistan from 1980 to 2022.

3.2 Core Infrastructure in Pakistan

Core infrastructure encompasses transport, energy, and communication networks, playing a critical role in the production of goods and services within a country. In Pakistan, the transport, storage, and communication services sector directly contributes to approximately 13 percent of the country's gross domestic output over the last decade (2010-2020)[3].

3.2.1 The transport infrastructure in Pakistan

Transport infrastructure plays a vital role in facilitating the movement of goods and services, reducing transaction costs, and promoting economic integration. It acts as a crucial component, enhancing mobility and the efficiency of factor inputs within an economy. Efficient transport infrastructure not only directly contributes to infrastructural-related services but also boosts the productivity of other related services. Its significance lies in enabling people's mobility, integrating geographically dispersed markets, and reducing social marginalization. Moreover, it, directly and indirectly, contributes to the national output, with indirect effects observed through linking rural areas to markets and connecting local markets to global supply chains, thereby facilitating value-addition processes.

By 2030, urbanization in Pakistan is expected to lead to a significant increase in travel demand, estimated at 1000 billion per kilometer (GoP, 2018). To address this, the Government of Pakistan (GoP) is investing substantial financial resources in the transport sector through major projects such as the Belt and Road Initiative (BRI), the Central Asia Regional Economic Cooperation (CAREC) program, and other initiatives. These projects aim to improve geographical connectivity and strengthen the existing transport network. Under the China-Pakistan Economic Corridor (CPEC), road construction projects have been completed and operationalized, while others are in progress or the planning phase. Due to limited fiscal space, road infrastructure development is carried out with the assistance of development financiers such as the Asian Development Bank (ADB), Korean Exim Bank Assistance (KEBA), and the World Bank (WB). Furthermore, there are projects for road infrastructure upgrade and construction under the Build-Operate-Transfer (BOT) and Public-Private Partnership (PPP) agreements (Ministry of Finance, 2021).

Over the past six decades (1961-2020), Pakistan has experienced relatively lower average growth rates of GDP and GDP per capita, approximately 5% and 2.3%, respectively, compared to China and India (Table 2.1). The lack of adequate infrastructure has been a major hindrance to achieving sustained economic growth in Pakistan in recent decades (SBP, 2007). In response to this

[3] Author's calculation based on the data available in various issues of the Pakistan's economic .

challenge, the National Transport Policy (NTP) was promulgated in 2018 to address the structural problem of inadequate transport infrastructure in the country (GoP, 2018).

Table 3. 1 *Growth rates of real GDP and real GDP per capita (percent)*

	Pakistan		India		China	
	GDP	GDP per capita	GDP	GDP per capita	GDP	GDP per capita
1961-70	7.24	4.53	4.03	1.88	4.96	2.71
1971-80	4.72	1.68	3.08	0.74	6.24	4.33
1981-90	3.96	1.09	5.60	3.61	10.45	9.28
1991-00	3.96	1.09	5.60	3.61	10.45	9.28
2001-10	4.28	1.89	6.75	5.10	10.57	9.93
2011-20	3.85	1.71	5.02	3.85	6.84	6.34
1961-2020	**5.06**	**2.30**	**5.01**	**3.07**	**8.07**	**6.73**

Source: Author's calculations, the data source is WDI.

There are four types of physical infrastructure for transportation services, including land (roads, rail), air, and seaports (maritime), utilized in Pakistan for passenger and freight movement. Each type of infrastructure is discussed in detail in the following sub-sections.

3.2.1.1. The infrastructure of the road transport network

Road infrastructure plays a crucial role in facilitating inland transportation services throughout Pakistan. However, the road classification system is not standardized across the country. Different road management authorities have developed their classification systems based on their specific requirements over time. For example, the National Highway Authority (NHA) is responsible for maintaining federal roads, which include Motorways, Expressways, and National Highways, including Strategic Highways as shown in Figure 3.1. On the other hand, provincial authorities in each province adopt various road classifications, as outlined in Table 3.2. These classifications are based on the type of road material used and road operational ownership (NTRC, 2020).

Table 3. 2 *Road Classification in the Provinces of Pakistan*

Punjab	Sindh	KPK	Baluchistan	AJ&K	GB
Provincial Highways R&B sector Farm-to-market road Sugar-Cess roads District council roads	High type Low type	Provincial Highways High type Low type	High type Low type	Public Works Dept. Roads Double Lane Roads Major Roads Links Roads Local Government roads	Blacktop roads Shingle Roads

Source: NTRC (2020). *The digitalization of Roads Directory in the Country, Final Report*.

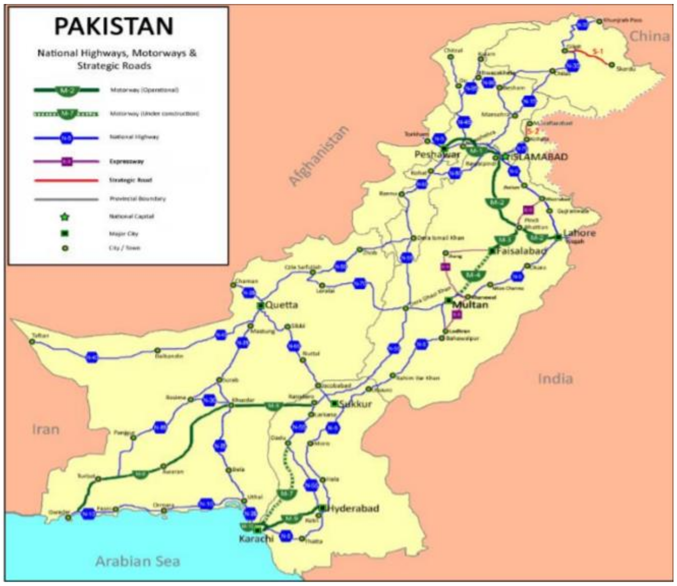


Figure 3. 1 The Federal Road network

Source: NTRC (2020). The digitalization of Roads Directory in the Country Final Report.

Following the initial classification, the entire road network is further divided into two categories: low type and high type. The National Transport Research Centre (NTRC) has gathered data on road length, encompassing Kacha roads, gravel/shingle roads, and unsurfaced roads as low-type roads, while metaled roads are considered high-type roads. However, in 2019-2020, the previous classification of low and high types was replaced with different categories during the digitalization of the road directory and the development of the "Pakistan Geo Directory Road Portal." This change was implemented after ground verification, resulting in the following new classification developed by NTRC (2020);

- Motorways: These high-speed roads provide uninterrupted travel with controlled access. They feature dual carriageways and are designed for speeds typically ranging from 100 to 120 kilometers per hour (KPH). Examples include M1, M2, etc.
- Expressways: These multiple-lane high-speed toll highways are upgraded versions of National highways and have fewer access restrictions.
- Highways: Highways offer largely uninterrupted travel between cities and districts with full access. They are designed for speeds ranging from 70 to 100 KPH.
- Primary Roads: These roads collect traffic from Motorways/Highways and distribute it to Secondary Roads. They have a moderate speed range of 60 to 70 KPH.

- Secondary Roads: Secondary Roads connect local Roads with Primary Roads and vice versa, with speed limits ranging from 40 to 60 KPH.
- Local Roads: Local Roads have the lowest speed limits and carry low traffic volumes. They connect with secondary or primary roads at speeds of 40 KPH or less, and in some areas, they may be shingle or unpaved.

These new road classifications provide a more comprehensive framework for understanding the different types of roads in Pakistan and their corresponding speed ranges.

Table 3. 3 *Road infrastructure: network size and the utilizations (1947-2022)*

FY	**Total road**	**High type road**	**Low type road**	**Vehicles on Road** (000)
1947-48	50367	9809	40558	-
1950-51	55376	11730	43646	-
1960-61	66236	16860	49376	-
1970-71	73006	24776	48230	191.7
1980-81	93960	38035	55925	760
1990-91	170823	86839	83984	2120
2000-01	249972	144652	105320	4471
2010-11	259463	180866	78597	10443.8
2022-23*	501,165	127,640	373,525	35499.10

Data source: Pakistan's Economic Survey, various issues.

* *For comparison with the previous year's data, the road length of expressways, highways, metro*

roads, motorways, national highways, primary roads, and secondary roads as high type, and local roads as low type roads.

3.2.1.2 The Infrastructure of the Rail Transport Network

Pakistan possesses a substantial railway network (Figure 2.2), but it has not fully capitalized on its potential for integration into the national transport network (GoP, 2018). Unfortunately, the railway network has undergone a reduction in size over time (Table 3.3). Presently, the total length of the railway network stands at 7,791 kilometers, accounting for a mere 2% and 5% share of the freight market and passenger market, respectively. This underperformance can be attributed primarily to insufficient investment in infrastructure services and operational mismanagement within the Railways sector (GoP, 2018).

Table 3. 4 *Rail infrastructure: network size and the utilization: 1948-2022*

FY	Routes (km)	Locomotive (nos)	Freight wagon (nos)	Passenger travelled (mln)	Fright carried (mln ton)
1948-49	8553	821	23849	72	6
1951-60	7711	767	22490	84	10
1961-70	6871	806	27439	103	12
1971-80	7930	895	33221	127	12
1981-90	7897	786	32145	86	10
1991-00	7799	587	24377	60	6
2001-10	7012	503	18339	70	6
2022-23	7791	467	14,448	28.40	8.20

Data source: Pakistan's Economic Survey, various issues.

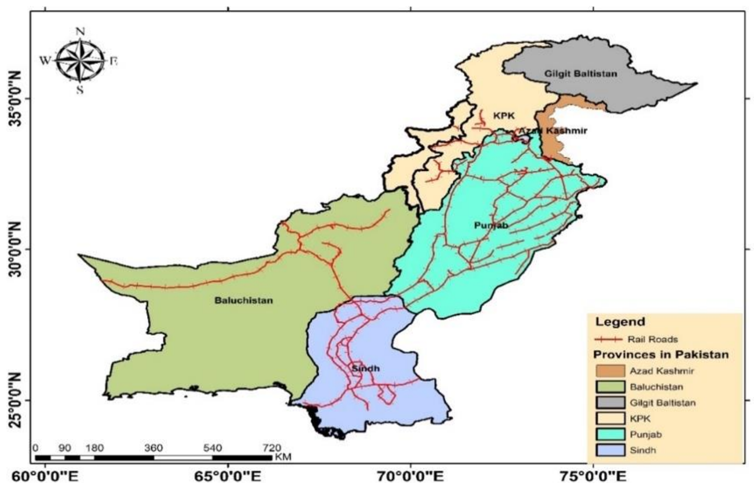


Figure 3. 2 **The existing rail network in Pakistan**

Data Source: Developed in DIVA-GIS software (by author)

3.2.1.3 The infrastructure of maritime transport

Maritime transport encompasses port infrastructure, maritime infrastructure, and shipping lines, all of which contribute significantly to facilitating international trade. Presently, Pakistan has three functional ports: Karachi Port, Port Qasim, and Gwadar Port. The utilization of maritime transport has progressively grown over time (Table 3.5), indicating advancements and increased investments in maritime infrastructure. Additionally, the Pakistan National Shipping Corporation (PNSC) has observed an increase in the transportation of deadweight (liquid cargo) over the past seven decades, indicating an expansion in maritime transportation capacity.

Table 3. 5 *Utilization of Port Infrastructure:1960-2022*

FY	Shipping: Deadweight (Tonnes)	Total cargo handled (Metric tons)
1960-61	4574.03	293256
1970-71	9450	679692
1980-81	14654	580225
1990-91	18710	494956
2000-01	39569	261836
2010-11	68075	633273
2022-23	82484.7	861669

Data Source: Pakistan's Economic Survey, (various issues).

3.2.1.4 The infrastructure for Air-transportation

The Pakistan Civil Aviation Authority (PCAA) serves as the governing body responsible for the provision of essential infrastructure for air transport services and operations in Pakistan. The authority makes a significant contribution to the national exchequer through direct and indirect taxes, amounting to approximately Rs 15-20 billion annually (Ministry of Finance, 2021). In order to enhance air transport services, several ongoing projects are aimed at improving the country's air transport facilities. Recently, the Federal Cabinet approved the National Aviation Plan (NAP) 2019, which outlines strategic initiatives for the sector. The data analysis indicates an upward trend in the utilization of air transport, particularly for passenger travel (Table 3.6).

Table 3. 6 *Utilization of Air-transport Infrastructure:1970-2022*

Year	freight (Million ton-km)	Passengers carried (Nos)	Registered carrier departures (Nos)	PIA route kilometer (000)
1970-71	70	1335900	44300	90555
1980-81	235	3029200	45300	205996
1990-91	421	5180200	66100	255336
2000-01	340	5293541	63956	317213
2010-11	333	6588114	64932	424570
2022-23	193	7420378	49749	374,054

Data Source: WDI and Pakistan's Economic Survey (various issues).

3.2.2 The Infrastructure for energy

Energy infrastructure plays a crucial role in meeting the energy requirements of a country. In Pakistan, energy infrastructure encompasses the necessary facilities for electricity generation and the extraction of gas, oil, petroleum products, and coal. While the exact physical count of plants and equipment involved in these extensive energy networks is unavailable, we utilize the generation and extraction capacity of these energy sources as a proxy for energy infrastructure in the country. The energy network comprises three primary components: generation, transmission, and distribution.

The generation system encompasses the physical facilities that convert various energy resources such as coal, oil, gas, uranium, hydro, and renewables into electricity and other usable forms. The transmission system transports the generated electricity to the load centers, ensuring efficient distribution. The distribution system facilitates the actual connection to individual customers, enabling them to consume electricity on demand.

The management of the electric system in Pakistan has undergone evolutionary changes. The Water and Power Development Authority (WAPDA), along with power generating companies known as Generation Companies (GENCOs) and Independent Power Producers (IPPs), are responsible for electricity generation in the country. With the exception of Karachi Electric (K-Electric), the National Transmission & Dispatch Company (NTDC) is the public entity entrusted with electricity transmission. As of the fiscal year 2020-21, NTDC manages 61 grid stations encompassing a transmission line spanning 17,292 km with a capacity of 56,486 MVA.The distribution of electricity to end-users is carried out by distributing companies (DISCOs), which purchase electricity from WAPDA, GENCOs, and IPPs and distribute it within their respective areas. While all GENCOs are publicly owned, K-electric is privately owned.

Pakistan's energy demand is experiencing substantial growth while also facing supply-side shortages. These energy bottlenecks have had a detrimental impact on the country's economic performance (Ministry of Finance, 2020). Over time, Pakistan has made significant improvements to its electricity generation, transmission, and distribution infrastructure. The installed capacity for electricity generation has increased from 1,862 MW in 1971-72 to 41,557 MW in 2022-23 (Table 3.7). Thermal power generation constitutes approximately 60% of the installed capacity, followed by hydro-generation at around 30%, with nuclear power contributing only 4%.[4]

To address the severe energy shortfall experienced during the fiscal years 2013-18, Pakistan added a cumulative capacity of 12,230 MW to the energy generation system. However, this sudden capacity expansion resulted in an upsurge in energy prices due to increased demand and supply shortfalls. Furthermore, despite the increase in generation capacity, congestion in transmission and distribution systems, along with system losses and inefficiencies, have hindered the consistent delivery of energy services (Ministry of Finance, 2020).

---

[4] *Author's calculation, data source; NTDC, Power system statistics, 45 editions.*

Table 3. 7 *Electricity generation capacity and network: 1980-2022*

FY	**Electricity**		**Electricity Transmission & Distribution lines (T&D)**	
	Installed Capacity (MW)	Generation (GwH)	Length (km)	loss (GwH)
1980-81	4105	16062	15847	-
1990-91	8356	41042	29007	6688
2000-01	17498	68117	42544	13762
2010-11	22477	94653	45509	18237
2022-23	41,557	144,099	56,749	20,822

Data Source: Power System Statistics 46th and 45th edition, NTDC and Ministry of Energy

Pakistan relies on the extraction of coal, natural gas, and crude oil to fulfill its energy requirements. The country possesses substantial natural gas reserves, estimated to be around nineteen trillion cubic feet (Tcf), primarily located in the province of Baluchistan. Additionally, Pakistan has developed a comprehensive gas infrastructure network, as indicated in Table 3.8, which encompasses 20,768 km of electricity transmission and distribution lines, a 12,971 km gas distribution network, 139,827 km of gas transmission lines, and 37,058 km of service gas pipelines. Although Pakistan has the necessary infrastructure for extracting local energy deposits, including gas reservoirs, coal, and oil, its natural gas resources are depleting due to the escalating and ever-increasing energy demand within the country. Consequently, Pakistan is compelled to import natural gas, coal, oil, and petroleum products to bridge this demand-supply gap.

Table 3. 8	*Production of Crude Oil, Gas, Petroleum Products and Coal: 1980-2022*			
	Crude Oil	**Gas**	**Petroleum Products**	**Coal**
FY	Local Extraction (000 barrels)	local Production (mcf)	local Production (000 tons)	local Production (000 tons)
1980-81	3554	299803	3994	1577
1990-91	23487	518483	6036	3054
2000-01	21084	857433	8337	3095
2010-11	24041	1471591	8911	3450
2022-23	20768	962397	8181	8842

Data Source: Economics Survey of Pakistan, various issues.

The utilization of energy infrastructure in Pakistan has witnessed a significant increase over time. This growth is reflected in the rising consumption of coal, electricity, and gas. However, the average growth rates vary for each energy source across different decades.

Table 3. 9 *Utilization of Energy Infrastructure (commercial energy Consumption)*

	Oil/Petroleum (tons)	**Gas (mm mcf)**	**Electricity (GwH)**	**Coal (000 metric ton)**
1980-89	6972383	339221	19354	2209
1990-99	14714176	581645	40589	3325
2000-09	16574945	1094524	62912	6955
2010-19	21082230	1314918	91430	11635
		Growth rates		
1980-89	9.8	5.9	10.9	8.5
1990-99	6.7	4.9	4.2	0.9
2000-09	1.1	6.1	5.1	11.0
2010-19	-0.3	0.4	4.3	10.9

Data Source: Pakistan's Economic Survey, various issues.

3.2.3 The infrastructure for telecommunication services

Initially, Pakistan's basic telecom services were provided by the state-owned Telephone and Telegraph (T&T) Department, which also acted as the regulator. However, in 1994, the department was transformed into a corporation, and an independent regulatory authority, the Pakistan Telecommunication Authority (PTA), was established. Recognizing the need for investment in cellular mobile telephony services, the Government of Pakistan decided to promote competition in the telephone and cellular market starting from FY 2003-04 (Telecomm Industry Report, 2003-04). Since then, the communication network in Pakistan has expanded significantly and is primarily operated by the private sector. Detailed data trends can be found in Table 3.10.

Table 3. 10 *Utilization of the infrastructure for communication: 1980-2022*

	Number of Post Offices			Number of Subscribers	
	Urban	Rural	Telephone (000)	Mobile phones (000)	Broadband (000)
1970-71	1271	6635	160.1	-	-
1980-81	2445	8793	358.8	-	-
1990-91	1867	11546	1188	-	-
2000-01	2302	9932	3252	743	-
2010-11	1580	10455	5720	108895	1492
2022-23	1,742	8,282	2,540	191,625.9	114,341.3

Source: Pakistan's Economic Survey, various issues

3.3 Social Infrastructure in Pakistan

Social infrastructure, also referred to as soft infrastructure, encompasses the facilities that strengthen and support a nation's social fabric. Examples of soft infrastructure include educational facilities, healthcare services, social protection systems, and law enforcement agencies. In this study, we consider education and health services as proxies for Pakistan's stock of social infrastructure.

3.3.1. The infrastructure for education

Since independence, there has been a gradual increase in the stock and utilization of social infrastructure for educational services. The data trends can be found in the tables presented below (Table 3.10).

Table 3. 11 *Physical and utilization of the infrastructure for formal education services*

	Primary schools (I-V)		Middle School (VI-VIII)		High School (IX-X)		Technical Vocational Institutions	
	No's (000)	Enrollment (000)	No's (000)	Enrollment (000)	No's (000)	Enrollment (000)	No's	Enrollment (000)
1947-48	8	770	2	221	0	58	46	4
1960-61	21	2060	2	449	1	160	109	15
1970-71	44	3960	4	933	2	336	206	35
1980-81	59	5474	5	1412	3	509	231	40
1990-91	114	10837	9	2821	8	1004	725	90
2000-01	148	14105	25	3759	15	1565	630	83
2010-11	156	18063	42	5644	25	2630	3224	281
2022-22	181	23704	47	7642	32	4015	3873	444

	Higher Secondary/Inter colleges		Degree Colleges		Universities	
	No's	Enrollment (000)	No's	Enrollment	No's	Enrollment
1947-48	40	14	0	-	2	644
1960-61	131	71	42	12921	4	5084
1970-71	314	199	73	37245	7	17057
1980-81	433	270	99	55897	19	42688
1990-91	612	630	99	75786	22	61857
2000-01	1710	582	366	305200	59	124944
2010-11	3435	1188	1558	431180	135	1107682
2022-23[5]	5770	1742	1667	602681	207	1782853

Data Source: Handbook of Statistics, SBP and Pakistan's Economic Survey 2021.

3.3.2 Infrastructure for health services

The total stock of social infrastructure for health services has also witnessed a gradual increase since independence. The data trends can be found in the tables 3.12 provided below. However, it is important to note that despite the growth in the number of hospitals and dispensaries, the basic infrastructure, such as the number of beds per population, has been decreasing. This indicates a relatively low investment rate in comparison to the population growth rate of Pakistan.

Table 3. 12 *Health infrastructure in Pakistan*

	Hospitals (No)	Dispensaries (No)	MCH Centers (No)	Total Beds (No)	Bed per population
1947-48	292	722	91	13769	2564
1950-51	304	807	107	14524	2431
1960-61	342	1195	348	22394	2038
1970-71	411	1875	668	28976	2061
1980-81	602	3466	812	47412	1716
1990-91	756	3795	1050	72997	1444
2000-01	876	4635	856	93907	1456
2010-11	972	4842	909	104137	1701
2019-20	1279	5671	747	132227	1608
2022-23(P)	1,276	5832	781	146,053	1584

Data Source: Handbook of Statistics, SBP and Pakistan's Economic Survey, 2022.

3.4 Financial infrastructure in Pakistan

Financial infrastructure refers to the institutions dedicated to handling financial transactions with the public. The commercial banking network plays a crucial role in facilitating these transactions

[5] Extrapolated (average of previous three years)

through lending and deposit facilities. Financial integration fosters economic cohesion and promotes productivity growth. Following the separation of East Pakistan, the ruling regime at the time, the Pakistan People's Party (PPP), conducted a nationwide nationalization process from 1971 to 1977, which included the banking sector. Since the 1990s, the banking sector in Pakistan has been a regulated industry, with attempts made to address structural inefficiencies through regulation and privatization (Iimi, 2003).

Over time, the network of financial infrastructure has expanded, particularly after privatization, deregulation, and foreign direct investment (FDI) in the telecommunication and banking sectors. This expansion has led to greater accessibility of financial services for the public, thanks to the installation of automated teller machines (ATMs) and real-time banking transactions. These advancements have made basic financial transactions, such as cash withdrawals and fund transfers, easier and available 24/7. Consequently, the average growth of the financial infrastructure has increased exponentially between 2004 and 2020. The number of bank branches has increased by over 4 percent, while the number of ATMs has increased by more than 22 percent, in contrast to the negative average growth observed in the banking network during FY 1981-2003. Furthermore, the utilization of financial infrastructure has also experienced a significant increase during the period of 2004-2020 compared to 1981-2003. Please refer to Table 3.13 for detailed data trends.

Table 3. 13 *State of the infrastructure for financial intermediation in Pakistan*

FY	**Bank Branches (no)**	**ATM machines (no)**	**Account holders (no mln)**
1980	7076		15.4
1990	7404	-	26.8
2000	7949	-	28.4
2010	9,362	5,200	28.1
2020	14849	16175	62.0
2022-23	17,516	17,678	67.52
Average Growth rates (%)			
1981-2003	-0.05	-	2.89
2004-2020	4.93	22.39	5.38

Source: Statistical Handbook of SBP, Payment Review statistics 2019, 2020, 2023

CHAPTER 4

THEORETICAL FRAMEWORK

4.1 Introduction

Economic growth is a prerequisite for modern long-run economic development. The nations with more pro-poor and redistributive economic growth policies make it possible to uplift the standard of living of their inhabitants. Therefore, public sector policies and programs directly and indirectly, impact the production capacity and livelihood of inhabitants of a nation. Governments had to enhance their productive capacity to generate output over time for a pro-poor economic growth agenda.

The theoretical literature had segregated capital further into physical capital and human capital as the source of long-run economic growth, besides the traditional factor inputs of the physical capital and labor. There is an ample literature available that had measured the long-run impact of public capital on economic growth. However, the debate spurs due to the estimation techniques and the indicators used to calibrate the public capital. Earlier studies used public capital stock data in monetary terms i.e., the book value of capital expenditures. However, recent studies using the physical indicators for measuring public capital are commonly termed "infrastructure" instead of public capital stock. This study aims to employ broader indicators to bring forth a better and deeper understanding of growth-infrastructure linkages in Pakistan. In this connection, this chapter provides a detailed explanation and theoretical construct for this research study.

4.2 Theoretical model for the Growth effect of Infrastructure

Arrow & Kurz (1970) and Weitzman (1970) initiated the theoretical construction of the infrastructure-growth relationship. They formally incorporate irreversible public capital investment into their models. Later on, Barro (1990) analyzed the role of tax-financed government investment in the endogenous growth model. The infrastructure had a constructive or foundational role in economic development and is an empirical question. The debate exists based on the outcomes of the findings of the empirical research, hence bringing forth the mixed results.

Aschauer debunked the debate of role of infrastructure in the productivity growth. The Aschauer (1989b) had revealed the correlation between the 'core infrastructure' (highways, parks, water and sewer systems…) with the private-sector productivity in United States. Moreover, Aschauer (1989a) found complementarity of public investment in infrastructure and evidence for the crowding-in instead of crowding-out of private investment. Despite the importance of infrastructure and its inextricably connectedness with economic growth, the impact of the direct and indirect effect on the infrastructure remained a conundrum. It is because soon after publication of Aschauer ground-breaking work, the academic debate heated over the quantification of impact of government expenditures on the output and productivity (Munnell, 1990). The large part of criticism was on the econometric estimation techniques of aggregate time series with the non-stationary data, endogeneity issue and measurement of infrastructure in spending terms (Felipe, 2001).

The theoretical underpinning of the public capital in the structural growth model framework has been developed by as component for the economics has been developed by Holtz-Eakin & Schwartz (1994) using neoclassical growth framework. Instead of using stock value of public capital in monetary units Canning (1999) measured infrastructure in physical units in per worker terms. Using insight from Canning (1999) and Canning & Pedroni (2008) of measuring infrastructure in physical units, and this study is using Mankiw et al. (1992) growth model framed for human capital which is an extended model of Solow-Swan model (1956) in the standard neoclassical growth paradigm. The neoclassical models are widely used for the empirical validation of the cross-country convergence hypothesis (Barro & Sala-i-martin, 1992).

We include public (core) infrastructure as an input for the production of goods and services, besides other forms of capital i.e. private capital stock and human capital, and labor.

$$Y_t = K_t^{\alpha} . \acute{H}_t^{\beta} . G_t^{\gamma} . (A_t L_t)^{(1-\alpha-\beta-\gamma)} \tag{4.1}$$

The model is based on a standard Cobb-production function. The symbols Y_t represents the real GDP, G_t represents the stock of core infrastructure, K_t represents the stock of private physical capital and $\dot{H}_t$ represents stock of human capital (conditional on the efficiency and provision of the social infrastructure). Hall & Jones (1999) defines the social infrastructure as the entity comprised of set of institutions and government regulations and policies to enable individuals and firms to enterprise the ideas to innovate to spur the productivity growth. Therefore, the interplay of social infrastructure and human capital critical are the driver of productivity growth. We define human capital based on Hall & Jones (1999) approach as $\acute{H}_t = e^{\emptyset(E_t)}H_t$ where $\emptyset$ is representing the efficiency of social infrastructure while E_t is the physical stock of the social infrastructure. While A_tL_t represents the labor augmented technology and 't' represent time. The production functions exhibit constant returns to scale in all its inputs capital, infrastructure, human capital and technology augmented labor force. Whereas α is income share of capital, β is the income share of human capital and γ is the income share of infrastructure in output. While, A_t represents the labor augmenting technological change, that's follows the path as;

$$A_t = A_o e^{\lambda t} \tag{4.2}$$

$$L_t = L_o e^{nt} \tag{4.3}$$

Where λ is exogenous technological growth and n is. population growth rate at which labour-force grows. The production function (equation 4.1) can be expressed as in effective-labor intensive units.

$$\frac{Y_t}{A_tL_t} = \left(\frac{K_t}{A_tL_t}\right)^{\alpha} \left(\frac{\acute{H}_t}{A_tL_t}\right)^{\beta} \left(\frac{G_t}{A_tL_t}\right)^{\gamma} \tag{4.4}$$

To express the production function in labor intensive units:

$$y_t = \frac{Y_t}{A_tL_t}, k_t = \frac{K_t}{A_tL_t}, \acute{h}_t = \frac{\acute{H}_t}{A_tL_t}, g_t = \frac{G_t}{A_tL_t}$$

The production function in per effective worker form is;

$$y_t = {k_t}^{\alpha} {\acute{h}_t}^{\beta} {g_t}^{\gamma} \tag{4.5}$$

Dynamics of the Model:

The accumulation of private physical capital, human capital, and infrastructure occurs gradually over time due to positive net investment (new investments minus depreciation). Existing literature extensively illustrates the significant correlation between financial development and economic growth, as documented by scholars such as Goldsmith (1969), McKinnon (1973) and Shaw (1973). Furthermore, Levine (1997) delved the functional approach of the financial system and the nexus with economic growth. He outlined five key functional channels that facilitate capital accumulation and technological growth. These channels encompass the mobilization of savings, efficient resource allocation, the exertion of corporate controls, risk management, and the facilitation of trading and contracts. It is therefore, we include the financial infrastructure as a critical driver of financial system, that facilitates trade and transaction leading efficiency improvements of existing capital stock in the model. $I = s_f S$ whereas s_f denotes the share of new investment in the financial infrastructure to create economic cohesion in the economy. Therefore, in this study we introduce investment in the financial infrastructure (s_f) as a driver to promote the net investment by private sector and infrastructure (see equation 4.6 & 4.8). Additionally, the net additional investment in the human capital stock occurs via two key investment channels ($s_{\acute{h}} = s_s + s_h$) using Hall & Jones (1999) approach. First, is through new investment in the physical form of social infrastructure and soft form for such as spending on the education, skills, and training in the social sector to develop human capital stock.

$$\dot{K} = s_f . s_k Y_t - \delta K_t \quad (4.6)$$

$$\dot{\acute{H}} = s_{\acute{h}} Y_t - \delta \acute{H}_t \quad (4.7)$$

$$\dot{G} = s_f . s_g Y_t - \delta G_t \quad (4.8)$$

Inserting (4.4) into (4.6);

$$\dot{k_t} = \frac{\dot{K_t}}{A_t L_t} - \frac{K_t}{A_t L_t}\left(\frac{\dot{A}}{A} + \frac{\dot{L}}{L}\right) \quad (4.9)$$

$$\dot{k_t} = \dot{h_t} = \dot{g_t} = 0, \ \frac{\dot{A}}{A} = \lambda, \frac{\dot{L}}{L} = n$$

Private capital: $s_f . s_k y_t = (\delta + \lambda + n)k$ (4.10)

Human capital: $s_{\acute{h}} y_t = (\delta + \lambda + n)\acute{h}$ (4.11)

Infrastructure: $s_f . s_g y_t = (\delta + \lambda + n)g$ (4.12)

By Substituting y_t into equation 4.10-4.12

$$s_f . {k_t}^{-(1-\alpha)} . {\acute{h}_t}^{\beta} . {g_t}^{\gamma} = \frac{\delta+\lambda+n}{s_k} \quad (4.13)$$

$$ {k_t}^{\alpha} . {\acute{h}_t}^{-(1-\beta)} . {g_t}^{\gamma} = \frac{\delta+\lambda+n}{s_{\acute{h}}} \quad (4.14)$$

$$s_f . {k_t}^{\alpha} . {\acute{h}_t}^{\beta} . {g_t}^{-(1-\gamma)} = \frac{\delta+\lambda+n}{s_g} \quad (4.15)$$

By solving,

$$k_t = \left(\frac{s_f \, s_k \, {\acute{h}_t}^{\beta} . {g_t}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\alpha)}} \quad (4.16)$$

$$\acute{h}_t = \left(\frac{s_h {k_t}^{\alpha} . {g_t}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\beta)}} \quad (4.17)$$

$$g_t = \left(\frac{s_f\, s_g\, k_t{}^{\alpha} \acute{h}_t{}^{\beta}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\gamma)}} \tag{4.18}$$

We get the steady state level of per worker capital, human and infrastructure.

$$k_t{}^* = \left(\frac{s_f\, s_k{}^{1-\beta-\gamma}\, s_{\acute{h}}{}^{\beta} . s_g{}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\alpha-\beta-\gamma)}} \tag{4.19}$$

$$\acute{h}_t{}^* = \left(\frac{s_k{}^{\alpha}\, s_{\acute{h}}{}^{1-\alpha-\gamma} . s_g{}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\alpha-\beta-\gamma)}} \tag{4.20}$$

$$g_t{}^* = \left(\frac{s_f . s_k{}^{\alpha}\, s_{\acute{h}}{}^{\beta} . s_g{}^{1-\alpha-\beta}}{\delta+\lambda+n}\right)^{\frac{1}{(1-\alpha-\beta-\gamma)}} \tag{4.21}$$

The steady state level of output per worker using equation (4.19-21) into the equation (4.4)

$$y_e{}^* = \left(\frac{s_f\, s_k{}^{1-\beta-\gamma}\, s_{\acute{h}}{}^{\beta} . s_g{}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{\alpha}{(1-\alpha-\beta-\gamma)}} \left(\frac{s_k{}^{\alpha}\, s_{\acute{h}}{}^{1-\alpha-\gamma} . s_g{}^{\gamma}}{\delta+\lambda+n}\right)^{\frac{\beta}{(1-\alpha-\beta-\gamma)}} \left(\frac{s_f\, s_k{}^{\alpha}\, s_{\acute{h}}{}^{\beta} . s_g{}^{1-\alpha-\beta}}{\delta+\lambda+n}\right)^{\frac{\gamma}{(1-\alpha-\beta-\gamma)}} \tag{4.22}$$

By simplifying; we get steady state level of output per effective worker.

$$y_e{}^* = \left(s_f{}^{\alpha+\gamma}\, s_k{}^{\alpha}\, s_{\acute{h}}{}^{\beta} . s_g{}^{\gamma}\right)^{\frac{1}{(1-\alpha-\beta-\gamma)}} (\,\delta+\lambda+n)^{\frac{-(\alpha+\beta+\gamma)}{(1-\alpha-\beta-\gamma)}} \tag{4.23}$$

The steady state level of output per worker

$$\hat{y}_e{}^* = A_t y_e{}^* = A_o e^{\lambda(t)} \left(s_f{}^{\alpha+\gamma}\, s_k{}^{\alpha}\, s_{\acute{h}}{}^{\beta} . s_g{}^{\gamma}\right)^{\frac{1}{(1-\alpha-\beta-\gamma)}} (\,\delta+\lambda+n)^{\frac{-(\alpha+\beta+\gamma)}{(1-\alpha-\beta-\gamma)}} \tag{4.24}$$

By log-transformation,

$$\ln \hat{y}_e{}^* = \ln A_o + \lambda(t) + \frac{\alpha+\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_f + \frac{\alpha}{(1-\alpha-\beta-\gamma)} ln\, s_k + \frac{\beta}{(1-\alpha-\beta-\gamma)} ln\, s_{\acute{h}} + \frac{\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_g - \frac{(\alpha+\beta+\gamma)}{(1-\alpha-\beta-\gamma)} ln(\delta+\lambda+n) \tag{4.25}$$

As described earlier, as $s_{\acute{h}} = s_h + s_s$ we can express this equation 4.25 as;

$$\ln \hat{y}_e{}^* = \ln A_o + \lambda(t) + \frac{\alpha+\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_f + \frac{\alpha}{(1-\alpha-\beta-\gamma)} ln\, s_k + \frac{\beta}{(1-\alpha-\beta-\gamma)} ln\, s_h + \frac{\beta}{(1-\alpha-\beta-\gamma)} ln\, s_s + \frac{\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_g - \frac{(\alpha+\beta+\gamma)}{(1-\alpha-\beta-\gamma)} ln(\delta+\lambda+n) \tag{4.26}$$

Equation 4.26 characterizes the long-run relationship in GDP per worker. For the analysis of the speed of convergence (transition path to steady-state), the time period of t_0 to $t_{0+\theta}$ has to be incorporated and η has been used as a symbol for the speed of adjustment following Mankiw et al., (1992) and Knight et al., (1993).

$$\frac{d\widehat{y_t}}{dt} = \eta(ln\widehat{y_t}^{\,e} - ln\hat{y}_t) \tag{4.27}$$

The parameter η is the speed of convergence. It shows how fast output per worker reaches its steady state. η=$(\delta + \lambda + n)(1 - \alpha - \beta - \gamma)$

On balanced growth path output per worker, grows at the rate of speed of convergence. At the steady state the growth rate of output per worker $[ln\hat{y}_t - ln\hat{y}_t{}^e]$ becomes nearly constant.

By integrating the eq (25) for time intervals, $t = t_0$ and $t = t_0 + þ$

$$ln\hat{y}_{t,o+þ} = \left(1 - e^{-\eta þ}\right) ln\hat{y}_t{}^e + \; e^{-\eta þ} ln\hat{y}_{t,0} \tag{4.28}$$

By subtracting the output per worker at time t_0, we get

$$ln\hat{y}_{t,o+þ} - ln\hat{y}_{t,0} = \left(1 - e^{-\eta þ}\right) ln\hat{y}_t{}^e + \; \left(1 - e^{-\eta þ}\right) ln\hat{y}_{t,0} \tag{4.29}$$

$$ln\hat{y}_{t,o+þ} - ln\hat{y}_{t,0} = \left(1 - e^{-\eta þ}\right)[\ln A_o + \lambda(t) + \frac{\alpha+\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_f + \frac{\alpha}{(1-\alpha-\beta-\gamma)} ln\;\, s_k + \frac{\beta}{(1-\alpha-\beta-\gamma)} ln\, s_h + \frac{\beta}{(1-\alpha-\beta-\gamma)} ln\, s_s + \frac{\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_g - \frac{(\alpha+\beta+\gamma)}{(1-\alpha-\beta-\gamma)}\; ln(\delta + \lambda + n)] + e^{-\eta þ}[lny_{t,0}] \tag{4.30}$$

By simplifying, we get,

$$ln\hat{y}_{t,o+þ} - ln\hat{y}_{t,0} = \left(1 - e^{-\eta þ}\right)[\frac{\alpha}{(1-\alpha-\beta-\gamma)} lns_k + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns_h + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns_s + \frac{\gamma}{(1-\alpha-\beta-\gamma)} lns_g + \frac{\alpha+\gamma}{(1-\alpha-\beta-\gamma)} ln\, s_f - \frac{\alpha+\beta+\gamma}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n) + lnA_0 + \lambda t] + \; e^{-\eta þ}[lny_{t,0}] \tag{4.31}$$

$$dln\hat{y}_t = \left(1 - e^{-\eta þ}\right)[lnA_0 + \lambda t] + \frac{\alpha\left(1-e^{-\eta þ}\right)}{(1-\alpha-\beta-\gamma)} lns_k + \frac{\beta\left(1-e^{-\eta þ}\right)}{(1-\alpha-\beta-\gamma)} lns_h + \frac{\beta\left(1-e^{-\eta þ}\right)}{(1-\alpha-\beta-\gamma)} lns_s + \frac{\gamma\left(1-e^{-\eta þ}\right)}{(1-\alpha-\beta-\gamma)} lns_g + \frac{\alpha+\gamma\,\left(1-e^{-\eta þ}\right)}{(1-\alpha-\beta-\gamma)} ln\, s_f - \frac{\alpha+\beta+\gamma((1-e^{-\eta\theta})}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n) + \; e^{-\eta\theta}[lny_{t,0}] \tag{4.32}$$

By further simplifying; $\left(1 - e^{-\eta\theta}\right) = \varphi$; $(\delta + \lambda + n) = \psi$; $e^{-\eta\theta} = B_\theta$

$$dln\hat{y}_t = \varphi lnA_0 + \varphi\lambda t + \frac{\alpha\varphi}{\psi} lns_k + \frac{\beta\varphi}{\psi} lns_h + \frac{\beta\varphi}{\psi} lns_s + \frac{\gamma\varphi}{\psi} lns_g + \frac{\alpha+\gamma}{(\eta} ln\, s_f -$$

$$\frac{\varphi(\alpha+\beta+\gamma)}{\eta} ln(\delta + \lambda + n) + B_\theta lny_{t,0} \quad (4.33)$$

We can express the estimable equation as;

$$dln\hat{y}_t = \alpha_0 lnA_0 + \alpha_1 t + \alpha_1 lns_k + \alpha_2 lns_h + \alpha_3 lns_s + \alpha_4 lns_g + \alpha_5 lns_f - \alpha_6 ln(\delta + \lambda + n) +$$

$$\alpha_7 lny_{t,0} + \varepsilon \quad (4.34)$$

The equation (4.34) is the estimable equation to be employed in the subsequent chapter for the empirical estimation of growth-effect at aggregate timeseries analysis. The coefficient α_3, α_4, $\alpha_5, \alpha_6, \alpha_6$ measures the elasticities of per-worker output-growth as result of the investment in the private capital stock, human capital stock, social, core infrastructure and financial infrastructure, respectively.

4.2 Theoretical model for the Spillover effect of Infrastructure

We developed a detailed framework developed to analyze the direct and indirect spatial effects of infrastructure over the Pakistan's regional development using the neoclassical growth framework of MWR model, which is an extension of the neoclassical growth paradigm developed by Arrow & Kurz (1970), Weitzman (1970) and Barro (1990). In this model, infrastructure is added as an additional factor of production as a public capital besides the labor and private capital. The Mankiw et al. (1992) growth model framed for human capital in the standard neoclassical growth of Solow (1956) paradigm and we includes infrastructure as an additional driver of long-run economic growth in this study.

$$Y_{it} = K_{it}^{\alpha}.H_{it}^{\beta}.G_{it}^{\gamma}.(A_{it}L_{it})^{(1-\alpha-\beta-\gamma)} \quad (4.35)$$

The model (equation 4.35) is based on a standard Cobb-production function. The symbols Y_{it} represents the output in regions 'i' over time 't', G_{it} represents the stock of infrastructure, K_{it} represent the stock of private physical capital and H_{it} represents the stock of human capital, $A_{it}L_{it}$ represent the labor augmented technology and 't' represent time. The subscript 'i' represents cross-sections, and *'t'* represents the time dimension. The production functions exhibit constant returns to scale in all its inputs capital, infrastructure, human capital, and technology augmented labor force. Whereas α is the income share of capital, β is the income share of human capital, and γ is the income share of infrastructure in output.

Furthermore, following Ertur & Koch (2007) we included the spatial externalities in the growth accounting through the technological interdependence. We further advanced Ertur & Koch (2007) by including not only the regional spatial externalities of physical capital and but also included the regional effects of investing in human capital and core infrastructure. Therefore, A_{it} represents labor-augmenting technological change including temporal spatial externalities, as well as the exogenous effects that a region receives from the other nearby regions due to investment in the human capital, private capital, and infrastructure in nearby locations. The * represents the neighbouring (foreign) provinces.

$$A_{it} = A_o . e^{\lambda t} . y_{it}^{\theta 1*} . y_{i,t-1}^{\theta 2*} S\text{k}_{it}^{\theta 3*} S\text{h}_{it}^{\theta 4*} S\text{g}_{it}^{\theta 5*} (\delta + \lambda + n)_{it}^{\theta 6*} \quad (4.36)$$

Here, λ is the exogenous rate of technological progress. The investment in the neighbouring geographic location exogenously impacts the long-run development through indirect effects. For this study we define neighbouring regions as provinces of Pakistan. Therefore, we assume the level of investment in human capital, private capital, and stock of infrastructure in the neighbourhood is exogenously impacting the neighbouring region's economy via indirect spatial spillover effects. The production function can be expressed as in effective-labor intensive units.

$$\frac{Y_{it}}{A_{it}L_{it}} = \left(\frac{K_{it}}{A_{it}L_{it}}\right)^{\alpha} \left(\frac{H_{it}}{A_{it}L_{it}}\right)^{\beta} \left(\frac{G_{it}}{A_{it}L_{it}}\right)^{\gamma} \quad (4.37)$$

To express the production function in labor-intensive units:

$$y_{it} = {k_{it}}^{\alpha} {h_{it}}^{\beta} {g_{it}}^{\gamma} \quad (4.38)$$

We can transform the production function into the natural log; we get

$$lny_{it} = \alpha lnk_{it} + \beta lnh_{it} + \gamma lng_{it} \quad (4.39)$$

Dynamics of the Model:

The stock of private physical capital, human capital and infrastructure accumulates over time.

$$\dot{K}_{it} = s_{ik} Y_{it} - \delta_{it} K_{it} \quad (4.40)$$

Where $\dot{K}$ is the stock of private capital increasing over time due to the new investment by the private sector. The symbol s_{ik} represent the investment/ or saving rate (as a fraction of the total national income devoted for capital accumulation). The stock of private capital declines depreciates at the same time, as result of the wear and tear that is represented by the symbol δ_{it}.

Human capital accumulates over time as well. Human capital is the investment in labor in form of educations, skills, and knowledge.

$$\dot{H}_{it} = s_{ih} Y_{it} - \delta_{it} H_{it} \quad (4.41)$$

The $\dot{H}$ is the growth of human capital over time as a result of investment over time. The s_{ih} represent the saving ratio rate relative to total income designated for the human capital development and δ_{it} depicting the depreciation rate, which reduces the rate of human capital accumulation.

Infrastructure accumulates over times as;

$$\dot{G_{it}} = s_{ig}Y_{it} - \delta_{it}\, G_{it} \tag{4.42}$$

The $\dot{G}$ is overtime growth in the physical stock of infrastructure as result of new capital-intensive investments as result of saving. The s_{ig} which is the investment rate in an economy especially for devoted of the infrastructural development. The rate of depreciation in infrastructural stock is represented by δ_{it} .

The growth of private capital stock in the labour-intensive (effective) term as.

$$\dot{k_t} = \frac{\dot{K_{it}}}{A_t L_t} - \frac{K_{it}}{A_{it}L_{it}}\left(\frac{\dot{A_{it}}}{A_{it}} + \frac{\dot{L_{it}}}{L_{it}}\right) \tag{4.43}$$

Inserting (4.32) into (4.38);

$$\dot{k_{it}} = s_{ik}y_{it} - (\delta + \frac{\dot{A_{it}}}{A_{it}} + \frac{\dot{L_{it}}}{L_{it}})k_{it} \tag{4.44}$$

The long-run level of capital stock in neighbouring regions would be subject to the capital stock accumulation in their own regions. The steady state (SS) is an equilibrium state where physical, human capital, and infrastructure per effective worker growth, technology, and population grow at the same rate, and mathematically we can express the SS as when,

$$\dot{k_t} = \dot{h_t} = \dot{g_t} = 0,\ \frac{\dot{A_{it}}}{A_{it}} = \lambda, \frac{\dot{L_{it}}}{L_{it}} = n$$

Private capital: $s\text{k}_{it}y_{it} = (\delta + \lambda + n)k_{it}$ (4.45)

Human capital: $s\text{h}_{it}y_{it} = (\delta + \lambda + n)h_{it}$ (4.46)

Infrastructure: $s\text{g}_{it}y_{it} = (\delta + \lambda + n)k_{it}$ (4.47)

By simplifying and using the log properties, we get the production function at SS;

$$lny_t{}^e = \frac{\alpha}{(1-\alpha-\beta-\gamma)} lns\text{k}_{it} + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns\text{h}_{it} + \frac{\gamma}{(1-\alpha-\beta-\gamma)} lns\text{g}_{it} - \frac{\alpha+\beta+\gamma}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n_{it}) \tag{4.48}$$

Here, $\frac{\alpha}{1-\alpha-\beta-\gamma}$ and $\frac{\beta}{(1-\alpha-\beta-\gamma)}$ and $\frac{\gamma}{(1-\alpha-\beta-\gamma)}$ represent the output elasticities with respect to the investment in physical capital, human capital, and infrastructure. On the balanced growth path, the exogenous technological (or TFP) growth and indirect spatial spillover effects impact the productivity growth.

$\frac{Y_{it}}{A_{it}L_{it}} = y : \frac{Y_{it}}{L_{it}} = A_{it}y , \frac{Y_t}{L_T} = y_t{}^e A_o . e^{\lambda t} . y_{it}^{\theta 1*} S\text{ƙ}_{it}^{\theta 2*} S\text{h}_{it}^{\theta 3*} S\text{g}_{it}^{\theta 4*} (\delta + \lambda + n)_{it}^{\theta 5*}$

By taking the log, we have output in per worker term;

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = ln\widehat{y_{it}} = lny_{it}{}^e + lnA_0 + \lambda t + \theta_1 \ln y_{it}^* + \theta_2 \ln S\text{ƙ}_{it}^* + \theta_3 \ln S\text{h}_{it}^* + \theta_4 \ln S\text{g}_{it}^* - \theta_5\, ln(\delta + \lambda + n)_{it}^* \quad (4.49)$$

$$ln\widehat{y_{it}} = \frac{\alpha}{(1-\alpha-\beta-\gamma)} lns\text{ƙ}_{it} + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns\text{h}_{it} + \frac{\gamma}{(1-\alpha-\beta-\gamma)} lns\text{g}_{it} - \frac{\alpha+\beta+\gamma}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n)_{it} + lnA_0 + \lambda t + \theta_1 \ln y_{it}^* + \theta_2 \ln y_{it-1}^* + \theta_3 \ln S\text{ƙ}_{it}^* + \theta_4 \ln S\text{h}_{it}^* + \theta_5 \ln S\text{g}_{it}^* - \theta_6\, ln(\delta + \lambda + n)_{it}^* \quad (4.50)$$

Eq (9.15) characterizing the long-run relationship in GDP per worker. For the analysis of the speed of convergence (transition path to steady-state), the time-period of t_0 to $t_{0+\theta}$ has to be incorporated, and η has been used as a symbol for the speed of adjustment following Mankiw et al. (1992) and Knight et al. (1993).

$$\frac{d\widehat{y_{it}}}{dt} = \eta(ln\widehat{y_{it}}{}^e - ln\hat{y}_{it})$$

The parameter η is the speed of convergence. It shows how fast output per worker reaches its steady state.

η=$(\delta + \lambda + n)(1 - \alpha - \beta - \gamma)$

On balanced growth path, output per worker grows at the rate of speed of convergence. In the steady state, the growth rate of output per worker $[ln\hat{y}_t - ln\hat{y}_t{}^e]$ becomes nearly constant.

By integrating for time intervals, $t = t_0$ and $t = t_0 + \text{þ}$

$$ln\hat{y}_{it,o+\text{þ}} = \left(1 - e^{-\eta \text{þ}}\right) ln\hat{y}_{it}{}^e + e^{-\eta \text{þ}} ln\hat{y}_{it,0} \quad (4.51)$$

By subtracting the output per worker at initial time t_0, we get

$$ln\hat{y}_{it,o+\text{þ}} - ln\hat{y}_{it,0} = \left(1 - e^{-\eta \text{þ}}\right) ln\hat{y}_{it}{}^e + \left(1 - e^{-\eta \text{þ}}\right) ln\hat{y}_{it,0} \quad (4.52)$$

$$ln\hat{y}_{it,o+\text{þ}} - ln\hat{y}_{it,0} = \left(1 - e^{-\eta \text{þ}}\right)\Big[\frac{\alpha}{(1-\alpha-\beta-\gamma)} lns\text{ƙ}_{it} + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns\text{h}_{it} + \frac{\gamma}{(1-\alpha-\beta-\gamma)} lns\text{g}_{it} - \frac{\alpha+\beta+\gamma}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n) + lnA_0 + \lambda t + \theta_1 \ln y_{it}^* + \theta_2 \ln y_{it-1}^* + \theta_3 \ln Sk_{it}^* + \theta_4 \ln Sh_{it}^* + \theta_5 \ln Sg_{it}^* - \theta_6\, ln(\delta + \lambda + n)_{it}^*\Big] + e^{-\eta \text{þ}}[ln\hat{y}_{it,0}] \quad (4.53)$$

By simplifying, we get

$$ln\hat{y}_{it,o+\text{þ}} = \left(1 - e^{-\eta\text{þ}}\right)\Big[\frac{\alpha}{(1-\alpha-\beta-\gamma)} lns\text{ƙ}_{it} + \frac{\beta}{(1-\alpha-\beta-\gamma)} lns\text{h}_{it} + \frac{\gamma}{(1-\alpha-\beta-\gamma)} lns\text{g}_{it} - \frac{\alpha+\beta+\gamma}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n) + lnA_0 + \lambda t + \theta_1 \ln y^*_{it} + \theta_2 \ln y^*_{it-1} + \theta_3 \ln Sk^*_{it} + \theta_4 \ln Sh^*_{it} + \theta_5 \ln Sg^*_{it} - \theta_6\, ln(\delta + \lambda + n)^*_{it}\Big] + (e^{-\eta\text{þ}} + 1)[ln\hat{y}_{it,0}] \quad (4.54)$$

$$dln\hat{y}_{it} = \left(1 - e^{-\eta\text{þ}}\right)[lnA_0 + \lambda t] + \frac{\alpha(1-e^{-\eta\text{þ}})}{(1-\alpha-\beta-\gamma)} lns\text{ƙ}_{it} + \frac{\beta(1-e^{-\eta\text{þ}})}{(1-\alpha-\beta-\gamma)} lns\text{h}_{it} + \frac{\gamma(1-e^{-\eta\text{þ}})}{(1-\alpha-\beta-\gamma)} lns\text{g}_{it} - \frac{\alpha+\beta+\gamma((1-e^{-\eta\theta})}{(1-\alpha-\beta-\gamma)} ln(\delta + \lambda + n) + \theta_1\left(1 - e^{-\eta\text{þ}}\right)\ln y^*_{it} + \theta_2\left(1 - e^{-\eta\text{þ}}\right)\ln y^*_{it-1} + \theta_3\left(1 - e^{-\eta\text{þ}}\right)\ln Sk^*_{it} + \theta_4\left(1 - e^{-\eta\text{þ}}\right)\ln Sh^*_{it} + \theta_5\left(1 - e^{-\eta\text{þ}}\right)\ln Sg^*_{it} - \theta_6\left(1 - e^{-\eta\text{þ}}\right) ln(\delta + \lambda + n)^*_{it} + (e^{-\eta\theta} + 1)[lny_{it,0}] \quad (4.55)$$

By further simplifying; $\left(1 - e^{-\eta\theta}\right) = \varphi$; $(\delta + \lambda + n) = \eta$; $e^{-\eta\theta} + 1 = B_\theta$

$$ln\hat{y}_{it} = \varphi lnA_0 + \varphi\lambda t + \frac{\alpha\varphi}{\eta} lns\text{ƙ}_{it} + \frac{\beta\varphi}{\eta} lns\text{h}_{it} + \frac{\gamma\varphi}{\eta} lns\text{g}_{it} - \frac{\varphi(\alpha+\beta+\gamma)}{\eta} ln(\delta + \lambda + n)_{it} + \theta_1\varphi\, lny^*_{it} + \theta_2\varphi\, lny^*_{it-1} + \theta_3\varphi \ln Sk^*_{it} + \theta_4\varphi \ln Sk^*_{it} + \theta_5\varphi \ln Sh^*_{it} + \theta_6\varphi \ln Sg^*_{it} - \theta_7\varphi\, ln(\delta + \lambda + n)^*_{it} + B_\theta lny_{it,0} \quad (4.56)$$

We can express the estimable equation as;

$$ln\hat{y}_{it} = \alpha_0 lnA_0 + \alpha_1 t + \alpha_2 lns\text{ƙ}_{it} + \alpha_3 lns\text{h}_{it} + \alpha_4 lns\text{g}_{it} - \alpha_5 ln(\delta + \lambda + n) + \alpha_6 \ln y^*_{it} + \alpha_7 \ln y^*_{it-1} + \alpha_8 \ln Sk^*_{it} + \alpha_9 \ln Sh^*_{it} + \alpha_{10} \ln Sg^*_{it} - \alpha_{11}\, ln(\delta + \lambda + n)^*_{it} + \alpha_{12} lny_{it,0} + \varepsilon_{it} \quad (4.57)$$

The equation (4.57) is the estimable equation to be employed to estimate the provincial spillover effect, whereas the coefficient α_3, α_4, α_4 measures the elasticities of the output of investment in the private capital stock, human capital, and stock of core infrastructure, respectively and $\alpha_8, \alpha_9, \alpha_{10}$ measuring indirect effects bordered effects arise from investment in stock of the human capital, private capital, and infrastructure in nearby provinces. We expect the positive sign of $\alpha_8, \alpha_9, \alpha_{10}$ when the indirect border effects are positively reinforcing networking and economic integration. On the other hand, the negative sign of $\alpha_8, \alpha_9, \alpha_{10}$ would signal negative indirect spatial effects arises because of factor outmigration from the locations of lower relative stock of physical, human and infrastructure towards to the locations with higher stock of physical, human capital, and infrastructure. On the other land the α_6 will gauge the spatial externalities or spillover or the size of the spatial autocorrelation that a region received unintended from another region. On the other land α_7 will capture the temporal regional spatial externalities, and α_{12} will gauge the temporal effect (autoregressive) impact of productivity growth of a reason.

CHAPTER 5

GROWTH EFFECT OF INFRASTRUCTURE: ESTIMATION

5.1 Introduction

This chapter is dedicated to the empirical examination to ascertain long run and short run growth effect of infrastructure on the productivity growth of Pakistan. The existing literature uses public capital investment (measured in monetary terms) and infrastructure (measured in physical units) as proxies to assess the impact of public-sector development expenditure on long-run economic growth. However, recent studies preferred the use of infrastructure (measured in physical units) over public investment, as it places more emphasis on the tangible presence of infrastructural services rather than solely relying on monetary terms. This study employs indicators that quantify infrastructure physically and consider their utilization, encompassing a comprehensive array of infrastructure variables and developing various indices for both hard and soft forms of infrastructures. This chapter presents an estimation approach and estimated results. The empirical model is constructed based on the theoretical construct, based on the infrastructure-augmented production function that establishes the theoretical connections of output per worker growth as function of the investment drivers of the private capital stock, human capital, core, social and financial infrastructure.

This chapter focuses on the empirical models to test the contribution of core, social and financial infrastructure in the productivity growth of Pakistan. This chapter has been arranged into twelve broad sub-sections. The section 5.2-7 elucidates, econometric model specification, indicators, indices, data, estimation methodology, pre-estimation tests, and results. the section 5.8 presents the estimated models, discussion about the key findings. Lastly, component-wise estimation for each type of infrastructure is provided in the sections from 5.9 to 5.12.

5.2 Model Specification and Indicators

The previous chapter provides the theoretical linkages and foundation for constructing an empirical model. The model encompasses the infrastructure as a key determinant and other key determining factors, including human capital and physical stock of capital. This estimable equation elaborates the interconnected relationship between real output per worker, with the share of capital stock, human capital, and components of infrastructure (core, social and financial). The foremost component is the core infrastructure is the facilitator for all economic and social activities. It has three main functions (Baldwin & Dixon, 2011). First, transportation and communication networks create spatial connections of economic activities between geographically distant locations. Second, transportation, communication, water and sewage facilitate the urbanised lifestyle. Third, electricity is an input used for producing goods and services.

Concomitantly, literature for the social infrastructure elucidates its critical role in raises capital accumulation, human capital development (educational attainment), and productivity (output per worker) and differences in income-level across countries (Chin & Chou, 2004; Hall & Jones, 1999). Moreover, social infrastructure for schools are not just limited to classrooms teaching but also engaged in creating socially inclusive civic places, with equal opportunities (Ralls, 2019). Moreover literature highlighted a key channels of social infrastructure of school is via positive impact on student's learning outcomes (Barrett et al., 2019; Schneider, 2002). Pertinently, the indoor school physical environment positing children's short-term and long-term health effects and affect learning ability and productivity (Bluyssen, 2017). Likewise, the social infrastructure for healthcare is crucial for the economic development, sustained economic growth and well-being (Bloom, 2019). In less developed countries, investment in health-care infrastructure is low that results into reduces female labour force participation, potentially reduces the children future's productivity. While in more developed countries, investments in health-care system primarily led to rising longevity and workforce productivity.

Theoretical as well as empirical literature supporting the argument of positive contribution of the financial development and productivity growth. The financial structure through its functional relationship of economic integration facilitates the process and efficiency of capital accumulation (Levine, 1997). The cross-country empirical studies for example by Kumbhakar & Mavrotas (2005) elucidate eminent role of institutional reforms for constructing development of financial institutions to improvise the productivity growth, however conditional on the countries state of economic development. Additional Valickova et al. (2014) using meta-analysis approach and by comparing 1334 estimates from 67 studies, this study by found the role of financial structure critical for the pace of economic development. In this purview, we framed an estimable empirical model which is developed based mathematical model (for details, equation 4.32) that framed each components of infrastructure of core, social and financial infrastructure in the neoclassical growth model of MWR which is advanced by Solow.

$$dln\widehat{y_t} = \varphi_0 + \varphi_1 lnSk_t + \varphi_2 lnSh_t + \varphi_3 \ln Sg_t + \varphi_2 lnSs_t + \varphi_2 lnSf_t + \varphi_4 t - \varphi_6 ln(\delta + \lambda + n) + \varphi_6 (ln\widehat{y_{t-1}}) + \varepsilon_t \quad (5.1)$$

This estimable equation elaborates the interconnected relationship between growth rate of real output per worker, with the share of capital stock, human capital, components of infrastructure, set of control variables and required breakeven level of investment. The Sk represents share of resources devoted by private sector for capital accumulation. We used gross capital formation (private sector) as percentage of GDP as indicator of share of reinvested total factor income in new capital assets. The Sh_t represent share of resources of devoted for human capital development. We used indicator of expenditure on the health and education sector expressed as a percentage of GDP as a proxy for the new investment for the human capital. The Sg_t represents the share of core infrastructure, and we are sub-comparting the core infrastructure based on the physical quantity and utilization. The core infrastructure constitutes the transport, energy, and communication infrastructure. We define core infrastructure as transport, energy, and communication infrastructure stock variables. We further bifurcate the core infrastructure index into the quantification, and utilization terms will be computed for the following indicators using the principal components analysis (PCA) technique. The indicator of high type roads per worker, length of electricity transmission lines per worker, and teledensity per worker had used to gauge per-worker investment in the transport, energy, and communication stock. However, including the production indicators of various energy inputs reflects shadow/indirect indicators that measure the size of energy infrastructure present in the country. Likewise, the inclusion of energy consumption indicators is to gauge the size of the utilization of core infrastructure in the country. We also analyze the impact of the utilization of transport infrastructure, energy, and communication infrastructure in segregation.

Indicators for Quantification Index	Indicators for Utilization Index
▪ Length of high-type roads ▪ Electricity generation ▪ Crude oil extraction ▪ Production of Gas ▪ Production of Petroleum Products ▪ Production of Coal ▪ Total Length of electricity transmission lines ▪ Landline telephone and mobile-phone subscribers	▪ Vehicles on road ▪ Passengers traveled by rail. ▪ Cargo carried by rail. ▪ Total Cargo handled at the Seaports of Karachi, Port Qasim and Gwadar ▪ Freight, Air Transport ▪ Electricity consumption ▪ Consumption of oil/Petroleum ▪ Consumption of Gas ▪ Consumption of Coal ▪ Landline telephone and mobile-phone subscribers

We defined social infrastructure (Ss_t) as share of resources devoted for the social infrastructure of education and health services. We are calibrating in terms of physical as well as utilization terms both. The educational infrastructure comprises all educational facilities that help increase the stock of knowledge and education. Therefore, we are developing an equi-weighted index using the following indicators for the educational infrastructure.

Quantification (per mln population)	Utilization (mln)
Number of primary schools	Enrollment in primary schools
Number of secondary schools	Enrollment in secondary schools
Number of high schools	Enrollment in High schools
Number of technical and vocational centers	Enrollment in technical vocational centers
Number of higher secondary school	Enrollment in higher secondary schools
Number of degree colleges	Enrollment in degree colleges
Number of universities	Enrollment in universities

The health infrastructure facilitates nations to build and maintain healthy human capital therefore we are using the number of hospital beds as an indicator of the health facilities available in the country. Unfortunately, direct data of health services utilization is unavailable, such as the number of patients who attended OPD, hospitalized, etc. However, immunization data is available, which is a utilization indicator of health infrastructure. Therefore, we use the utilization indicator of the percentage of children receiving the BCG immunization in the country.

Additionally, Ss_t representing financial infrastructure network for the financial services. We define the total number of Bank Branches (BBR) and ATM machines (ATM) as essential indicators to measure size of financial structure. The number of ATMs holds significance because it facilitates fundamental banking services such as withdrawals and inter- and intra-bank cash transfers, in addition to the key services available at bank branches. Furthermore, we adopt the usage indicator 'number of bank accounts per population as a proxy to assess the utilization of financial infrastructure in the country. We utilize this indicator of the number of bank accounts to measure the extent of financial infrastructure utilization since having a basic bank account is a prerequisite for accessing financial services in Pakistan. As such, it acts as a proxy or shadow variable to evaluate the level of financial infrastructure utilization in the country. The variable $\delta + \lambda + n$ measures the breakeven level of capital investment to keep-up the existing per worker capital stock, we follow Mankiw et al., (1992) approach and computed annual population growth rate and 5 percent rate of depreciation rate.

5.3 Pre-Estimation tests

A large body of research unfolded a positive correlation between economic growth and infrastructure; however, it remained a puzzle and gained focus in the academic discussion, largely due to methodological and estimation technique differences. The size of the impact of infrastructure investment was estimated by Aschauer (1989) using the OLS estimation technique for the United States, which many researchers later challenged, including Holtz-Eakin & Schwartz (1994) and Munnell (1990). The subsequent studies by Barro (1991) found a positive but insignificant impact of public investment on the economic growth of 98 developing countries during 1960-1985. The study of Devarajan et al. (1996) used 20 years of data from 43 developing countries and found that per capita growth is negatively associated with public expenditure. They highlighted the possible reason for the misallocation of public expenditures favouring capital expenditures at the expense of current spending. A debate has existed among researchers regarding the measurement of output elasticity of the public capital and this problem was addressed sorted out by Bom & Ligthart (2014) through a meta-regression approach. He used a sample of 67 studies for 1983–2008 and found 'meta-output estimates' 0.146, and the actual effect is subject to the heterogeneity between cross sections. They showed that the high output elasticities found in the time-series literature are subject to the cointegration of the economic variables, so studies based on the time-series data inflated the output elasticity due to the bidirectional publication bias, if researcher ignores the long-term association as the time series data and mostly follows 'an autoregressive' data generation process. The current values of the variables are partly depending on its previous values.

This study is particularly interested in measuring the integrated relationship of infrastructure vis-à-vis productivity growth. Most time series data series are non-stationarity, showing an increasing trend over time. Likewise, the infrastructure-growth relationship is no exception. The core infrastructure, private capital, and human capital variables are expected to increase and co-move over time. The appropriate empirical model for estimating output elasticities for each type of infrastructure depends on the stationarity test of the data series. Therefore, an appropriate estimation method is subject to the order of integration and tested for the presence of unitroot.

5.3.1 Testing stationarity (the Unit root test) and detection of the structural breaks

The graphical analysis is helpful for data visualization, but formal unit root diagnostic tests are recommended in empirical literature to detect non-stationarity data series. Without incorporating the formal stationarity test of variables, the results of estimated models provide spurious results—the spurious regressions show a strong association between economic variables than actual. Therefore, we are using the following formal tests for stationarity testing.

5.3.1.1 An Augmented Dickey-Fuller (ADF) test

By using a generalized description for all variables of the model, a data series m_t follow an autoregressive data-generating process.

$$m_t = \alpha_0 + \sigma_1 m_{t-1} + \sigma_2 m_{t-2} + \cdots \sigma_n m_{t-n} + \varepsilon_t \qquad (5.2)$$

The ADF test is formal empirical testing based on the following empirical equation for data series m_t .

$$\Delta m_t = \mu + \delta m_{t-1} + \boldsymbol{\partial}_i \sum_{i=1}^{k} m_{t-i} + e_t$$

Whereas, Δm_t is the first difference and e_t is the error term assuming constant zero mean and normally distributed error. while $\delta = \sigma_1 - 1$ and null hypothesis using ADF test is $\delta = 0$ against the alternative hypothesis of $\delta < 0$. If we accept the null hypothesis, the series has a unit root (non-stationary). If we reject the null hypothesis, the series has no unit root depicting that the series is stationary at level.

5.3.1.2. Philip-Perron (PP) test

The PP test was developed by Phillips & Perron (1988). The formal empirical equation for stationarity testing is;

$$\Delta m_t = \varpi m_{t-1} + \boldsymbol{\rho}_i D_i + e_t$$

Where D_i is the deterministic trend and e_t is the error term and is $I(0)$ means stationary at level. The null hypothesis of the PP test tested is $\varpi = 0$, and this test makes a 'nonparametric correction' in testing hypotheses by correcting the unspecified autocorrelation and heteroscedasticity of error terms (Phillips & Perron, 1988).

5.3.1.3 Kwiatkowsky, Phillips, Schmidt and Shin (KPSS) test

The KPSS test was developed by (Kwiatkowski et al., 1992) as an alternative method for stationarity testing. The test decomposes the series into a deterministic part, a random walk, and an error term component.

$$m_t = \gamma t + \delta Z_t + e_t$$

$$Z_t = Z_{t-1} + \mu_t$$

The hypothesis testing is based on the μ_t – using Lagrange Multiplier (LM).

The empirical literature recommends performing all three tests to conclude the order of integration (Shrestha & Bhatta, 2018).

5.3.1.4 Testing the structural breaks

The conventional test for the unit root in the presence of structural break does not reach the right conclusion (Perron, 1989). The structural breaks are sudden changes in time series data due to any major shift in policy, regime, and unexpected external shock (Shrestha & Bhatta, 2018). (Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989)(Perron, 1989) Using an exogenous time for a structural break, Perron (1989) showed that the power to reject a unit root (non-stationarity) decreases when the alternative stationarity is true and when a structural break is ignored. Using the idea of Perron (1989), Zivot and Andrews (1992) constructed the data-dependent algorithm is used to determine the unknown breakpoints. The Zivot and Andrews test uses the three models to test a unit root with a structural break. Model (1) permits a one-time change in the level of the series; Model (2) allows for a one-time change in the slope of the trend function, and model (3) allows for the change in both in level and slope of trend.

$$m_t = c + \beta m_{t-1} + \gamma t + \theta DU_t + \sum_{l=1}^{k} \Delta m_{t-l} + e_t \dots model\ 1$$

$$m_t = c + \beta m_{t-1} + \gamma t + \delta DT_t + \sum_{l=1}^{k} \Delta m_{t-l} + e_t \dots Model\ 2$$

$$m_t = c + \beta m_{t-1} + \gamma t + \theta DU_t + \delta DT_t + \sum_{l=1}^{k} \Delta m_{t-l} + e_t \dots Model\ 3$$

$$DU_t = \langle \begin{matrix} 1 \dots if\ t > BD \\ 0 \dots otherwise \end{matrix} \rangle$$

While

$DT_t = \langle {t - TB \ldots if \ t > BD \atop 0 \ldots otherwise} \rangle$

The null hypothesis for these models is β=0, while the alternative hypothesis β <0. In the Zivot-Andrews method, every point is a potential break-date (BD) and runs a regression sequentially for every possible BD. The selection of the possible breakpoints (BD) is based on the minimised the t-statistic for testing $\hat{\beta} = \beta - 1 = 1$. The Zivot-Andrews suggests that the 'trimming region' needs to be specified as (0.15T, 0.85T) due to the presence of the endpoints causing the asymptotic distribution of the statistics that diverges towards infinity. For this study, we will use model (3), based on the recommendations of Sen (2003) that suggest that model 3, 'mixed model', is superior to model 1 and model 2. Based on these comments we choose model C for testing the structural break analysis.

5.4 Estimation Methodology

For this study we are proposing two widely used method of estimation (instead of relying on single technique). First, we are using ARDL bound testing estimation techniques and error correction model, as this technique is highly recommended when variables are integrated of mixed order (I) and (0). This method will help us to obtain the long run output elasticity of each type of infrastructure but also helpful to analyse the short run dynamics. Second, we are estimating the long run coefficients using fully modified ordinary least square method (FMOLS). This method provides reliable estimates if model suffers from the issue of serial autocorrelation. This estimation method also takes endogeneity into the account and estimators are consistent with the variables of mixed order of integration (1) & (0) (Phillips, 1993).

5.4.1 Cointegration tests: ARDL bound testing approach.

Using the ARDL estimation technique, F-bound test is based on the joint F-statistic under the null hypothesis of no cointegration. The unrestricted vector error correction model (VECM) representation of the model as $ARDL(p, q_1, q_2, q_3, q_4, q_5)$ model for equation as eq (5.2)

$$\Delta ln\widehat{y_t} = \varphi_0 + t + \varphi_1 ln\widehat{y_{t-1}} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgq_{t-1} + \varphi_5 lnSgs_{t-1} + \varphi_6 lnSgf_{t-1} - \varphi_7 ln(\delta + \lambda + n)_{t-1} + \sum_{i=1}^{p} \delta_{1i} \Delta\, ln\widehat{y_{t-i}} + \sum_{i=1}^{q1} \delta_{2i}\, \Delta lnSk_{t-i} + \sum_{i=1}^{q2} \delta_{3i}\, \Delta lnSh_{t-i} + \sum_{i=1}^{q3} \delta_{4i}\, \Delta lnSg_{t-i} + \sum_{i=1}^{q4} \delta_{4i}\, \Delta lnSs_{t-i} + \sum_{i=1}^{q5} \delta_{4i}\, \Delta lnSf_{t-i} + \sum_{i=1}^{q6} \delta_{3i}\, \Delta ln(\delta + \lambda + n)_{t-i} + \varepsilon_t \qquad (5.2)$$

The F-bound test using the joint-testing hypothesis to test null hypothesis of no-cointegration ($\varphi_0 = \varphi_1 = \varphi_2 = \varphi_3 = \varphi_4 = \varphi_5 = \varphi_6 = \varphi_7 = 0$) against the alternative hypothesis, we can only conclude of the existence of cointegration if the F-stat value is greater than upper bound value I(1), if value fall below the lower bound value than we can conclude for no cointegration and if value fall between the critical bound I(1) and I(0) values then we can conclude an inconclusive evidence to support for cointegration. The critical level of I(1) and I(0) are developed by the Pesaran et al. (2001). While, if results concluded for the cointegration than we can estimated restricted VECM.

$$\Delta ln\widehat{y_t} = \varphi_0 + t + \sum_{i=1}^{p} \delta_{1i} \Delta\, ln\widehat{y_{t-\iota}} + \sum_{i=1}^{q1} \delta_{2i}\, \Delta lnSk_{t-i} + \sum_{i=1}^{q2} \delta_{3i}\, \Delta lnSh_{t-i} + \sum_{i=1}^{q3} \delta_{4i}\, \Delta lnSg_{t-i} + + \sum_{i=1}^{q4} \delta_{4i}\, \Delta lnSs_{t-i} + \sum_{i=1}^{q5} \delta_{4i}\, \Delta lnSf_{t-i} + \sum_{i=1}^{q6} \delta_{3i}\, \Delta ln(\delta + \lambda + n)_{t-i} - \gamma ECT + \varepsilon_t \quad (5.3)$$

The ECT represents the error correction term, shows the speed of adjustment to reach the long-run equilibrium in case of any shock in the system. The coefficients of negative and must be statistically significant as necessary condition, while sufficient condition, is that the absolute value of coefficient should less than 1, for monotonic restoring equilibrium.

We are following standard diagnostics of the Autoregressive distributed lag (ARDL) model for testing the reliability of estimates.

1. Testing normality: We are using the Jarque–Bera test to measure the normal distribution of residuals (Jarque & Bera, 1987). It is a test that measure a goodness-of-fit, measuring whether data have the skewness and kurtosis matching a normal distribution. If residual when plot away from zero, depicts the data do not have a normal distribution.
2. Detection of serial autocorrelation: In order detect the problem of autocorrelation in the errors in a regression model, we are using the Breusch–Godfrey test. This test is based on the Lagrange multiplier (LM) testing mechanism that’s reason also commonly referred to the LM test for serial correlation. We are testing the null hypothesis of no serial correlation for the lag order of (2).
3. Detection of the Heteroscedasticity: In order to detect the presence of heteroscedastic variance, we are using the autoregressive conditional heteroskedasticity (ARCH) LM test. We are testing the null hypothesis of no ARCH effects against the alternative, for detecting the presence of heteroscedasticity.

4. Stability of the model: The regressions coefficients are expected to stable over time, if not; the time-series model has been suffering from the issue of the structural break. Therefore, the Cusum and Cusum Square test that inspects the stability of the coefficients over time (Ploberger & Krämer, 1992). The test states to be stable if the Cusum graph and Cusum square graph remained within the graphical of 5% significance level.
5. Functional form specification: Besides, all the other diagnostics, the test for specification of correct functional form is important. In order to confirm the model is correct specified, we are using the Ramsey Regression Equation Specification Error Test (RESET) test. This joint test of the significance of coefficients and accepting the null hypothesis represents that the model is specified correctly. This test was developed by J. B. Ramsey, (1969) to test the specification issues of Classical Ordinary Linear Least-Squares (OLS) Regression Analysis. We are using ARDL estimation technique, and this model is built up on the OLS regression estimation technique; therefore, RESET test would be useful in this scenario.

5.4.2 Cointegration tests: FMOLS estimation

For this study, we employ Fully Modified Ordinary Least Squares (FMOLS) too, preferably for robustness check and keeping the post-regression diagnostics of the ARDL estimation technique in view. The post estimation regression test for the presence of cointegration if statistically significant, then coefficients obtained from the FMOLS regression will be interpreted as long-run estimates or the output-elasticity.

The cointegration test, namely 'Engle & Granger cointegration test' (1987) and 'Phillips & Ouliaris cointegration test' (1990) and 'Hansen's instability test' developed by (Hansen, 1992) would be used for the detection of a cointegrating relationship between growth and infrastructure. The Engle-Granger and Phillips-Ouliaris residual-based tests for cointegration based on the unit root tests of the residuals obtained from FMOLS estimation. Under assumption, the series, and all linear combinations of dependent and independents must not be cointegrated, including the residuals. Therefore, the null hypothesis of no cointegration against the cointegration alternative is to be tested. The Engle-Granger cointegration test uses a parametric approach of the augmented Dickey-Fuller (ADF), while the Phillips-Ouliaris test uses the nonparametric approach based on the Phillips-Perron (PP) methodology. While Hansen (1992) developed a langrage multiplier test to test cointegration given the condition of stability of parameters, this test follows the null hypothesis of cointegration against the alternative of no cointegration.

5.5 Empirical Model

The literature extensively asserts a positive role of infrastructure as a growth-promoting input. However, an important policy question of what components of the infrastructure are more important for productivity-driven growth in Pakistan is yet to answer. The primary aim of this research exercise is to answer this key policy-oriented question and empirically validate the postulation and corollary of the positive association of infrastructure components with economic growth. In this context, this study developed indices using an extensive set of economic, social, and financial infrastructure indicators for Pakistan. Moreover, this study incorporates gauging the growth effect of the physical stock of infrastructure and considering infrastructure utilization for policy-driven analysis. In this connection we developed an empirical model (equation 5.1), based on the mathematical model developed in the previous chapter 4.

$$dln\widehat{y_t} = \varphi_0 + \varphi_1 lnSk_t + \varphi_2 lnSh_t + \varphi_3 \ln Sg_4 + \ lnSs_t + \varphi_4 \ lnSf_t - \varphi_5 ln(\delta + \lambda + n) + \varphi_6 (ln\widehat{y_{t-1}}) + \varepsilon_t \quad (5.4)$$

Where Sg_t represents the share of core infrastructure, $Sc_t\, represents$ share of social infrastructure which is defined as the social infrastructure for education and health. While Sf represent share of resources devoted to the financial infrastructure. In this study we have further categorized these components of infrastructure into physical quantity and infrastructure utilization terms. Therefore, we can express the empirical models into two equations 5.5 & 5.6.

$$dln\widehat{y_t} = \varphi_0 + \varphi_1 lnSk_t + \varphi_2 lnSh_t + \varphi_3 \ln Sgcq_t + \varphi_4 lnSgedq_t + \varphi_5 lnSghthq_t + \varphi_6 \ lnSgfq_t - \varphi_7 ln(\delta + \lambda + n) + \varphi_8 (ln\widehat{y_{t-1}}) + \varepsilon_t \quad (5.5)$$

$$dln\widehat{y_t} = \varphi_0 + \varphi_1 lnSk_t + \varphi_2 lnSh_t + \varphi_3 \ln Sgcu_t + \varphi_4 lnSgedu_t + \varphi_5 lnSghthu_t + \varphi_6 \ lnSgfu_t - \varphi_7 ln(\delta + \lambda + n) + \varphi_8 (ln\widehat{y_{t-1}}) + \varepsilon_t \quad (5.6)$$

5.6 Indicators, indices, and Data

The data have been attained from the various issues of the Pakistan's Economic Survey published by the Government of Pakistan yearly and Payment Review--Statistical Reports published by the State Bank of Pakistan, and the World Development Indicators (WDI) by the World Bank for FY 1980-81 to 2021-22. A brief description of variables is provided in Table 5.1. We are using the set of controls variables, including domestic credit to the private sector ($lnDCRPS_t$), Current account balance (CAB_t), FDI net inflows ($lnFDI_t$), trade intensity index ($lnTI_t$), real effective exchange rate index($lnREER_t$), and structural break (SB) dummy for year 2004. The structural break dummy of 2004, which takes on values of 1 from 2004 and onwards years is to control the structural transformation effects of post deregulation of Musharaf-Aziz era and privatization of the telecommunication/banking sector and preferential-trade agreements with the neighbouring nation China. All variables are transformed into natural logarithms except for the indices for core infrastructure and CAB.

Table 5. 1 *Description of indicators*

Variable	Description
y_t	Output (Real GDP) per worker measuring productivity of a worker
Sk_t	Gross fixed capital formation by private sector expressed as a percentage of GDP measuring share of private sectors investment in physical capital.
Sh_t	Sum of expenditure on the health and education sector expressed as a percentage of GDP as a proxy for the new investment in the human capital
$Sgcq_t$	Index for quantification of core infrastructure
$Sgcu_t$	Index for measuring share of core infrastructure's utilization
$Sgedq_t$	Index for quantification of Infrastructure for education
$Sgedu_t$	Index for utilization of educational infrastructure
$Sghthq_t$	Total hospital beds per (million) workers
$Sghthu_t$	BCG immunization (% of children)
$Sgfq_t$	Number of bank branches and ATM machines
$Sgfu_t$	Number of bank accounts (million)
$n_t+\delta_t + \lambda$	Breakeven level of investment

5.6.1 Construction of Indices

5.6.1.1. Indices of Core Infrastructure

Infrastructure is multidimensional by definition. Fourie (2006) elaborated two widely used definitions of infrastructure. The first definition is the 'social overhead capital (SOC)'; the second is a list of all possible infrastructural goods. Such a transport infrastructure, communications infrastructure, energy supply infrastructure, water infrastructure, environmental infrastructure, education infrastructure, etc., Besides the theoretical underpinnings, the empirical literature used the four major but different approaches to measure infrastructure. These are financial flows, financial stock, physical stock and common inventory methods (Gianpiero, 2010). There are merits and demerits of each empirical definition adopted. However, in Pakistan, using physical units is a better indicator than monetary units, as Pritchett (2000) elaborates the major loophole, especially in developing countries, that all public investment or ‘social overhead capital’ doesn’t transform into physical capital. We are constructing indices, and we are using the Principal Component Analysis (PCA) technique to calibrate the physical stock and utilization of types/components of infrastructure. The statistical method of PCA has been used to reduce the number of dimensions/indicators. Literature defines the infrastructure for transport, energy, and communication services as the core infrastructure. The infrastructure for transportation is a precursor for spatial connectivity and economic integration. And the infrastructure for energy is a fundamental driver to provide the essential ingredient of power, which is a necessary non-substitutable input for the production of goods and services. At the same time, the infrastructure for communication services is essentially integrating and reducing the transaction cost. For estimation of the nexus of infrastructure-productivity growth, we further compartmentalize each type's two aspects. First for the quantification and second to calibrate the role of utilization of core infrastructure. The brief method of construction of these indices are as under;

The index for quantification of core infrastructure ($Sgcq_t$) that represents the share of core infrastructure comprises the following variables/indicators of transport, energy, and communication infrastructure.

- Length of high-type (HTR) roads (km)
- Electricity generation (EGE)(Gw/h)
- Crude oil extraction (COE)(000 barrels)
- Production of Gas (PoG)(mcf)
- Production of Petroleum Products (PoPP) (000 tons)
- Production of Coal (PoCO) (000 tons)
- Total Circuit Length (km) of electricity transmission lines (LTL)
- Landline telephone and mobile-phone subscribers (TELMOB)(000 Nos)

The data of these indicators were obtained from various issues of the Economic Survey of Pakistan, published by the Ministry of Finance, except Total Circuit Length (km) of electricity transmission lines, that Power System Statistics published by NTDC.

All indicators are in different units of measurement. Therefore, the Z-score method for data normalisation transforms all data series into a single unit of measure. The z-score is employing the following Eq 5.7;

$$z = \frac{x_i - \mu}{\sigma} \quad (5.7)$$

The PCA technique reduces the number of dimensions for $Sgcq_t$. The PCA technique assign appropriate weights to the indicators. This method allows various data series into to single variable, preserving the information of the original dataset.

$$Sgcq = \omega_1 HTR + \omega_2 COE + \omega_3 PoG + \omega_4 PoPP + \omega_5 PoCO + \omega_6 EGE + \omega_7 LTL + \omega_8 TELMOB \quad (5.8)$$

The $Sgcq$ is constructed based on the estimation of equation (1.32), where $\omega_k' s$ represents the respective weights of each indicator, given by the respective eigenvector of the selected principal component. The eigenvalues and eigenvectors of the index for quantification of core infrastructure are in table 5.2 given below. The main principle component (PC)-1, showing 82.1% variation, with the highest Eigenvalue of 6.573. The (decision criteria) rule of thumb for selecting PC is to choose the Eigen value is equal and or greater than one.

Table 5.2 *The correlation matrix of indicators (Sgcq)*

	Eigen Vectors							
Indicators	**PC 1**	**PC 2**	**PC 3**	**PC 4**	**PC 5**	**PC 6**	**PC 7**	**PC 8**
HTR	0.349	-0.429	0.013	-0.246	0.354	-0.201	0.549	0.408
COE	0.352	-0.232	0.650	0.588	0.039	0.211	-0.091	-0.017
POG	0.366	-0.263	-0.392	-0.193	0.009	0.745	-0.223	-0.034
POPP	0.370	-0.188	-0.288	0.193	-0.743	-0.352	-0.062	0.166
POCO	0.286	0.639	0.321	-0.303	-0.307	0.270	0.256	0.288
EGE	0.390	0.117	-0.065	-0.038	0.022	-0.119	0.379	-0.819
LTL	0.383	0.034	0.229	-0.450	0.195	-0.371	-0.648	-0.051
TELD	0.321	0.485	-0.426	0.473	0.434	-0.110	-0.087	0.221
Eigen Values	6.573	0.742	0.338	0.160	0.099	0.048	0.023	0.014
Variation	0.821	0.092	0.042	0.020	0.012	0.006	0.003	0.001

We normalized these weights to of PC-1 into 1, and constructed the 'index for core infrastructure's quantification $Sgcq_t$' using these normalized weights by using equation (5.7)

$$Sgcq = 0.12\, HTR + 0.12\, COE + 0.13\, PoG + 0.13 PoPP + 0.11 PoCO + 0.14 EGE + 0.13\, LTL + 0.12 TELMOB \quad (5.9)$$

Besides its physical presence is important to analyze the growth effect of utilization of core infrastructure. The index is index for measuring share of core infrastructure's utilization ($Sgcu_t$) includes the following indicators of transport infrastructure, energy, and communication infrastructure utilization;

- Vehicles on the road (VOR) (000 Nos)
- Passengers traveled by rail (PTR) (mln Nos)
- Cargo carried by rail (CCR) (mln tons)
- Freight, Air Transport (FAT) (million ton-km)
- Total Cargo handled at the Seaports of Karachi, Port Qasim and Gwader (CHP) (000 tons)
- Consumption of Electrcity (COEL) (GwH)
- Consumption of oil/Petroleum (COOP)(tons)
- Consumption of Gas (COG) (mm mcf)
- Consumption of Coal (COCO) (000 metric tons)
- Teldensity (TELD); Landline telephone, mobile-phone subscribers, and broadband subscribers (mln Nos)

These indicators were measured in different units of measurement; therefore, we normalized data series using z-score method before applying PCA techniques for assigning weights for $Sgcu_t$.

$$Sgcu_t \ = \omega_1\text{VOR} + \omega_2\text{PTR} + \omega_3\text{CCR} + \omega_4\text{FAT} + \omega_5\text{CHP} + \omega_6 COEL + \omega_7\ COOP + \omega_8 COG + \omega_9 COCO + \omega_{10} TELD + \quad (5.10)$$

Where, $\omega_k's$ represents the respective weights of each indicator, given by an eigenvector of principal component. The eigenvalues and eigenvectors of the $Sgcu_t$ are provided in table given below.

The principle-component (PC)-1 showing 72.41 % variation and PC 1 has the highest Eigen value of 7.24. The correlation matric and scree plot (the ordered Eigen values). Therefore, we are using PC-1 for assigning normalize weights of PC-I are used to calculate $Sgcu_t$ using equation (5.9).

$$Sgcu_t \ = 0.22\ \text{VOR} - 0.19\text{PTR} - 0.12\text{CCR} + 0.20\text{FAT} - 0.15\text{CHP} + 0.20 COEL + 0.21 COOP + 0.22 COG + 0.19 COCO + 0.22 TELD \quad (5.11)$$

Table 5. 3 *The correlation matrix of indicators of Sgcu*

					Eigen vector					
Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10
VOR	0.36	0.19	-0.03	0.03	0.03	0.14	-0.12	0.49	-0.57	0.49
RPT	-0.31	0.23	0.54	-0.46	0.40	-0.05	0.33	0.28	-0.02	-0.04
CCR	-0.20	0.61	0.22	0.43	-0.01	0.55	-0.12	-0.17	0.10	0.00
FAT	0.32	0.13	0.53	0.37	-0.40	-0.49	0.26	0.00	0.02	-0.03
CHP	-0.24	-0.50	0.18	0.61	0.42	-0.01	-0.06	0.30	0.07	0.09
COEL	0.32	-0.32	0.06	-0.04	-0.13	0.59	0.62	0.03	0.17	0.04
COOP	0.34	-0.19	0.43	-0.18	0.28	0.06	-0.35	-0.52	0.10	0.38
COG	0.36	-0.08	0.18	0.01	0.17	0.18	-0.25	0.07	-0.31	-0.78
COCO	0.31	0.29	-0.35	0.21	0.61	-0.22	0.40	-0.26	-0.02	-0.01
Eigen values	7.24	1.56	0.41	0.34	0.19	0.12	0.08	0.04	0.01	0.00
variation	72.41	15.63	4.12	3.45	1.87	1.23	0.79	0.37	0.11	0.02

The correlation test score are given in table 5.4 for the indicators measuring core infrastructure stock. They test exhibits a positive correlation amongst all variables.

Table 5. 4 *Correlation Analysis: Indicators of core infrastructure*

	HTR	EGCOIL	EGGAS	EGPP	EGCOAL	ELEG	TCLKM	TELD
HTR	1.00	0.87	0.94	0.89	0.36	0.83	0.87	0.50
EGCOIL	0.87	1.00	0.82	0.85	0.52	0.84	0.85	0.59
EGGAS	0.94	0.82	1.00	0.93	0.49	0.89	0.88	0.65
EGPP	0.89	0.85	0.93	1.00	0.54	0.91	0.87	0.69
EGCOAL	0.36	0.52	0.49	0.54	1.00	0.79	0.75	0.85
ELEG	0.83	0.84	0.89	0.91	0.79	1.00	0.96	0.87
TCLKM	0.87	0.85	0.88	0.87	0.75	0.96	1.00	0.77
TELD	0.50	0.59	0.65	0.69	0.85	0.87	0.77	1.00

At the same time, in Table 5.5 presented the correlation test scores of indicators of transport infrastructure utilization such as vehicles on roads (VOR), Passengers traveled by rail, Cargo carried by rail, Freight, via Air Transport, Total Cargo handled at the Seaports. The negative correlation test score between VOR and RPT, CCR and FAT represents the substitution effects between road and rail transportation, and air transport.

Table 5. 5 *Correlation Analysis: Indicators of transport infrastructure*

	VOR	RPT	CCR	CHP	FAT
VOR	1.00	-0.74	-0.33	0.86	-0.77
RPT	-0.74	1.00	0.64	-0.63	0.34
CCR	-0.33	0.64	1.00	-0.27	-0.02
CHP	0.86	-0.63	-0.27	1.00	-0.59
FAT	-0.77	0.34	-0.02	-0.59	1.00

At the same time, utilization indicators of energy infrastructure are positively correlating depicting each energy type is associated in a complementing fashion. The results are self-explanatory and provided in table 5.6.

Table 5. 6 *Correlation Analysis: Indicators of energy infrastructure utilization*

	ECONOIL	ECONGAS	ECONELECTRICITY	ECOCOAL
ECONOIL	1.00	0.88	0.89	0.56
ECONGAS	0.88	1.00	0.96	0.62
ECONELECTRICITY	0.89	0.96	1.00	0.77
ECOCOAL	0.56	0.62	0.77	1.00

5.6.1.2. Indices for social infrastructure (education)

The educational infrastructure comprises all educational facilities that help increase the stock of knowledge and education. Therefore, we are developing an equi-weighted index using the following indicators using equation 5.10 & 5.11.

Quantification (per mln population)	**Utilization (mln)**
▪ Number of primary schools (NPS)	▪ Enrollment in primary school (EPS)
▪ Number of secondary schools (NSC)	▪ Enrollment in secondary school (ESC)
▪ Number of high schools (NHS)	▪ Enrollment in High school (EHS)
▪ Number of technical and vocational centers (NTVC)	▪ Enrollment in technical vocational centers (ETVC)
▪ Number of higher secondary school (NHHS)	▪ Enrollment in higher secondary school (EHHS)
▪ Number of degree colleges (NDC)	▪ Enrollment in degree college (EDC)
▪ Number of universities (NUN)	▪ Enrollment in universities (EUN)

$lnSgedq_t = 1/7 * (NPS) + 1/7 * (NSC) + 1/7 * (NHS) + 1/7 * (NTVC) + 1/7 * (NHHS) +$
$1/7 *$ (NDC)+1/7(NUN) (5.10)

$lnSgedu_t = 1/7 * (EPS) + 1/7 * 4(ESC) + 1/7 * (EHS) + 1/7 * (NTVC) + 1/7 * (EHHS)$
$+1/7 *$ (EDC) +1/7*((EUN) (5.11)

5.7 Unit-root tests: Results

Moreover, following the traditions of time series econometric analysis, at the outset of empirical estimation testing for the presence of presence of unit root in data series is conducted. We are using the Augmented Dicky Fuller test (ADF) test, Philip Perron Test (PP test) Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test for the detection of a unit root in series and Z-Andrew (ZA) test for testing the presence a structural break in data series. The empirical literature recommends to perform three test of ADF, PP and KPSS for decision of order of integration (Shrestha & Bhatta, 2018). A summary is given in table 5.7.

Table 5. 7 *Order of Integration: A Summary*

	ADF test	**PP test**	**KPSS**	**Year of SB**
Variable				
lny_t	I(1)	I(0)	I(1)	1991
$lnSk_t$	I(1)	I(1)	I(0)	2009***
$lnSh_t$	I(0)	I(1)	I(0)	2004***
$Sgcq_t$	I(0)	I(1)	I(0)	2004***
$Sgcu_t$	I(1)	I(1)	I(1)	2001
$lnSgedq_t$	I(0)	I(0)	I(0)	1995***
$lnSgedu_t$	I(1)	I(0)	I(1)	1999**
$lnShthq_t$	I(1)	I(1)	I(0)	1988
$lnSghthu_t$	I(0)	I(0)	I(1)	2004
$lnSgfq_t$	I(1)	I(1)	I(1)	2001***
$lnSgfu_t$	I(1)	I(1)	I(0)	2004**
$ln(\delta + \lambda + n)_t$	I(1)	I(0)	I(0)	2005***
$lnFDI_t$	I(1)	I(1)	I(0)	2006
$lnCPI_t$	I(0)	I(0)	I(0)	1998*
$lnDCRPS_t$	I(0)	I(1)	I(0)	2003***
$lnTI_t$	I(1)	I(1)	I(0)	1998
$lnREER_t$	I(0)	I(0)	I(1)	1989
$lnCAB_t$	I(0)	I(0)	I(0)	1998

5.8 Estimations and key findings

From the previous section, we conclude that the variables are integrated of I(1) and I(0) and none of the variables is I(2), therefore we are employing the ARDL and FMOLS estimation technique techniques for estimation of parameters. We developed an empirical model for the ARDL estimation technique and testing for cointegration using the bound testing approach as described in equation 5.5.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcq_{t-1} + \varphi_4 lnSgedq_{t-1} + \varphi_4 lnSghthq_{t-1} + \varphi_4 lnSgfq_{t-1} - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgcq_t + \delta_4 \Delta lnSgedq_t + \delta_5 \Delta lnSghthq_t + \delta_6 \Delta lnSgfq_t - \delta_7 \Delta ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.7)$$

We estimated the equation (5.7), but estimated coefficients for all control variables were insignificant at 5 percent therefore estimated the model again without control variables and tested the post estimation diagnostics. The F-bound test value is greater than I(1) bound that confirms the cointegration amongst variables of the model and we selected model ARDL(1,1) based on lowest AIC, BIC and HQIC. The bound-test and information criterion for ARDL(1,1), ARDL(2,2) is given below in table 5.8. Based on the lowest AIC, SIC and HQIC, we estimated model eq 5.7 using (1,1) lag order and the estimated results along with post estimation diagnostics are presented in model 5/1-A in table 5.10.

Table 5.8 *Lag-order Selection of Model 5(A)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test:	4.049***	3.188	3.342
5% critical value I(1) upper bound: 3.723			
AIC	-6.170***	-5.151	-5.414
SIC	-4.819***	-4.489	-4.411
HQIC	-5.682***	-4.908	-5.049

In addition to this, we are estimated growth effect of utilization of infrastructure the ARDL estimation technique based on equation 5.8.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcu_{t-1} + \varphi_4 lnSgedu_{t-1} + \varphi_4 lnSghthu_{t-1} + \varphi_4 lnSgfu_{t-1} - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgcu_{t-1} + \delta_4 \Delta lnSgedu_t + \delta_5 \Delta lnSghthu_t + \delta_6 \Delta lnSgfu_t - \delta_7 \Delta ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.8)$$

The F-bound tests and results of the model selection criterion are given in table 5.9. Based on lowest SIC and HQIC, we estimated model eq 5.8 using (1,1) lag order. and the estimated results along with post estimation diagnostics are presented in model 5-B in table 5.10.

Table 5.9 *Lag-order Selection of Model (B)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test:	4.798***	2.175	2.105
5% critical value I(1) upper bound: 3.723			
AIC	-5.313	-5.169	-5.543***
SIC	-4.651***	-4.166	-4.192
HQIC	-5.070***	-4.804	-5.055

For the purpose of conducting robustness checks, we employed the Fully Modified Ordinary Least Squares (FMOLS) model estimation technique, using equation 8.6 to assess the long-run growth effect of the physical quantity and utilization of core, social, and financial infrastructure (using equations 5.9 and 5.10, respectively). Initially, we ran the model with only the intercept term and performed the 'parked added variable test' to include trend and trend square as additional regressors in the model.

$$ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgcq_t + \varphi_4 lnSgedq_t + \varphi_4 lnSghthq_t + \varphi_4 lnSgfq_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnFDI_t + \vartheta_2 lnCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 D_{2004} + \varepsilon_t \quad (5.9)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgcu_t + \varphi_4 lnSgedu_t + \varphi_4 lnSghthu_t + \varphi_4 lnSgfu_t - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 D_{2004} + \varepsilon_t \quad (5.10)$$

The results of the post-estimation analysis revealed the presence of cointegration, indicating a long-run association among the variables. The coefficients, along with their corresponding p-values, model summary statistics (R-Square and adjusted R-Square), and the results of all co-integration tests, are presented in Table 5.11. Moreover, the post-estimation test for co-integration, the Engle-Granger and Phillips-Ouliaris tests, confirmed the existence of a long-run association among the variables. The p-values allowed us to reject the null hypothesis of no cointegration in favor of the alternative hypothesis of cointegration. These results are also provided in Table 5.11. Additionally, Table 5.11 contains estimated coefficients, their corresponding p-values, model summary statistics (R-Square and adjusted R-Square), and the outcomes of co-integration tests.

5.8.1 Key findings

The results demonstrate a positive contribution of public investment in human capital and physical health infrastructure to the long-run productivity growth of Pakistan during 1980-2022. Across all estimated models, productivity growth is impeded by the effects of increasing population growth and depreciation of physical capital stock in Pakistan. The coefficient of the variable $ln(\delta + \lambda + n)_t$ is consistently negative and highly significant in all estimated models. This coefficient measures the break-even level of investment required to maintain the labor force and counteract capital depreciation. Furthermore, the productivity growth does not increase significantly as a result of the increase in output per worker. The coefficient of lny_{t-1} measures the level effect of output per worker on productivity growth in Pakistan, and the estimated parameters show a negative and statistically significant relationship. Based on estimated models C and D, we can conclude that the level effect of output per worker is not sufficiently high to generate long-run productivity growth in Pakistan. In terms of private sector investment, the coefficient of investment in physical capital is only significant in model B, while the coefficient of investment in human capital is significant at the 5 percent level in model C.

The estimated models have revealed all crucial components, particularly the highly statistically significant LR estimates for physical infrastructure and health infrastructure, while the utilization indicators of education, health, and financial infrastructure have proven to be critical factors that contributed to the LR productivity growth of Pakistan. By comparing the estimated coefficients of aforementioned models of LR productivity growth, it can be concluded that the stock of physical infrastructure has contributed to LR productivity growth; however, the utilization of core infrastructure is insufficient to generate LR productivity growth. An increase in the stock of core infrastructure by one standard deviation (0.900) results in an average LR productivity growth increase of 0.238 percentage points, while holding the effects of other variables in the models constant. Meanwhile, the physical infrastructure of the educational network does not directly impact LR productivity growth, but its utilization has shown a positive contribution. Holding the effects of investment in physical core, health, and financial infrastructure, as well as other investment drivers and control variables, constant, a one percent increase in the utilization of infrastructure for education or the size of educational services leads to a 0.195 percentage point increase in LR productivity growth. Similarly, both physical infrastructure and utilization indicators of health infrastructure have contributed to LR productivity growth in Pakistan. Holding the effects of other variables in the models constant, a one percent increase in the number of hospital beds per million workers leads to a 0.267 percentage point increase in output per worker growth in LR. Additionally, a one percent increase in BCG immunization coverage has contributed to a 0.087 percent increase in LR productivity. Furthermore, keeping the effects of infrastructure components and control variables constant, an increase in the utilization of financial infrastructure has also contributed to the productivity growth of Pakistan. A one percent increase in the number of bank accounts (in millions) has been associated with a 0.105 percent increase in LR productivity growth.

Estimated long-run coefficients of infrastructure's types

	0.238*** (A)	-0.001 (B)	
$Sgcq_t$	0.048** (C)	-0.008 (-D)	$Sgcu_t$
$lnSgedq_t$	-0.005 (A)	0.195** (B)	$lnSgedu_t$
	-0.059 (C)	0.007 (D)	
$lnSghthq_t$	0.267***(A)	0.087*** (B)	$lnSghthu_t$
	0.291*** (C)	0.021*(D)	
$lnSgfq_t$	-0.041(A)	0.139 (B)	$lnSgfu_t$
	0.013(C)	0.105***(D)	

In SR, the physical infrastructure of core infrastructure, utilization of infrastructure for education, and physical infrastructure for health and financial services are the significant components that impact the productivity growth in SR. At the same time, the high population growth rate is hampering the productivity growth in SR by 0.011 percentage points. The ECT is negative and absolute value is less than one therefore we can conclude for the convergence. Amongst all control variables, log of the real effective exchange rate is negative and significant, indicating a negative impact on productivity growth due to changes in relative competitiveness resulting from currency depreciation against major trading currencies. The depreciation implies a loss of trade competitiveness in the global market, as exports become more expensive, and imports become cheaper. In the long run (LR), Pakistan experiences a reduction of 0.09 percentage points in productivity growth due to the loss of trade competitiveness resulting from exchange rate depreciation.

Table 5.10 *Growth effect: an ARDL estimation technique*

Dependent Variable: Real GDP per worker (labour productivity)				
	Model 5-A		**Model 5-B**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	-0.181	0.318	0.257**	0.035
$lnSh_t$	0.032	0.769	-0.207	0.131
$lnSgcq_t$	0.238***	0.001		
$lnSgcu_t$			-0.001	0.979
$lnSgedq_t$	-0.207	0.570		
$lnSgedu_t$			0.195**	0.029
$lnSghthq_t$	0.653*	0.063		
$lnSghthu_t$			0.087***	0.012
$lnSgfq_t$	-0.085	0.431		
$lnSgfu_t$			0.139	0.142
$ln(\delta+\lambda+n)_t$	-0.041	0.206	-0.033	0.177
C	10.565***	0.000	11.470***	0.000
Short run dynamics				
$\Delta lnSk_t$	-0.046	0.179	0.056	0.089
$\Delta lnSh_t$	0.006	0.715	-0.020	0.208
$\Delta Sgcq_t$	0.040**	0.056		
$\Delta Sgcu_t$			0.011	0.252
$\Delta lnSgedq_t$	-0.005	0.917		
$\Delta lnSgedu_t$			0.182***	0.000
$\Delta lnSghthq_t$	0.267***	0.002		
$\Delta lnSghthu_t$			-0.016	0.173
$\Delta lnSgfq_t$	0.072**	0.040		
$\Delta lnSgfq_t$			0.129***	0.001
$\Delta ln(\delta+\lambda+n)_t$	-0.011***	0.000	-0.009***	0.000
ΔECT_{t-1}	-0.293***	0.000	-0.342***	0.000
Diagnostics				
Normality (Jarque–Bera test)	0.548	0.760	0.799	0.671
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	1.165	0.329	1.278	0.297
Heteroskedasticity: ARCH test	0.455	0.504	0.350	0.557
Functional form: Ramsey RESET Test (t-stat)	0.366	0.718	0.455	0.653
Cusum/Cusum square graph	Stable		Stable	

Table 5.11 *Growth effect: FMOLS estimation technique*

Dependent Variable: Growth rate of Real GDP per worker (Δlny_t)				
	Model 5-C		**Model 5-D**	
Variable	**Coefficient**	**P-value**	**Coefficient**	**P-value**
lny_{t-1}	-0.373***	0.001	-0.450***	0.001
$lnSk_t$	0.005	0.852	0.042	0.198
$lnSh_t$	0.041**	0.036	-0.022	0.398
$Sgcq_t$	0.048**	0.044		
$Sgcu_t$			-0.008	0.280
$lnSgedq_t$	-0.059	0.234		
$lnSgedu_t$			0.007	0.861
$lnSghthq_t$	0.291***	0.000		
$lnSghthu_t$			0.021*	0.070
$lnSgfq_t$	0.013	0.445		
$lnSgfu_t$			0.105***	0.000
$ln(\delta + \lambda + n)_t$	-0.013***	0.000	-0.012***	0.001
SBD_2004	0.060***	0.005	5.448***	0.002
$lnFDI_t$	-0.001	0.840	0.012	0.103
$lnCPI_t$	-0.008	0.387	-0.008	0.331
$lnDCRPS_t$	0.024	0.276	0.003	0.899
$lnTI_t$	0.022	0.468	0.005	0.889
$LnREER_t$	-0.033	0.417	-0.089**	0.031
CAB_t	0.002	0.165	0.000	0.915
Constant	2.851**	0.023	5.448***	0.002
R-squared	0.718		0.708	
Adjusted R-squared	0.548		0.532	
Cointegration Tests: -				
Engle-Granger tau-statistic	-8.419	0.000	-8.419	0.000
Engle-Granger z-statistic	-52.686	0.000	-52.686	0.000
Phillips-Ouliaris tau-statistic	-8.973	0.000	-8.973	0.000
Phillips-Ouliaris z-statistic	-47.724	0.003	-47.724	0.003
Hansen Parameter Instability-Test	2.751	< 0.01	2.751	< 0.01

5.9 Component-wise Growth Effect: Core Infrastructure

Core Infrastructure and productivity growth nexus is one of globally studied topic (Aschauer, 1989b; Bom & Ligthart, 2014; Canning, 1999; Cantos et al., 2005; Devarajan et al., 1996; Esfahani & Ramírez, 2003; Fan & Chan-Kang, 2005; Holtz-Eakin & Schwartz, 1994; Hulten et al., 2006; Jiang et al., 2016; Munnell, 1990). The studies for Pakistan highlighting the key significance and causality of energy and transport infrastructure in the economic development (K. Ahmed et al., 2021; Javid, 2019; Mehmood et al., 2020; Rehana Siddiqui, 2004; Rizwana Siddiqui & Pant, 2007). This section is particularly presenting the analysis of role of core infrastructure by further enriching the understanding and by examining the asymmetric and non-linear association.

5.9.1 ARDL and NARDL bound testing approach.

We employ two estimation techniques for measuring the growth effect, i.e., ARDL Bounds Testing approach and FMOLS estimation technique to precisely estimates. The variables are integrated of a mixed order I(1) and I(0), and there is no I(2) variable in the dataset. Therefore, we are employing the ARDL bound testing approach to determine long-run and short-run relations amongst the variables. This approach uses joint hypothesis testing to test cointegration. And we are using the FMOLS technique as this technique performs better when the model possesses the characteristics of serial autocorrelation.

We estimated the dynamics of the growth effect of the core infrastructure's physical availability and its utilization first by ARDL bound testing approach, followed by the FMOLS. Using index of core infrastructure quantification along with a set of aforementioned control variables and structural break dummy we estimate equation (5.11) which provides a general representation of the unrestricted VECM.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcq_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \sum_{i=1}^{m} \delta_i \Delta ln\hat{y}_{t-i} + \sum_{i=1}^{n} \delta_i \Delta lnSk_{t-i} + \sum_{i=1}^{n} \delta_i \Delta lnSh_{t-i} + \sum_{i=1}^{n} \delta_i \Delta lnSgc_{t-i} + \sum_{i=1}^{n} \delta_i \Delta ln(\delta + \lambda + n)_{t-i} + \varepsilon_t \quad (5.11)$$

By following Enders (2014), we are using a maximum of three lags for lag-order selection even though the empirical literature suggests that lag (1) and or lag (2) are appropriate for the ARDL estimation technique for time-series datasets. We are estimating the model with all the control variables and estimating the model again with only significant control variables. We are selecting the appropriate lag length based on the criteria of the lowest Akaike information criterion (AIC), Schwarz information criterion (SIC), and Hannan-Quinn information criterion (HQIC) to select the parsimonious model to follow the criteria mentioned above. By using equation (5.1), we are testing general model at lag order of 1 to 3. Results are provided in Table 5.14.

Table 5. 12 *Lag order selection for Model 5/1-1(1.1)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	7.211***	3.57	1.72
AIC	-5.15***	-4.99	-4.94
SIC	-4.74***	-4.36	-4.10
HQIC	-5.00***	-4.76	-4.64

The ARDL (1,1) model concludes for the existence of the cointegration, as the F-bound test stat 5.584 is greater than the critical upper bound values. Moreover, we select lag order 1 to assess the model's short-run and long-run dynamics. It is because the ARDL (1,1) has the lowest values of AIC, SIC, and HQIC compared to the other two models of ARDL (2,2) and ARDL (3,3). The unrestricted VECM at the lag order (1) of the estimable model is represented in equation (5.12).

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcq_{t-1} - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgq_t - \delta_4 \Delta ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.12)$$

We have estimated the model using all control variables with the fixed lag order of lag (1). However, the structural break dummy and the control variables were highly insignificant; therefore, we have estimated the model by excluding the insignificant control variables. According Pesaran et al., (2001), the developer of this approach, the ratio of coefficients of $-\varphi_2/\varphi_1$, $-\varphi_3/\varphi_1$, $-\varphi_4/\varphi_1$ and $-\varphi_5/\varphi_1$ are the long run elasticities that are determining the long run output growth per worker. The results are provided in Table 5.10. As we detect the cointegration, we can estimate the error correction term to test the convergence hypothesis. The convergence is a state and tendency of the dependent variable (i.e., output per worker moves toward a steady-state equilibrium level of output per worker (Cellini, 1997). Therefore, we estimate the short-run dynamics by estimating the restricted VECM at lag order (1) of equation (5.13).

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgq_t - \delta_4 \Delta \ln(\delta + \lambda + n)_t + \tau ECT_t + \varepsilon_t \quad (5.13)$$

The error correction mechanism (ECM) is widely used in econometrics analysis to measure the convergence hypothesis. The residual of the long run cointegration functions is an essential determinant of ECM. These residuals are generally also known as disequilibrium estimates or error correction terms. It measures the convergence from the long-run equilibrium level and provides the speed of adjustment toward a steady-state level of output-per-worker. The necessary condition for convergence toward equilibrium is the negative sign of the coefficient of the lagged level of error correction term, and the sufficient condition is that the size coefficient must lie between -1 (minus one) and 0 (zero). We are estimating the restricted VECM to calculate the error correction term and analyse the short-run dynamics of the model. The error correction term is -0.152, which is statistically significant, so we conclude that the model is dynamically stable and converges back to long-run equilibrium. The details of the estimation and ECT are in Table 5.14.

To increase our understanding, we estimate cointegrating nonlinear ARDL (NARDL) model to analyse the short- and long-run nonlinearities are introduced via positive and negative partial sum decompositions of the explanatory variables following (Shin et al., 2014). Under NARDL approach the data series is decomposed into positive and negative series around a threshold of zero, thereby distinguishing between positive and negative changes in the rate of growth in data series. This approach is following the Pesaran et al., (2001) approach bound testing approach for testing cointegration and to obtains the LR estimates relationship amongst variables. For estimated equations 5.14 and following F-bound test significant ECM is estimated using equation 5.15 to analyse asymmetric growth effects of core infrastructure in SR.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcq^{+}{}_{t-1} - \varphi_4 lnSgcq^{-}{}_{t-1} - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgq_t - \delta_4 \Delta ln(\delta + \lambda + n)_t + \varepsilon_t \quad (5.14)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgq^{+}{}_t - \delta_4 \Delta lnSgq^{-}{}_t - \delta_5 \Delta \ln(\delta + \lambda + n)_t + \tau ECT_t + \varepsilon_t \quad (5.15)$$

The coefficients of long-run elasticities and error correct term and results of the diagnostics test of the model are given as model 5-1(B) in table 5.14.

Similarly, we are estimating the growth effect using index of core infrastructure's utilization. Using equation 5.16 we estimated model using lag (1), (2), and (3) to select the parsimonious model. We have selected the model ARDL (1,1) as AIC, SIC, and HQIC are the lowest in this case. The unrestricted VECM at the lag order of 1, is as under.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcu_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgu_t - \delta_4 \Delta \ln(\delta+\lambda+n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.16)$$

Table 5. 13 *Lag-order selection for model 5/1-2(1.1)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	9.423***	3.031	2.702
AIC	-5.006***	-4.790	-4.920
SIC	-4.593***	-4.163	-4.076
HQIC	-4.855***	-4.562	-4.615

The coefficients of the control variables were highly insignificant; therefore, we have estimated the model by excluding the insignificant control variables. The bound test results confirmed the evidence of a long-run relationship as the F-Stat value of all models is greater than the critical value of 4.01(Table 5.13). The estimates represent the long-run elasticities and determinants of output per worker, as shown in Table 5.15. Furthermore, we estimated equation (5.14) restricted (VECM) to test the convergence hypothesis and short-run dynamics.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgu_t - \delta_4 \Delta ln(\delta+\lambda+n)_t + \tau ECT_t + \varepsilon_t \quad (5.17)$$

We test the LR and SR asymmetric effect using Shin et al (2014) and Pesaran et al. (2001) bound testing approach for testing cointegration and to obtains the LR estimates. For estimated equations 5.18 and following F-bound test significant ECM is estimated using equation 5.19 to analyse asymmetric growth effects of core infrastructure's utilization in the SR.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgcu^{+}{}_{t-1} - \varphi_4 lnSgcu^{-}{}_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgq_t - \delta_4 \Delta ln(\delta+\lambda+n)_t + \varepsilon_t \quad (5.18)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgu^{+}{}_t - \delta_4 \Delta lnSgu^{-}{}_t - \delta_5 \Delta \ln(\delta+\lambda+n)_t + \tau ECT_t + \varepsilon_t \quad (5.19)$$

The coefficients of long-run elasticities and error correct term and results of the diagnostics test of the model are given as model 5-2(B) in table 5.15.

5.9.2 Estimations: FMOLS estimation technique.

We estimated equation (5.20) to test the cointegration amongst the variables and obtain the long-run estimates of growth effects due to the physical availability of core infrastructure.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 Sgcq_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + t + \varepsilon_t \quad (5.20)$$

Firstly, we have estimated the model (baseline model) without control variables and have tested for cointegration that concluded the existence of cointegration. Secondly, we have estimated the model with all control variables; however, the Hansen instability test was not significant enough to accept the cointegration with the parameter's stability. Therefore, we used the 'Park Added Variables' to determine the inclusion of additional deterministic restrictions of trend and trend-square in the model. The test result suggests that the models must include linear and quadratic trends. After estimation, we used Engle-Granger, and Phillips-Ouliaris cointegration tests to detect a long-run relationship among the variables. We further estimated the non-linear impact of core infrastructure quantification by including square term ($Sgcq\hat{}2_t$) in the equation 5.20. The estimations and test results mentioned above have presented in Table 5.16.

Moreover, we estimated equation (5.21) for growth effects using utilization indicators of core infrastructure.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 Sgcu_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.21)$$

Similar to the previous estimation, we have estimated the baseline model as well as with a set of control variables. For the second estimation, we have used the 'Park Added Variables-cointegration test' to determine whether to include the linear and quadratic trends in the model. The test suggests incorporating the linear and quadratic trends into the estimation for better results. The test results and the long-run estimates are provided in Table 5.17. We also included square term of utilization index $Sgcu\hat{}2_t$ in equ. 5.21 to further analyse for the non-linear LR association.

5.9.3 Interpretation of estimated models and key findings

The estimated model using ARDL and NARDL bound testing approach depicts that core infrastructure stock is the critical investment driver of Pakistan's long-run and short-run productivity growth. We conclude that on average an increase in stock of core infrastructure by one standard deviation (SD), the long-run (LR) productivity growth increased by (0.163*SD+AVG) 0.148 percentage points approx., keeping the effects of other variables of models effect constant. On the other hand, an increase in the index core infrastructure quantity/stock by one SD point increases productivity growth by 0.053 percentage points in the short-run SR. Moreover, LR's coefficient size of 0.163 is larger than SR's estimated coefficient of 0.053, depicting the key importance of infrastructure in the economic development of Pakistan.

The physical and human capital investment coefficients are insignificant in the LR and SR estimated models 5/1-(1) and 5/1-(2). In addition, the asymmetric effect has also been estimated and based on the results of the NARDL bound testing approach (model 2), we can conclude that the asymmetric effects exist in Pakistan's economy in LR and SR. The coefficient for Sgq_pos_t is positive and highly significant, while the coefficient of Sgq_neg_t is statistically insignificant in the LR and SR. The ECT term is negative and highly significant, and its absolute value is less than 1; therefore, we evidence convergence to LR equilibrium. The estimated model 3 & 4 results are based on ARDL and NARDL bound testing to estimate the productivity growth effect using an index core infrastructure utilization (Sgu). The estimated coefficients for the index of the utilization of the core infrastructure ($Sgcu_t$) turned out statistically insignificant at 5 percent significance level in LR and SR. However, the asymmetric effects is significant as the estimated coefficient for $Sgcu\_pos_t$ is statistically significant while $Sgcu\_neg_t$ is insignificant, therefore we can conclude for the prescence of assymtric effect.

The FMOLS estimation technique evaluates the long run cointegrating relationship, keeping reverse causality and endogeneity issues in view. The post-estimation Engle-Granger test and Phillips-Ouliaris test the residuals based on demonstrated cointegration among the variables. We included the square terms of index in the regression to inquire further whether non-linear relationship relation exists. The coefficient of index for core infrastructure quantification is positive and statistically significant, and the coefficient of square terms of index of core infrastructure quantification is negative and significant, therefore, we can conclude that an inverted U-shape-like relation exist. The long-run marginal effect of investment will be 0.09-2*0.021 ($Sgcq_t$) and the turning point value of the index is 2.14. The results given in model 5/1-3 & 5/1-4, shows that coefficients of $Sgcu_t$ is significant and negative, while $Sgcu\text{^}2_t$ is positive and significant, depicting a parabolic graphic relationship between core infrastructure utilization and productivity growth. The long-run marginal effect of core infrastructure utilization over the productivity growth is -0.011+2*0.002($Sgcu_t$) and the turning point value is 2.75 value of index.

In all estimated models the increasing working force (growth rate of the labour force) and depreciation of the capital appeared to be a key impacting variable of SR and LR productivity growth. Based on ECM estimates, on average, an increase in one percent point in labor force growth decreases productivity growth by 0.017 percent approx. in LR and SR. Therefore, we can conclude Pakistan's high population growth is offsetting the impact of investment drivers of physical capital, human capital, and infrastructure or the investment drivers are not high enough to offset the dampening effects of high population and capital stock depreciation.

All estimated models elucidate that the coefficients for investment in private capital and human capital development are statistically insignificant. Henceforth, we can conclude that investment in capital stock and human capital is insufficient to generate productivity growth in Pakistan. The primary reason behind these trends is that the share of the investment in development private capital stock and human capital in percentage terms is not increasing over time. The figure Below is self-explanatory, presenting the data trends in the productivity growth and other investment drivers of private capital, infrastructure, and human capital.

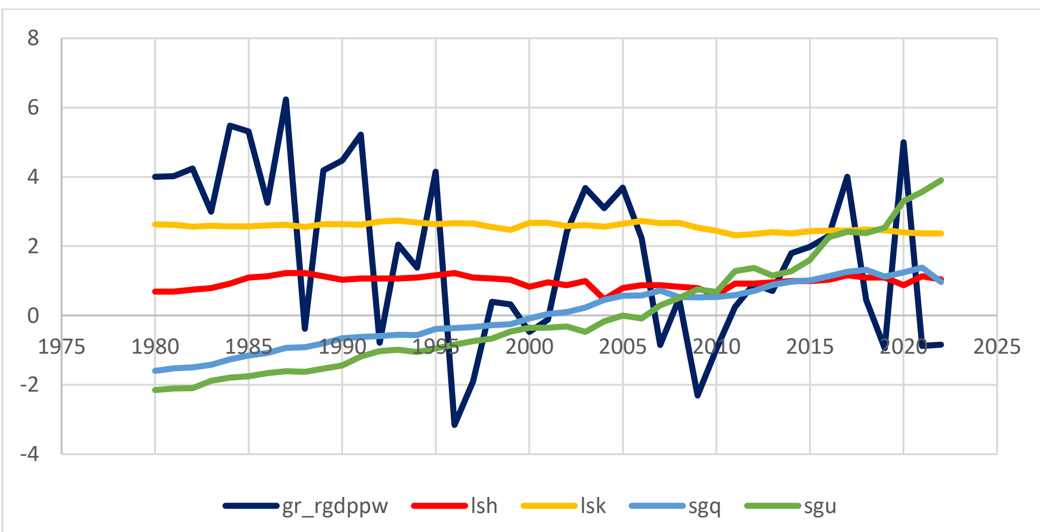


Figure 5.1: **Data Trends: Productivity Growth and investment drivers of core Infrastructure, Private Capital and Human Capital in Pakistan (FY 1980/81-22/23)**

Source: *Author's construct*

Table 5. 14 *Growth effect of the Core infrastructure: ARDL and NARDL bound testing approach*

Dependent Variable: Growth rate of Real GDP per worker(Δlny)				
	Model 5/1-1(1.1)		**Model 5/1-1(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long-run estimates				
$lnSk_t$	0.103	0.650	0.146	0.651
$lnSh_t$	-0.011	0.949	-0.057	0.783
$Sgcq_t$	0.163***	0.000		
$Sgcq\_pos_t$			0.173***	0.006
$Sgcq\_neg_t$			0.154	0.673
$ln(\delta + \lambda + n)_{t-1}$	-0.114	0.154	-0.106	0.183
Short-run dynamics				
C	1.973***	0.000	2.005***	0.000
$\Delta lnSk_t$	0.002	0.954	-0.004	0.928
$\Delta lnSh_t$	-0.009	0.630	-0.005	0.789
$\Delta Sgcq_t$	0.053**	0.028		
$\Delta Sgcq\_pos_t$			0.115**	0.027
$\Delta Sgcq\_neg_t$			0.009	0.821
$\Delta ln(\delta + \lambda + n)_t$	-0.015***	0.000	-0.013***	0.000
SBD2004	-0.007	0.198	-0.006	0.287
ΔECT_t	-0.152***	0.000	-0.161***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.343	0.511	0.945	0.623
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.174	0.841	0.462	0.635
Heteroskedasticity: ARCH test	2.238	0.143	2.463	0.125
Functional form: Ramsey RESET Test (t-stat)	0.956	0.347	1.317	0.199
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 15 *Growth effect of the Core infrastructure: ARDL & NARDL Bound testing approach*

Dependent Variable: Growth rate of Real GDP per worker(Δlny)				
	Model 5/1-2(1.1)		**Model 5/1-2(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long-run estimates				
$lnSk_t$	-0.447	0.693	0.043	0.949
$lnSh_t$	-0.592	0.385	-0.316	0.421
$Sgcu_t$	0.001	0.990		
$Sgcu\_pos_t$			-0.014	0.908
$Sgcu\_neg_t$			-0.320	0.663
$ln(\delta+\lambda+n)_{t-1}$	-0.420	0.266	-0.271	0.211
C	15.306***	0.000	13.469***	0.000
Short-run dynamics				
$\Delta lnSk_t$	0.002	0.954	0.014	0.723
$\Delta lnSh_t$	-0.019	0.321	-0.016	0.391
$\Delta Sgcu_t$	0.005	0.658		
$\Delta Sgcu\_pos_t$			0.020*	0.091
$\Delta Sgcu\_neg_t$			-0.090	0.108
$\Delta ln(\delta+\lambda+n)_t$	-0.017***	0.000	-0.016***	0.000
ΔECT_t	-0.052***	0.000	-0.077***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.298	0.523	1.359	0.507
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.127	0.881	0.056	0.945
Heteroskedasticity: ARCH test	0.319	0.575	0.660	0.422
Functional form: Ramsey RESET Test (t-stat)	1.379	0.178	1.321	0.198
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 16 *Growth effect of the Core infrastructure utilization: FMOLS*

Dependent Variable: Output per worker Growth ($\Delta\ lny_t$)						
	Model 5/1-3(1.1)		**Model 5/1-3(1.2)**		**Model 5/1-3(1.3)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}	-0.065	0.354	-0.096	0.357	-0.080	0.265
$lnSk_t$	-0.024	0.225	-0.013	0.603	-0.013	0.645
$lnSh_t$	-0.008	0.522	-0.012	0.379	-0.007	0.602
$Sgcq_t$	0.099***	0.000	0.105***	0.000	0.090***	0.000
$Sgcq\hat{}2_t$			-0.024**	0.050	-0.021*	0.066
$ln(\delta + \lambda + n)$	-0.016***	0.000	-0.013***	0.000	-0.013***	0.000
$lnFDI_t$	-0.005	0.187	-0.006	0.244		
$lnCPI_t$	0.012**	0.029	0.007	0.297		
$lnDCRPS_t$	0.030**	0.038	0.018	0.287		
$lnTI_t$	0.007	0.723	-0.007	0.795		
$LnREER_t$	0.032	0.144	0.060*	0.083		
CAB_t	0.004***	0.000	0.003***	0.007	0.002***	0.012
SBD_{2004}	-0.015	0.132				
C	-0.028	0.870	0.065	0.760	1.302	0.146
@TREND	-0.009***	0.000	-0.012***	0.001	-0.013***	0.001
@TREND SQUARE	0.000***	0.015	0.000***	0.018	0.000***	0.007
R-squared	0.652192		0.667		0.670	
Adjusted R-squared	0.49071		0.513		0.575	
Cointegration Tests: -						
Engle-Granger tau-statistic	-9.825***	0.000	-8.535***	0.000	-7.789	0.002
Engle-Granger z-statistic	-47.777***	0.001	-52.935***	0.000	-49.330	0.002
Phillips-Ouliaris tau-statistic	-9.825***	0.000	-9.590***	0.000	-8.255	0.001
Phillips-Ouliaris z-statistic	-47.777***	0.001	-45.355***	0.004	-43.905	0.014
Hansen Parameter Inst.Test	2.973	< 0.01	2.206	< 0.01	1.046*	0.062

Table 5. 17 *Growth effect of the Core infrastructure utilization: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker (Δlny_t)						
	Model 5/1-4(1.1)		**Model 5/1-4(1.2)**		**Model 5/1-4(1.3)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}			0.150	0.229	0.009	0.154
$lnSk_t$	0.015	0.704	-0.020	0.622	-0.029	0.314
$lnSh_t$	-0.004	0.872	-0.019	0.455	-0.015	0.368
$Sgcu_t$	-0.002	0.893	-0.029**	0.041	-0.011***	0.004
$Sgcu\text{^}2_t$			0.005**	0.028	0.002***	0.018
$ln(\delta + \lambda + n)$	-0.017***	0.000	-0.018***	0.000	-0.019***	0.000
$lnFDI_t$	0.004	0.542	-0.008	0.252		
$lnCPI_t$	-0.006	0.539	-0.016	0.100		
$lnDCRPS_t$	0.005	0.833	-0.014	0.603		
$lnTI_t$	-0.008	0.836	0.014	0.718		
$LnREER_t$	-0.005	0.907	0.006	0.901		
CAB_t	0.003	0.115	0.001	0.477		
SBD_{2004}			0.034*	0.028	0.007	0.451
C	0.004	0.991	-1.692	0.316		
@TREND	-0.004	0.209				
@TREND SQUARE	0.000	0.315				
Cointegration Tests: -						
Engle-Granger tau-statistic	-7.866***	0.000	-7.452***	0.001	-5.779***	0.013
Engle-Granger z-statistic	-50.075***	0.000	-47.975***	0.001	-38.006***	0.010
Phillips-Ouliaris tau-statistic	-8.528***	0.000	-7.876***	0.000	-5.848***	0.011
Phillips-Ouliaris z-statistic	-43.029***	0.002	-41.926***	0.007	-37.808***	0.011
Hansen Parameter Inst. Test	1.597	< 0.01	1.534	< 0.01	0.913	0.054

5.10 Component-wise Growth effect: Social Infrastructure (Education)

This section is presenting the analysis of role of social infrastructure of education on the LR productivity growth of Pakistan and this section also presenting to analysis of the asymmetric effect and non-linear association of investing in educational infrastructure.

5.10.1 ARDL & NARDL bound testing approach.

The educational infrastructure is defined as the educational facilities that help increase the stock of knowledge and education. The details of equi-weighted index for education's physical infrastructure and its utilization is provided in the previous section of data and indicators, in detail. We employ ARDL bound testing approach as none of the variables were I(2) and were mixed order of integration of I(1) and I(0), the unrestricted VECM representation of empirical model given in equation (5.22).

$$\Delta ln\widehat{y}_t = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgedq_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgedq_t - \delta_4 \Delta ln(\delta+\lambda+n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.22)$$

We follow the standard practices in the time series analysis, for the selection of the lag order. We estimated the model at the lag order (1), (2) and (3), and selected model with lag order (1), based on the lowest AIC, SIC and HQIC (Table 5.18). The long run coefficients are provided table 5.20. The bound test of ARDL (1,1) conclude for the cointegration as the F-Stat value is greater than the critical value of I(1) bound of 4.01. Therefore, we can further estimate the restricted VECM model using equation (5.23) to analyse the short-run dynamics. The results are presented in Table 5.20.

$$\Delta ln\widehat{y}_t = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgedq_{t-1} - \delta_4 \Delta \ln(\delta+\lambda+n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.23)$$

The ECT is negative, and the absolute value is less than one, and highly significant and depicts the convergence hypothesis. The short-run estimates are provided in Table 5.20

Table 5. 18 *Lag order selection for model 5/2-1(1.1)*

	ARDL (1,1)	ARDL (2,2)	ARDL (3,3)
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	9.312***	2.749	2.215
AIC	-5.011***	-4.922	-4.970
SIC	-4.597***	-4.295	-4.125
HQIC	-4.859***	-4.693	-4.664

We further analyse the asymmetric effect we used nonlinear ARDL (NARDL) model in which short-and long-run nonlinearities are introduced via positive and negative partial sum decompositions of the explanatory variables following (Shin et al., 2014). Under NARDL approach the data series is decomposed into positive and negative series around a threshold of zero, thereby distinguishing between positive and negative changes in the rate of growth in data series. This approach is following the Pesaran et al., (2001) approach bound testing approach for testing cointegration and to estimate LR and SR relationship amongst variables. For NARDL estimation approach equations 5.24 and 5.25 will be estimated the LR and SR productivity growth effect as result of investing in the educational infrastructure.

$$\Delta ln\widehat{y}_t = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgedq^{+}{}_{t-1} + \varphi_5 lnSgedq^{-}{}_{t-1_4} - \varphi_6 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgedq_t - \delta_4 \Delta ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.24)$$

The F-bound test confirmed the cointegration relationship. We estimated ECM using equation 5.25 to analyse the effects of shocks in SR and its impact on LR growth. The estimated output is presented in table 5.20.

$$\Delta ln\widehat{y}_t = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgedq^{+}{}_{t-1} + \delta_4 \Delta lnSgedq^{-}{}_{t-1} - \delta_5 \Delta \ln(\delta+\lambda+n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.25)$$

Furthermore, to calibrate the growth effect using the utilization index of infrastructure of education in Pakistan, we estimated equation (5.26).

$$\Delta ln\widehat{y}_t = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgedu_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgedu_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.26)$$

However, we estimate the model again with only significant control variables at the lag order (1), (2) and (3). We selected the lag order (1), based on lowest AIC, SIC and HQIC.

Table 5. 19 *Model Selection Lag order; model 5/2-2(1.1)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	5.214***	2.501	3.177
AIC	-5.116***	-4.914	-5.115
SIC	-4.700***	-4.287	-4.271
HQIC	-4.962***	-4.685	-4.810

The equation (5.26) is the unrestricted VECM representation at lag order (1) and estimating this equation to measure the growth effect of infrastructure (quantity) available for education. The long run coefficients are provided in Table 5.21, and we can estimate restricted VECM model to estimate the error correction term by estimating equation (5.27) as the F-bound test signifies the cointegration (Table 5.19).

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgedu_{t-1} - \delta_4 \Delta \ln(\delta + \lambda + n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.27)$$

The ECT is negative and absolute value is less than one, and highly significant and depicting the convergence hypothesis. The estimated ECT and short run adjusting variables that converges to the long run equilibrium, are given in Table 5.21. Furthermore, we analyse the asymmetric effect using NARDL approach of cointegration and estimate the equation 5.28.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgedu^{+}{}_{t-1} + \varphi_5 lnSgedu^{-}{}_{t-1} - \varphi_6 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSgedq_t - \delta_4 \Delta ln(\delta + \lambda + n)_t + \varepsilon_t \quad (5.28)$$

The F-bound test signifies the cointegration therefore we estimate equation 5.29 to estimate restricted VECM model for analyzing the SR dynamics.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgedu^{+}{}_{t-1} + \delta_4 \Delta lnSgedu^{+}{}_{t-1} - \delta_5 \Delta \ln(\delta + \lambda + n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.29)$$

5.10.2 FMOLS: estimation of cointegrating equations

We have estimated equation (5.30) to estimate the growth effect of educational infrastructure by using the indices measuring physical infrastructure for education services.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgedq_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.30)$$

We used the ‘Park added variable’ cointegration test to include the additional deterministic and quadratic time trends. The results are presented as model 5/2-3(1.1) in Table 5.22. we also included the square term of $lnSgedq_t$ to capture the non-linear effects of investing in education as model 5/2-3(1.2).

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgedq_t + \varphi_4 lnSgedq^2{}_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.31)$$

We have estimated equation (5.32) to estimate the growth effect of educational infrastructure utilization by size and hypothesis testing for significance through indices developed to measure educational infrastructure usage, an index measuring the magnitude of the utilization of education services in Pakistan.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgcedu_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.32)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgedq_t + \varphi_4 lnSgedq^2{}_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.33)$$

We used the ‘Park added variable’ cointegration test to include the additional deterministic and quadratic time trends. The results of the estimated models using equation 5.32 and 5.33 are presented as model 5/2-4(1.1) and 5/2-4(1.2) results in Table 6.7.

5.10.3 Interpretation

The finding reveals that social infrastructure for education is the critical drivers of LR productivity growth in the long run. Based on the estimated model 5/1-1(1.1) we can conclude that investment in in the physical infrastructure for education per worker is not statistically significant in LR as well as in SR productivity growth in Pakistan. Based on model 5/2-1(1.2) we conclude that the impact of increase in infrastructure for education is positive and significant while the impact decreasing in the infrastructure of education is not negative and significant. Therefore, we support the argument for the existence of asymmetric in LR as the effect of increases vs decrease is not similar. Based on estimated model 5/2-2(1.1), we can conclude that utilization of educational services generates the productivity growth in LR as well as SR. Keeping the effects of other variables constant, a percent increase in index of education utilization increases LR and SR productivity growth by 0.284 and 0.155 percentage points. It is pertinent to note that size of coefficient in LR is greater than SR, therefore we can conclude that increase in the educational services creates larger impact on productivity growth in LR compared to SR. Moreover, the asymmetric effect exists in SR depicting the role of utilization to increase productivity growth.

Moreover, based on the estimated model 5/2-3(1.1) we can conclude that increase in education infrastructure increases productivity growth in nonlinear manner. The marginal effect of investing in stock of educational infrastructure increases productivity growth by 0.233-0.020 ($lnSgedq_t$).

Table 5. 20 *Growth effect of the infrastructure for education: ARDL approach*

Dependent Variable: Growth of Real GDP per worker (Δlny_t)				
	Model 5/2-1(1.1)		**Model 5/1-1(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	-0.715	0.201	0.003	0.989
$lnSh_t$	-0.620	0.161	-0.123	0.403
$lnSgedq_t$	0.309	0.559		
$lnSgedq\_pos_t$			0.432***	0.012
$lnSgedq\_neg_t$			-0.323	0.169
$ln(\delta+\lambda+n)_{t-1}$	-0.394*	0.054	-0.102	0.123
C	14.006***	0.000	12.870***	0.000
Short run dynamics				
$\Delta lnSk_t$	-0.012	0.772	0.019	0.629
$\Delta lnSh_t$	-0.023	0.231	-0.019	0.318
$\Delta lnSgedq_t$	0.014	0.817		
$\Delta lnSgedq\_pos_t$			0.021	0.779
$\Delta lnSgedq\_neg_t$			0.097	0.465
$\Delta ln(\delta+\lambda+n)_{t-1}$	-0.018***	0.000	-0.016***	0.000
ΔECT_{t-1}	-0.060***	0.000	-0.227***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.402	0.496	1.653	0.438
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.103	0.902	0.251	0.780
Heteroskedasticity: ARCH test	0.458	0.503	0.405	0.529
Functional form: Ramsey RESET Test (t-stat)	0.8259	0.4160	0.372	0.713
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 21 *Growth effect of the infrastructure for education: ARDL approach*

Dependent Variable: Growth of Real GDP per worker (Δlny_t)				
	Model 5/2-2(1.1)		**Model 5/2-2(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	0.074	0.803	0.101	0.746
$lnSh_t$	-0.153	0.598	-0.028	0.906
$lnSgedu_t$	0.284***	0.015		
$lnSgedu\_pos_t$			0.135	0.654
$lnSgedu\_neg_t$			-0.945	0.603
$ln(\delta + \lambda + n)_{t-1}$	-0.154	0.274	-0.153	0.330
C	12.792***	0.000	12.683***	0.000
Short run dynamics				
$\Delta lnSk_t$	0.008	0.827	0.016	0.671
$\Delta lnSh_t$	-0.002	0.895	0.004	0.827
$lnSgedu_t$	0.155***	0.004		
$lnSgedu\_pos_t$			0.131**	0.023
$lnSgedu\_neg_t$			0.094	0.530
$\Delta ln(\delta + \lambda + n)_{t-1}$	-0.015***	0.000	-0.015***	0.000
ΔECT_{t-1}	-0.135***	0.000	-0.134***	0.000
Diagnostics				
Normality (Jarque–Bera test)	3.373	0.185	3.214	0.201
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.096	0.909	0.102	0.903
Heteroskedasticity: ARCH test	0.319	0.576	1.027	0.317
Functional form: Ramsey RESET Test (t-stat)	1.099	0.280	0.979	0.336
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. *22* *Growth Effect of physical infrastructure for Educational Services: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker (Δlny_t)				
	Model 5/2-3(1.1)		**Model 5/2-3(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}	-0.127	0.199	-0.046***	0.013
$lnSk_t$	0.051	0.206	-0.009	0.757
$lnSh_t$	0.027	0.311	-0.028***	0.037
$lnSgedq_t$	0.016	0.767	0.233***	0.004
$lnSgedq\text{^}2_t$			-0.020***	0.001
$\ln(\delta+\lambda+n)$	-0.017***	0.000	-0.018***	0.000
$lnFDI_t$	0.001	0.861		
$lnCPI_t$	0.002	0.814		
$lnDCRPS_t$	-0.004	0.874		
$lnTI_t$	-0.036	0.397		
$LnREER_t$	-0.004	0.930		
CAB_t	0.002	0.273		
SBD_{2004}	0.029	0.124		
C	1.531	0.289		
@TREND				
@TREND Square				
R-squared	0.579		0.517	
Adjusted R-squared	0.398		0.448	
Cointegration Tests: -				
Engle-Granger tau-statistic	-6.694	0.002	-5.729	0.014
Engle-Granger z-statistic	-43.317	0.001	-37.200	0.013
Phillips-Ouliaris tau-statistic	-6.823	0.001	-5.789	0.013
Phillips-Ouliaris z-statistic	-40.524	0.004	-36.408	0.017
Cointegration Test-Hansen Parameter Instability	1.74	< 0.01	0.835	0.078

Table 5. *23* *Growth Effect as result of the utilization of the Educational Services: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker (Δlny_t)				
	Model 5/2-4(1.1)		**Model 5/2-4(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}	-0.153	0.302	-0.066	0.697
$lnSk_t$	0.057	0.227	-0.001	0.979
$lnSh_t$	0.029	0.351	-0.014	0.666
$lnSgedu_t$	0.007	0.885	-0.062	0.495
$lnSgedu\text{^}2_t$			0.028	0.300
$\ln(\delta + \lambda + n)$	-0.017***	0.000	-0.020***	0.000
$lnFDI_t$	0.002	0.826	0.001	0.861
$lnCPI_t$	0.004	0.636	-0.007	0.540
$lnDCRPS_t$	0.002	0.941	0.016	0.637
$lnTI_t$	-0.049	0.237	-0.011	0.829
$LnREER_t$	-0.017	0.722	-0.033	0.608
CAB_t	0.002	0.258	0.001	0.829
SBD_{2004}	0.027*	0.084		
C	2.024	0.300		
@TREND	1.813	0.362	1.057	0.632
@TREND Square				
R-squared	0.573		0.550	
Adjusted R-squared	0.391		0.358	
Cointegration Tests: -				
Engle-Granger tau-statistic	-6.464	0.003	-6.008	0.020
Engle-Granger z-statistic	-41.950	0.002	-38.977	0.018
Phillips-Ouliaris tau-statistic	-6.568	0.002	-6.076	0.017
Phillips-Ouliaris z-statistic	-39.348	0.006	-37.967	0.024
Cointegration Test-Hansen Parameter Instability	1.373	< 0.01	0.817	0.153

5.11 Component-wise Growth Effect: Social Infrastructure (Health)

This section is presenting estimations of productivity growth-effect relationships of health infrastructure. we are estimating the direct, asymmetric, and non-linear association of health infrastructure and we are calibrating into the physical stock as well in utilization terms.

5.11.1 ARDL and NARDL bound testing approach.

We measured productivity growth effect of health infrastructure using indicators of physical quantification and its utilization. We have used the indicator for total number of hospital beds per million workers to measure the stock of the physical infrastructure for health services. We have used the utilization indicator of the immunization (BCG vaccination) which is one of the primary and necessary vaccination administered to the neonatal, and it is considered an important vaccine to build immunity against Tuberculosis (TB) and it is administered in the developing countries with the incidence of TB6.

Following the steps, we estimated the ARDL (1,1), ARDL (2,2) and ARDL (3,3), and selected the ARDL (1,1) as a parsimonious model due to lowest SIC and HQIC. Details are provided in Table 5.24. We estimated the unrestricted VECM, as expressed in equation (5.33).

$$\Delta ln\hat{y}_t = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSghthq_{t-1} - \varphi_5 \ln(\delta + \lambda + n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthq_{t-1} - \delta_4 \Delta ln(\delta + \lambda + n)_{t-1} + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.34)$$

Following the steps for the estimation, we estimated the model with all control variables however, coefficients of the control variables were highly insignificant, so we estimated model with only significant control variables. The estimated model established the cointegration, therefore, we estimated restricted VECM as per equation (5.35).

$$\Delta ln\hat{y}_t = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthq_{t-1} - \delta_4 \Delta \ln(\delta + \lambda + n)_{t-1} + \tau ECT_{t-1} + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.35)$$

The long run estimates and the error correction term and the short run adjusting coefficients are provided in Table 5.26. Moreover, to analyse whether asymmetric effects exist in LR and SR we used NARDL bound testing approach. Therefore, we estimate equations 5.34 and the post estimation F-bound test signalled cointegration therefore we estimated ECM using equation 5.35.

6 For further details: https://www.who.int/teams/health-product-policy-and-standards/standards-and-specifications/vaccines-quality/bc

Table 5. 24 *Lag order selection for model 5/3-1(1.1)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test	9.442***	2.548	1.617
5% critical value I(1) upper bound: 4.01			
R-squared	0.992	0.992	0.993
Adjusted R-squared	0.990	0.988	0.986
AIC	-5.051*	-4.884	-4.891
SIC	-4.637*	-4.257	-4.047
HQIC	-4.899*	-4.655	-4.586

To analyse the asymmetric effect of investment in education we used NARDL approach of cointegration, for doing so we estimated equation 5.36 and 5.37 as model 5/3-1(1.2).

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSghthq^{+}{}_{t-1} + \varphi_4 lnSghthq^{-}{}_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \gamma_1 SBD_{2004} \quad (5.36)$$

We concluded for cointegration therefore we estimated ECM using equation 6.17.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSghthq^{+}{}_t - \delta_4 \Delta lnSghthq^{-}{}_t - \delta_4 \Delta \ln(\delta+\lambda+n)_t + \tau ECT_t + \varepsilon_t \quad (5.38)$$

To analyse the impact of utilization of health infrastructure on the productivity the ARDL (1,1), ARDL (2,2) and ARDL (3,3) estimated by estimating the unrestricted VECM. The F-bound test revealed cointegration are given in table 5.25. Based on the results we selected the ARDL (1,1) as a better model as indicated by the lowest AIC, SIC and HQIC. Therefore, we estimated the unrestricted VECM equation (5.39).

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSghthu_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.39)$$

The coefficients of $lnCPI_t$ amongst all control variables was highly insignificant and to get a better model with a greater degree of confidence, we estimated model by dropping the insignificant control variables. The F-bound test results conclude for the cointegration (Table 5.25) and therefore, estimating the restricted VECM at lag (1) as expressed in the equation (5.40).

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_t + \delta_2 \Delta lnSh_t + \delta_3 \Delta lnSghthu_t - \delta_4 \Delta \ln(\delta+\lambda+n)_t + \tau ECT_t + \varepsilon_t \quad (5.40)$$

The long run estimates and the error correction term along with the results of the diagnostics tests are given in the Table 5.27.

Table 5. 25 *Lag order selection for model 6-6(A)*

	ARDL (1,1)	ARDL (2,2)	ARDL (3,3)
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	6.488***	3.003	2.029
R-squared	0.993	0.993	0.994
Adjusted R-squared	0.990	0.988	0.986
AIC	-5.041*	-4.890	-4.886
SIC	-4.544*	-4.180	-3.958
HQIC	-4.859*	-4.632	-4.551

Additionally, we estimated the model using NARDL bound testing approach to analyse the asymmetric LR and SR effects on productivity growth using equation 5.41. The F-bound test supported the argument of existence of cointegrating relationship therefore we estimated VECM to analyse the SR dynamics/ transition to LR equilibrium using equation 5.42. The results are presented as model 5/3-2(1.2) in table 5.27.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSghthu^{+}{}_{t-1} + \varphi_5 lnSghthu^{-}{}_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.41)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSghthu^{+}{}_{t-1} + \delta_4 \Delta lnSghthu^{-}{}_{t-1} - \delta_4 \Delta \ln(\delta+\lambda+n)_{t-1} + \tau ECT_{t-1} + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.42)$$

5.11.2 FMOLS estimation technique

The previous chapter briefly oriented to the merits of using the FMOLS, we are estimating long-run elasticities of the productivity growth using this technique as an alternative and to bring forth more reliable and informed policy-oriented results. We estimated the equation (5.43) to measure the long run growth effect of health infrastructure's quantification.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSghthq_t - \varphi_5 ln(\delta+\lambda+n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.43)$$

We estimated the model with and without control variables as model 5/3-3(1.1) and model 5/3-3(1.2) respectively. Estimated output along with the diagnostics, that detected the long run cointegration are provided in Table 5.28. We also estimated model using square term of $lnSghthq_t$

to test whether productivity growth and health infrastructure is linearly associated or not. The estimated model 5/3-3(1.3) is estimated output for equation 5.44.

$$\Delta ln\hat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSghthq_t + \varphi_4 lnSghthq_t^{\ 2} - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.44)$$

To estimate the growth effect health infrastructure utilization, we estimated equations 5.45 and 5.46 by using FMOLS approach to estimate cointegration approach. The equation 6.46 included the square term of $lnSghthu_t$.

$$\Delta ln\hat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSghthu_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.45)$$

$$\Delta ln\hat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSghthu_t^{\ 2} - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.46)$$

We estimated outputs of models without and without all control's variables are provided as model 5/3-4(1.1) and 5/3-4(1.2) respectively and model with the square term as model 5/3-3(1.3) is presented in Table 5.29.

5.11.3 Interpretation and key findings

The estimated model 5/3-1(1.1) and 5/3-2(1.1) shows insignificant growth effect of investing in physical infrastructure for health models and its utilization both in LR as well as in SR. However, the estimated models of 5/3-3(1.1) and 5/3-3(1.2) depicts positive and significant coefficient of $lnSghthq_t$. Therefore, we can conclude that increase in the physical health infrastructure iterms of numbers of hospitals beds per million workers by one percentage points contributes to productivity growth of Pakistan positively by 0.188 percentage points. The estimated model 5/3-1(1.2) supported the argument for the presence of an asymmetric effect in LR and SR for indicators of numbers of hospitals beds per million workers. It is because the coefficient for $lnSghthq\_pos_t$ is highly significant while $lnSghthq\_neg_t$ is insignificant. The control variables of $lnCPI_t, lnREER_t$ and SBD_{2004} are significant factors in Model 5/3-1(1.2) and Model 5/3-2(1.2), it therefore we conclude that increases in inflation rate and reduction in the relative trade competitiveness with the rest of the world decreases the SR productivity growth by 0.013 and

0.077 percentage points. Also, the structural changes in 2003-4 and beyond had increased the SR productivity growth of Pakistan.

The impact of immunization coverage as measure of health infrastructure's utilization is symmetric both in LR as well as SR. Based on the estimated model 5/3-4(1.3) we can conclude that increase in the immunization coverages increases productivity growth non-linearly as the coefficient of $lnSghthu_t$ and $lnSghthu^\wedge 2_t$ are highly significant. Therefore, the LR marginal effect would be 0.027+(-0.015) $lnSghthu_t$.

Table 5. 26 *Growth effect of the physical infrastructure for health services: Bound testing approach*

Dependent Variable: Growth of Real GDP per worker (Δlny_t)				
	Model 5/3-1(1.1)		**Model 5/3-1(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	-0.967	0.262	0.037	0.643
$lnSh_t$	-0.900	0.174	0.101**	0.029
$lnSghthq_t$	0.688	0.488		
$lnSghthq\_pos_t$			0.852***	0.000
$lnSghthq\_neg_t$			0.148	0.327
$ln(\delta+\lambda+n)_{t-1}$	-0.409*	0.086	-0.029**	0.057
C	11.722**	0.044	13.178***	0.000
Short run dynamics				
$\Delta lnSk_t$	-0.009	0.805	0.013	0.707
$\Delta lnSh_t$	-0.023	0.216	0.053***	0.004
$\Delta lnSghthq_t$	0.140	0.135		
$\Delta lnSghthq\_pos_t$			0.404***	0.004
$\Delta lnSghthq\_neg_t$			0.058	0.750
$\Delta ln(\delta+\lambda+n)_{t-1}$	-0.015***	0.000	-0.014***	0.000
ΔECT_{t-1}	-0.048***	0.000	-0.584***	0.000
$lnCPI_t$			-0.013***	0.005
$lnREER_t$			-0.077***	0.000
SBD_{2004}			0.089***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.152	0.562	1.376	0.503
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.047	0.955	0.139	0.711
Heteroskedasticity: ARCH test	0.067	0.797	0.139	0.711
Functional form: Ramsey RESET Test (t-stat)	0.082	0.935	1.280	0.211
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 27 *Growth effect of the infrastructure for health's utilization: bound testing approach*

Dependent Variable: Growth of Real GDP per worker (Δlny_t)				
	Model 5/3-2(1.1)		**Model 5/3-2(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	-0.214	0.614	-5.270	0.900
$lnSh_t$	-0.394	0.359	-7.457	0.900
$lnSghthu_t$	0.128	0.201		
$lnSghthu\_pos_t$			1.073	0.887
$lnSghthu\_neg_t$			10.056	0.904
$ln(\delta + \lambda + n)_{t-1}$	-0.210	0.164	-2.132	0.897
C	13.814	0.000	40.153	0.855
Short run dynamics				
$\Delta lnSk_t$	-0.015	0.705	-0.035	0.373
$\Delta lnSh_t$	-0.023	0.220	-0.037**	0.053
$\Delta lnSghthu_t$	-0.004	0.775		
$\Delta lnSghthu\_pos_t$			-0.008	0.592
$\Delta lnSghthu\_neg_t$			0.098	0.210
$\Delta ln(\delta + \lambda + n)_{t-1}$	-0.008***	0.013	-0.015***	0.000
ΔECT_{t-1}	-0.092***	0.000	-0.008***	0.000
$lnCPI_t$	-0.008***	0.013	-0.012***	0.000
SBD_{2004}	0.007***	0.203		
Diagnostics				
Normality (Jarque–Bera test)	0.912	0.634	0.645	0.724
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.224	0.801	0.287	0.753
Heteroskedasticity: ARCH test	0.742	0.394	0.148	0.702
Functional form: Ramsey RESET Test (t-stat)	1.338	0.191	1.091	0.285
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 28 *Growth Effect of physical infrastructure for health: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker						
	Model 5/3-3(1.1)		**Model 5/3-3(1.2)**		**Model 5/3-3(1.3)**	
Variable	**Coefficient**	**p-val**	**Coefficient**	**p-val**	**Coefficient**	**p-val**
lny_{t-1}	-0.132*	0.092	-0.045***	0.007	-0.045***	0.006
$lnSk_t$	0.007	0.831	-0.005	0.888	-0.022	0.488
$lnSh_t$	0.045**	0.037	-0.019	0.243	-0.014	0.375
$lnSghthq_t$	0.188***	0.001	-0.008	0.817	7.273	0.112
$lnSghthq\text{^}2_t$					-0.476	0.112
$\ln(\delta + \lambda + n)$	-0.014***	0.000	-0.020***	0.000	-0.019***	0.000
$lnFDI_t$	0.000	0.964				
$lnCPI_t$	-0.002	0.792				
$lnDCRPS_t$	0.012	0.563				
$lnTI_t$	-0.038	0.217				
$LnREER_t$	0.036	0.306				
CAB_t	0.004***	0.015				
SBD_{2004}	0.072***	0.000				
C	0.130	0.905	0.702**	0.025	-27.106	0.122
R-squared	0.665		0.509		0.539	
Adjusted R-squared	0.522		0.439		0.458	
Cointegration Tests: -						
Engle-Granger tau-statistic	-8.482	0.000	-5.649	0.020	-5.945	0.023
Engle-Granger z-statistic	-52.858	0.000	-36.626	0.017	-38.775	0.019
Phillips-Ouliaris tau-statistic	-8.935	0.000	-5.717	0.018	-6.012	0.020
Phillips-Ouliaris z-statistic	-48.877	0.000	-36.500	0.018	-38.044	0.023
-Hansen Parameter Instability	1.707	< 0.01	0.693	0.155	0.769	0.184

Table 5. 29 *Growth Effect of health Services utilization: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker						
	Model 5/3-4(1.1)		**Model 5/3-4(1.2)**		**Model 5/3-4(1.3)**	
Variable	**Coefficient**	**p-val**	**Coefficient**	**p-val**	**Coefficient**	**p-val**
lny_{t-1}	-0.170*	0.099	-0.063**	0.002	-0.103***	0.000
$lnSk_t$	0.052	0.190	-0.032	0.104	-0.008	0.658
$lnSh_t$	0.012	0.681	-0.037***	0.009	-0.035***	0.004
$lnSghthu_t$	0.012	0.388	0.008	0.370	-0.169**	0.023
$lnSghthu^\wedge 2_t$					0.027***	0.012
$\ln(\delta+\lambda+n)$	-0.015***	0.001	-0.018***	0.000	-0.015***	0.000
$lnFDI_t$	0.004	0.640				
$lnCPI_t$	0.007	0.378				
$lnDCRPS_t$	-0.012	0.681				
$lnTI_t$	-0.054	0.164				
$LnREER_t$	-0.010	0.825				
CAB_t	0.002	0.404				
SBD_{2004}	0.023	0.149				
C	2.202	0.132	0.928**	0.001	1.627	0.000
R-squared	0.578		0.596		0.603	
Adjusted R-squared	0.397		0.497		0.533	
Cointegration Tests: -						
Engle-Granger tau-statistic	-6.473	0.003	-5.827	0.013	-6.258	0.012
Engle-Granger z-statistic	-41.935	0.002	-37.791	0.010	-40.872	0.010
Phillips-Ouliaris tau-statistic	-6.590	0.002	-5.893	0.011	-6.340	0.010
Phillips-Ouliaris z-statistic	-38.773	0.007	-37.161	0.013	-38.530	0.020
Hansen Parameter Instability	1.360	< 0.01	0.836	0.114	0.741	> 0.2

5.12 Component-Wise Growth Effect: Financial Infrastructure

This section presenting estimating the direct, asymmetric, and non-linear effects resulting from investing in the financial infrastructure. The financial infrastructure facilitates financial transactions and economic integration. Our focus lies in evaluating the impact of both the quantity and utilization of financial infrastructure on growth. Therefore, to gauge the size of the physical infrastructure network for financial services (denoted as $Sgfq_t$), indicators of total number of Bank Branches (BBR) and ATM machines (ATM) are used. The number of ATMs holds significance because it facilitates fundamental banking services such as withdrawals and inter- and intra-bank cash transfers, in addition to the key services available at bank branches. As $Sgfq_t$, is measured in numerical terms, we transform it into natural logarithm form to aid in estimation. Furthermore, we adopt the usage indicator 'number of bank accounts per population (NBAP) (in millions)' as a proxy to assess the utilization of financial infrastructure ($Sgfu_t$) in the country. We utilize this indicator of the number of bank accounts to measure the extent of financial infrastructure utilization since having a basic bank account is a prerequisite for accessing financial services in Pakistan. As such, it acts as a proxy or shadow variable to evaluate the level of financial infrastructure utilization in the country. For statistical estimation, we transform this variable into log form.

5.12.1 ARDL & NARDL bound testing approach.

We estimated ARDL (1,1), ARDL (2,2), and ARDL (3,3) by estimating the unrestricted VECM using equation 5.46. The F-bound test revealed cointegration as given in table 5.30. We choose the ARDL model at the large order (1), which has the lowest AIC, SIC and HQIC.

Table 5. 30 *Lag-order Selection Model 5/4-1(1.1)*

	ARDL (1,1)	ARDL (2,2)	ARDL (3,3)
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	6.182***	4.259***	4.292***
R-squared	0.993	0.993	0.995
Adjusted R-squared	0.991	0.988	0.988
AIC	-5.080***	-4.905	-5.055
SIC	-4.615***	-4.222	-4.150
HQIC	-4.912***	-4.660	-4.733

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgfq_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.46)$$

We estimated the model with all sets of control variables, however, the control variables turned out to be highly insignificant. And, to provide a greater degree of freedom, we drop the insignificant control variables and estimated the model again. The value of the F-Stat bound test result of the models is 6.18 which is greater than I(1) bound at 5% significance level. Therefore, we failed to accept the null hypothesis of no co-integration and accepts the alternative hypothesis of co-integrating relationships. We can now estimate restricted VECM by estimating the equation (5.47) to analyse the short-run dynamics for the long run convergence.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq_{t-1} - \delta_4 \Delta \ln(\delta+\lambda+n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.47)$$

The coefficients of long-run elasticities and error correct term and results of the diagnostics test of the model are given in Table 5.32 as model 5/4-1(1).

In additional we estimate cointegrating nonlinear ARDL (NARDL) model in which short- and long-run nonlinearities are introduced via positive and negative partial sum decompositions of the explanatory variables following (Shin et al., 2014). Under NARDL approach the data series is decomposed into positive and negative series around a threshold of zero, thereby distinguishing between positive and negative changes in the rate of growth in data series. This approach is following the Pesaran et al., (2001) approach bound testing approach for testing cointegration and to estimate LR and SR relationship amongst variables. For NARDL estimation approach equations 5.48 and 5.49 will be estimated the LR and SR productivity growth effect of financial infrastructure.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgfq^{+}{}_{t-1} + \varphi_5 lnSgfq^{-}{}_{t-1} - \varphi_6 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.48)$$

The nonlinear conditional ECM is expressed in eq. 5.49.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq^{+}{}_{t-1} + \delta_4 \Delta lnSgfq^{-}{}_{t-1} - \delta_4 \Delta \ln(\delta + \lambda+n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.49)$$

The coefficients of long-run elasticities and error correct term and results of the diagnostics test of the model are given in Table 5.32 as model 5/4-1(1.2).

In addition, we also estimated the model's using indicator for the financial infrastructure utilization along with critical determinants of growth such as private sector investment in physical capital, human capital, population growth and depreciation of physical capital stock, and the set of control variables (already mentioned). We estimated ARDL (1,1), ARDL (2,2) and ARDL (3,3) by estimating the unrestricted VECM using equation 5.50.

Table 5. 31 *Lag-order Selection, Model 5/4-2(1.1)*

	ARDL (1,1)	**ARDL (2,2)**	**ARDL (3,3)**
F-Stat Bound test 5% critical value I(1) upper bound: 4.01	7.418***	5.301***	4.568***
R-squared	0.997	0.998	0.998
Adjusted R-squared	0.994	0.995	0.994
AIC	-5.776*	-5.675	-5.528
SIC	-4.979*	-4.737	-4.785
HQIC	-5.329*	-5.339	-5.423

We estimated the model at lag (1), and estimated equation (5.50) at the lag order (1) given the lowest SIC.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgfu_{t-1} - \varphi_5 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.50)$$

The F-bound best signals the cointegration therefore, unrestricted VECM equation (table 5.31)

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfu_{t-1} - \delta_4 \Delta \ln(\delta+\lambda+n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.51)$$

The long run output elasticities, error correction term and the post-estimation diagnostics of the model are in Table 5.34 as model 5/4-2(1.1). We also used NARDL estimation technique to analyse the asymmetric impact of financial infrastructure on productivity growth of the country. Equations 5.52 and 5.53 estimated to assess the asymmetric growth effect of financial infrastructure in LR and SR.

$$\Delta ln\widehat{y_t} = \varphi_4 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_{t-1} + \varphi_3 lnSh_{t-1} + \varphi_4 lnSgfu^{+}{}_{t-1} + \varphi_5 lnSgfu^{-}{}_{t-1} - \varphi_6 \ln(\delta+\lambda+n)_{t-1} + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfq_{t-1} - \delta_4 \Delta ln(\delta+\lambda+n)_{t-1} + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.52)$$

The nonlinear conditional ECM is expressed in eq. 5.53.

$$\Delta ln\widehat{y_t} = \varphi_0 + \delta_1 \Delta lnSk_{t-1} + \delta_2 \Delta lnSh_{t-1} + \delta_3 \Delta lnSgfu^{+}{}_{t-1} + \delta_4 \Delta lnSgfu^{-}{}_{t-1} - \delta_4 \Delta \ln(\delta + \lambda + n)_{t-1} + \tau ECT_{t-1} + \varepsilon_t \quad (5.53)$$

The LR and SR estimates are reported in the table 5.34 as Model 5/4-2(1.2).

5.12.2 FMOLS estimation techniques

We estimated the equation (5.54) to measure the long-run growth effect due to physical availability of the financial infrastructure using FMOLS. We also included the square term in the estimated model to assess to assess non-linear impact of investing in the financial infrastructure and estimated model using equation (5.55).

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgfq_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.54)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgfq_t + \varphi_4 lnSgfq_t{}^2 - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.55)$$

The significance of φ_4 would support for the argument for existence of non-linear relationship of financial infrastructure in the productivity growth of Pakistan. The estimated Model 5/4-3(1.1) with all control variables and the diagnostics test of Engle-Granger and Phillips-Ouliaris supported for the hypothesis of existence of cointegration and presented in table 5.34. The control variables were insignificant therefore we estimated Model 5/4-3(1.2) and estimated Model 5/4-3(1.3) by including square terms of financial infrastructure. Moreover, the long-run productivity growth effect has been estimated using equation (5.56) the utilization indicators of financial infrastructure using FMOLS estimation technique. Additionally, to estimate non-linear effect by including square term estimated model using equation 5.57.

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgfu_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.56)$$

$$\Delta ln\widehat{y_t} = \varphi_0 + \varphi_1 ln\hat{y}_{t-1} + \varphi_2 lnSk_t + \varphi_3 lnSh_t + \varphi_4 lnSgfu_t - \varphi_5 ln(\delta + \lambda + n)_t + \vartheta_1 lnLFDI_t + \vartheta_2 lnLCPI_t + \vartheta_3 lnDCRPS_t + \vartheta_4 lnTI_t + \vartheta_5 lnREER_t + \vartheta_6 lnCAB_t + \gamma_1 SBD_{2004} + \varepsilon_t \quad (5.57)$$

We estimated the model using equation 5.56 are provided as model 5/4-4(1.1) and 5/4-4(1.2) with all control variables and without control variables and while estimated output of equation 7.17 is presented as model 5/4-4(1.3) and the diagnostics test supported for the cointegration, sand details are provided in the Table 5.35.

5.12.3. Key Findings

The long-run (LR) effect on productivity growth resulting from investing in physical infrastructure for financial services is statistically insignificant in model 5/4-1(1.1), model 5/4-3(1.1), and model 5/4-3(1.2). In contrast, estimated model 5/4-3(1.3) provides support for the evidence of a non-linear effect of investing in the physical stock of financial infrastructure. Consequently, we conclude that investing in financial infrastructure generates a positive impact on productivity growth, and this impact occurs at an increasing rate. The marginal effect of investing in financial infrastructure is calculated as 0.668 + 0.060($lnSgfq_t$). Moreover, the estimated model 5/4-1(1.2) provides support for the hypothesis that there exists an asymmetric effect of physical infrastructure for financial system. Specifically, the coefficient $lnSgfq\_pos_t$ is positive and highly significant, while the coefficient for $lnSgfq\_neg_t$ is statistically insignificant. This leads us to conclude that an increase in the stock of financial infrastructure has a positive impact on productivity growth, but this impact differs is unequal when compared to a reduction in the physical stock of financial infrastructure.

Based on the estimated model, we can deduce that an increase in the utilization of financial infrastructure augments LR productivity growth in Pakistan. The coefficient of $lnSgfu_t$ is 0.334 and statistically significant at the 10 percent in estimated model 5/4-2(1.1). At the same time an estimated coefficient in estimated model 5/4-4(1.1) is 0.081 which is highly significant at the 1 percent level. On the other hand, estimated model 5/4-2(1.2) does not support the existence of an asymmetric growth effect resulting from the utilization of financial infrastructure. The coefficient for $lnSgfu\text{^}2_t$ is statistically insignificant, thus we cannot conclude the presence of a non-linear association between the utilization of financial infrastructure and LR productivity growth in Pakistan.

Based on the estimated models presented above, it can be concluded that productivity growth is adversely affected by high population growth rates both in the long run (LR) and short run (SR). Additionally, the coefficients for lny_{t-1} and $lnSh_t$ are negative and statistically significant, indicating a poor translation of the output per worker's level effect and limited returns on investment in human capital development for the long-run productivity growth rate of Pakistan.

Regarding the control variables, $LnREER_t$ and SBD_{2004} exhibit significance at the 1 percent level in LR model 5/4-4(1.1), underlining the critical importance of relative trade competitiveness and structural changes for long-run productivity growth. Specifically, based model 5/4-4(1.1) we can conclude that Pakistan experienced a decline in productivity growth by 0.103 percentage points due to currency depreciation in the LR. Furthermore, we infer that productivity growth has increased by 0.037 percent in the years 2004-2022 compared to the duration of 1980-2003. Hence, the significance of the structural break dummy of 2004 highlights the key importance of structural changes that occurred during and after the Musharraf-Aziz Era.

Table 5. 32 *Growth effect of physical infrastructure for financial services: an ARDL estimation technique*

Dependent Variable: Growth rate of output per worker				
	Model 5/4-1(1.1)		**Model 5/4-1(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	-0.882	0.183	-0.004	0.990
$lnSh_t$	-0.321	0.507	0.110	0.605
$lnSgfq_t$	0.264	0.337		
$lnSgfq\_pos_t$			0.505***	0.005
$lnSgfq\_neg_t$			0.029	0.824
$ln(\delta + \lambda + n)_t$	-0.388*	0.087	-0.137	0.146
C	17.560***	0.000	12.865***	0.000
Short run dynamics				
$\Delta lnSk_t$	-0.021	0.580	-0.002	0.969
$\Delta lnSh_t$	-0.013	0.479	0.009	0.616
$\Delta lnSgfq_t$	-0.033	0.457		
$\Delta lnSgfq\_pos_t$			-0.097	0.279
$\Delta lnSgfq\_neg_t$			0.024	0.674
$\Delta ln(\delta + \lambda + n)_t$	-0.017***	0.000	-0.015***	0.000
ΔECT_t	-0.057***	0.000	-0.143***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.302	0.522	1.081	0.582
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.061	0.941	0.434	0.652
Heteroskedasticity: ARCH test	0.171	0.995	0.932	0.340
Functional form: Ramsey RESET Test (t-stat)	0.352	0.727	0.717	0.480
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 33 *Growth effect of financial infrastructure utilization: an ARDL estimation technique*

Dependent Variable: Growth rate of output per worker				
	Model 5/4-2(1.1)		**Model 5/4-2(1.2)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Long run estimates				
$lnSk_t$	0.298	0.567	0.270	0.695
$lnSh_t$	-0.599*	0.075	-0.947	0.508
$lnSgfu_t$	0.334*	0.071		
$lnSgfu\_pos_t$			0.435	0.302
$lnSgfu\_neg_t$			0.832	0.661
$ln(\delta+\lambda+n)_t$	-0.254*	0.084	-0.294	0.424
C	12.015***	0.000	13.444***	0.000
Short run dynamics				
$\Delta lnSk_t$	0.030	0.440	0.028	0.478
$\Delta lnSh_t$	-0.024	0.194	-0.031	0.114
$\Delta lnSgfu_t$	0.082*	0.065		
$\Delta lnSgfu\_pos_t$			0.089	0.105
$\Delta lnSgfu\_neg_t$			0.116	0.194
$\Delta ln(\delta+\lambda+n)_t$	-0.016***	0.000	-0.016***	0.000
ΔECT_{t-1}	-0.074***	0.000	-0.062***	0.000
Diagnostics				
Normality (Jarque–Bera test)	1.096	0.578	0.704	0.703
Serial Autocorrelation: Breusch-Godfrey Serial Correlation LM Test	0.612	0.549	0.831	0.447
Heteroskedasticity: ARCH test	0.061	0.806	0.000	0.997
Functional form: Ramsey RESET Test (t-stat)	1.925	0.063	1.650	0.110
Cusum square graph	Stable		Stable	
Cusum Graph	Stable		Stable	

Table 5. 34 *Growth effect of financial infrastructure's physical quantity: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker (labour productivity)						
	Model 5/4-3(1.1)		**Model 5/4-3(1.2)**		**Model 5/4-3(1.3)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}	-0.011	0.618	-0.046***	0.012	-0.069***	0.000
$lnSk_t$	0.060	0.117	-0.010	0.756	-0.026	0.357
$lnSh_t$	0.013	0.639	-0.027	0.141	-0.017	0.271
$lnSgfq_t$	0.006	0.817	0.002	0.877	0.668**	0.020
$lnSgfq\text{^}2_t$					0.060**	0.021
$\ln(\delta + \lambda + n)$	-0.018***	0.000	-0.021***	0.000	-0.021***	0.000
$lnFDI_t$	-0.004	0.573				
$lnCPI_t$	0.006	0.526				
$lnDCRPS_t$	0.012	0.639				
$lnTI_t$	-0.048	0.223				
$LnREER_t$	0.048	0.076				
CAB_t	0.003	0.068				
SBD_{2004}	0.025	0.080				
C	0.548***	0.002	0.691***	0.013		
@TREND						
@TREND Square						
R-square	0.542		0.506		0.587	
Adjusted R-square	0.369		0.436		0.514	
Cointegration Tests: -						
Engle-Granger tau-statistic	-6.630	0.002	-5.716	0.015	-6.331	0.009
Engle-Granger z-statistic	-42.952	0.002	-37.064	0.012	-41.825	0.006
Phillips-Ouliaris tau-statistic	-6.758	0.001	-5.783	0.013	-6.410	0.007
Phillips-Ouliaris z-statistic	-39.929	0.005	-36.812	0.013	-40.163	0.010
Hansen Parameter Instability-Test	1.319	< 0.01	0.860	0.070	0.576	> 0.2

Table 5. 35 *Growth effect of financial infrastructure's utilization: FMOLS*

Dependent Variable: Growth rate of Real GDP per worker (Δlny_t)						
	Model 5/4-4(1.1)		**Model 5/4-4(1.2)**		**Model 5/4-4(1.3)**	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
lny_{t-1}	-0.411***	0.001	-0.061***	0.014	-0.087***	0.008
$lnSk_t$	0.053	0.115	-0.003	0.892	-0.005	0.841
$lnSh_t$	0.002	0.931	-0.039***	0.018	-0.047***	0.010
$lnSgfu_t$	0.081***	0.001	0.015	0.296	0.176	0.313
$lnSgfu\char`^2_t$					-0.021	0.358
$\ln(\delta + \lambda + n)$	-0.015***	0.000	-0.018***	0.000	-0.018***	0.000
$lnFDI_t$	0.010	0.136				
$lnCPI_t$	-0.008	0.279				
$lnDCRPS_t$	0.018	0.423				
$lnTI_t$	0.003	0.931				
$LnREER_t$	-0.103***	0.018				
CAB_t	0.001	0.700				
SBD_{2004}	0.037***	0.006				
C	5.147***	0.001	0.807***	0.004	0.864***	0.004
@TREND						
@TREND Square						
R-square	0.676		0.528		0.541	
Adjusted R-square	0.537		0.460		0.460	
Cointegration Tests: -						
Engle-Granger tau-statistic	-6.280	0.005	-5.729	0.015	-5.647	0.036
Engle-Granger z-statistic	-84.267	0.000	-37.461	0.011	-36.704	0.029
Phillips-Ouliaris tau-statistic	-7.240	0.000	-5.779	0.013	-5.701	0.032
Phillips-Ouliaris z-statistic	-40.157	0.004	-35.954	0.017	-35.700	0.039
Hansen Parameter Instability-Test	1.102	0.026	0.707	0.140	0.767	0.107

CHAPTER 6

SPATIAL SPILLOVERS OF INFRASTRUCTURE: ESTIMATION

6.1 Introduction

Infrastructure reshapes the economic geography of through regional geographical connectivity. It reduces the trade costs and smooths the trade flows, so positively influencing the economic development (Cohen & Paul, 2004). Infrastructure is one of the critical drivers of long-run regional development and as it creates economic cohesion and networking interconnected via multiple direct and indirect channels (Chen & Haynes, 2015). Infrastructure development requires extensive financial resources to build. Therefore, it is imperative to estimate the direct and indirect effects for an informed policy formulation.

Earlier models of the regional development models can only detect the presence of spatial dependence by detecting the cross-sectional dependence in panel-data models. But the recent advancements in the body of literature on spatial econometrics allowed to estimate estimated these direct and indirect effects of infrastructure, therefore very useful to determine the actual size of impact. Therefore, this chapter presents a model to carefully analyse the role of core infrastructure as the critical long-run driver of regional productivity growth. Therefore, this chapter presenting the analysis that focuses on the role of investment in core infrastructure as one of the critical drivers of the long-run regional development and other traditional drivers of human capital and private capital stock in Pakistan.

The remaining chapter is organized into seven-sections. Section 6.2-5 includes empirical model and its specification, indicators and Data source, proposed methodology and lastly stationarity test and cointegration tests results. Details of index construction and estimations and key Findings are provided in section 6.6 & 6.7, respectively.

6.2 Empirical model and its specification

We developed a detailed theoretical framework to analyse the direct and indirect spatial effects of infrastructure over the Pakistan's regional development using the extended version of the neoclassical growth framework of (Ertur & Koch, 2007; Mankiw et al., 1992).

$$ln\hat{y}_{it} = \alpha_0 lnA_0 + \alpha_1 t + \alpha_2 lns\mathrm{k}_{it} + \alpha_3 lns\mathrm{h}_{it} + \alpha_4 lns\mathrm{g}_{it} - \alpha_5 ln(\delta + \lambda + n) + \alpha_6 \ln y^*_{it} + \alpha_7 \ln y^*_{it-1} + \alpha_8 \ln Sk^*_{it} + \alpha_9 \ln Sh^*_{it} + \alpha_{10} \ln Sg^*_{it} - \alpha_{11}\, ln(\delta + \lambda + n)^*_{it} + \alpha_{12} lny_{it,0} + \varepsilon_{it} \quad (6.1)$$

The estimable equation (6.1) represented the temporal relationship between output per worker, and the share of capital stock, human capital, and core infrastructure. The core or economic infrastructure has intrinsic networking characteristics and the capacity to generate externalities. While investing in human capital generates positive returns on education (Krueger & Lindahl, 2001) and spillovers (McMahon, 1979) and regional productivity (Andersson et al., 2007). For the calibration of the infrastructure, we are using the physical unit of measurement instead of the book-value of public capital. The core infrastructures include the transport networks, communication, and energy networks (Gianpiero, 2009). The core infrastructure is widely acknowledged growth-promoting input, via multiple direct and indirect channels (Baldwin & Dixon, 2011). Based on the previous chapter 4, section 2, empirical model equation (6.1) is based on the mathematical model constructed. The sg_t represents the share of resources invested in building and developing the core infrastructure. Literature defines the major components of core infrastructure as transport, energy, and communication infrastructure. Therefore, we analyze the component-wise and in aggregate terms (by constructing an index using PCA). The Sg^*_{it} represents the share of resources in developing core infrastructure in the neighbouring provinces. The α_8 are the indirect benefits or losses that arise as a result of the regional connectedness.

6.3 Indicators and Data Sources

Following indicators are used;

1. y_{it}: **Gross National Income Per Capita (GNIPC):** ***Dependent Variable***

The data of GDP at the provincial level is not available, but the data of gross national income per capita (GNIPC[7]) in Purchasing Power Party (PPP) measured in prices of (2011) measured in USD, therefore we are using GNIPC as a proxy to measure the size of output per person. The data has been obtained from the global data lab (GDL)[8]. The global data lab developed sub-national

[7] The link to the methodology https://datahelpdesk.worldbank.org/knowledgebase/articles/378831-why-use-gni-per-capita-to-classify-economies-into

[8] *Global Data Lab (GDL), Institute for Management Research, Radboud University, Global Data Lab (GDL) Institute for Management Research, Radboud University*

statistics for 131 countries across the globe. The GNIPC is an indicator depicting the state of development as income is closely correlated with the indicators measuring the quality of life.

2. $S\text{k}_{it}$: **Energy consumption data**

The data of capital stock is not available; therefore, we use the unit of electricity sold(used) for commercial and industrial purposes as a proxy to measure the size of private capital stock. However, the literature recommends data energy consumption as a better proxy of the book value of the capital stock (Frank, 1959), as it does help to estimate the size of the actual economy using shadow measures, particularly in developing economies. The data has been obtained from various provincial development statistics publications issues.

3. $S\text{h}_{it}$: **Mean years of schooling:**

we use years of schooling(mean) to proximate the share of resources devoted to building the human capital. The data has been obtained from the GDL.

4. $S\text{g}_{it}$: **Core Infrastructure**

We developed an index to measure core infrastructure using PCA techniques road length of high type roads (road network), length of transmission lines (energy distribution network), and the number of telephone and mobile subscribers (communication network). We use population weights to segregate province-wise data of communication infrastructure. The road length data and teledensity have been obtained from provincial development statistics, while the length of transmission lines data has been obtained by a report of NTDC (2021). The literature highlights road, energy, and communication are highly correlated and jointly create a joint effect. Therefore, we developed an index representing the core infrastructure (also known as economic infrastructure). We employ the z-score method to normalize, as the unit of measurement differs, and weights have been assigned using PCA. We used the normalized weights using the principle-component 1, as it shows the largest variation in the dataset.

5. $(\delta + \lambda + n)$: **breakeven level investment per capita**

We constructed this variable by calculating the population growth rate using the province-wise population statistics provided by GDL and added a breakeven investment level of 0.05 (assumed 5% depreciation percentage of capital stock). This is the desired level of investment necessary to keep up the falling stock of capital per capita.

6. $S\text{h}_{it}^{*}$: **Mean years of schooling in neighboring provinces**

We estimated to gauge the indirect effect of investing in human capital in neighboring province. We developed a spatial lag variable using a spatial weight matrix.

7. $S\text{k}_{it}$: **Energy consumption for commercial and industrial purposes in the neighboring province**

In order to incorporate the indirect effects of private capital stock, we developed a spatial lag variable by the matrix operations using the spatial weight matrix and the unit of electricity sold for commercial and industrial purposes.

8. $S\text{g}_{it}^{*}$: **Core Infrastructure in a neighboring province**

In order to indirect effects of core infrastructure, we developed a spatial lag variable using core infrastructure for the neighboring locations via matrix operations using spatial weight matrix [W].

9. $(\delta + \lambda + n)_{it}^{*}$: **Per-capita Breakeven investment in the neighboring province**

We developed a spatial lag variable using matrix operations using the spatial weight matrix [W].

10. [**W**]: **Spatial weight matrix**

We employed a row-standardized spatial weight matrix [W] of 4 x 4 dimensions by using defining contiguity neighborhood of queen continuity of order 1. The connectivity maps are constructed using GeoDa software for four provinces of Punjab, KPK, Sindh and Baluchistan only due to the data limitation.

6.4 Econometric Methodology

The panel-data includes both cross-sections (*i*=4) and autoregressive (*t*=31). Therefore, we need to test for panel unit-root and slope homogeneity tests to select an appropriate estimation model to unfold the key drivers of regional economic development in Pakistan.

6.4.1 Testing stationarity (the panel Unit root tests)

The panel unit root diagnostic tests are recommended in empirical literature to detect non-stationarity data series. However, incorporating the non-stationarity of variables in the estimated models can lead to spurious results, showing a stronger association between economic variables than actual. Therefore, it is imperative to conduct a unit-root analysis. The literature has categorized panel-unit test into two broad types. First-generation panel unit root tests and second-generation panel unit root tests.

6.4.1.1 First Generation Panel Unit root tests:

These tests assume each cross-section*(i)* are cross-sectionally independent. There are various tests under this category. These are the Levin–Lin–Chu (LLC) test, Im–Pesaran–Shin (IPS) Test, Breitung unit-root test, Harris–Tsavalis and Hadri LM test. We are using testing LLC, IPS, Breitung and Hadri LM unit-root test.

$m_{it} = \delta_i m_{i,t-1} + \boldsymbol{z'_{it}}\partial_i + e_{it}$

where *i* =1,..N, number of panels and, t =1,…T indexes time: m_{it} is the variable being tested; and e_{it} is stationary error term. the $\boldsymbol{z_{it}}$ represents panel specific fixed effects. The panel unit-root tests are used to test the null hypothesis H0 : $\delta_i = 1$ for all *i* versus the alternative Ha : $\delta_i < 1$. However, the statistical estimation software uses an alternative approach by using the first difference of dependent variable which is the same equation using a difference operator.

$\Delta m_{it} = \emptyset_i m_{i,t-1} + \boldsymbol{z'_{it}}\partial_i + e_{it}$

The null hypothesis is then Ho: $\emptyset_i = 0$ for all *i* against the alternative hypothesis of $\emptyset_i < 0$. Almost all panel unit test such as LLC, IPS, Breitung test use the same hypothesis (with different set of assumptions) testing except Hadri LM test. We conducted panel unit test using Stata software using xtunitroot command. These tests are slightly different assumptions but the LLC (xtunitroot llc), and Breitung (xtunitroot breitung) tests make the simplifying assumption that all panels share the same autoregressive parameter so that $\delta_i = \delta$ for all *'i'*. However, Hadri LM test for panel stationarity instead assumes the null hypothesis that all panels are stationary against the alternative that at least some of the panels non-stationary.

6.4.1.2 Second-Generation Panel Unit root tests:

The second generation of panel unit root tests relaxes the assumptions and include the cross-sectional dependence. The cross-sectional dependency (CSD) in macro-panel data arises mainly due to a type of correlation arising from the outcome of the common shocks and the local spillover effects between regions or countries. There are a variety of tests for cross-section dependence in the literature. These are Breusch-Pagan (1980) LM, Pesaran (2004) scaled LM, Baltagi, Feng, and Kao (2012) bias-corrected scaled LM and Pesaran (2004) CD test. The null hypothesis is no autocorrelations between disturbances terms across cross-section units, against the correlation of disturbance terms across cross-sections.

6.4.2 Slope homogeneity test

The standard panel-data regression models like fixed effects (FEs) and random effects (RE) panel models assume that the parameters are homogeneous. Hence, ignoring or incorrectly specifying the model by ignoring slope heterogeneity incurs bias in the results (Pesaran & Smith, 1995). We conduct the slope homogeneity hypothesis test of the Pesaran and Yamagata (2008) and used the robust standard errors for macro-dynamic panel by Blomquist and Westland (2013) which takes serial autocorrelation into account.

6.4.3 First and Second-Generation Panel Cointegration Tests

Macro timeseries usually are autoregressive, therefore we need to test for the long run association of the variables. When CSD is absent in model, the first generation cointegration test cointegration tests of Pedroni (1999, 2004) and Kao (1999) are recommended (Lau et al., 2019). On the other hand when CSD presents in the model the Pedroni and Kao cointegration tests results remained non-robust. In this case, the Westerlund (2007) test is highly recommended. This test is developed based on testing an error correction mechanism given CSD (also known as the second-generation panel cointegration test). This test examines the absence of cointegration by determining if error correction exists among the individual panel members or the whole panel.

6.4.4 Model-selection criterion

Scenario-I: If the test found no cross-sectional dependence and slope are homogeneous, and variables are I(0), then pooled OLS model, FE models, or RE models are appropriate.

Scenario-II: If cross-sectional dependence and slope are heterogeneous and variables are integrated of higher order. Then, the cross-sectional dependence model with error correction mechanisms such as Panel ARDL models such as cross-sectional autoregressive distributed lag model (CS-ARDL) or cross-sectional common correlated effects (CS-CCE) models are recommended.

Scenario-III: If cross-sectional dependence exists, slopes are homogenous, and variables are integrated of a higher order than I(0). Then, we need to incorporate the cross-sectional dependence in the autoregressive model. In this scenario, spatial panel (auto-regressive) models are more appropriate, which take the spatial spillover effect into account. These models are preferred because the detected cross-dependence is mainly developed due to the regional interdependence.

6.4.5 Spatial Panel data models

The panel data consist of both time-series and cross-sectional characteristics, and in this study, the number of time periods (T) is greater than the number of cross-sections (N). Therefore, we need to employ a dynamic panel data model for estimation. We will estimate both dynamic and static models (for comparison of estimates and robustness check). And we will estimate dynamic spatial models if data is non-stationary and CSD exists in the data. More details are available in the next chapter of estimations. We are using the provincial data of Pakistan, and we expect cross-sectional dependence due to multiple factors such as a common federal regime, easy mobility, and geographic connectedness. Therefore, spatial models are more appropriate for regional-level analysis.

A general representation of the dynamic spatial autoregressive panel data model is as given in the equation. This model is also known as Spatial Durbin Model (SDM) (Anselin, 2007; Lesage, 1999). The SDM model can be represented in the dynamic Spatial autoregressive (SAR) model form when post estimation test reveals ρ & ψ & $\tau \neq 0$ and $\theta = 0$

$$y_{it} = \tau y_{i,t-1} + \rho \boldsymbol{W} * y_{it} + \psi \boldsymbol{W} * y_{i,t-1} + X_{it}\beta + \boldsymbol{W} * X_{it}\theta + \alpha_i + \varepsilon_{it} \quad \text{SDM model} \quad (6.2)$$

$$y_{it} = \tau y_{i,t-1} + \rho \boldsymbol{W} * y_{it} + \psi \boldsymbol{W} * y_{i,t-1} + X_{it}\beta + \alpha_i + \varepsilon_{it} \quad \text{SAR model} \quad (6.3)$$

The spatial model includes the spatial rho representing the intra-regional spatial spillover, and we estimate these direct, indirect, and spillover effects using the 'xsmle' command in Stata. The dynamic models are time-space-lagged dependent and compute the bias-corrected Quasi-Maximum Likelihood (QML) approach. This approach constructs maximum likelihood estimates, treating the lagged variables as exogenous regressors. The bias corrections are computed for each coefficient to adjust the initial maximum likelihood estimates. The coefficients' default asymptotic variance-covariance (VC) matrix is obtained from the observed information matrix. To address the potential dangers of unknown serial correlation in the errors for each panel unit, the robust standard errors are calculated. We are using the loglikelihood ratio to select the appropriate spatial model following the Debarsy (2012) recommendations.

Step 1: Estimation of Spatial Panel Data Model

Using equation 6.4 we estimated the FE and RE spatial data model and test the restrictions for the spatial variables.

$$lny_{it} = \rho[W] * lny_{it} + \tau lny_{i,t-1} + \psi[W] * lny_{i,t-1} + \beta_0 + \beta_1 lnSk_t + \beta_2 lnSh_t + \beta_3 lnSg_t - \beta_4 ln(\delta + \lambda + n) + \theta_1[W] * Sk_t + \theta_2[W] * Sh_t + \theta_3[W] * lnSg_t - \theta_4[W] * ln(\delta + \lambda + n) + \alpha_i + \varepsilon_{it} \quad (6.4)$$

The fixed-effect model take-into account the unobserved heterogeneity and estimate α_i as non-random allows errors to correlate it with explanatory variables $Covar(X_{it,}\alpha_i) \neq 0$. While the random effect model estimates the parameter α_i by taking as random variables, assuming $Covar(X_{it,}\alpha_i) = 0$.

The Hausman specification test the null hypothesis of the difference in coefficients is not systematic against the alternative of system difference in coefficient. This test evaluates the standard errors obtained from fixed and random effects regression. A lower p-value allowed researchers to accept the alternative hypothesis, and therefore, we prefer a fixed effect panel data model over a random effect model.

The restriction test decides whether the spatial panel data models are more appropriate than non-spatial dynamic models. For this study, utilization of the spatial model is subject to the post estimation significant test $\rho = 0$, ψ=0 and $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$ (non-spatial panel data model) against the alternative of $\rho \neq 0, \psi \neq 0$ and $\theta's \neq 0$ (for SDM) and $\rho \neq 0$, $\psi \neq 0$ and $\theta = 0$ (SAR model). Model selection would be made based on the LR ratio and robust standard errors to be used to counter the possible serial autocorrelation issue.

Step 2: **Estimation of the long-run and short-run direct and indirect effects**

The dynamic short-run effects of infrastructure would be estimated $(I - \rho\,[W])^{-1} * (\beta_3 I + \theta_3[W])\}^d$ and short-run indirect effects will be $\{(I - \rho\,[W])^{-1} * (\beta_3 I + \theta_3[W])\}^{rsum}$.

The long-rum direct effects are estimated $(I - \tau)I - (\rho + \psi)\,[W])^{-1} * (\beta_3 I + \theta_3[W])\}^d$ and indirect effects are measured as $(I - \tau)I - (\rho + \psi)\,[W])^{-1} * (\beta_3 I + \theta_3[W])\}^{rsum}$.

d is an operator that calculates the mean of diagonal elements of a matrix. rsum is an operator that calculates means row sum of non-diagonals elements. The estimated model is to be computed using delta standard errors. The delta method ensures that the results do not depend on the stochastic variability of data (Belotti et al., 2017).

6.5 Data, stationarity test and cointegration tests

This section provides the describe data and pre-estimation tests for panel unit root detection and cointegration test.

6.5.1 Description of the data

We are using panel data (provincial level data) for FY 1990 to 2020 for four provinces of Pakistan. The data description is given in table 6.1 below.

Table 6. 1 *Description of Data*

Variable	**Mean**	**Std. Dev**	**Min**	**Max**
lny_{it}	8.10	0.21	7.77	8.63
$lnSk_{it}$	7.67	1.69	4.22	10.15
$lnSh_{it}$	1.20	0.31	0.58	1.71
Sg_t	0.04	1.22	-2.75	2.68
$lhtr$	9.58	1.06	6.93	11.39
$lltl$	8.92	0.82	7.66	10.46
$ltelmob$	9.94	1.88	6.18	12.86
$ln(\delta + \lambda + n)$	1.02	0.11	0.73	1.23

6.5.2 Panel Stationarity tests

The time-series data is often nonstationary, and the selection estimation techniques largely depend on the order of integration., Therefore, we conducted the first and second-generation panel-unit tests.

6.5.2.1. First-Generation Panel Unit Root Tests

We use first-generation panel unit tests of the Levin, Lin, and Chu (LLC) unit root, Im-Pesaran-Shin (IPS) panel unit root, Breitung unit root, and Hadri LM test. The results of the first-generation panel unit-root tests are provided in table 6.2.

Table 6.2: *The Stationarity test (first-generation Panel Unit test)*

Variables	LLC Test		IPS Test		Breitung unit-root test		Hadri LM test	
	A. t-stat	*p-val*	*t-bar*	*p-val*	*L-Stat*	*p-val*	*Z-stat*	*p-val*
At level								
lny_{it}	2.62	0.99	1.62	1.00	3.20	0.99	17.71	0.00
$lnSh_{it}$	-2.82***	0.00	-3.25***	0.00	0.38	0.65	18.09	0.00
$lnSk_{it}$	-0.52	0.30	-1.05	0.85	1.88	0.97	17.97	0.00
Sg_t	-1.20	0.12	0.98	0.84	0.09	0.54	22.59	0.00
$lhtr$	-1.94	0.03	-1.58	0.45	1.67	0.95	15.56	0.00
$lltl$	-0.62	0.27	-1.21	0.76	2.90	1.00	18.73	0.00
$ltelmob$	-1.50	0.07	-1.26	0.69	0.21	0.58	18.52	0.00
$ln(\delta + \lambda + n)$	0.25	0.60	-0.98	0.91	1.80	0.96	7.39	0.00
First-difference								
$d.lny_{it}$	-2.34***	0.01	-3.10***	0.00	-2.99***	0.00	5.26	0.00
$d.lnSk_{it}$	-5.80***	0.00	-7.14***	0.00	-2.79***	0.00	-1.48***	0.93
$d.lltl$	-4.91***	0.00	-5.40***	0.00	-4.72***	0.00	0.34***	0.37
$d.ltelmob$	-1.57*	0.06	-1.57*	0.06	-3.28***	0.00	2.32	0.01
$d.ici$	-2.59***	0.00	-2.62**	0.02	-2.37***	0.01	0.65***	0.26
$d.ln(\delta + \lambda + n)$	-5.88***	0.00	-5.03***	0.00	-3.90***	0.00	0.91***	0.18
Order of Integration								
lny_{it}		I(1)						
$lnSh_{it}$		I(0)						
$lnSk_{it}$		I(1)						
Sg_t		I(1)						
$lhtr$		I(0)						
$lltl$		I(1)						
$ltelmob$		I(1)						
$ln(\delta + \lambda + n)$		I(1)						

6.5.2.2 Second-Generation Panel Unit Root Tests

The second-generation panel unit root tests detect the existence of the cross-section dependence. Therefore, we computed the tests of the Breusch-Pagan Langrage Multiplier (LM), the Pesaran scaled LM, the Bias-corrected scaled LM, and the Pesaran CD test (also known as the cross-sectionally augmented IPS (CIPS) test by Pesaran (2007)). The results are presented in Table 6.3.

Table 6.3: *Cross-sectional dependence test (second-generation Panel Unit test)*				
Variable	Breusch-Pagan LM	Pesaran scaled LM	Bias-corrected scaled LM	Pesaran CD
lny_{it}	163.19***	45.38***	45.31***	12.76***
$lnSh_{it}$	170.10***	47.37***	47.30***	13.03***
$lnSk_{it}$	145.92***	40.39***	40.32***	12.07***
Sg_t	172.79***	48.15***	48.08***	13.14***
$lhtr$	64.53***	16.90***	16.83***	6.42***
$lltl$	112.46***	30.73***	30.67***	10.19***
$ltelmob$	185.74***	51.89***	51.82***	13.63***
$ln(\delta + \lambda + n)$	99.07***	26.87***	26.80***	-1.75

The panel unit-root test results revealed that none of the variables of the models is I (2), and most of the variables are I(1), except the variable of $lnSh_{it}$. All test of cross-sectional (CS) dependence depicts cross-sectional dependence. Literature proves that neglecting the cross-sectional dependence can lead to biased estimates and spurious inferences (Chudik et al., 2011).

6.5.2.3 Panel Cointegration test and slope homogeneity test

We need a further test for cointegration and slope homogeneity (delta test) to select an appropriate estimation technique. In the given scenario of the CS dependence, the cointegration tests results of Pedroni and Kao panel-data cointegration test became irrelevant, and the literature recommends the Westerlund panel cointegration. The test results are provided in table 6.4.

Table 6.4 Results of Westlund: Panel cointegration test

	Gt	**Ga**	**Pt**	**Pa**
Model 1	-1.936	-2.573	-2.492	-3.083
Model 2	-1.705	-2.442	-1.02	-1.442
Model 3	-1.993	-4.888	-0.324	-0.365
Model 4	-1.505	-2.66	-1.664	-2.594

For selecting an appropriate estimation technique, we are testing for the slope homogeneity developed by Pesaran and Yamagata (2008), using robust standard errors. These robust standard errors are computed using Blomquist and Westland (2013) approach for the macro-dynamic panel, which considers the serial autocorrelation. The results of the test are given in table 6.5.

Table 6.5: *Result of the Pesaran and Yamagata 2008 (Delta test)*		
	Delta (HAC)	P-value
Model 1	-1.276	0.202
Model 2	-0.249	0.804
Model 3	0.362	0.717
Model 4	-1.132	0.258

The result of panel cointegration shows no cointegration amongst the variables. Therefore, we can use alternative estimation techniques from long-run models, such as Error correction models. In the presence of CS dependence and slope heterogeneity, the recent literature recommends CS-DL, CS-ARDL, and CS-CCE models (especially for large T and N). However, our models exhibit insignificant long-run cointegration and homogenous slopes with the cross-sectional dependence. Therefore, we will use spatial panel models given the reasons for the cross-sectional dependence due to regional interdependence and economic cohesion. These spatial panel models will segregate the direct and indirect effects and the unintended spillover effects.

6.6 Index construction and Estimation of models

We estimated four models. In the Model (1) represents infrastructure is defined as an aggregated index for the core infrastructure. The core infrastructure includes the components of transport, energy, and communication infrastructures. The index has been constructed using normalized variable using z-score standardization method, due to the difference in measurement unit and developed index using Principle-Compnent-1 as the eigen value is greater than 1 and it explains 84% variation.

	PC-1	PC-2	PC-3
$nlnhtr$	0.60	-0.29	-0.74
$nlnltl$	0.59	-0.47	0.66
$nlntel$	0.54	0.83	0.11
Eigenvalue	2.51	0.38	0.11
Proportion	0.84	0.13	0.04

We constructed index for core infrastructure using the normalized weights using PC-1 by using equation using equation 6.5.

$$lnSg = 0.35\, nlnhtr + 0.34 nlnltl + 0.31 nlntel \quad (6.5)$$

We estimated the Dynamic Spatial Durbin Model (SDM) using index of core infrastructure along-with other critical drivers of long-term economic growth of human capital, physical capital, and break-even level of investment per capita. We estimated the model (1) using random effect (equation 6.6) and fixed effect (equation 6.7) specifications.

$lny_{it} = \rho[W] * lny_{it} + \beta_0 + \beta_1 lnS_k + \beta_2 lnS_h + \beta_3 lnS_g - \beta_4 ln(\delta + \lambda + n) + \theta_1[W] * lnS_k + \theta_2[W] * lnS_h + \theta_3[W] * lnS_g - \theta_4[W] * ln(\delta + \lambda + n) + \varepsilon_{it}$ (mode1 1: RE) (6.6)

$lny_{it} = \rho[W] * lny_{it} + \tau lny_{i,t-1} + \psi[W] * lny_{i,t-1} + \beta_1 lnS_k + \beta_2 lnS_h + \beta_3 lnS_g - \beta_4 ln(\delta + \lambda + n) + \theta_1[W] * lnS_k + \theta_2[W] * lnS_h + \theta_3[W] * lnS_g - \theta_4[W] * ln(\delta + \lambda + n) + \alpha_i + \varepsilon_{it}$

(model 1: FE) (6.7)

We estimated the fixed effect SDM (using spatial-FE dynamic SDM specifications) and RE and conducted the Hausman test. The Hausman test the spatial FE SDM against the RE SDM. The SDM is a nested model therefore, we performed following restriction tests.

1: Test $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$ against, $\theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4 \neq 0$

2: Test $\rho = 0$, and $\psi = 0$ against Test $\rho \neq 0$, and $\psi \neq 0$

2: Test $\theta_1 = -\rho * \beta_2 = \theta_2 = -\rho * \beta_2 = \theta_3 = -\rho * \beta_3 = \theta_4 = -\rho * \beta_4$ against, $\theta_1 = -\rho * \beta_2 \neq \theta_2 = -\rho * \beta_2 \neq \theta_3 = -\rho * \beta_3 \neq \theta_4 = -\rho * \beta_4$

All test has a high p-value; therefore, we reject the null hypothesis against the alternative; therefore, we conclude dynamic spatial fixed effect SDM is a better model than SAR or SEM model. The output results are provided in Table 6.6. The marginal direct, indirect, and total effects using dynamic effects (3) and delta standard errors has been obtained the static and dynamic (short-run and long-run) and results are provided in Table 6.8.

Model 2 (based on Equation 6.6 and 6.7) included the variable of the road-infrastructure (log of high-type roads length km) in the model.

$lny_{it} = \rho[W] * lny_{it} + \beta_0 + \beta_1 lnSk_{it} + \beta_2 lnSh_{it} + \beta_3 lnhtr_{it} - \beta_4 ln(\delta + \lambda + n) + \theta_1[W] * lnS_k + \theta_2[W] * lnS_h + \theta_3[W] * lnhtr_{it} - \theta_4[W] * ln(\delta + \lambda + n) + \varepsilon_{it}$

(Model 2: RE) (6.8)

$lny_{it} = \rho[W] * lny_{it} + \tau lny_{i,t-1} + \psi[W] * lny_{i,t-1} + \beta_1 lnS_k + \beta_2 lnS_h + \beta_3 lnS_g - \beta_4 ln(\delta + \lambda + n) + \theta_1[W] * lnS_k + \theta_2[W] * lnS_h + \theta_3[W] * lnS_g - \theta_4[W] * ln(\delta + \lambda + n) + \alpha_i + \varepsilon_{it}$

(Model 2: FE) (6.9)

The post estimation Hausman specification test supported for the selection of RE compared to FE-SDM. The results are provided in the Table 6.7. The post-estimation Wald-coefficient test supported for the SDM than SAR and SEM.

In addition to this, energy infrastructure (log of length of transmission lines) is included in the model 3 (Equation 6.10 & 6.11).

$$lny_{it} = \rho[W] * lny_{it} + \beta_0 + \beta_1 lnSk_{it} + \beta_2 lnSh_{it} + \beta_3 lnltl_{it} - \beta_4 ln(\delta + \lambda + n)_{it} + \theta_1[W] * lnSk_{it} + \theta_2[W] * lnSh_{it} + \theta_3[W] * lnltlr_{it} - \theta_4[W] * ln(\delta + \lambda + n)_{it} + \varepsilon_{it}$$

(Mode1 3: RE) (6.10)

$$lny_{it} = \rho[W] * lny_{it} + \tau lny_{i,t-1} + \psi[W] * lny_{i,t-1} + \beta_1 lnSk_{it} + \beta_2 lnSh_{it} + \beta_3 lnltl_{it} - \beta_4 ln(\delta + \lambda + n)_{it} + \theta_1[W] * lnSk_{it} + \theta_2[W] * lnSh_{it} + \theta_3[W] * lnltl_{it} - \theta_4[W] * ln(\delta + \lambda + n)_{it} + \alpha_i + \varepsilon_{it}$$

(Model 3: FE) (6.11)

The post estimation revealed energy infrastructure is RE-SDM is better than FE-SDM based on Hausman specification test. Moreover, post-estimation Wald-coefficient test supported for the SDM than SAR and SEM.

Lastly communication infrastructure (teledensity) is included in the model 4. Model 4 is estimated using FE and FE specification using equations 6.12 and 6.13.

$$lny_{it} = \rho[W] * lny_{it} + \beta_0 + \beta_1 lnSk_{it} + \beta_2 lnSh_{it} + \beta_3 lntel_{it} - \beta_4 ln(\delta + \lambda + n)_{it} + \theta_1[W] * lnSk_{it} + \theta_2[W] * lnSh_{it} + \theta_3[W] * lntel_{it} - \theta_4[W] * ln(\delta + \lambda + n)_{it} + \varepsilon_{it}$$

(Mode1 4: RE) (6.12)

$$lny_{it} = \rho[W] * lny_{it} + \tau lny_{i,t-1} + \psi[W] * lny_{i,t-1} + \beta_1 lnSk_{it} + \beta_2 lnSh_{it} + \beta_3 lntel_{it} - \beta_4 ln(\delta + \lambda + n)_{it} + \theta_1[W] * lnSk_{it} + \theta_2[W] * lnSh_{it} + \theta_3[W] * lntel_{it} - \theta_4[W] * ln(\delta + \lambda + n)_{it} + \alpha_i + \varepsilon_{it}$$

(Model 4: FE) (6.13)

$\boldsymbol{\rho}$ spatial dependence parameter and $\boldsymbol{\tau}$ the autoregressive time dependence parameter and $\boldsymbol{\psi}$ the spatiotemporal diffusion parameter while $\boldsymbol{\theta' s}$ represents a linear combination of physical, human and core infrastructure in the neighbouring provinces weighted a common borders. (Debarsy et al., 2012). The estimated four models are presented in the table 6.6. Based on the all-estimated models, we found a positive and significant spatial dependence and positive autoregressive time-dependence parameter. However, the spatiotemporal diffusion parameter remained negative and significant except model 4.

Based on the FE-SDM for model 1, the diffusion parameter is negative and significant and the size of ψ is 0.244, therefore, we can support for the evidence of reverse-spatial-temporal diffusion, and we can state that a percent increase in the income per capita in neighbouring provinces is likely reduce by 0.244 by income per capita, keeping the effect of physical, human and core infrastructure effect in constant. The term s*pillover* to refer to contemporaneous cross-partial derivatives or the marginal effects (Debarsy et al., 2012). These are the partial derivatives measuring how region *i'*s dependent variable responded over time to changes in the initial period levels of the explanatory variables. These cross-partial derivatives that involve different time periods are referred to as diffusion effects, since diffusion takes time. The results of the short-run and long run marginal effects are provided in the Table 6.8.

6.7 Key Findings

The term s*pillover* refers to contemporaneous cross-partial derivatives or the marginal effects (Debarsy et al., 2012). These are the partial derivatives measuring how region *dependent variables* responded over time to changes in the initial period levels of the explanatory variables. These cross-partial derivatives that involve different time periods are referred to as diffusion effects, since diffusion takes time. The results of the short-run and long-run marginal effects are provided in Table 6.8. All estimated models from 1 to 4, spatial dependence-parameter (*spatial-rho*) is positive and highly significant. Based on the FE-SDM for model 1, the diffusion parameter is negative and significant and the size of $\boldsymbol{\tau}$ is 0.244, therefore, we can support for the evidence of reverse-spatial-temporal diffusion, and we can state that a percent increase in the income per capita in neighbouring provinces is likely reduce by 0.244 by income per capita, keeping the effect of physical, human and core infrastructure effect in constant. These effects arises due to relative attractiveness of provinces elucidating the resource flow from relatively poor provinces to the rich provinces of Pakistan. These resource flows are in in terms of labour resources and capital resources.

The marginal effects depicting the cross-partial effects and the direct effect is dependent on the spatial-dependence parameters spatial-rho and size of indirect effect dependent on the all parameters of spatial-temporal parameters and spatial dependence of ρ (spatial Rho), $\boldsymbol{\psi}$ and $\boldsymbol{\tau}$. These indirect effects represent the contemporaneous spatial spillovers plus diffusion over time. The estimated marginal effects for Model-1, the direct and indirect spillover both are negative and high (highly significant) in short run for core infrastructure. In short run, a point increase in the index for the core infrastructure is likely to reduce GNI per capita by 5.07 percent in short run while in long run the direct effect is positive and significant. In the long run, keeping the effect of other variables effects constant, an increase in index for core infrastructure by one standard-deviation point is likely to increase GNI per capita by 16.6 percent. On the other hand, the indirect short-run effects of investing in the core infrastructure are negative and statistically significant. On average, a point increase in the index for core infrastructure by one standard-deviation, GNI per capita reduces through negative spillovers by 7.9 percent in the short run. Therefore, the total effect is negative (12.4 percent). However, in the long-run indirect spillovers effects are positive in size (16.9 percent) but are statistically insignificant. Therefore, the long run total-effects is positive (33.4 percent) but statically insignificant.

We analysed the component wise impacts for doing so, therefore we estimated Dynamic SDM (model-2) using the indicators of natural log of high-type road (HTR). The estimated parameter is based on the results of Dynamic SDM (RE) supporting the evidence of positive networking-externalities. The β_3 is positive but the statistically insignificant, however θ_3 is positive and statistically significant. Keeping the effect of other variables of model constant on average, an increase in the road length in the neighbouring provinces by one percent point resulted into increase GNI per capita by 0.114 percentage. Likewise, we estimated model 3 for component wise analysis for energy infrastructure (natural log of length of transmission lines). The spatial dependence parameters (FE and RE both) are positive and significant. The estimated model of SDM (RE) supported for existence of positive and significant /networking effect of energy infrastructure. Keeping the direct and neighbouring factors of human, physical and break-even level of investment effects constant, an increase in one percentage point increase in the length of transmission in the neighbouring provinces resulted into increase GNI per capita by 0.173 percentage. Lastly, we estimated the impact of telecommunication infrastructure exclusively on the long run and short-run impacts on the income per capita. Compared to RE-SDM, FE-SDM is a well-specified model. Therefore, under FE specification we can estimate long run and short-run direct and indirect marginal effects.

Table 6.6 ***Results of the Dynamic SDM***

Dependent variable: log of real GNI per capita (USD)								
	Model (1)		***Model (2)***		***Model (3)***		***Model (4)***	
	RE	**FE**	**RE**	**FE**	**RE**	**FE**	**RE**	**FE**
$\rho\ (\boldsymbol{W}.\ ln\boldsymbol{y}_{it})$	**0.304*****	**0.277*****	**0.471*****	**0.522*****	**0.442*****	**0.473*****	**0.530*****	**0.567*****
	(0.093)	(0.083)	(0.0713)	(0.077)	(0.0741)	(0.0806)	(0.084)	(0.055)
$\boldsymbol{\psi}\ (ln\boldsymbol{y}_{i,t-1})$		**1.264*****		**0.983*****		**0.971*****		**0.174*****
		(0.040)		(0.037)		(0.0412)		(0.063)
$\boldsymbol{\tau}(\boldsymbol{W}.\ ln\boldsymbol{y}_{i,t-1})$		**-0.244*****		**-0.392*****		**-0.342*****		**0.296*****
		(0.093)		(0.087)		(0.0923)		(0.0785)
Constant	4.011***		0.96		0.205		1.622***	
	(0.772)		(0.892)		(0.719)		(0.494)	
$lnSk_{it}$	0.082***	-0.025*	0.074***	-0.0524***	0.0618**	-0.0500***	0.0596**	-0.004
	(0.0137)	(0.013)	(0.027)	(0.012)	(0.025)	(0.0132)	(0.030)	(0.007)
$lnSh_{it}$	-0.121	-0.236***	0.465***	-0.218***	0.477***	-0.130**	0.578***	0.555***
	(0.132)	(0.059)	(0.132)	(0.057)	(0.126)	(0.0597)	(0.093)	(0.054)
$lnSg_{it}$	0.006	**-0.043*****						
	(0.0161)	(0.006)						
$lnhtr_{it}$			**0.0042**	-0.0396*				
			(0.065)	(0.023)				
$lnltl_{it}$					**0.0792**	0.0453*		
					(0.051)	(0.0232)		
$lntel_{it}$							0.263	**0.953*****
							(0.163)	(0.067)
$ln(\delta+\lambda+n)_{it}$	-0.098	0.210***	0.263**	0.111**	0.255**	0.128**	0.228**	0.0797***
	(0.132)	(0.054)	(0.121)	(0.046)	(0.119)	(0.051)	(0.100)	(0.030)
$\mathbf{W}lnSk_{it}$	0.100***	-0.071***	0.190***	-0.0402**	0.160***	-0.0666***	0.194***	0.004
	(0.034)	(0.021)	(0.052)	(0.020)	(0.0502)	(0.0218)	(0.042)	(0.013)
$\boldsymbol{W}lnSh_{it}$	-0.0824	0.721***	-0.557***	0.414***	-0.587***	0.347***	-0.619***	-0.472***
	(0.146)	(0.064)	(0.148)	(0.062)	(0.142)	(0.0653)	(0.111)	(0.058)
$\boldsymbol{W}lnSg_{it}$	0.110***	**-0.046*****						
	(0.021)	**(0.009)**						
$\boldsymbol{W}lnhtr_{it}$			**0.114***	0.0601**				
			(0.065)	(0.026)				
$\boldsymbol{W}lnltl_{it}$					**0.173*****	0.0412*		
					(0.0557)	(0.0248)		
$\boldsymbol{W}lntel_{it}$							-0.259	**-0.963*****
							(0.162)	(0.066)
$\mathbf{W}ln(\delta+\lambda+n)_{it}$	0.534*	1.295***	0.0168	1.114***	0.152	1.417***	-0.001	0.129*
	(0.296)	(0.111)	(0.315)	(0.107)	(0.317)	(0.121)	(0.238)	(0.068)
Lgt-theta	-2.896		-2.662***		-2.406***		-3.486	
Variance	0.0040***	0.0004***	0.00294***	0.000349***	0.00287***	0.000434***	0.00196***	0.000152***

Table 6.7 ***Model Selection Criterion***

	Model (1)		***Model (2)***		***Model (3)***		***Model (4)***	
AIC	-303.63	-562.31	-311.73	-579.99	-318.37	-556.56	-353.47	-673.78
BIC	-272.60	-528.86	-277.88	-546.54	-284.53	-523.11	-322.45	-640.33
LLR	178.95	178.95	224.31	224.31	204.10	204.10	344.35	344.35
R-Sq (With-in)	0.86	0.95	0.89	0.94	0.89	0.93	0.91	0.99
R-Sq (Between)	0.98	0.63	0.40	0.93	0.60	0.29	0.99	0.92
R-Sq (Over-all)	0.88	0.88	0.66	0.63	0.75	0.80	0.61	0.35
Hausman Test	**71.13*** **		-81.39		-106.29		**23.13*** **	

Table 6.8 Short and Long Run: Direct and Spillover Effects

Dependent variable: log of real GNI per capita (USD)						
	Short-Run Effects			**Long-Run Effects**		
	Direct	*Indirect*	*Total*	*Direct*	*Indirect*	*Total*
$lnSk_{it}$	-0.146***	0.827***	0.681***	1.021*	-2.841***	-1.820
	(0.056)	(0.079)	(0.083)	(0.572)	(0.943)	(1.348)
$lnSh_{it}$	-0.0362**	-0.102***	-0.138***	0.089	0.266*	0.355*
	(0.015)	(0.031)	(0.041)	(0.068)	(0.149)	(0.195)
Sg_{it}	-0.0507***	-0.0736***	-0.124***	0.166**	0.169	0.334
	(0.007)	(0.014)	(0.018)	(0.067)	(0.205)	(0.267)
$ln(\delta + \lambda + n)_{it}$	0.397***	1.72***	2.12***	-0.670	-4.927*	-5.598
	(0.083)	(0.226)	(0.290)	(1.076)	(2.747)	(3.768)

CHAPTER 7

ECONOMIC GEOGRAPHY OF INFRASTRUCTURE AND SOCIAL WELFARE IN PAKISTAN

7.1 Introduction

In Pakistan, the stature of infrastructure in general and rural areas in particular is poor and inequitably distributed and literature highlighted these fractured regional socio-economic disparities within Pakistan. These unequal district-wise landscape of the socio-economic development in Pakistan has been reported during 1960s, 1970s, and 1980's (Hussain, 1996), and these differences are engendering (UNDP, 2016). It is because the statistics glorified the poverty reduction at aggregate level, but inter and intra-provincial disparities (Gazdar, 1999) and districts-wise disparities (UNDP, 2016) are non-uniform reduction of poverty at sub-national levels. The data reveals in some districts the district-wise differentials of poverty are increasing. For instance, the negative change has been observed in the Multidimensional Poverty Index (MPI) in the districts of Harnai, Killa Abdullah, Ziarat, Umerkot, Sherani, Kashmore, Panjgur, Chagai, Pishin, Tando Muhammad Khan, and Badin during 2004 to 2014 (UNDP, 2016). While a negative change in poverty incidence has been reported for the district Chagai, Tharparker, Pishin, Ziarat, Tando Muhammad Khan, Kashmore, Killa Abdullah, Panjgour, Harnai, and Umerkot during 2004 to 2014. In contrast to these negative trends, the districts Islamabad, Attock, Jhelum, Lahore, Karachi, Rawalpindi, and Sialkot had a greater reduction in poverty reduction during 2004-2014, compared to other districts.

Keeping the descriptive data trends and differential in the social welfare across districts, one of the potential causes is appeared to be the relative difference in the stock of infrastructure in Pakistan. Therefore, this study chapter of the study focuses on the research question of how the differences in stock of transport infrastructure had created this inequitable distribution of social welfare in Pakistan. This chapter presenting the estimates of the direct and the spillovers effects of transport infrastructure that can account for alleviating poverty and social deprivation. Amongst all types of infrastructure, the transport infrastructure is the critical in reducing economic marginalization via networking, connectively and economic cohesion in developing countries(Fan & Chan-Kang, 2005; Hulten et al., 2006; Marinho et al., 2017). Therefore, this chapter is to supplant the existing literature for the regional development of Pakistan and support to devise an informed regional transport policy to abridge the regional differences in the social welfare at the district level.

7.2 An Analytical Framework of Infrastructure and Social Welfare

The study has incorporated 'social welfare' in the domains of poverty and infrastructure increases the welfare directly as well as indirectly, in multiple ways (theoretical channels):

- Households and firms are the direct users of public infrastructure such as roads, sewerage, electricity. So, the infrastructure directly impacting the quality of life and improving the quality of life resonates with low poverty.
- Infrastructure reduces economic poverty due to accessibility through spatial connectivity and economic integration.
    1. The income earnings due to better market accessibility. Households can market their excess good and services and earn more due to the efficiency gains due to ease in travel and communication.
    2. Poverty reduces via income diversification. People shifts their sources of income from farm to non-farm and service-oriented jobs.
    3. Poverty reduces through spatial accessibilities to other infrastructural services such as educational facilities, financial services, and health services. The economic infrastructure acts as a foundation to other social infrastructure-related services that enhance development outcomes, such as maternal mortality, mother and child health, and nutrition, leading to low poverty.
- Infrastructure reduces poverty in the neighboring regions through its geographical spillover effects. Infrastructure helps to integrate geographically sparse locations provides an avenue to find more and discover new employment opportunities across the board by creating economic clusters.
- Infrastructure reduces inequality due to greater relative expected benefits to the poorest and geographically marginalized population in an economy. The average income earning is likely to rise as a result of the infrastructure provisioning, but a more significant income rise is expected in the lowest quintile (bottom 40%). Therefore, infrastructure reduces the

gap between the top and bottom income groups by benefiting more to the poor. Infrastructure not only reduces inequality in locations where it is provided but is also expected to rise relative income in the poor population compared to non-poor through its spillover effects, as the income-earning opportunities are expected to expand in the poorest section of society in other regions too.

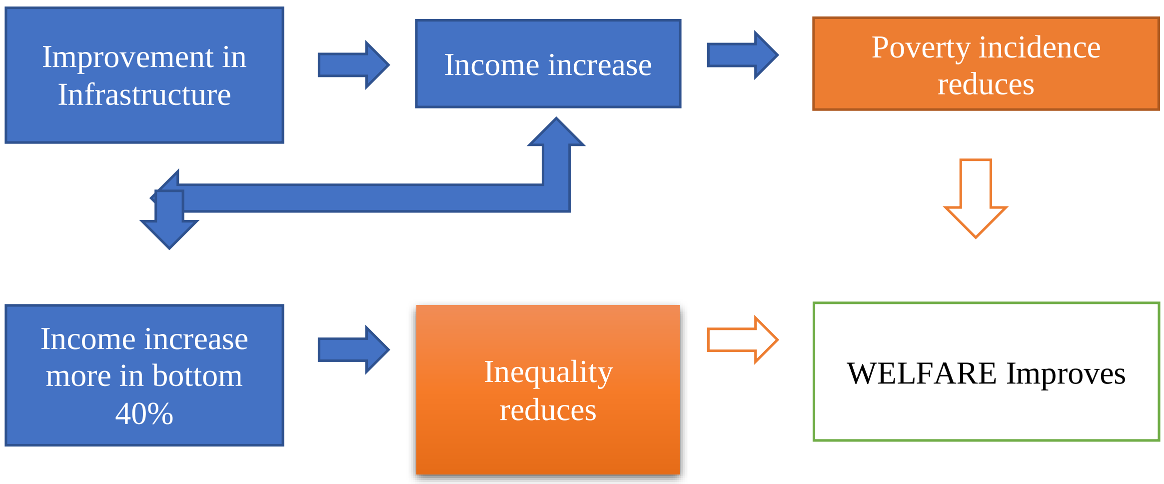


Figure 7. 1 Infrastructure and Welfare—theortical Linkages

Source: *Author's construct*

The empirical model construction/theoretical construction of social welfare depends on the approach of poverty measurement. Economic welfare is measured in terms of income or consumption. The measurement of the poverty is inherently embedded with the concept of utility. The theorical literature developed by Ravallion (1988) and Kurosaki (2006) used the constant relative risk aversion utility function, which is assuming poverty a continuous phenomenon where the welfare cost of poverty increases with the size of deprivation under the poverty line.

Let P be the aggregate measure of poverty for a population of size N and p_i be its individual score for person i, which is a function of his/her consumption c_i and an exogenously given poverty line z. Because of the scale invariance axiom, only $x_i(\equiv \frac{c_i}{z})$ matters. The class of poverty measures that are additively separable, symmetric, taking the value of zero for the consumption level exactly at z, and non-decreasing with the depth of poverty. Then,

$$P = \frac{1}{N}\sum_{i=1}^{N} p_i = \frac{1}{N}\sum_{i=1}^{N} p(x_i) \qquad (4.1)$$

Where $p(x_i) = 0$, when $x_i \geq 1, p(x_i) > 0$ when $x_i < 1$, and $\frac{\partial p}{\partial x_i} \leq 0$, when $x_i < 1$.

Assuming c_i is stochastic, the expected value of P can be decomposed into chronic and transient components as Ravallion (1988):

$$P^P = E\left[\frac{1}{N}\sum_{i=1}^{N} p_i\right] = \frac{1}{N}\sum_{i=1}^{N} E[p(x_i)] \qquad (4.2)$$

$$P^C = \frac{1}{N}\sum_{i=1}^{N} p[E(x_i)] \qquad (4.3)$$

$$P^T = P^P - P^C = \frac{1}{N}\sum_{i=1}^{N} E[p(x_i)] - \frac{1}{N}\sum_{i=1}^{N} p[E(x_i)] \qquad (4.4)$$

Where $E[.]$ represents an expectation operator. The expected value of P^P is total poverty, its components corresponding to the expected income P^C is chronic poverty, and the residual P^T reflecting the transient poverty. If there is no risk in income, the total poverty become equivalent to the chronic poverty so that the transient poverty become zero. An increase in risk will increase P^T is the function $p(x_i)$ belongs to the Atkinson (1987) poverty measures and is strictly convex in x_i when $x_i < 1$.

Poverty measurement by equation (4.1) can be representing as a social welfare function to aggregate the loss of individual welfare due to low income or consumption. The transient poverty component P^Tinterpreted as the welfare cost of fluctuations in income/consumption and the chronic poverty P^C can be interpreted as the welfare cost due to low level of expected income/or consumption.

The infrastructure is likely to reduce poverty by increasing permanent income as well as transient poverty directly and also the locational facilities provided by the neighbour. The locational impact welfare via networking. It impacts on the probability of households being poor or non-poor not only in the regions where infrastructure is located but the infrastructure impact on the permanent income in the other regions too. Therefore, we expected to include the spatial dimension into analysis, as expected the incidence of poverty to be affected due to the locational effects of infrastructure (direct and spatial spillover effects). The recent advancement in the spatial econometric[9] techniques to measure spatial dependence and spatial-lag regressions are helpful to estimates these indirect locational effects. We can present the model for infrastructure and poverty in eq (4.5).

$$Y_{it} = f\left(X_{it}, G_{it}, WG_{jt}, WY_{jt}\right) \qquad (4.5)$$

- Y_{it} social welfare in a location

[9] Detailed description of spatial weight is available on https://geodacenter.github.io/workbook/4a_contig_weights/lab4a.html

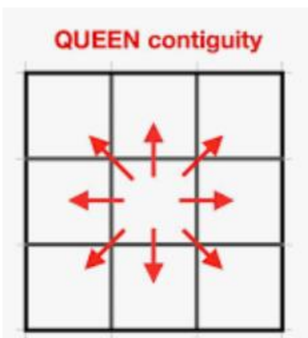
- X_{it} location specific other characteristics (such as land-use, cropping intensity)
- G Infrastructure
- W Spatial Weight matrix for the neighborhood

We are using to main definitions of the spatial weights [W] matrix for the spatial proximity relationship into the account as; It is (n x n) dimension matrix.

$$W_{nxn}=\begin{bmatrix} 0 & w_{12} & w_{13} & \cdots & w_{1m} \\ w_{11} & 0 & w_{23} & & w_{2m} \\ & \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \cdots & 0 \end{bmatrix}$$

We are using two main definitions to developed spatially dependent variables.

1. Contiguity-based matrix is to develop spatially lag variable of WY . The queen's contiguity criterion more encompassing and defines neighbors as the spatial units that shares a common edge or a common vertex.



$$W_{nxn} = \begin{Bmatrix} = 1, for\ neighbours \\ = 0, otherwise \end{Bmatrix}$$

$$WY_{ij} = \sum_{j=1}^{n} W_{ij}Y_j$$

The contiguity matrix for an assessment of localized effect when neighborhood spillovers are more important. The neighborhood is expressed as a dummy and takes the value 1 for pair region, 0 for otherwise, and these weights are row-standardized horizontally.

2. The inverse-distance (IDW) weight matrix is to develop spatially lag variable of WG . The inverse-distance matrix assigs weights for the distance $(1/d_{ij})$ for Euclidian. However, this matrix gives a linear, means strength of relationship increases with increase in distance. This matrix is better to taking into global effects into account without taking

into local clusters. We are using this matrix; because we interested to captures the global effects or spatial spillover effects across the regions.

$$WG_{ij} = \sum_{j=1}^{n} W_{ij} G_j$$

7.3 Empirical Model Specification

Infrastructure potentially reduces poverty through better geographical connectedness, which results in greater access to other social infrastructure services (e.g., education, health) and diversification of sources of income-earning (for example, establishment of the local markets for farm and non-farm products). In addition to the direct impacts, the geographic spillovers arise across neighboring regions, impacting social welfare. Literature for many countries documented the contribution of these effects of geographical spillover as positive, negative, or neutral in the regional productivity growth (Boarnet, 1998; Chen & Haynes, 2015; Shi et al., 2017). The positive spillovers are generated due to the networking and economies of scale. In contrast, negative spillovers occur in a region due to the net outflow of the factors from the places of less attractive regions to the relatively dense infrastructural networks and services. In light of the discussion, theoretical literature acknowledges the direct and indirect impact of investing in infrastructure on social welfare indicators (Allen & Arkolakis, 2019; Chatterjee & Turnovsky, 2012). Therefore, such a study was required for Pakistan to examine the impact of road infrastructure on social welfare by calibrating the direct, indirect, and spillover effects.

We can specify the econometric model using equation 4.5. The $\boldsymbol{y_{it}}$ is representing the dependent variable, i.e., the indicators measuring the incidence and intensity of multidimensional poverty in each district and $\boldsymbol{X_{it}}$ represents the size of the transport infrastructure in each district, while Z_{it} represents controlling variables that can impact the incidence and intensity of multidimensional poverty. We use the controlling variable of land-use intensity, which measures the percentage of land used for cultivation purposes. The variables $W_c y_{it}$ is the variable representing the average level of multidimensional poverty in the adjacent districts, that represents the spatial spillovers that a region received from nearby regions, while $W_d X_{it}$ is variable representing the transport infrastructure in the neighboring districts. W_c and W_d the spatial weight matrices. We used two spatial weights as welfare seems to be clustered in close neighborhoods, while transport infrastructure is thought to be spatially correlated to nearby districts and the far. Therefore, we use multiple bandwidths to avoid the researcher's selection biases.

$$y_{it} = \alpha_0 + \alpha_1 W_c y_{it} + \alpha_2 X_{it} + \alpha_3 Z_{it} + \alpha_4 W_d X_{it} + \mu \qquad (4.5)$$

To estimate empirical models requires further inquiry, and it is further subject to the empirical validation of the spatial dependence of the social welfare and transport infrastructure. Therefore, as a first step, we are exploring the role of the economic geography of social welfare and transport infrastructure in Pakistan by conducting an exploratory spatial exploratory data analysis to test the spatial dependence.

7.4 Estimation Techniques

The estimation technique includes the tests for the spatial dependence and selection of appropriate estimation technique to assess the contribution of infrastructure provisioning in alleviating deprivation.

7.4.1 Detection of the Spatial dependence

The spatial dependence is a collection of sample data means that observations at location 'i' depend on other observations at locations 'j' when $i \neq j$ (Lesage, 1999, p.11). Spatial autocorrelation is synonymous to the spatial dependence in the literature of spatial econometrics. It is the degree of independent values observed in the geographic locations of the neighborhood. We are estimating Moran's I to assess the spatial autocorrelation. Moran's I statistic is based on the Pearson product-moment correlation coefficient, and geography is included via the spatial weight [W] matrix. It finds the correlation between two variables, the correlation of one variable with itself vis-à-vis a spatial weight matrix (Getis, 2010). Moran's-I focus on each observation as a difference from the mean of all observations.

We are using two leading indicators. The first indicator is the Global Moran's I, is the test for global spatial autocorrelation. The second indicator, the Local Moran's I or Local indicators of spatial association (LISA) cluster map, analyzes the clustering effect's details.

7.4.1.1 Global Moran's I (Global Spatial Autocorrelation)

The Moran's I is computed by (eq. 4.6).

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j W_{ij}.Z_i.Z_j}{\sum_i {Z_i}^2} \qquad (4.6)$$

Where I represent the Moran's statistic, $\boldsymbol{W_{ij}}$ are elements of the spatial weight matrix, $\boldsymbol{S_0} = \sum_{\boldsymbol{i}} \sum_{\boldsymbol{j}} \boldsymbol{W_{ij}}$ is the sum of all the weights, and n as the number of observations. It is a cross-product statistic between a variable and with its spatial lag, and the variable expressed in deviations from its mean. For an observation at location *i*, expressed as $Z_i = X_i - \bar{X}$ where $\bar{X}$ is the mean of a variable X . $Z_j = X_j - \bar{X}$ where $\bar{X}$ is the mean of a variable X .

Inference for Moran's I is based on a null hypothesis of spatial randomness. The statistic distribution under the null can be derived using an assumption of normality (independent normal random variates) (i.e., each value is equally likely to occur at any location).

$$p = \frac{R + 1}{M + 1}$$

$\boldsymbol{R}$ is the number of times the computed Moran's I from the spatial random data sets, and $\boldsymbol{M}$ equals the number of permutations, typically taken as 99, 999, etc., to yield a pseudo-p-value.

This software constructs the Moran Scatter Plot, and this tool of exploratory analysis is developed by Anselin (1996). When we use row-standardized weights, the sum of all the weights ($\boldsymbol{S_0} = \sum_{\boldsymbol{i}} \sum_{\boldsymbol{j}} \boldsymbol{W_{ij}}$) equals the number of observations ($\boldsymbol{n}$).

As a result, the expression for Moran's I simplify to:

$$I = \frac{\sum_i \sum_j W_{ij} . Z_i . Z_j}{\sum_i {Z_i}^2} = \frac{\sum_i (Z_i \times \sum_j W_{ij} Z_j)}{\sum_i {Z_i}^2}$$

Moran's scatter plot consists of a plot with the spatially lagged variable on the y-axis and the original variable on the x-axis. The slope of the linear fit to the scatter Plot equals Moran's I. The Plot is centered on the mean (of zero). All points to the right of the mean have $Z_i > 0$, and all points to the left have $Z_i < 0$. Visualization in the Moran scatter Plot is the classification of the *nature* of spatial autocorrelation into four categories.

Similarly, we can classify the values for the spatial lag above and below the mean as *high* and *low*. The scatter Plot is then decomposed into four quadrants. The upper-right and lower-left quadrants correspond with *positive* spatial autocorrelation (similar values at neighboring locations). We refer to them as respectively *high-high* and *low-low* spatial autocorrelation. The lower-right and upper-left quadrants correspond to *negative* spatial autocorrelation (dissimilar values at neighboring locations). They are referred to as *high-low* and or *low-high* spatial autocorrelation.

7.4.1.2 LISA Cluster Map (Local Moran's I)

The LISA Cluster maps are developed based on the pioneered work of Anselin (1995). The global spatial autocorrelation indicators are designed to reject the null hypothesis of spatial randomness in favor of an alternative of *clustering*. However, such *clustering* is a characteristic of the complete spatial pattern and does *not* indicate the *location* of the clustering. Therefore, LISA is seen as having two essential characteristics. First, it provides a statistic for each location with an assessment of significance. Second, it establishes a proportional relationship between the sum of the local statistics and a corresponding global statistic.

$$\boldsymbol{I_i} =. \frac{\sum_j \boldsymbol{W_{ij} Z_i Z_j}}{\sum_i \boldsymbol{Z_i}^2}$$

Here, the denominator is the same $\sum_i Z_i{}^2$ for every location. Therefore, we can consider it as a constant 'c'

$$\boldsymbol{I_i = c.Z_i \sum\nolimits_j W_{ij} Z_j}$$

The preferred approach for hypothesis testing is a conditional permutation method, which is similar to the permutation approach considered in the Moran scatter Plot, except that the value of each $\mathbf{Z_i}$ held fixed at its location $\mathbf{i}$. The remaining $\mathbf{n-1}$ Z-values are then randomly permuted to yield a reference distribution for the local statistic (one for each location).

7.5 Economic Geography of Transport Infrastructure in Pakistan: An Exploratory Spatial Data Analysis (ESDA)

This section presents a district-level exploratory spatial data analysis (ESDA) of transport infrastructure in Pakistan. The ESDA will help us to understand the spatial dependence of Pakistan's transport infrastructure and allow us to build empirical geography-based regression models. This analysis is based on the administrative boundaries of districts as the basic geographic unit and data of road length kilometrage as the definition of the road infrastructure. The road infrastructure is critically important in Pakistan, as it has been highly used in Pakistan compared to other types of transport (rail, air, and seaport) infrastructures. Moreover, this chapter uses the up-to-date dataset of road infrastructure that has been developed by NTRC using digital resources and on-ground assessments during 2020. This helps us to analyze the spatial dependence of road infrastructure at the national and provincial scales in Pakistan. Furthermore, this dataset is helpful for comparability at the provincial level. It is because, the mettled high type roads (HTR) data for the district of Punjab include the road length of the national highways, motorways, district roads (provincial highways, R&B Sector DR, Farm to Market Roads, Sugar-Cess Roads, and District Council Roads). Given this bifurcation based on the types, data is not available for the other provinces. The provinces of Sindh, KPK and Baluchistan used the definition of blacktop (mettled) road as the HTR, and the data for national highways and motorways is not available separately for each district. Therefore, for this chapter, we are using the road length data developed NTRC in 2020, that is a recent dataset and with ease in comparability within the nation and across nations.

For this study, we are estimating the Moran's-I's and local indicators of spatial association (LISA) cluster map to measure spatial autocorrelation to detect the spatial dependence of road infrastructure in Pakistan. The Moran's I is also known as a measure to estimate global autocorrelation, while LISA is also known as local spatial autocorrelation. It is called 'global' as Moran's I scatter plot illustrates the relationship between the values of the road infrastructure at each district and with the average value of the road infrastructure in the neighboring districts. While LISA provides the analysis of local effects especially the clustering effect detection. LISA analyses where road infrastructure is strongly positively or negatively associated with one another. The significance of the Moran's I is based on the significance test results for the null hypothesis of the spatial randomness of road infrastructure against the alternative of the spatial dependence. The hypothesis testing is conducted using 999 random computations. In addition to this, the literature signifies the importance of road infrastructure for spatial network externalities; therefore, the distance weight matrix (Euclidian distance) is a better definition for the construction of spatial weights.

We are comparting into five sub-sections. The first subsection provides an ESDA using dataset of all the districts in Pakistan at the national level of Pakistan, and the following subsections provide the districts level ESDA for each province of Punjab, Sindh, KPK and Baluchistan, respectively.

7.5.1 ESDA of Road Infrastructure in Pakistan: A districts level analysis

We conducted ESDA for the 109 districts of Pakistan using DWM of two bandwidths of 3 and 5. The connectivity graph and maps is provided in Figure 7.2.

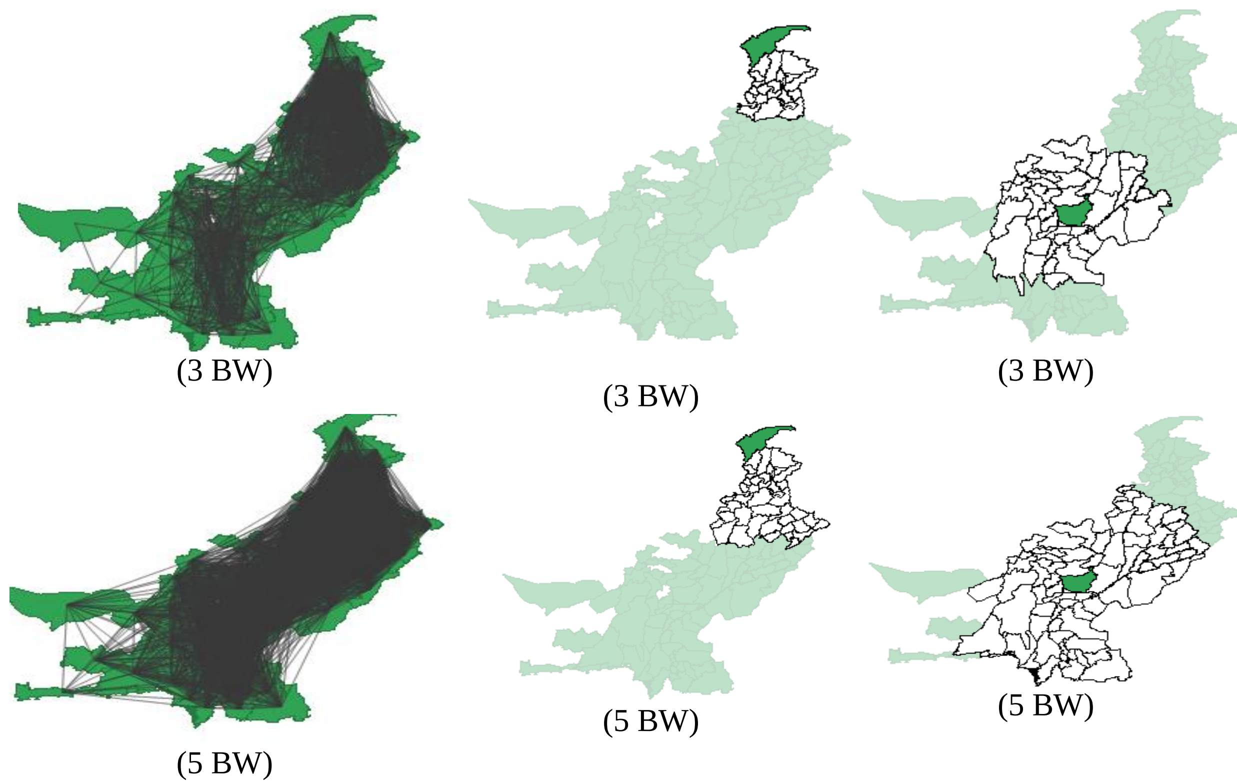
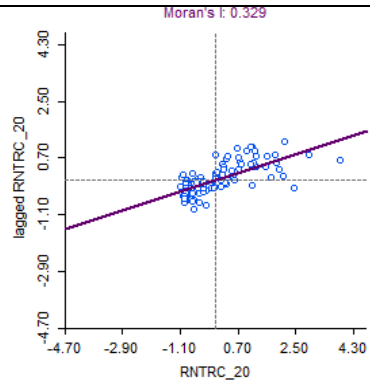
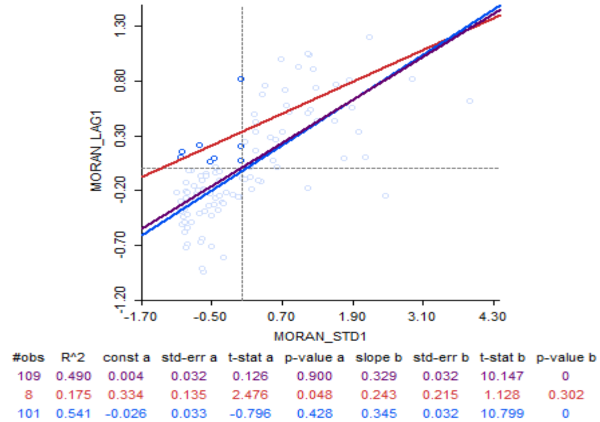


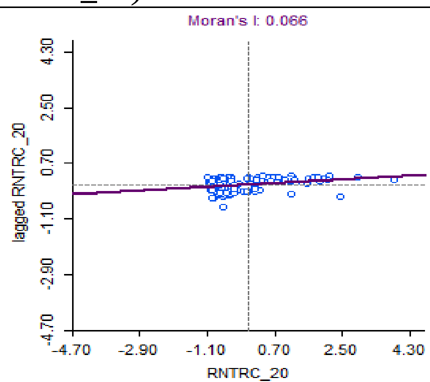
Figure 7. 2 **Connectivity Graphs and Maps: Districts of Pakistan**

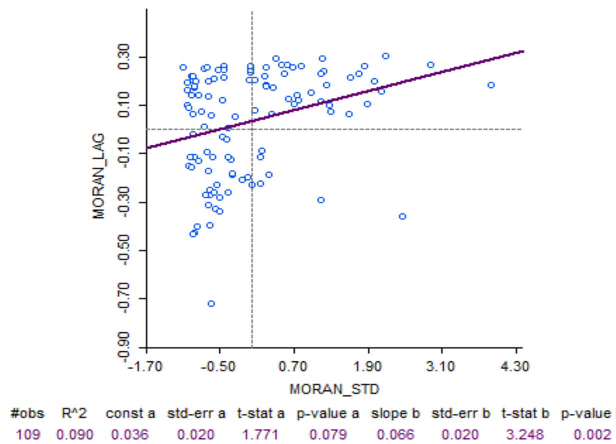
We are estimated two indicators for the detection of spatial dependence. First, the global Moran's I and second, is the LISA Cluster analysis, also known as 'local' spatial autocorrelation. We had estimated Moran's I and developed LISA cluster Map using IDWM of 3 BW and 5 BW, and results are provided in Figure 7.3 & 7.4, respectively.





I: 0.329; Pseudo P-value: 0.001 (IDWM_3B)





I: 0.06; Pseudo P-value: 0.001; (IDWM_5B)

Figure 7. 3 **Global Moran's I and Moran Scatter Plot: Road infrastructure in Pakistan**

The results signify the existence of spatial dependence and positive spatial autocorrelation (highly significant) of road infrastructure in Pakistan. The Moran's I scatter Plot elucidates that the road infrastructure is not uniformly distributed across districts but lies in the quadrants of high-high, low-high, and high-low regions. It seems a strong clustering effect exists, and therefore, the Local Moran's I, the LISA, provides a detailed analysis. The LISA cluster map depicts an extreme polarization and clustering of road infrastructure in Pakistan. The LISA cluster using SMW of 3 BW depicts 30 districts spatially clustered as high-value and 38 districts clustering as low values of road infrastructure compared to average with other districts of Pakistan. On the other using 5 BW, LISA cluster map highlights 21 districts with the least road infrastructure, and 32 districts with more extensive road networks are geographically clubbed as high-value clusters. These polarized clusters are statistically significant at 5% significance using 999 random computations.

The low-value clusters lie in Baluchistan, Sindh, and KPK, while the high-value cluster lies in the province of Punjab except the Nowshera, a district of KPK. More details are in table 7.1.

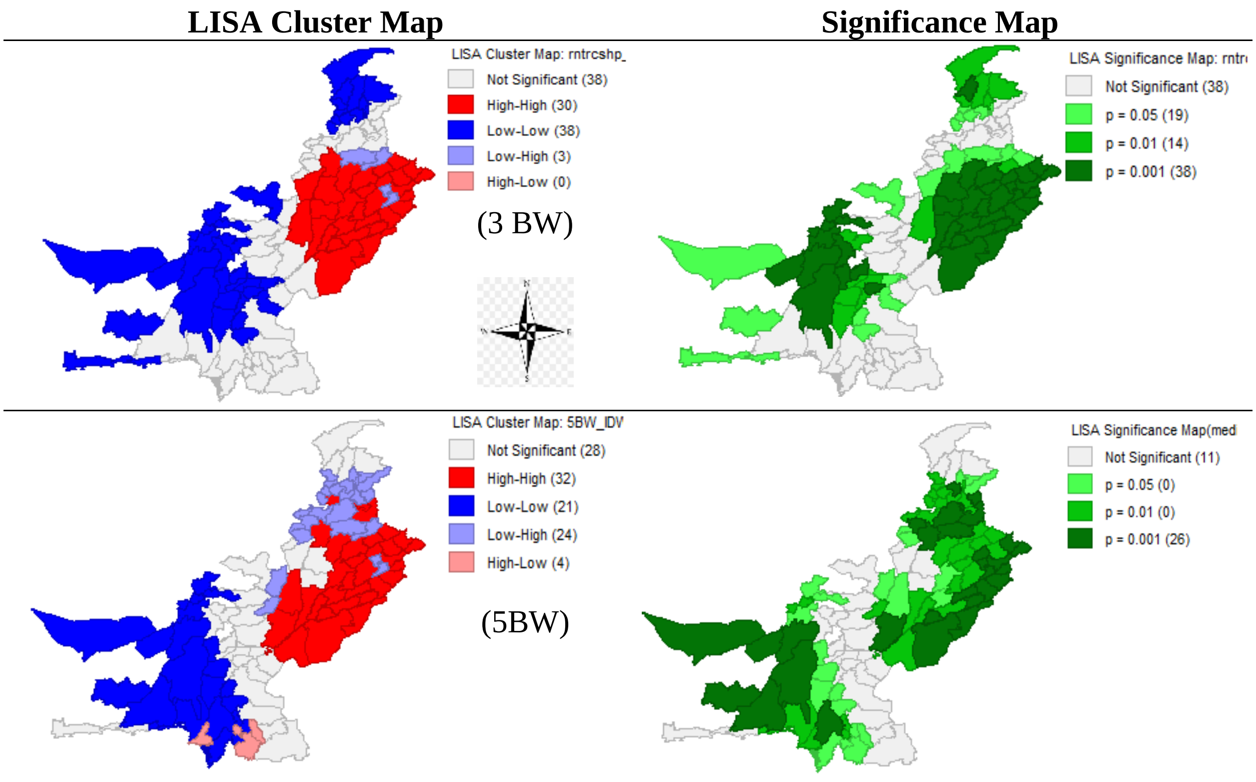


Figure 7. 4 **LISA Cluster Map: Road Infrastructure in Pakistan**

Table 7. 1 *Highly significant Spatial clusters of Road infrastructure in Pakistan*

IDWM (3 BW)	
Low-Low	**High-High**
KPK Buner, Charsada, Chitral, Kohistan, Lower Dir, Malakand PA, Mardan, Peshawar, Pishin, Shangla, Swat, Tor ghar, Upper Dir. **Sindh** Dadu, Jacobabad, Larkana, Naushero Feroze, Qambar Shahdatkot, SB Abad, Shikarpur, Sukkur, **Balochistan** Chagai, Gwader, Jaffarabad, Jhal Magsi, Kacchi, Kalat, Kashmore, Kharanm, Khuzdar, Mastung, Nasirabad, Panjgur, Quetta, Sibi, Sohbatpur, Zhob and Ziarat.	**Punjab** Bahawalnagar, Bahwalpur, Bhakkar, Chiniot, DG Khan, DI Khan, Faisalabad, Gujranwala, Gujrat, Hafizabad, Jhang, Kasur, Khanewal, Khusab, Lahore, Layyah, Lodhran, Mandi Bahauddin, Mianwali, Multan, Muzaffargarh, Narowal, Okara, Pakpattan, Sahiwal, Sarghoda, Sheikhupura, Sialkot, Toba tek Singh, Vehari,

IDWM (5 BW)	
Low-Low	**High-High**
Balochistan Awaran, Chagai, Jhal Magsi, Kalat, Kharan, Khuzdar, Lasbela, Mastung, Panjgur, Pishin, Qambar Shahdadkot, Quetta, Ziarat, **Sindh** Dado, Jamshoro, Matiari, Naushehro Feroze, SB Abad, Tando Allah Yar, Tando Muhammad Khan, Thatta.	**Punjab** Bahawalnagr, Bhalwalpur, Chiniot, Dera Ghazi Khan, Faisalabad, Gujranwala, Gujrat, Hafizabad, Islamabad, Jhang, Kasur, Khanewal, Khushab, Lahore, Lodhran, Mandi Bahauddin, Mianwali, Multan, Muzaffargarh, Narowal, Okara, Pakpattan, Rahm yar Khan, Rajanpur, Rawalpindi, Sahiwal, Sargodha, Sheikhpura, Sialkot, Toba Tek Sings, Vehari **KPK** Nowshera

7.5.2 ESDA of Road Infrastructure: A Districts Level Analysis of Punjab, Pakistan

We analysed the spatial dependence of road infrastructure within Punjab using road length data of 36 districts. The data of federal capital district Islamabad is included given the geographical adjacency. For analysing this spatial dependency, we developed the **W** using two bandwidths of 2 and 5. The connectivity graph and maps are provided in Fig 7.5 for further details. Moran's I and LISA cluster maps are provided in Figure 7.6 & 7.7, respectively.

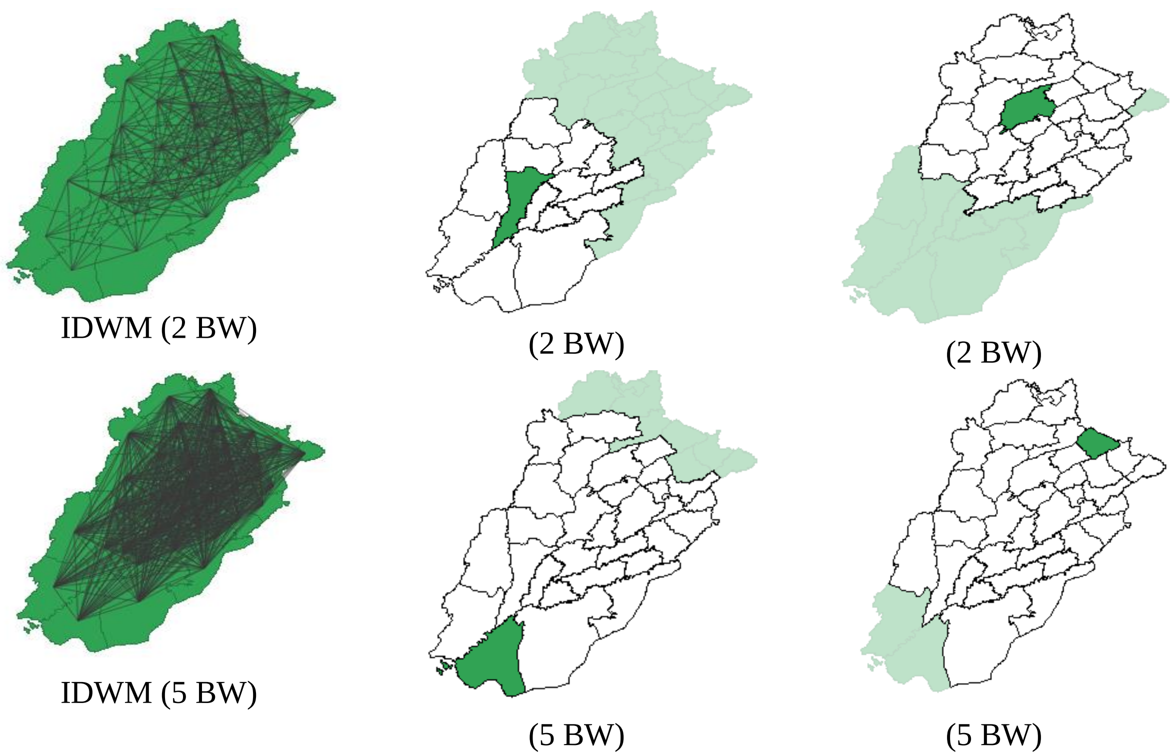
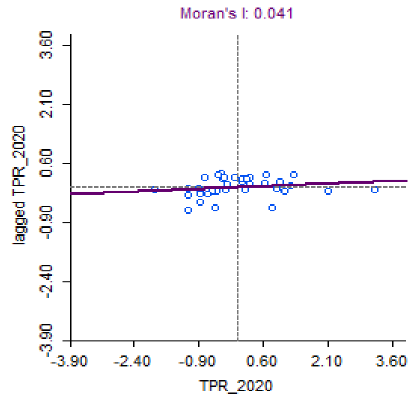
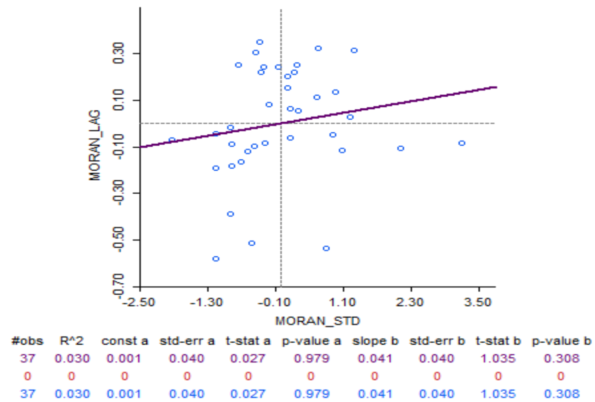


Figure 7. 5 **Connectivity Graphs and Maps: District of Punjab, Pakistan**

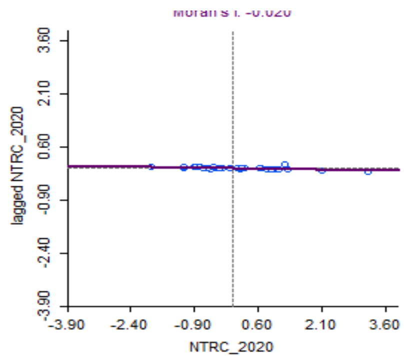
Using W with 2 BW and 5 BW, estimated Moran's I are 0.041 and 0.020, not statistically significant at 5 percent (using randomization option of 999) that depicts weak spatial dependence of road infrastructure in Punjab but not highly polarized. The LISA Cluster map analysis depicts districts in four colours, and red-coloured districts depict the high-high clusters. It means the districts with high-road infrastructure are neighbouring with high road infrastructure. At the same time, Blue coloured district depicts the statistically significant clusters of low values. It means districts neighbouring with a low-road network on average. The LISA depicts one significant cluster as high values for the district of Sahiwal and two low-value significant clusters of Mianwali and Attock.



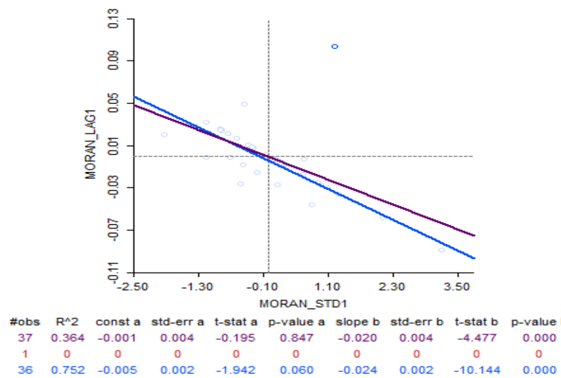
Pseudo p-values ranges (0.065 to 0.077):
IDWM (2BW)



IDWM (2BW)



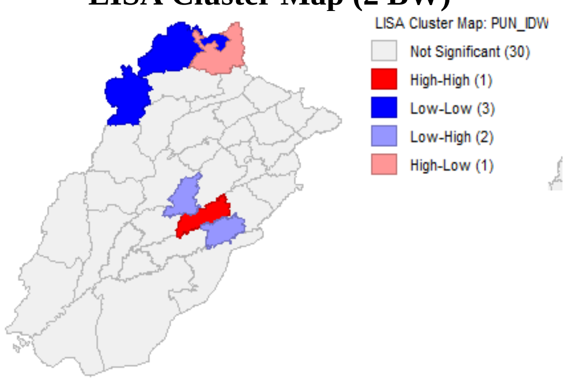
(Pseudo p-values ranges 0.040 to 0.070)
(IDWM 5BW)



(IDWM 5BW)

Figure 7. 6 **Global Moran's I: Road Infrastructure in Districts of Punjab and ICT**

LISA Cluster Map (2 BW)



Significance Map

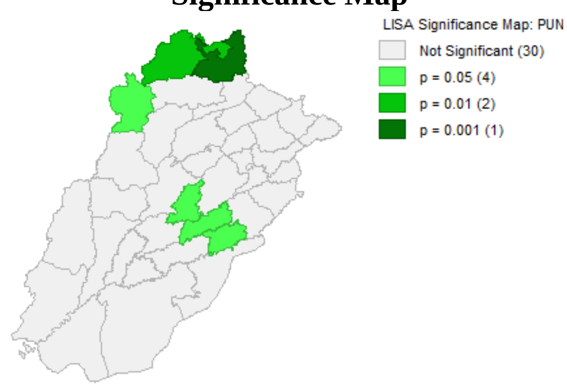


Figure 7. 7 **Local Moran's I: Road Infrastructure in Districts of Punjab and ICT**

7.5.3 ESDA of Road Infrastructure: A districts level analysis of Sindh, Pakistan

We analyzed the spatial dependence of road infrastructure within the 23 districts of Sindh province. We constructed spatial W of 2 and 5 BW. The connectivity graphs and maps are provided as Figure 7.8, while the results of Moran's I is given in Figure 7.9 & 7.10., respectively.

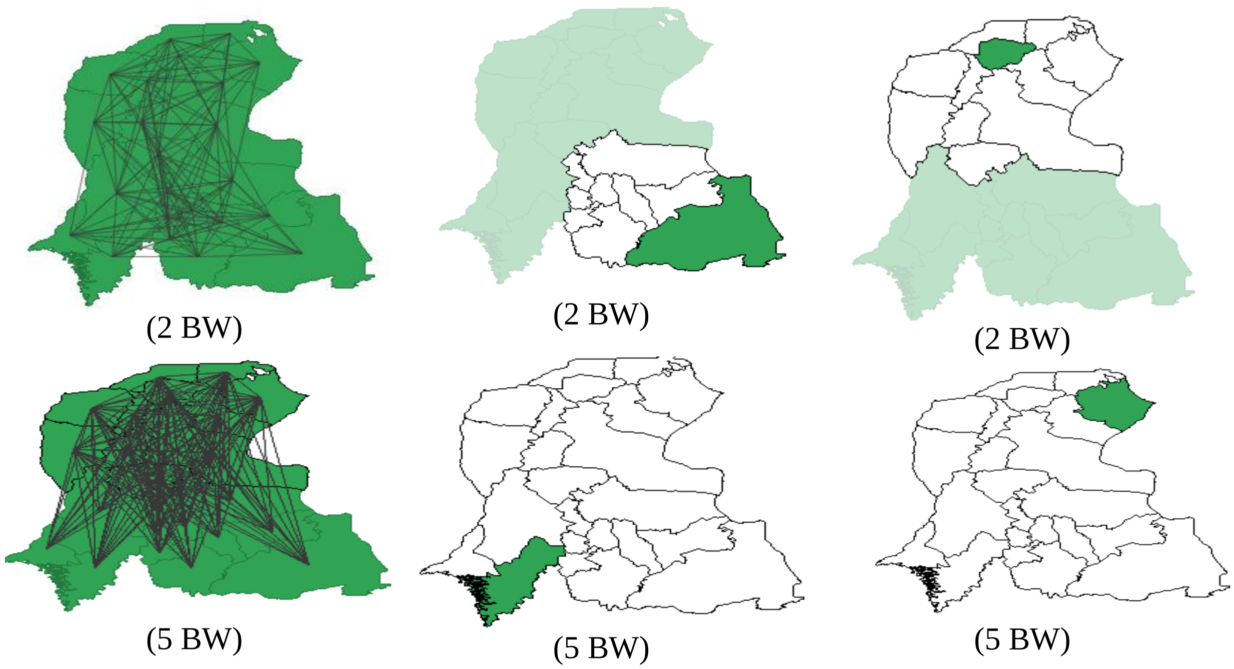
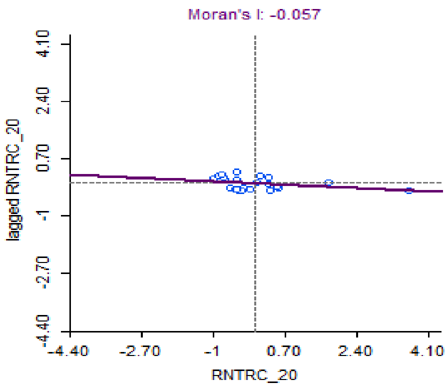
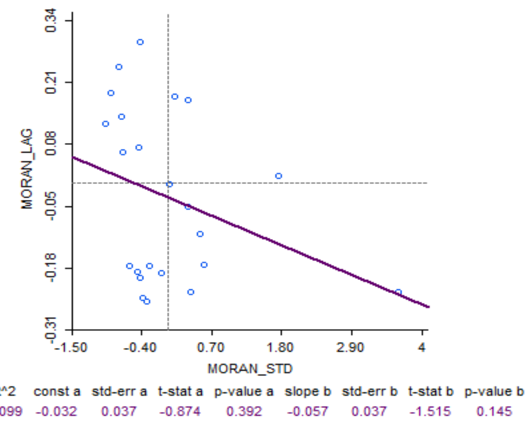


Figure 7. 8 **Connectivity Graphs and Maps: Districts of Sindh, Pakistan**

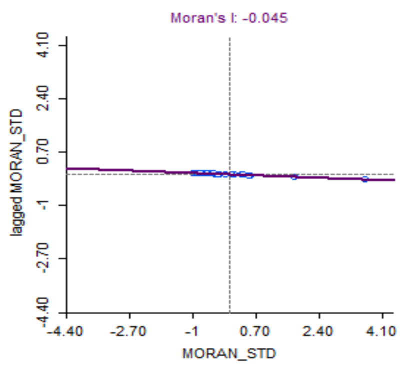
We estimated global Moran's I using W at BW 2 and 5 and concluded for insignificant positive spatial autocorrelation of road infrastructure within the province of Sindh (as Moran's I is not statistically significant at 5% using random computations of 999). The district with high-road networks are neighboring with districts with low (on avergae) road network, and its vice versa. Therefore, LISA cluster map is depicted no significant clustering effects using W of 2bw and 5bw.



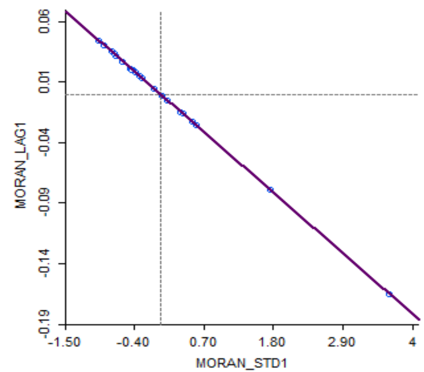
Pseudo p-values ranges (0.464 to 0.477):
DWM (2BW)



DWM (2BW)



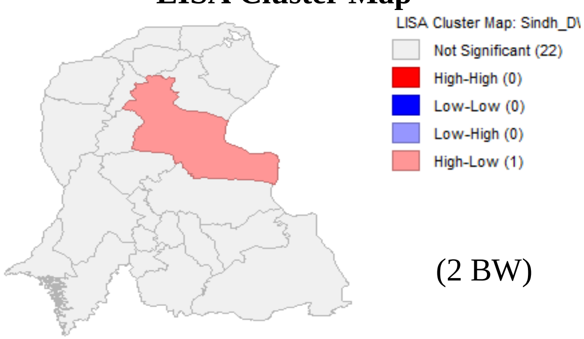
Pseudo p-values ranges (0.060 to 0.076)
(DWM 5BW)



(DWM 5BW)

Figure 7. 9 **Global Moran's I: Road Infrastructure in Districts of Sindh, Pakistan**

LISA Cluster Map



(2 BW)

Significance Map

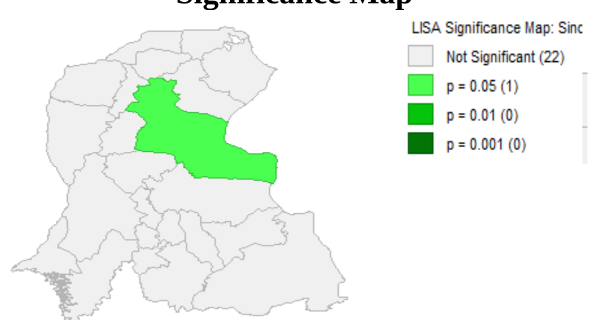
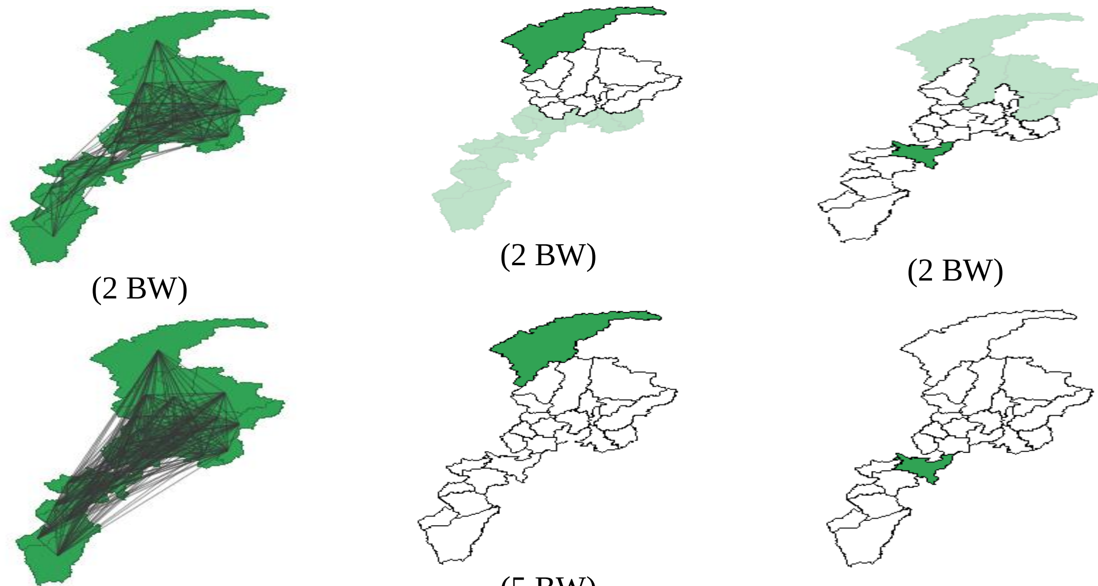


Figure 7. 10 **Local Moran's I: Road Infrastructure in Districts of Sindh**

7.5.4 ESDA of Road Infrastructure: A district-level analysis of KPK, Pakistan

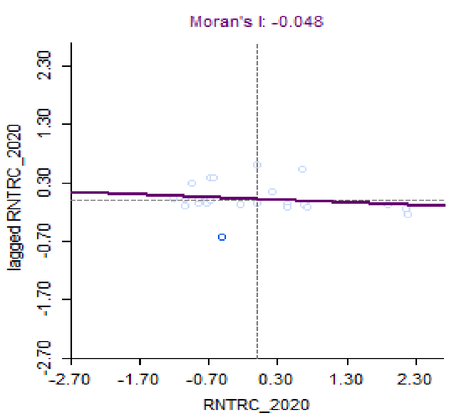
The test for detecting the spatial dependence of the road infrastructure has been conducted for the 25 districts of KPK. We developed W using 2 BW and 5 BW. The connectivity map and graphs are provided as figures (7.11). In addition, the estimated global Moran's and local Moran's I (LISA cluster map) to determine the spatial dependence are provided as Figure 7.12 & 7.13, respectively.



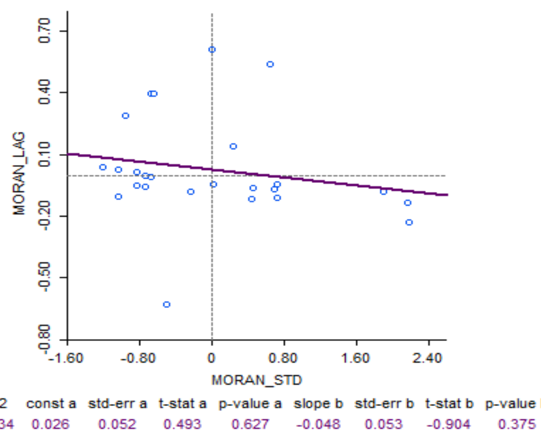
(2 BW) (2 BW) (2 BW)

(5 BW) (5 BW) (5 BW)

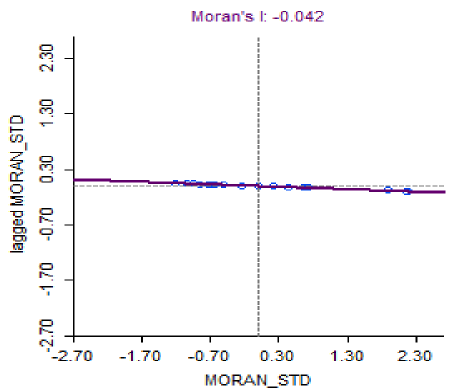
Figure 7. 11 **Connectivity Graphs and Maps: Districts of KPK, Pakistan**



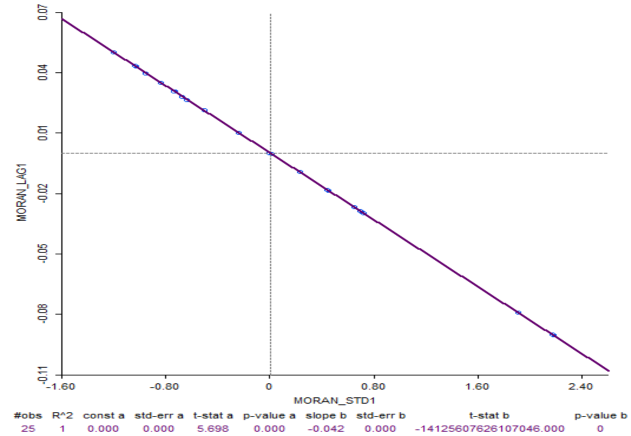
I: -0.048, Pseudo p-values (0.448 to 0.494)
IDWM (2BW)



IDWM (2BW)



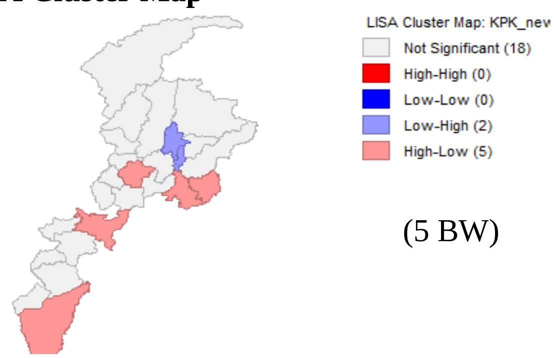
Pseudo p-values ranges (0.114 to 0.136)
(IDWM 5BW)



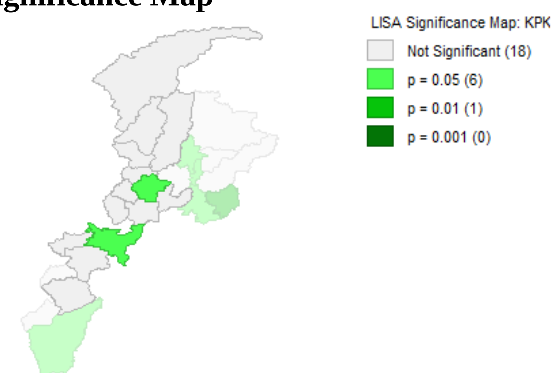
(IDWM 5BW)

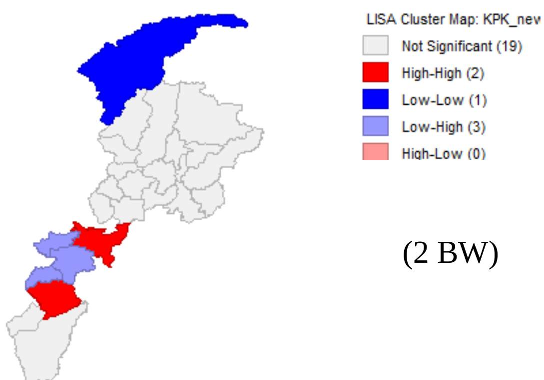
Figure 7. 12 **Global Moran's I: Road Infrastructure in Districts of KPK and ICT**

LISA Cluster Map **Significance Map**



(5 BW)





(2 BW)

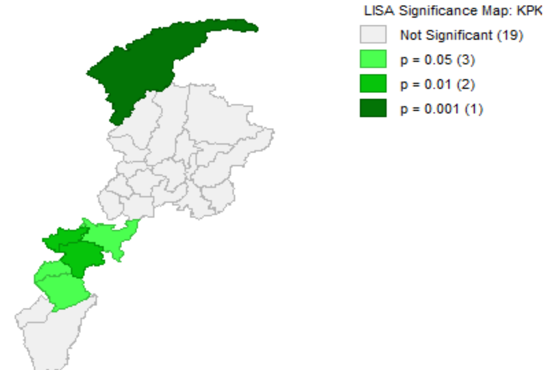
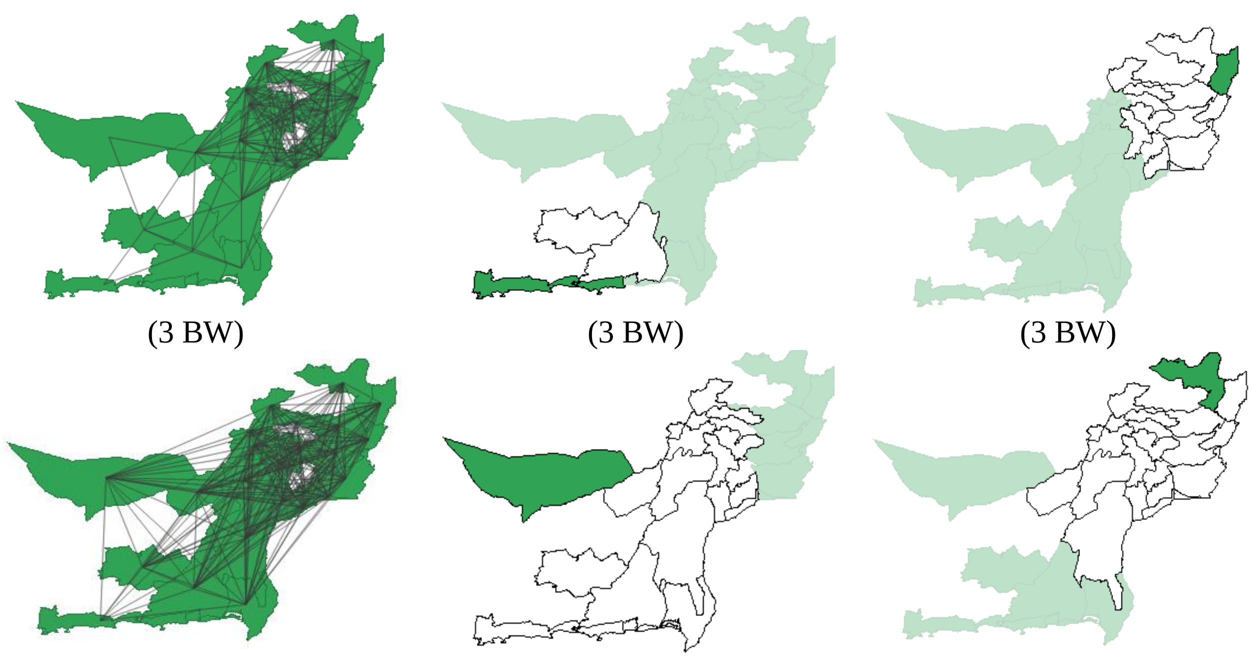


Figure 7. 13 **Local Moran's I: Road Infrastructure in Districts of KPK and ICT**

The estimated Moran's-I is -0.048 and -0.042, and Moran's graphs lie in the upper left (low-High) and lower right (High-low) quadrants depicting a weak clustering effect. The randomized permutation also signals an insignificant Moran's I, so we can accept the null hypothesis and conclude that the spatial randomness of road infrastructure in the districts of KPK. The LISA cluster map at 5BW shows two significant spatial outliers districts but doesn't find any significant spatial clusters. While using 2 BW, one significant spatial cluster of low-value and two significant spatial clusters of high value exist. The district of Chitral lies in the low-value cluster (on average), and the districts of Lakki-Mawat and Kohat are districts of high-values spatial cluster road infrastructure compared to their neighborhood (on average).

7.5.5 ESDA of road infrastructure: A districts level analysis of Baluchistan, Pakistan

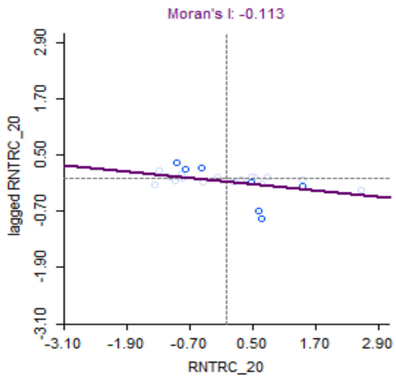
To conduct ESDA of road infrastructure to inquire about the spatial dependency of road infrastructure, data of 24 districts of Baluchistan, Pakistan. We developed spatial W using 3 bw, and 5bw. The connectivity graphs and maps are provided in Figure 7.14. The global and local Moran's I result are provided in Figure 7.15 and 7.16, respectively. The analysis revealed an insignificant positive spatial autocorrelation of road infrastructure within the districts of Baluchistan using 999 randomized computations. The LISA cluster map also signifies insignificant local spatial autocorrelation at 3 BW.

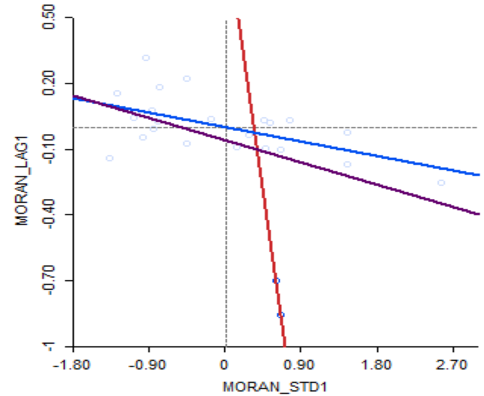


(3 BW) (3 BW) (3 BW)

(5 BW) (5BW) (5BW)

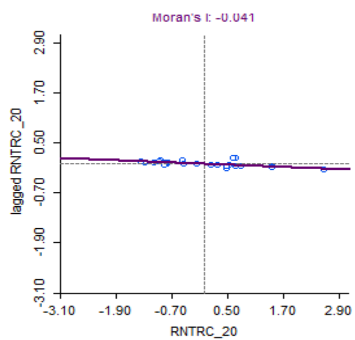
Figure 7. 14 **Connectivity Graphs and Maps: Districts of Baluchistan, Pakistan**

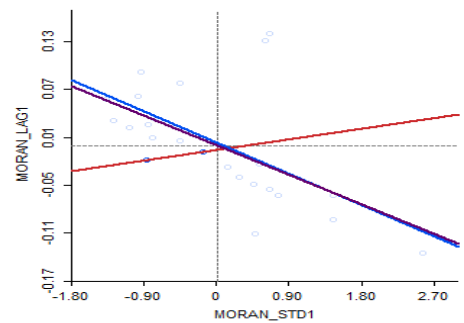




#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
24	0.195	-0.060	0.048	-1.252	0.224	-0.113	0.049	-2.307	0.031
2	1	0.931	0	0	0	-2.711	0	0	0
22	0.314	0.001	0.024	0.043	0.966	-0.073	0.024	-3.027	0.007

Pseudo p-values ranges (0.110 to 0.137)
DWM (3BW)



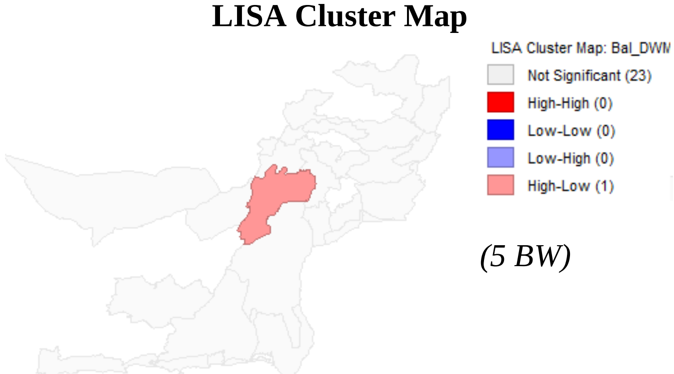


#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
24	0.333	0.000	0.012	0.018	0.986	-0.041	0.012	-3.316	0.003
2	1	-0.006	0	0	0	0.015	0	0	0
22	0.361	0.004	0.013	0.269	0.790	-0.044	0.013	-3.363	0.003

Pseudo p-values ranges (0.346 to 0.434)
(DWM 5BW)

Figure 7. 15 **Global Moran's I: Road Infrastructure in Districts of Baluchistan**

LISA Cluster Map **Significance Map**



(5 BW)

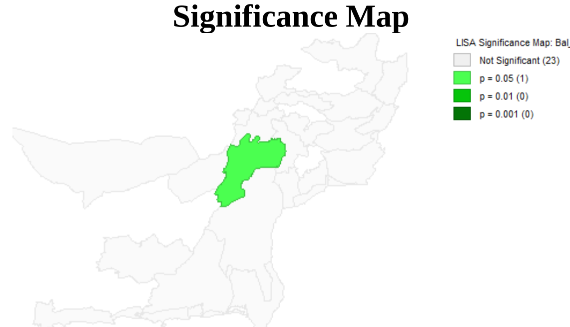


Figure 7. 16 **Local Moran's I: Road infrastructure in districts of Baluchistan**

7.5.6 A Summary: ESDA of road infrastructure

This exercise highlighted the presence of the spatial clusters of road infrastructure, depicting provisioning of road infrastructure is not unform across districts Pakistan. On the national level analysis, we can infer that the road infrastructure is not randomly distributed across the geographic space of Pakistan but spatially clustered. However, within each province, the distribution of road infrastructure is not clustered significantly.

The main results are;

- On the national scale, road infrastructure is significantly spatially dependent within 109 districts of Pakistan. ESDA revealed a significant positive spatial autocorrelation reveals the road infrastructure is clustered (low-low and high-high). These low-value clusters are located in the Baluchistan and North-East districts of KPK, while the high-value cluster lies in the Punjab and North-west districts of KPK.
- At the regional scale of Punjab, the ESDA of the road infrastructure revealed an insignificant positive spatial autocorrelation. At the same time, the LISA cluster map depicts the districts of Sahiwal as the significant low-low cluster and the Mianwali and Attock as the significant high-high clusters.
- At the regional scale of Sindh, the ESDA of road infrastructure depicts an insignificant positive spatial autocorrelation. Still, data reveals significant negative spatial autocorrelation that doesn’t depict any clustering effect of road infrastructure within districts of Sindh.
- At the regional scale of KPK, ESDA of road infrastructure signaled insignificant positive spatial autocorrelation at the regional scale. The LISA analysis depicts that the district Chitral has lower road infrastructure than the road infrastructure (on average) in its neighbors. In contrast, the districts of Lakki-Mawat and Kohat are located as high-value spatial clusters relatively.
- At the regional scale of Baluchistan, the ESDA of road infrastructure revealed an insignificant positive spatial autocorrelation and LISA clusters.

7.6 The Economic Geography of Social Welfare: An Exploratory Spatial Data Analysis (ESDA)

This section focuses to answers the research question, 'Whether location matters for the social welfare in Pakistan'? For empirical inquiry, we measure social welfare in terms of multidimensional poverty (MHCI and MPI). And we are estimating the indicators of global and local Moran's I. We conducted ESDA to measure the degree of the spatial dependence of social welfare. Therefore, we constructed a queen contiguity weight matrix of order 1 (Wc) that defines neighborhoods based on the nearby locations. The results of ESDA at the national level for all districts of Pakistan are presented following the ESDA for every four provinces, Punjab, Sindh, KPK, and Baluchistan, respectively.

7.6.1 An ESDA of Multidimensional Poverty Index: A District-level Analysis for Pakistan

We developed Wc for 110 districts of Pakistan, the connectivity graph and maps are provided in Figure 7.17 We estimated global Moran's I, which is an indicator widely used to assess the spatial dependence, and local Moran's I (LISA cluster map) using Wc. The results revealed significant positive spatial Autocorrelation, using 999 random computations. The LISA cluster maps highlighted the inequitable landscape of social welfare (multidimensional poverty) in Pakistan. There are 19 districts (high-high clusters) facing greater socio-economic deprivation (see LISA cluster map for MPI 2014/15) compared to other districts on average. These districts are mostly located in Kharan/Washuk, Awaran, Kech, Jhal-Magsi, Jaffarabad/Sohbatpur, Kohlu, Loralai, Sibi/Harnai/Lehri, MusaKhail, Rajanpur, Dera-Bhughti, Killa-Saifullah, Nasirabad, Pishin, Tharparkar, Badin, Mirpurkhas, Umerkot, and Kohistan. Out of these 19 districts, 14 districts are located in the province of Baluchistan. There are 27 districts on average, has a lower level of MPI compared to other districts on average. These districts are Abbottabad, Attock, Chakwal, Charsada, Chiniot, Faisalabad, Haripur, Islamabad, Gujranwala, Gujrat, Hafizabad, Jhang, Jhelum, Kasur, Khanewal, Khusab, Lahore, Mandi Bhawindin, Nankana Sahib, Nowshera, Okara, Rawalpindi, Sahiwal, Sargodha, Sheikhupura, Swabi, and Toba Tek Singh. The districts with lower multidimensional poverty are located in the province of Punjab, ICT Northern KPK. However, these districts have a greater level of poverty compared to other districts in Pakistan, on average.

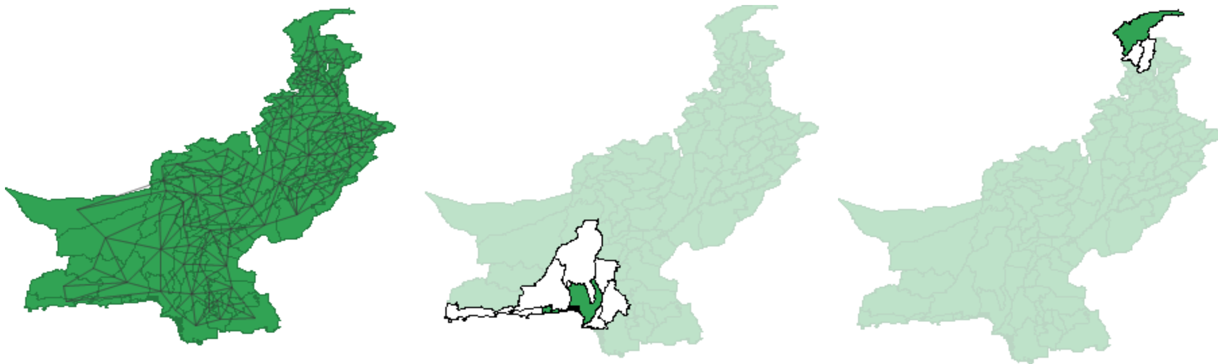
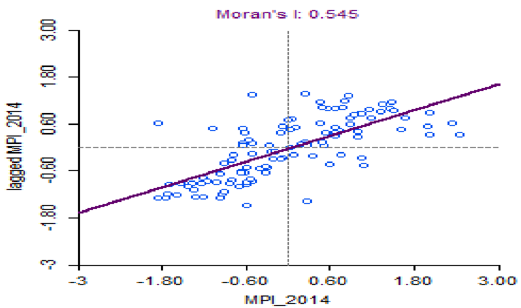
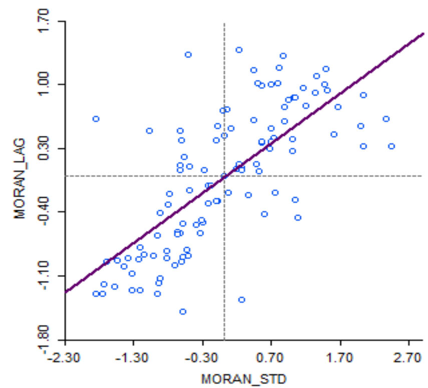


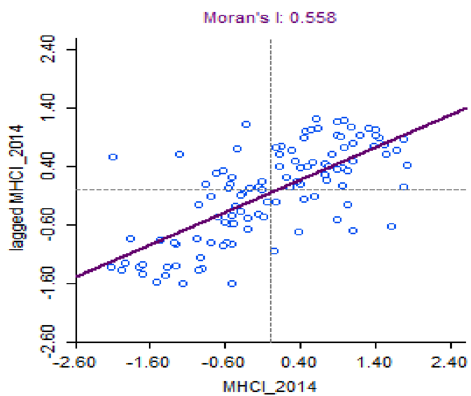
Figure 7. 17 **The connectivity Graphs and Maps**

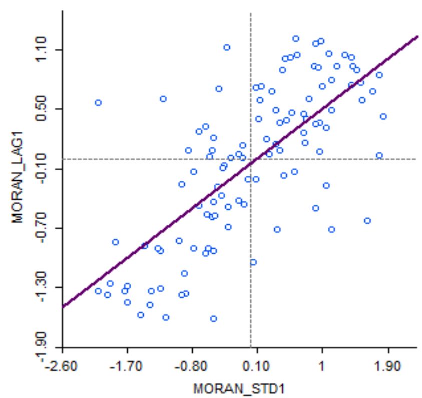


(I: 0.545: Pseudo P-value: 0.001)
MPI (2014/15)



#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
110	0.503	-0.024	0.052	-0.463	0.644	0.545	0.052	10.450	0



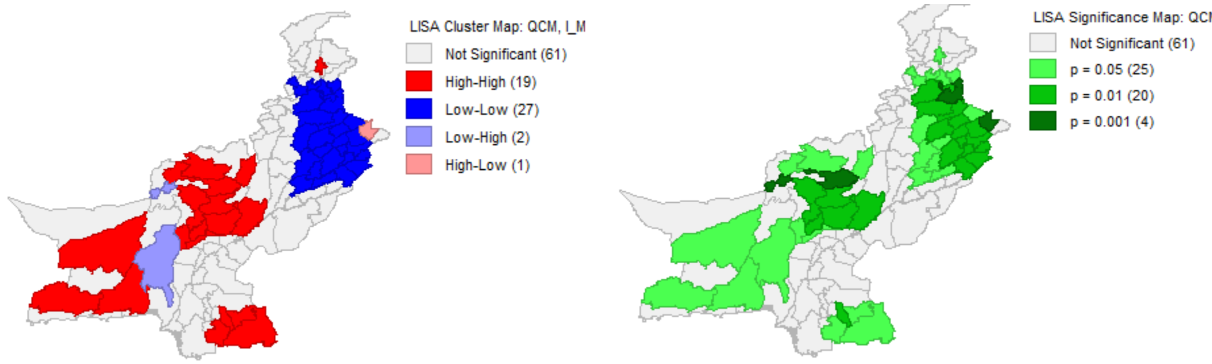


#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
110	0.494	-0.049	0.054	-0.905	0.368	0.558	0.054	10.277	0

(I:0.558, Pseudo P-value: 0.001:
MHCI: 2014/15

Figure 7. 18 **Moran's I Graph and Scatter Plot**

MPI (2014/15)



MHCI (2014/15)

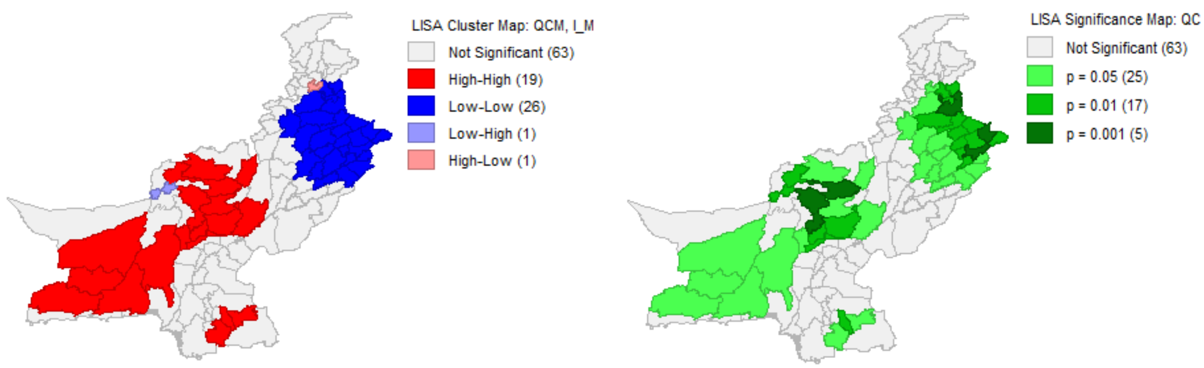


Figure 7. 19 **LISA cluster maps**

7.6.2 An ESDA of Multidimensional Poverty: A District-level Analysis for the Punjab

We estimated the queen contiguity spatial weight matrix (Wc) by editing the district-level shapefiles of Pakistan in Arcmap software. We developed Wc using the software GeoDa and developed the Moran's I for indicators of MPI and MCHI for the year 2014/15. The connectivity graphs and maps are provided in Figure 7.20. While the Moran's I graph, and scatter plot is in Figure 7.21 & 7.22.

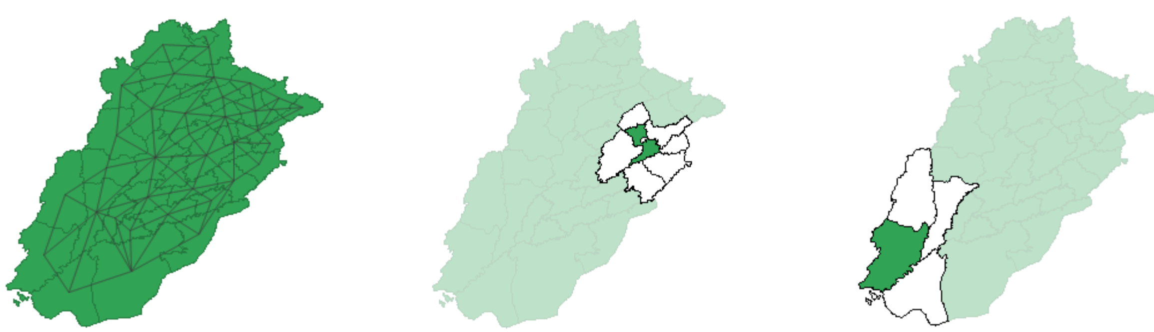
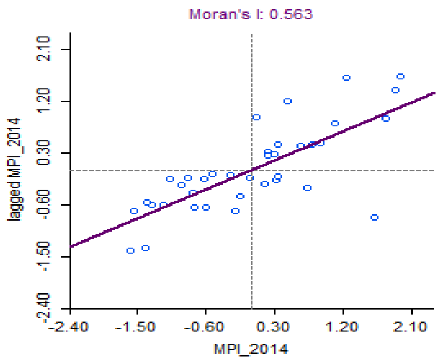
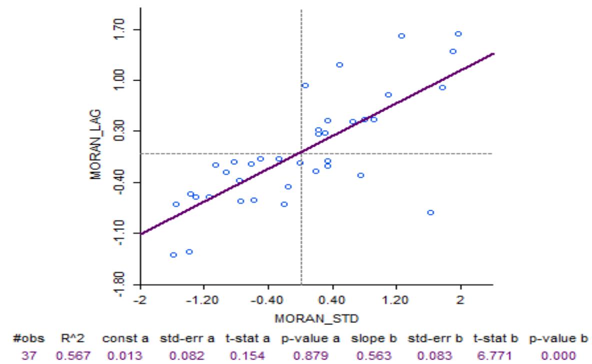


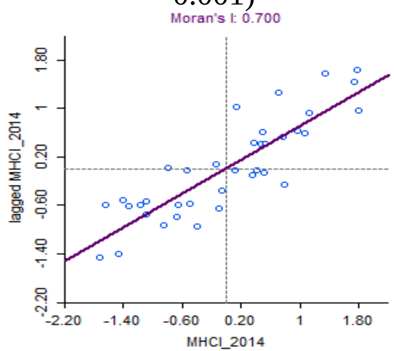
Figure 7. 20 **Connectivity Graph and Maps using QWM: Districts of Punjab**

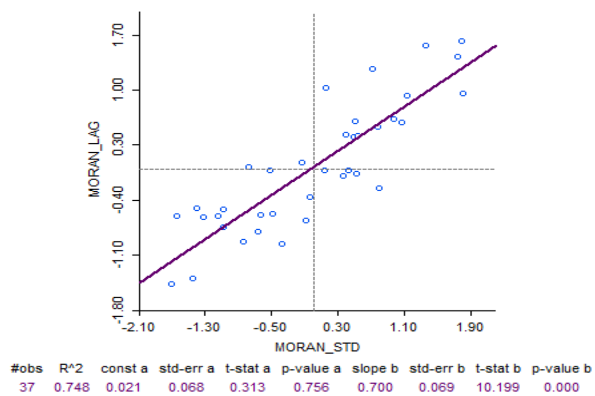
The Moran I is 0.583 for MPI and 0.700 for MHCI, which are highly significant using 999 random computations, depicting a positive spatial autocorrelation and spatial dependence. Moreover, the results of local Moran's I for MPI and MHCI as provided in Figures given below that graphically elaborate the statistically significant clusters of poor and rich districts. The high-high districts have a higher level of poverty, which is largely clustered in the districts of Southern Punjab. And the district with low multidimensional poverty levels is clustering in the northern part of Punjab.





MPI (2014/15) (I: 0.563, Pseudo p-value: 0.001)

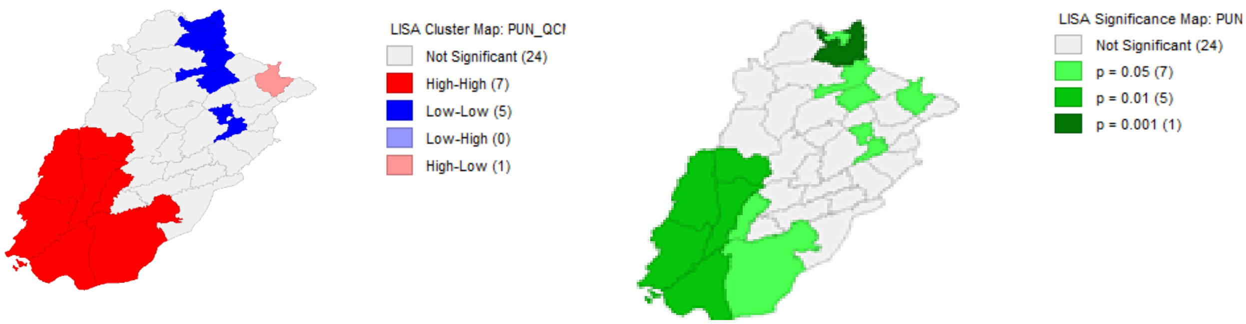




MHCI 2014/15 (I: 0.700, Pseudo p-value: 0.001)

Figure 7. 21 **Moran's I and Scattered Plot: Districts of Punjab Pakistan**

MPI 2014/15



MHCI 2014/15

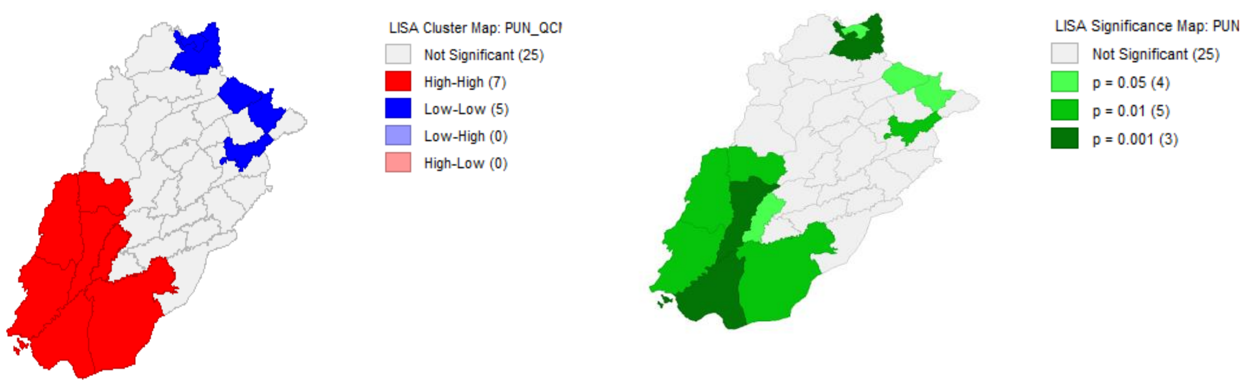


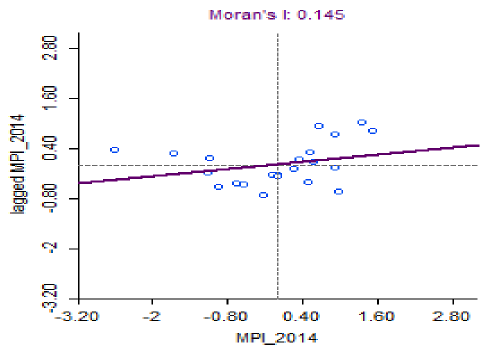
Figure 7. 22 **LISA Cluster Map: Districts of Punjab**

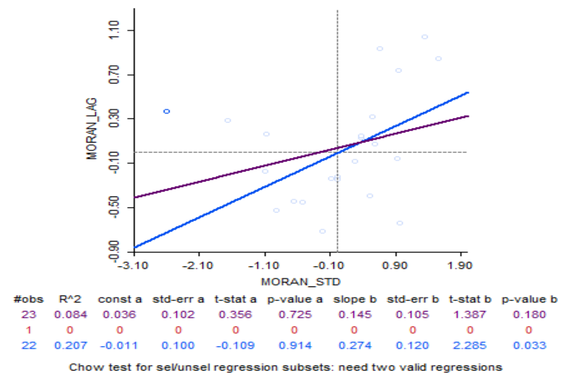
7.6.3 Economic geography of social welfare in the districts of Sindh, Pakistan

To carry out an ESDA for the districts of Sindh the spatial Wd using 2 BW and 5 BW had developed to estimate and visualize the global and local Moran's I and the Moran's scattered plot for indicators of MPI (2014/15) and MHCI (2014/15).

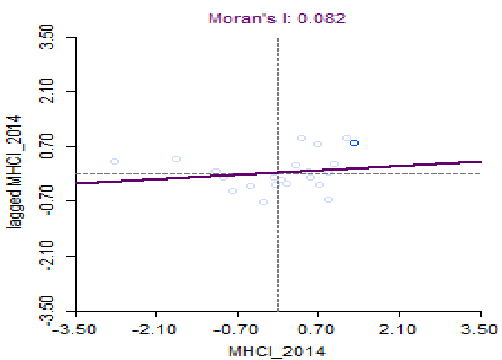
Figure 7. 23 **Connectivity Graph and map: districts of Sindh**

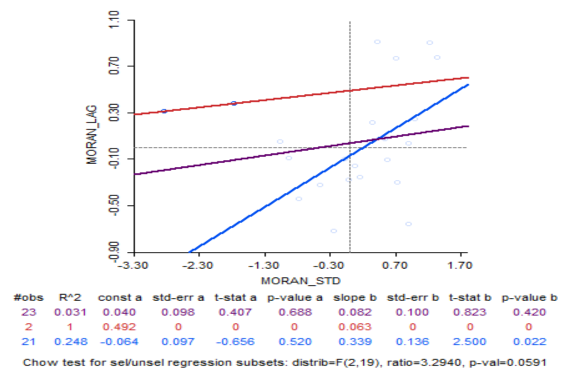
Using [**Wc]** we tested for the null hypothesis of spatial randomness against the alternative of spatial dependence of indicators. The results are provided in Figure given below.





(I: 0.145, pseudo p-value:0.05 to 0.08)





(I: 0.082; Pseudo P-value: 0.146-0.187)

Figure 7. 24 **Global Moran's I: MPI and MHCI (2014/15) districts of Sindh**

The results for the indicators of MPI (2014/15) and MHCI (2014/15) reveal a positive spatial autocorrelation in the districts of Sindh province except for the Karachi district. The Moran's I is weakly significant but the slope of Moran's I is significant depicting high-high and low-low spatial autocorrelation in the districts of that depict similar values at neighboring locations, except Karachi district. The LISA cluster map also depicts similar significant clusters of high-high and low-low. The (red) color symbolizes the districts of significant poorest clusters when districts within the province of Sindh are compared.

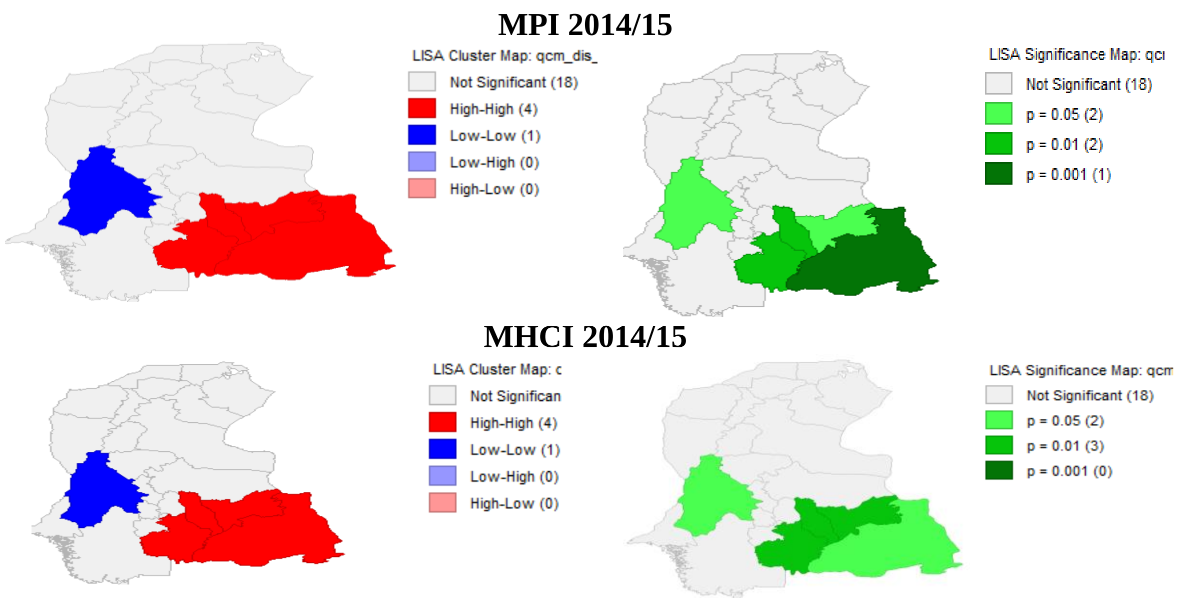


Figure 7. 25 **LISA Cluster Map: Districts of Sindh**

7.6.4 Economic geography of social welfare in the districts of KPK, Pakistan

We conducted an ESDA using the districts data for MPI and MHCI using the spatial weights of Wc order 1. the connectivity graph, global Moran's I and LISA cluster map is given in figures provided below.

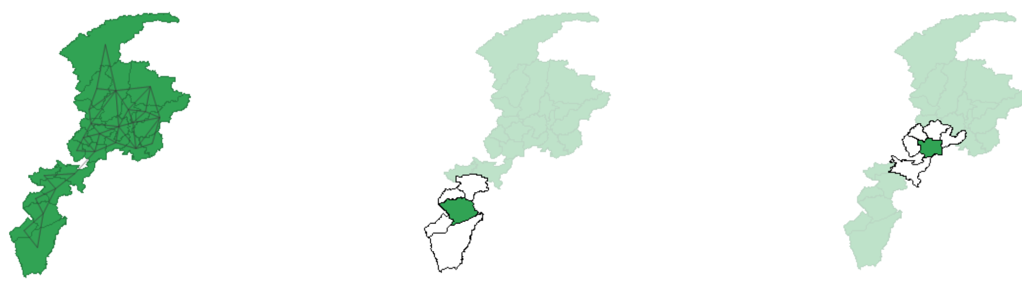
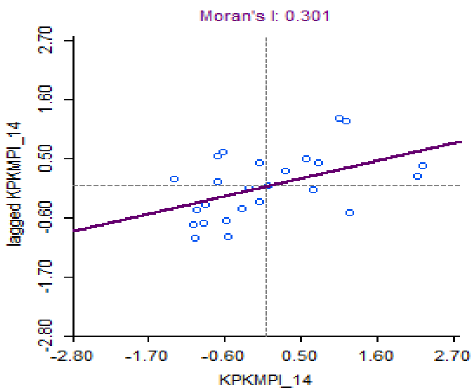
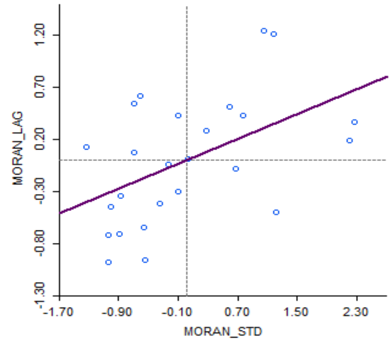


Figure 7. 26 **Connectivity Graphs Districts of KPK, Pakistan**

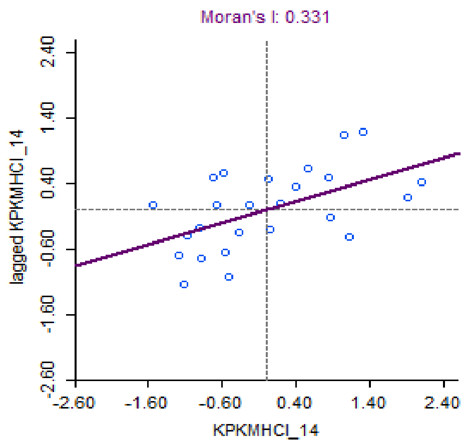


(I:301, pseudo p-values: 0.011 to 0.022)

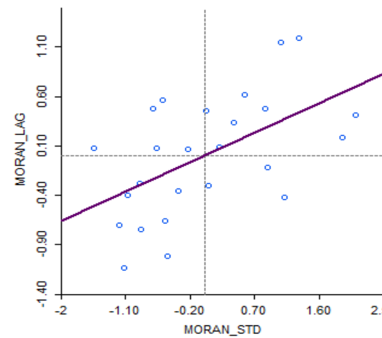


#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
25	0.248	-0.004	0.107	-0.036	0.972	0.301	0.109	2.753	0.011
0	0	0	0	0	0	0	0	0	0
25	0.248	-0.004	0.107	-0.036	0.972	0.301	0.109	2.753	0.011

Chow test for sel/unsel regression subsets: can't compute



(I:331, pseudo p-values: 0.008 to 0.011)



#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
25	0.296	0.002	0.104	0.017	0.986	0.331	0.107	3.108	0.005
0	0	0	0	0	0	0	0	0	0
25	0.296	0.002	0.104	0.017	0.986	0.331	0.107	3.108	0.005

Chow test for sel/unsel regression subsets: can't compute

Figure 7. 27 **Global Moran's I: MPI and MHCI districts of KPK**

The results reveal the significant positive spatial autocorrelation that depicts the clustering effect of social welfare in districts of KPK. The districts of Shangla and Batagram have a higher level of multidimensional poverty (incidence and intensity) compared to the average level of multidimensional poverty in the districts of KPK that have higher poverty compared to other districts of KPK. While the districts Abbottabad, Nowshera and Charsada has a significantly lower level of multidimensional poverty compared to other districts of KPK.

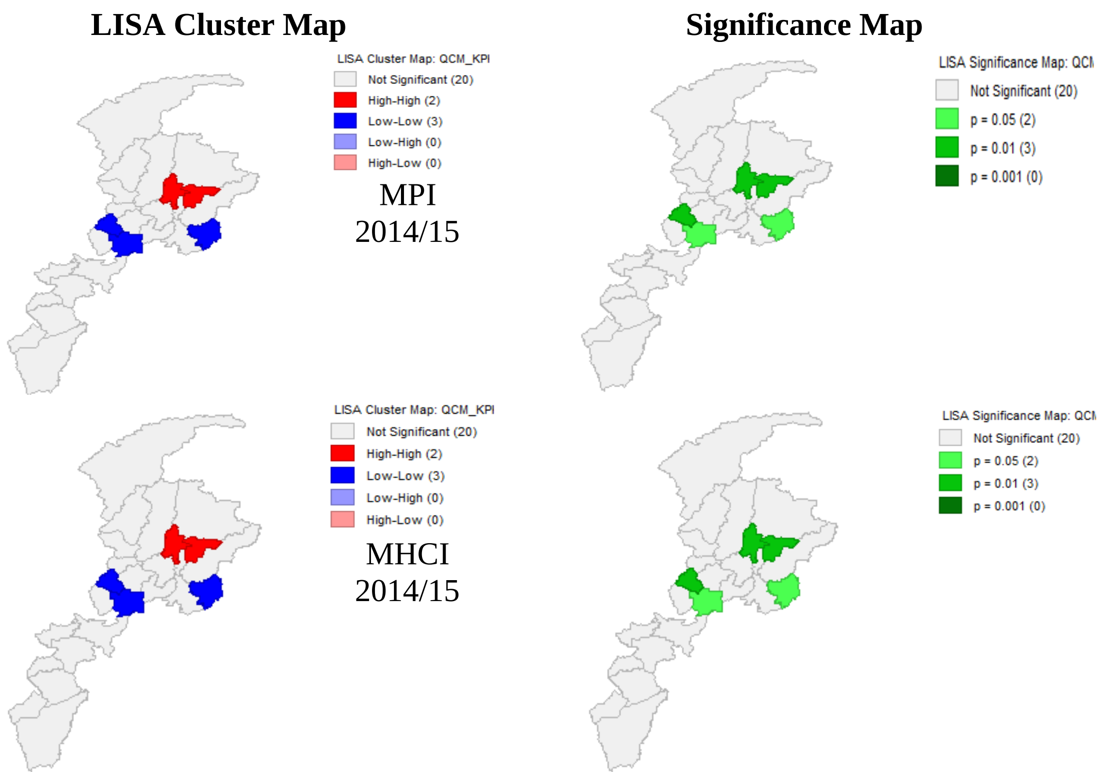


Figure 7. 28 **Local Moran's I of MPI and MHCI in district of KPK**

10.6.5 Economic geography of the welfare in the districts of Baluchistan

We conducted ESDA for indicators are MPI and multidimension headcount index (MHCI) for the district of Baluchistan for FY2014-15 using queen contiguity weight matrix. The results of Moran and connectivity graph for the districts of Baluchistan are presented in Figure 7.29 & 7.30 The ESDA results depicts insignificant spatial positive spatial autocorrelation for indicators of multidimensional poverty. The only Mastung district shows low-low spatial local autocorrelation compared to the average incidence within the districts of Baluchistan.

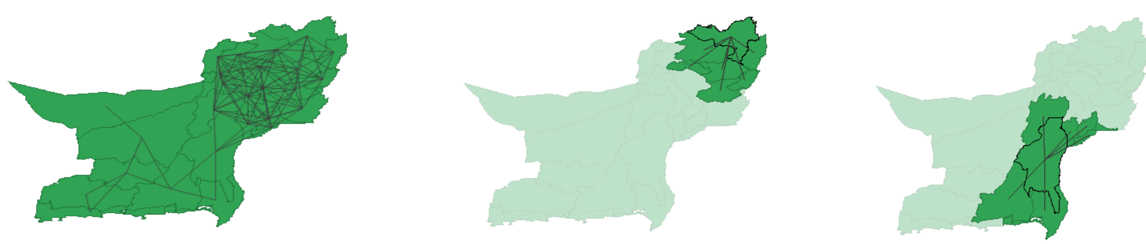
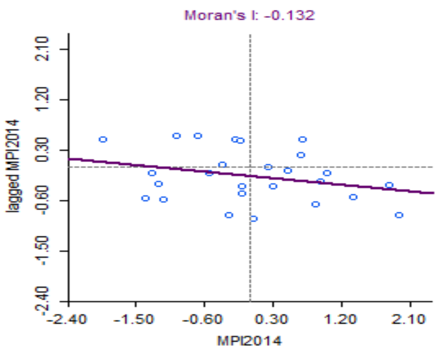
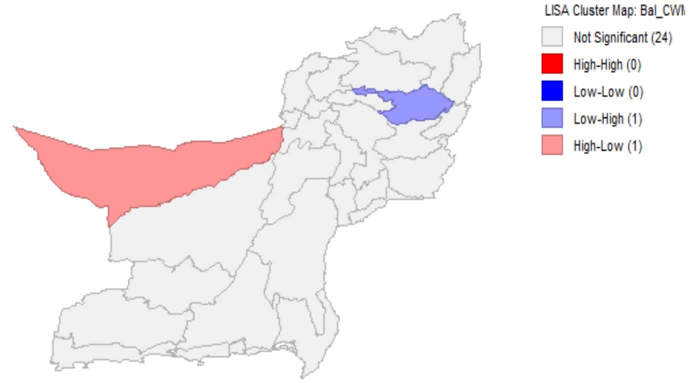
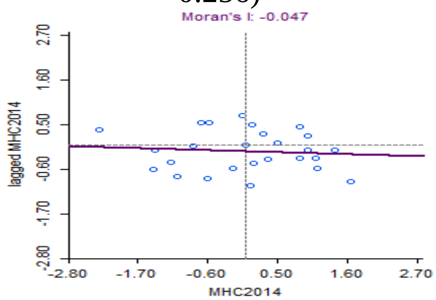


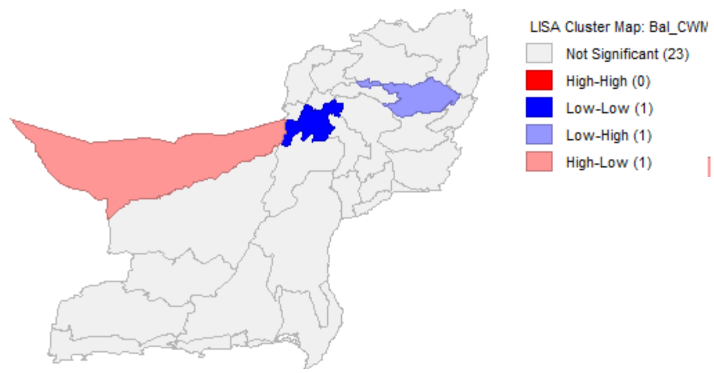
Figure 7. 29 **Connectivity Graphs and Moran-I for MPI (2014/15) using QWM**





(I: 0.132, pseudo p-value: 0.122 to 0.256)





(I; -0.047, pseduo p-value: 0.487, 0.500)

Figure 7. 30 **Global and local Moran's I: MPI and MHCI (2014/15)**

The district-level analysis revealed that the global Moran's I graph located in the quadrants of high-high and low-low quadrants of Moran's scattered plot, depicting the clustering effects at the national level as well as provincial-level except Baluchistan. By random computations, this polarization and the clustering effect is statistically highly significant at 1 percent confidence the national level and in the province of Punjab and weakly significant at 5 percent in the province of Sindh and KPK. These results elucidates that the social welfare is a space-dependent phenomenon in Pakistan. Therefore, utilization of spatial techniques of estimation are imperative to infer and devise better and more informed policies and targeted programs.

CHAPTER 8

REGIONAL CONNECTIVITY AND SOCIAL WELFARE: A DISTRICT LEVEL SPATIAL DATA ANALYSIS

8.1 Introduction

The previous chapter 7 presenting the ESDA of road and indicators of social welfare. In this chapter is presenting the regression analysis by conducting panel as well cross-sectional regression analysis. The transport infrastructure is critical input for the regional connectivity, therefore in this section present regression analysis ascertaining the role of transport infrastructure in reducing the socio-economic marginalization. This chapter is classified into two main sections. First section encapsulated the regression analysis has been conducted panel-data settings to capture the direct as well as indirect and spillover effects for the data of 110 districts during the years from 2004/5 to 2014/15. While, in the second section the regression analysis has been conducted for the year 2019-20 to capture the spatial spillover effect.

8.2 Panel-Data Regression Analysis (2004/5 to 2014/15)

This section provides the model estimation results based on the regression analysis aiming to capture the role of transport infrastructure in social welfare by estimating the direct and indirect, and spillover effects. The previous section presented an ESDA of the transport infrastructure and social welfare as the global Moran's I results signify the spatial dependence of road infrastructure and social welfare in Pakistan. In this purview, this section is presenting an attempt to estimate the direct and indirect effects of road infrastructure on the social welfare of Pakistan via estimating panel-data regressions using data of 110 districts during the years from 2004/5 to 2014/15. This section presenting brief outline of the empirical model, the second section orienting to the spatial weights and the descriptive statistics, and the third section provides the results of the estimated model, tests, and interpretation.

8.2.1 Empirical Model and Estimation methodology

The selection of the estimable model or dependent on the significance of the test of spatial dependence i.e., global Moan's I result. For example, if the test for the spatial dependence for social welfare and road infrastructure indicators are significant in the model, the spatial panel data would be estimated by adding the spatially lag variable in panel data.

$$y_{it} = \alpha_0 + \alpha_1 W_c y_{it} + \alpha_2 X1_{it} + \alpha_3 X2_{it} + \alpha_4 W_d X1_{it} + \alpha_i + \varepsilon_{it}$$

The proposed model is an SDM model using two different spatial weights. The decision to use the fixed effect and random effect would be based on the estimated outcome of the 'Robust Hausman test' model specification test. The post-estimation diagnostics tests will detect cross-sectional dependence, heteroscedasticity, and serial autocorrelation in the estimated model. If cross-sectional dependence is found, then Spatial Error Model (SEM) will also be estimated, as this model includes spatial autocorrelated error terms separately. Finally, if the problem of heteroscedasticity and serial autocorrelation is present in the estimated model only, we will use the panel corrected standard errors (PCSE) model.

8.2.2 Data and the operational definition

We use data from 110 districts of Pakistan (Cross-sections; N) and six data points (T) with a time interval for the years 2004/5, 2006/7, 2008/9, 2010/11, 2012/13, and 2014/15. In this study, the number of cross-sections is more significant than time; therefore, it seems cross-sectional data-properties will be more prominent than temporal autoregressive data properties. Detailed elaboration is provided in Table 8.1. For this study, we are not using income-based poverty instead we are using MPI as indicators of social welfare, as literature evidences that income of the poor households do not well-matched with the households socio-economic deprivation such as the basic requirements for a quality life such as malnutrition, no access to education (Alkire et al., 2014).

Table 8. 1 *Description of Variables* for Panel-Data Regression

Variable description	Indicator	Data source
y_{it} Represents the dependent variable, indicator reflecting the stature of social welfare. We are defined as social welfare in terms of the multidimensional poverty. We define regions as the administrative units of districts in Pakistan 'The time duration for 6. For years of a survey of PSLM was conducted during FY 2004-5, 2006-7, 2008-9, 2010-11, 2012-13, and 2014-15.	We are using the two leading indicators as the dependent variables. 1. The Multidimensional Poverty Index *(mpi)* 2. Incidence of Multidimensional poverty, i.e. multidimensional headcount index(*mhci*)	A joint report of GoP and UNDP, published by UNDP (2016), was developed under Dr. Alkire's[10] team lead.
$W_c y_{it}$ Spatially lag-variable of dependent variable	We are constructed spatial lag-variable of dependent variable for the respective dependent varaible.	Authors' construct using GeoDa
X_{it}	1. land utilization (the cultivated land area) 2. Road length of the high type roads[11] (provincial roads) of each district.	Various issues of the Provincial Development Statistics Reports (Punjab, Sindh, KPK, and Baluchistan)
$W_d X1_{it}$ Spatial Lag variable	Spatially lag-variable, which is weighted average of the provincial high-type road network of neighboring districts	Authors' construct using GeoDa
W Represent the Spatial weight matrix	We are using two main types of the spatial matrix (contiguity W_c and inverse distance matrix W_d).	Shapefile source: United Nation Office for coordinating Humanitarian Affairs (OCHA)[12]. We used layer (2) the district-level polygons.

10 Dr. Sabina Alkire, Director of the Oxford Poverty and Human Development Initiative (OPHI) at the University of Oxford

11 High type road are classified as the black-top mettled roads.

12 detailed for dataset and downloading is available on https://data.humdata.org/dataset/cod-ab-pak.

8.2.3 Construction of spatial lag variables and descriptive Statistics

For gauging the direct and indirect impact of road infrastructure investment in the development outcomes space of the social welfare. We collected data from provincial development statists reports for various years for the key indicators of high types (mettled) road length at the administrative level of districts for 2004/5, 2006/7, 2008/9, 2010/11, 2012/13, and 2014/15. The data of multidimensional poverty index from obtained from UNDP report published by Government of Pakistan. To club data in geographical space, we use the district level shapefiles. The working boundaries in some new districts of Pakistan have been merged into their parent districts for comparative analysis, except for the district Chiniot and Nankana. We use the Arcmap to merge the districts of Jaffarabad/Sohbatpur, Kharan/Washuk, Mansehra/Torghar, Zhob/Sherani, Sibi/Haranai/Lehri, Thatta/Sujawal, and Chagai/Noshhki. We extrapolated data for Nankana and Chiniot, keeping the data trends in view. We construct the inverse distance weights matrix [Wd] for the road infrastructure and queen contiguity matrix [Wc] for indicators of MPI and MHCI using GeoDa. We developed two types of spatial weights. First, queen contiguity matrix (QCM) (order 1), and second, inverse Euclidean distance weight matrix of 5 and 15 bandwidth. The connectivity graphs of these three weight matrices are provided in Figure 8.1

In order to estimate, we conducted a panel data analysis. We define cross-section of 110 districts and 6-time durations. We used Stata-software for panel data estimation, and we developed spatially lag variables in GeoDa software in cross-sectional data settings. We constructed the spatial lag variables using the ‘calculator’ tab and exported the same for all six-time points in excel readable csv format to prepare data in panel-data settings. This is manual computation, mainly to avoid errors. This software builds new variables (mx1) through matrix multiplication operation using the spatial weights matrix(mxn) with the vector matric(nx1) of road infrastructure and social welfare. We used row-standardized spatial weights to develop the spatial-lag variables using cross-sectional data, using Geoda’s calculator interface.

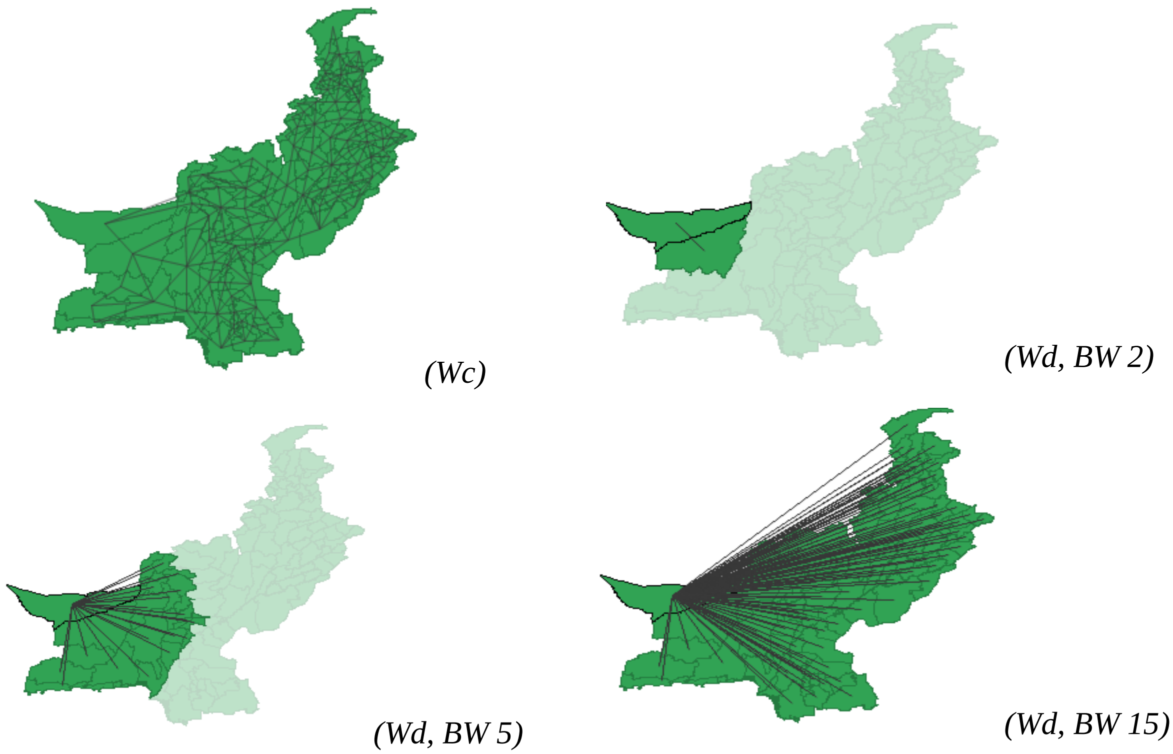


Figure 8. 1 **Connectivity Graphs and maps: Districts of Pakistan**

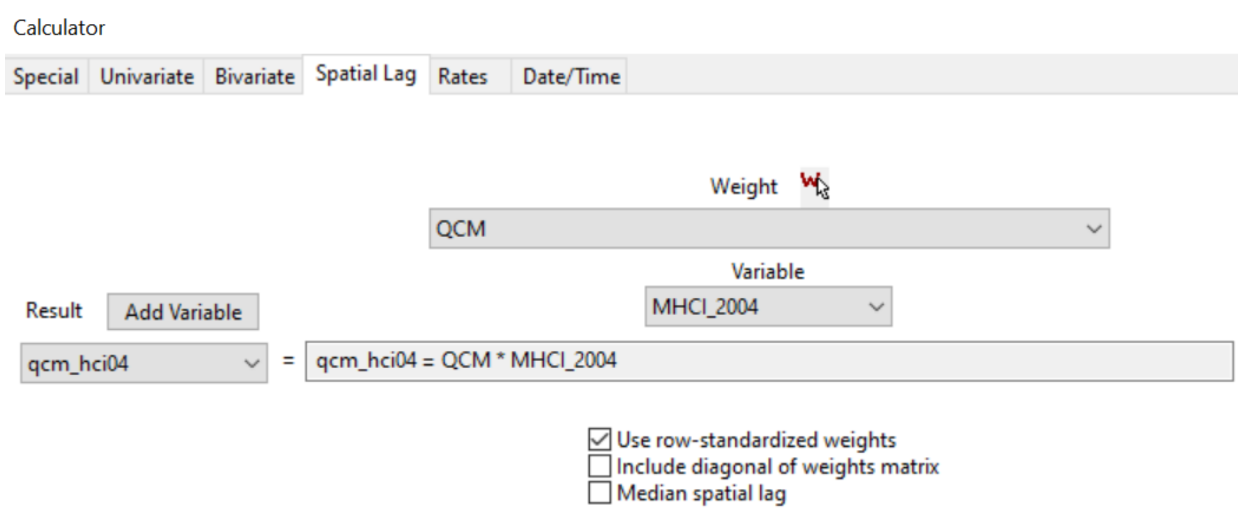


Figure 8. 2 **Geoda Interface for developing spatial lags for each period**

We computed the spatial lag variables for the panel analysis instead of using direct commands for matrix import in the Stata. This manual computation in the matrix multiplication when two types of W matrix are used. Therefore, we compute these spatial lag variables using spatial weights in the GeoDA separately and compute constructed variables in the panel-data format. Therefore, we developed spatial lag variables for road infrastructure of IDW2*LHTR, IDW5*LHTR, and IDW15*LHTR using row-standardized weights of IDWM's [Wd] of 5 BW and 15 BW to the indirect spatial effect of road infrastructure. At the same time, we developed spatial lag variables of MPI as 'Wc*MPI' and LMHCI as 'Wc*LMHCI using QCM of order 1, with row standardization. A details description of the spatial lag variables is provided with other variables in table 8.2.

Table 8. 2 *Descriptive Statistics for Panel Data Regression*

Variables	**Descriptions**	**Mean**	**STD**	**Min**	**Max**
ID	Districts are cross-sections	55.50	31.78	1.00	110.00
T	Six years data, 2004/5, 2006/7, 2008/9, 2010/11, 2012/12 and 2014/15	3.50	1.70	1.00	6.00
mpi_{it}	Multidimensional poverty index (0 to 1)	0.330	0.15	0.02	0.68
$Lmhci_{it}$	Natural log of Multidimensional headcount index (0 to 100)	4.014	0.515	1.45	4.60
$lhtr_{it}$	Natural log of Total length of high type road (km)	6.537	0.985	2.77	9.15
$llui_{it}$	Natural log of Land-use intensity (percentage)	3.008	1.399	-1.68	4.78
$Wc * mpi_{it}$	Spatial lag: QCM *MPI	0.328	0.117	0.07	0.62
$Wc * lmhci_{it}$	Spatial lag: QCM*LMCHI	4.008	0.397	2.61	4.55
$Wid_{2B} * lhtr_{it}$	Spatial lag: IDWM2*LHTR	6.527	0.599	5.24	7.76
$Wid_{5B} * lhtr_{it}$	Spatial lag: IDWM5*LHTR	6.547	0.329	5.79	7.21
$Wid_{15B} * lhtr_{it}$	Spatial lag: IDWM15*LHTR	6.555	0.262	6.05	7.10

8.2.4 An Empirical models, estimations, and discussion of results

We developed a panel data model by incorporating spatial dependence in the model. In this context, the spatial Durbin model is an appropriate method for our study, as poverty and road infrastructure both are space-dependent, and we use two different spatial weights to define neighbourhoods in a single panel regression.

$$mpi_{it} = \beta_o + \beta_1\, Wc * mpi_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it} \qquad (8.1)$$

$$lmchi_{it} = \beta_o + \beta_1\, Wc * lmhci_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it} \qquad (8.2)$$

We estimate two main models, using the definition of queen continuity and inverse distance weights. We use three bandwidths of an inverse distance of 2 BW, 5 BW, and 15BW. Therefore, for this study, we are estimating six models to estimate the indirect effect of infrastructure.

$$mpi_{it} = \beta_o + \beta_1\, Wc * mpi_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(2b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 1)

$$mpi_{it} = \beta_o + \beta_1\, Wc * mpi_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(5b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 2)

$$mpi_{it} = \beta_o + \beta_1\, Wc * mpi_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(15b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 3)

$$lmchi_{it} = \beta_o + \beta_1\, Wc * lmhci_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(2b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 4)

$$lmchi_{it} = \beta_o + \beta_1\, Wc * lmhci_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(5b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 5)

$$lmchi_{it} = \beta_o + \beta_1\, Wc * lmhci_{it} + \beta_2\, lhtr_{it} + \beta_3\, Wid_{(15b)} *\ lhtr_{it}\ + \beta_4\, llui_{it}\ + \varepsilon + \mu_{it}$$

(model 6)

We regress for these six models using the data of 110 districts (cross-sections) for 6-time units of 2004/5, 2006/7, 2008/9, 2010/11, 2012/13, and 2014/15. We estimated model regress using fixed and random effects and used the Hausman test to select the appropriate model. The Hausman test recommends the fixed-effect model for all models. We performed a post-regression test for the cross-sectional dependence, heteroscedasticity, and serial autocorrelation. The result for all models reveals insignificant cross-sectional dependence and serial autocorrelation but significant heterogeneity. Therefore, we are using the panel-corrected standard errors (PCSE) model that accounts for the issues of heteroscedasticity by correcting the standard errors of panel regression. The PCSE model is recommended in empirical literature when the number of cross-sections is larger than time (N>T). The results of all six models are provided in table 8.3.

All six models revealed the critical importance of road infrastructure in the poverty reduction and welfare improvement in Pakistan. The coefficient measuring the direct effect of road infrastructure in all the six models is highly significant and negative. The results show that MPI is lowered by 0.03 when high-type road length increases by one percentage point. At the same time, MHCI is reduced by 0.12 percentage points on average when the road network in each district is increased by one percent. Therefore, these results reinforce the empirical validation of the theoretical linkage of transport infrastructure in welfare improvement through geographic connectivity and integration.

Table 8. 3 *Results of Estimated Model (Method: Panel Corrected Standard Error)*

Dependent Variable	Model 1	Model 2 mpi_{it}	Model 3	Model 4	Model 5 $lmhci_{it}$	Model 6
$Wc * mpi_{it}$	0.861*** (0.040)	0.834*** (0.042)	0.817*** (0.048)			
$Wc * lmhci_{it}$				0.804*** (0.572)	0.779*** (0.056)	0.767*** (0.058)
$lhtr_{it}$	-0.032*** (0.007)	-0.029*** (0.007)	-0.027*** (0.007)	-0.133*** (0.036)	-0.118*** (0.037)	-0.114*** (0.034)
$Wid_{(2b)} * lhtr_{it}$	0.022*** (0.010)			0.077 (0.482)		
$Wid_{(5b)} * lhtr_{it}$		0.015 (0.0194)			0.048 (0.086)	
$Wid_{(15b)} * lhtr_{it}$			0.005 (0.024)			0.025 (0.106)
$llui_{it}$	-0.009*** (0.002)	-0.008*** (0.003)	-0.008*** (0.003)	-0.016 (0.036)	-0.012 (0.010)	-0.011 (0.011)
Constant	0.145*** (0.059)	0.171 (0.118)	0.232 (0.148)	1.200*** (0.362)	1.380*** (0.571)	1.556** (0.696)
N of Cross-sections	110	110	110	110	110	110
Time	6	6	6	6	6	6
R-Square	0.6147	0.6162	0.6116	0.5147	0.5115	0.5112
Wald χ^2	1132.85***	1143.34***	1145.50***	492.11***	469.01***	482.98***

Standard Errors are given in ()
**** significance at 1%, ** significance at 5%, *significance at 10% confidence interval*

The sign of the coefficient $Wid_{(2b)} * lhtr_{it}$ is positive and significant (according to model 1), depicting the resource mobility from the regions with high poverty to lower ones. The sign of the coefficient of $Wid_{(5b)} * lhtr_{it}$ and $Wid_{(15b)} * lhtr_{it}$ (in the models 2-6) are positive but statistically insignificant, depicting the positive but insignificant indirect effects of infrastructure. The results of all six models signal an adverse indirect effect of the relative road infrastructure that signals the factor mobility from the districts with high deprivation to lower regions. The results revealed that keeping the effect of other variables constant, a percent increase in the road length in the neighbouring districts can potentially increase the incidence and intensity of multidimensional poverty in a district by 0.02 points.

8.3 Cross-sectional Spatial Regression Analysis (2019-20)

This section presents the estimation of spatial spillover effects that are hypothesized to be developed as the results of the investment in the transport infrastructure with the aim to generate regionally connectivity and cohesiveness. We had used MPI in the previous section as indicators of assessing the regional welfare however due to the issues of comparability of NTRC data of road infrastructure with the previous provincial statistics and officially also GoP has not published district level MPI. The MPI required cut-off points and these cut-off scores essential to measure deprivation are decided by the GoP and UNDP in Pakistan, due to these two issues we couldn't incorporate the 2019-20 data in panel-data regression (in previous section). Henceforth, to analyze the welfare effect of transport infrastructure using cross-sectional data for the year 2019-20 we developed an indicator for social welfare. This index is for assessing the regional welfare using the three main dimensions of education, health and quality of life or standard of living at district level dataset of PSLM 2019-20. We used the latest district profiles of road network that has been developed with a great accuracy under the digitization project of NTRC except for very few districts (provincial data has been used for those district). We employed the control variables of land-use intensity for cultivation, population density per square-km, proportion of the urban area in a district, geographical coverage of a district in each province, average population growth and average household size. We collected data of these control variables from the various published report of development statistics of each province and census-survey data that has been conducted in 2017.

8.3.1 Econometric Model and Estimation technique

We are estimating two models. First with all controlling variables and second with provincial dummies. The dependent variable (y_i) includes index to social welfare and dependent variables includes a set of independent variables.

$$lny_i = \beta_0 + \beta_1 lnroad_i + \beta_2 lnpden_i + \beta_3 lnurpop_i + \beta_4 lnlui_i + \beta_5 lnrgca_i + \beta_5 apg_i + \beta_6 lnahhs_i + \varepsilon_i \quad \text{(Model 1)}$$

$$lny_i = \beta_0 + \beta_1 lnroad_i + \beta_2 D\_pun_i + \beta_3 D\_Sindh_i + \beta_4 D\_KPK_i + \varepsilon_i \quad \text{(Model 2)}$$

The estimation OLS estimation technique may infer into wrong conclusion due to existence of spatial dependence. Therefore, in order to select an appropriate spatial model, we estimated OLS regression model at outset for each regression and based on the test for the detection (significance) of the spatial dependence (Moran's I) in the error term of regression. There are many spatial model(s) but two representative models for the model with the special dependency. These are Spatial Error Model (SEM) and Spatial Lag Model (SLM). These computed using maximum likelihood Estimation (MLE) technique. SLM into spatial lag therefore autocorrelation among observed variables, and while SEM takes the autocorrelation of error terms into account therefore, SEM estimates are interpretable.

$$lny_i = \rho \boldsymbol{W}\, lny_i + \beta_0 + \beta_1 lnroad_i + \beta_2 lnpden_i + \beta_3 lnurpop_i + \beta_4 lnlui_i + \beta_5 lnrgca_i + \beta_5 apg_i + \beta_6 lnahhs_i + \varepsilon_i \quad \text{(SLM: Model 1)}$$

$$lny_i = \rho \boldsymbol{W}\, lny_i + \beta_0 + \beta_1 lnroad_i + \beta_2 D\_pun_i + \beta_3 D\_Sindh_i + \beta_4 D\_KPK_i + \varepsilon_i \quad \text{(SLM: Model 2)}$$

The spatial rho ρ captures the local scale spillover effects. The $\rho > 0$ representing positive spatial autocorrelation depicting a clustering effect depicting similar values are neighbouring to each other. On the other hand, $\rho < 0$ represents negative spatial autocorrelation depicting no clustering effect representing dissimilar values are neighbouring to each other. While SEM model, also known as Simultaneous Autoregressive Model (SAR) or SAR Error Model is useful to capture the global-scale spillover effects or global spatial autocorrelation. These captures the effect of a shock at a certain point spread to all other points.

$$lny_i = \beta_0 + \beta_1 lnroad_i + \beta_2 lnpden_i + \beta_3 lnurpop_i + \beta_4 lnlui_i + \beta_5 lnrgca_i + \beta_5 apg_i + \beta_6 lnahhs_i + u_i,\ u_i = \lambda u_i + \varepsilon_i \quad \text{(SEM: Model 1)}$$

$$lny_i = \beta_0 + \beta_1 lnroad_i + \beta_2 D\_pun_i + \beta_3 D\_Sindh_i + \beta_4 D\_KPK_i + \varepsilon_i, u_i = \lambda u_i + \varepsilon_i \quad \text{(SEM: Model 2)}$$

The SLM coefficient are not interpretable except the spatial rho (ρ) as the OLS assumption strict exogeneity of independent variables, and in the presence of spatial-lag variable this assumption violates, however the SEM model assumes error -terms are spatial correlated therefore, following the assumptions of exogeneity of independent variables. Therefore, we can interpret the coefficients of SEM or SAR model. If value of $\lfloor\lambda\rfloor < 1$ and statistically significant then we can deem the presence of the global spatial multiplier effects (and the positive network-externalities) are being generated.

8.3.2 Data and Sources

The details of index construction and other indicator and source is given in table 8.4.

Table 8. 4 *Indicators, Data-Sources, and its Description*

Indicator	Description	Main-Source
lny_i	Welfare Index (WIN): Equi-weighted Index in the three main domains of; 1. Education (1/3 weight) using the indicators of • Population percentage that ever-attended school (1/2 weight) • Gross Enrolment Rate at The Primary Level (Age 5-9) (1/2 weight) 2. Health (1/3 weight) using indicator of immunization of Children (12-23 months age) 3. Indicators of quality of Life (1/3 weight) using indicator of • Percentage of households using RCC, RBC, Sheet, Iron and or Cement for Roof (1/6 weight) • Percentage of households using burnt bricks and blocks for walls (1/6 weight) • Percentage of households using electricity for lighting purposes (1/6 weight) • Percentage of households using gas and oil for cooking purposes (1/6 weight) • Percentage of household using tap-water as main source of drinking water (1/6 weight) • Percentage of household using flush for sanitary purposes (1/6 weight)	PSLM district level report by Government of Pakistan (2021)
$lnroad_i$	Length of roads in kilometres	NTRC (2020)
$lnpden$	Population density per square kilometres	Population Census Report (2017)
$lnurpop_i$	Percentage of population living in urban locality	Population Census Report (2017)
$lnlui_i$	Land-use intensity as a Percentage of land-used for cultivation purposes	Provincial Development Reports
$lnrgca_i$	Relative geographic covered area (rgca) Percentage of covered area of each district relative to its province geographic size/area.	Census (2017)
apg_i	Average (annual) population growth during 1998-2017	Census (2017)
$lnahhs_i$	Average Household Size	Census (2017)
W	Spatial weight matrices constructed based on district level shapefiles	

* All variables are transformed into natural log except apg_i, as it contain some negative values for some districts.

8.3.3 Estimation and Results

Following the standard procedure, we estimated OLS model for Model 1 and 2 both and the Moran's I test is highly significant signaling the spatial dependence and requirement of spatial-regression model. The appropriate Spatial Model selection is dependent on Lagrange Multiplier tests however, we estimated both SEM and SLM models to avoid the research subjective selection biasedness. We constructed two spatial weights' matrices $\boldsymbol{W}$. First, is the queen-contiguity weight (QCW) matrix of order 1 and second is the inverse-distance weight matrix at 3 bandwidths. Row-standardized weights were computed for the analysis. The estimated output for model (1) and (2) using OLS is presented the table 1. The results of test of Moran's I for error supporting the argument for spatial dependence and suggesting supporting for utilization of the spatial models. The output of SEM and SLM using QCW is provided in the table 2. Likewise, the output of SEM and SLM using IDW is provided in the table 3 and 4 respectively. The model related statistics and diagnostics-test are also provided in its respective tables.

The estimated SEM (or SAR) and SLM model supporting the argument for generation of clustering as well as global spillover effects as result of investing in road infrastructure, while keeping the effects of size population density, urbanization, land-use intensity for cultivation, household-size, and population growth effect constant. In addition, model 2 the result seconds to the analysis when provincial dummies where incorporated. In model 2 the controlling variables were excluded, because these effects are being controlled using provincial level effects and it also pertinent to keep enough degree of freedom for statistical analysis. Moreover, results of SEM or SLM usually exist the issue of heteroskedastic error-variance and the condition of homoscedasticity can only be met on the lattice points, therefore, is inevitable becomes heteroskedastic, so the stationarity of the covariance is not satisfied. Moreover, the coefficients of SEM are interpretable while, SLM requires the analysis of the marginal effects.

The estimated results confirmed the presence of the clustering effect as well global spatial effect. Using QCW estimated spatial rho $\hat{\rho}$ in both models are 0.31 and 0.40 suggesting that a percent increase in a welfare of neighbouring region likely to increase the welfare approximately by 0.3 and or 0.4 percent through clustering effect. In addition to this, based on IDW the spatial rho $\hat{\rho}$ is approximately 0.46 and 0.55 supporting for the evidence of clustering effect of social welfare at district level in Pakistan. While the global-networking effect or spillover effect is highly significant and the absolute is less than 1, all estimated models. Using QCM, estimated value spatial lambda $\hat{\lambda}$ is 0.37 and 0.56, while using IDW the value of $\hat{\lambda}$ is 0.73 and 0.82 respectively. Therefore, we can conclude that the global spillover or networking effect that generates from one region to another are approximately ranges up to 0.8 percent.

The results revealed that investing in road infrastructure is promising to elevate the social and economic depreciation and welfare improving. The estimated coefficient of SEM using both QCW and IDW for road infrastructure is significant and positive. Therefore, we can conclude that an increase in road-length by one percentage point is likely to increase WIN by 0.06 percentage point by keeping the effect of population density, degree of urbanization, geographic cover of district, population growth and household size effect in constant. Moreover, a percent increase in road length increasing WIN by 0.13 percent when provincial effect are controlled using provincial dummies.

This study also found that the controlling variables of population density, degree of urbanization, geographic cover of district, population growth, household size effect also determine the regional welfare of Pakistan at district level. The coefficient measuring elasticity of population density is positive and significant and size of coefficient is 0.05 (QCM) and or 0.06 (IDW) keeping the other variables of the model effects constant. Likewise, proportion of population residing in the urban settings also significantly impact the regional welfare. The estimated coefficient is 0.09 (QCM) and 0.12 (IDW) and significant depicting the fact that in urban area has greater access to educational, health services and standard of life services compared to rural area. In addition to this, the district with greater land utilization for cultivation (agriculture purposes) has improved quality of life and social welfare. The result revealed the coefficient for land-use intensity for cultivation is significant (and positive) impact and an increase in land-use intensity for cultivation by one percent, the social welfare of a district is likely to increase by 0.04(QCW) and 0.05(IDW). This study has added very important and least discussed controlling variable and important determinant of regional geography in Pakistan. The Author constructed a variable measuring the relative geographic size of each district given the geographic coverage of each province. And this study reveals the key importance of geographic size of each district as an important determinant of regional social welfare landscape of Pakistan. The estimated size of coefficient is negative (significant at 5 percent) in sign depicting the district with greater geographic size tend to have low social welfare attainment. The results revealed that an increase in one percent of relative geographic covered area of each district tend to reduce WIN by 0.05 percent (IDW) while keeping the effect of other control variables constant. Moreover, the control variables of average population growth and average household size coefficient are negative however these estimated coefficients are statistically significant in all estimated SEM using QCW and IDW.

The estimated model (2) the model with the dummy variables revealed that on average the social welfare in the province of Punjab and KPK is positive and significant and estimated average coefficient using SEM for β_2 and β_4 are approx. 0.377 and 0.365, respectively. However, the provincial dummy for province of Sindh is not statistically significant. Therefore, we can conclude the percentage change in WIN is Punjab province is 45.7 percent (100($e^{0.377}$ -1)) higher compared to the other province. Moreover, the percentage change in WIN is 42.8 percent (100($e^{0.356}$ -1)) when a district is located in KPK province compared to the other provinces.

Table 8. 5 *Estimated Output: OLS Estimation technique*

		(1)	(2)
Constant	2.503*** (0.216)	2.025*** (0.452)	2.716*** (0.257)
$lnroad_i$	0.191*** (0.027)	0.065* (0.036)	0.135*** (0.037)
$lnpden$		0.043*** (0.035)	
$lnurpop_i$		0.099*** (0.028)	
$lnlui_i$		0.055*** (0.022)	
$lnrgca_i$		-0.056 (0.030)*	
apg_i		-0.0287 (0.034)	
$lnahhs_i$		0.495*** (0.172)	
D_pun_i			0.380*** (0.089)
D_Sindh_i			0.045 (0.076)
D_Kpk_i			0.369*** (0.061)
Moran's I (QCW)	**8.174 *****	**0.295*****	**4.35*****
Moran's I(IDW)	**20.749*****	**14.44*****	**7.6237*****
R-Square	0.293	**0.565**	0.538
Multi-condition number	16.491	70.11	31.535
Akaike info criterion	41.6261	**-4.178**	-2.96555
Schwarz criterion	47.1844	18.05	10.9301
Diagnostics			
White Specification Test	7.21(0.027)	30.95(0.663)	119.00(0.000)
Jarque-Bera test	16.491(0.027)	3.326(0.189)	8.081(0.017)
Breusch-Pagan test	7.0402(0.007)	20.005(0.005)	12.812(0.012)

Standard Errors are given in ()

**** significance at 1%, ** significance at 5%, *significance at 10% confidence interval*

Table 8. 6 *Estimated Spatial Models: QCW*

	(1)		**(2)**	
Estimation technique	**SEM**	**SLM**	**SEM**	**SLM**
Constant	2.358***	0.787	2.696***	1.540***
	(0.4401)	(0.487)	(0.244)	(0.470)
$lnroad_i$	**0.067***	0.037	**0.137*****	0.134***
	(0.0362)	(0.033)	**(0.034)**	(0.035)
$lnpden$	0.055*	0.045		
	(0.0336)	(0.031)		
$lnurpop_i$	0.085***	0.099***		
	(0.0251)	(0.024)		
$lnlui_i$	0.044**	0.030		
	(0.0213)	(0.020)		
$lnrgca_i$	-0.039	-0.027		
	(0.030)	(0.027)		
apg_i	0.0031	-0.024		
	(0.0293)	(0.030)		
$lnahhs_i$	0.256	0.410**		
	(0.191)	(0.151)		
D_Pun_i			0.380***	0.219**
			(0.101)	(0.103)
D_Sindh_i			0.064	0.021
			(0.093)	(0.072)
D_Kpk_i			0.369***	0.269
			(0.083)	(0.067)
Spatial Rho (**Wy)**		**0.404*****		**0.314*****
		(0.089)		(0.105)
Spatial-Lamda ($\boldsymbol{\lambda}$)	**0.560*****		**0.373*****	
	(0.091)		(0.112)	
R-Square	**0.669**	0.640	**0.59**	0.576
Akaike info criterion	**-27.27**	-20.08	**-13.43**	-8.421
Schwarz criterion	**-5.044**	4.92486	**0.46**	8.252
Log Likelihood	21.34	19.04	11.71	10.211
Diagnostics				
Breusch-Pagan test	27.01(0.000)	31.47(0.000)	21.19(0.000)	22.32(0.000)
Likelihood Ratio Test	23.09(0.000)	17.90(0.000)	10.46(0.001)	7.45(0.006)

Standard Errors are given in ()

**** significance at 1%, ** significance at 5%, *significance at 10% confidence interval*

Table 8. 7 *Estimated Output: Spatial Error Model:IDW*

		(1)		(2)
Constant	3.035***	3.041***	2.712***	2.704***
	(0.450)	(0.250)	(0.470)	(0.262)
$lnroad_i$	0.049	0.065*	0.137***	0.137***
	(0.037)	(0.035)	(0.035)	(0.036)
$lnpden$	0.044		0.061**	
	(0.031)		(0.028)	
$lnurpop_i$	0.120***	0.123***		
	(0.024)	(0.024)		
$lnlui_i$	0.028	0.048***	0.018	
	(0.021)	(0.015)	(0.019)	
$lnrgca_i$	-0.024	-0.051**	-0.025	
	(0.029)	(0.022)	(0.027)	
apg_i	-0.004	0.017		
	(0.029)	(0.023)		
$lnahhs_i$	-0.020		-0.073	
	(0.172)		(0.177)	
D_pun_i				0.374***
				(0.096)
D_Sindh_i				0.053
				(0.082)
D_Kpk_i				0.344***
				(0.073)
Spatial-Lamda (λ)	0.822***	0.821***	0.731***	0.282
	(0.093)	(0.0944)	(0.132)	(0.279)
R-Square	0.653	0.647	0.579	0.542
Akaike info criterion	-25.85	-27.86	-8.58	-3.63
Schwarz criterion	-3.62	-11.18	8.09	10.25
Log Likelihood	20.92	19.93	10.29	
		Diagnostics		
Breusch-Pagan test	27.01(0.000)	23.57(0.000)	26.22(0.000)	15.61(0.003)
Likelihood Ratio Test	21.67(0.000)	29.23(0.000)	13.47(0.000)	0.673(0.411)

Standard Errors are given in ()

**** significance at 1%, ** significance at 5%, *significance at 10% confidence interval*

Table 8. 8 *Estimated Spatial Lag Model(s) :IDW*

Constant	0.358	0.922	0.489***	2.431***
	(0.668)	(0.635)	(0.651)	(0.235)
$lnroad_i$	0.035	0.114	0.041	0.136***
	(0.034)	(0.0282)	(0.034)	(0.0367)
$lnpden$	0.042	0.0578		
	(0.0315)	(0.0339)		
$lnurpop_i$	0.119***		0.124***	
	(0.025)		(0.025)	
$lnlui_i$	0.025	0.009	0.046***	
	(0.020)	(0.0223)	(0.012)	
$lnrgca_i$	-0.031	-0.033	-0.0524***	
	(0.027)	(0.028)	(0.021)	
apg_i	-0.014	0.001		
	(0.030)	(0.031)		
$lnahhs_i$	0.293**		0.294**	
	(0.155)		(0.145)	
D_pun_i				0.346***
				(0.134)
D_Sindh_i				0.040
				(0.075)
D_Kpk_i				0.342***
				(0.134)
Spatial Rho (Wy)	0.559***	0.46566***	0.553***	0.0731
	(0.151)	**(**0.158)	(0.152)	0.235
R-Square	0.620728	0.547967	0.614427	0.53893
Akaike info criterion	-16.6179	-0.34392	-18.6988	-1.068
Schwarz criterion	8.39424	19.1099	0.755057	15.606
Log Likelihood	17.3089	7.17196	16.3494	6.534
Diagnostics				
Breusch-Pagan test	24.85(0.000)	29.08(0.000)	20.78(0.00089)	14.09(0.006)
Likelihood Ratio Test	14.43(0.000)	10.38(0.001)	14.21(0.000)	0.102(0.748)

Standard Errors are given in ()

**** significance at 1%, ** significance at 5%, *significance at 10% confidence interval*

CHAPTER 9

CONCLUSION, POLICY RECOMMENDATION, AND LIMITATIONS

9.1 Introduction

This chapter presents the conclusion, policy recommendations and broad limitations of the study.

9.2 Conclusion

Academic literature extensively explores the significant influence of infrastructure on an economy, serving as a cornerstone for both economic activities and social mobility. By alleviating development constraints and bottlenecks, infrastructure plays a direct role in stimulating economic growth and enhancing productivity. The impact of infrastructure on economic growth operates through diverse channels. Additionally, infrastructure serves as a key factor in determining the distribution of wealth, income, and welfare. However, this influence is contingent upon various factors such as infrastructure spillovers, the time frame under consideration, and the source of financing for infrastructure projects (Chatterjee & Turnovsky, 2012). Infrastructure improved access to productive opportunities and reduced production and transaction costs resulting from enhanced infrastructure contribute to the development of industries and agro-industries, consequently raising the value of assets among the impoverished (Bajar, 2013; Bajar & Rajeev, 2016). Furthermore, a well-developed transport infrastructure facilitates labour mobility through improved geographical access, transportation services, and information flow. This connectivity and networking benefit disadvantaged individuals and marginalized communities by granting them access to productive opportunities. As a result, infrastructure plays a vital role in reducing poverty, a noticeable trend observed in numerous developing countries.

Additionally, the geographical distribution of economic activities is influenced by infrastructural improvements, as observed in Ottaviano's work in 2008. Infrastructure not only boosts local economic activity in its immediate vicinity but also generates benefits in neighbouring areas through spatial diffusion processes, leading to spillover effects (Ottaviano, 2008). These spillover effects can be either positive or negative. Notably, the positive spatial spillover effects of transport infrastructure have been extensively studied and empirically estimated at both national and regional levels in numerous countries. For example Shi et al. (2017); Yu et al. (2013); Zhang & Ji (2019) had validated this phenomenon of the spillover effects for China. Likewise, Cantos et al.

(2005); Cohen (2010); Cohen & Paul (2004) estimated the size of inter-state spillovers of transport infrastructure in Spain. The literature documented positive as well as negative spillover effects. The positive spillovers are caused by the connectivity characteristics of infrastructures. At the same time, negative regional spillover effects are generated due to factor outmigration from the region with inadequate infrastructure to higher due to relative economic attractiveness. In addition, infrastructure is one of the important factors that describe the geospatial patterns of economic development, but conditional upon the infrastructure spillover or externalities, time of consideration, and mode of financing (Chatterjee & Turnovsky, 2012).

In this connection, a detailed a comprehensive study was required to uncover the role of investment in infrastructure vis-a vis with productivity growth and the geo-spatial patterns of the regional development in Pakistan. Therefore, this study examined the direct but indirect economy wide effects of infrastructure and its key component. This study analysed the role of investment in infrastructure at three geographic levels of national and sub-national level of provinces and district level. Further to cater the spatial-temporal effects, the spatial econometric technique for estimation and techniques were employed.

The national level analysis is conducted using time-series econometric techniques for the key components of core, social and financial infrastructure. These components were further categorized based on the physical quantity and its utilization. The study developed model in the neoclassical growth framework to analyse the long-run impact on the steady-state level of growth. For long run estimate estimated multiple regressions models and used two estimated techniques of ARDL/NARDL bound testing approach and FMOLS to validate the true size of parameters. The short-run estimates were obtained by VECM using ARDL/NARDL estimation technique. Based on the estimated model we can conclude that productivity growth in both the long run (LR) and short run (SR) in Pakistan is significantly influenced by various components of infrastructure. A summary of the estimated coefficients for each type of infrastructure are provided in tables 9.1 to 9.4. To gain a comprehensive understanding, we present a summary of the estimated coefficients measuring the influence of each type of infrastructure on productivity growth in SR in table 9.5. Additionally, our analysis accounts for the asymmetric effects of each infrastructure component, and these results are presented in table 9.6.

Table 9. 1 *LR Growth effect of Core Infrastructure: Estimated coefficients: A Summary*

	$LnSgcq_t$	$LnSgcq_t$^2		$LnSgcu_t$	$LnSgcu_t$^2
5/1-1(1.1)	0.163***		5/1-2(1.1)	0.001	
5/1-3(1.1)	0.099***		5/1-4(1.1)	-0.002	
5/1-3(1.2)	0.105***	-0.024**	5/1-4(1.2)	-0.029**	0.005**
5/1-3(1.3)	0.090***	-0.021*	5/1-4(1.3)	-0.011***	0.002***
5-A	0.238***		5-B	-0.001	
5-C	0.048**		5-D	-0.008	

The core infrastructure played a critical role in generating both LR and SR productivity growth in Pakistan. Regarding physical infrastructure and its utilization, the core infrastructure increased productivity growth in a non-linear manner. However, the utilization of infrastructure was not yet sufficient to significantly boost productivity growth in Pakistan, as indicated by the insignificant coefficients in models (Table 9.1). On the other hand, the results of models 5/1-4(1.2) and 5/1-4(1.3) revealed a parabolic relationship between core infrastructure utilization and productivity growth. Based on estimated insignificant coefficients we can conclude that Pakistan had not yet reached the turning point, implying that policy measures to optimize the utilization of core infrastructure were crucial to drive productivity growth through utilization channels effectively. Similarly, the relationship between physical infrastructure stock and productivity growth followed an inverted parabola U-shape type pattern. Based on estimated significant coefficients we can infer that Pakistan had not yet reached the turning point. It is because estimated Models demonstrated a positive increase in productivity growth due to investments in building physical infrastructure for core infrastructure during the last 42 years. Henceforth the study emphasizes the importance of optimal utilization of core infrastructure and the significance of investing in physical infrastructure to drive productivity growth in Pakistan.

Table 9. 2 *LR Growth effect of Social infrastructure (Education): A Summary of Estimates*

	$LnSedq_t$	$LnSgedq_t$^2		$LnSedu_t$	$LnSgedu_t$^2
5/2-1(1.1)	0.309		5/2-2(1.1)	0.284***	
5/2-3(1.1)	0.016		5/2-4(1.1)	0.007	
5/2-3(1.2)	0.233***	-0.020***	5/2-4(1.2)	-0.062	0.028
5-A	-0.207		5-B	0.195**	
5-C	-0.059		5-D	0.007	

The estimated models revealed the crucial role of physical infrastructure for education in augmenting the educational services in Pakistan. It is pertinent to note that the growth of educational infrastructure stock does not keep pace with the increasing population. The estimated coefficients for physical infrastructure in various models (5/2-1(1.1), 5/2-3(1.1)6-3(A), 5-A, and 5-C) indicate an insignificant impact on productivity growth. However estimated model 5/2-3(1.2) shows significant coefficients reflecting a positive association of physical infrastructure for educational services resembling an inverted parabola graphically. This indicates that there is a positive impact at the outset but at later stages of development investing too much in later stages might lead to a wastage of resources. Based on the insignificance of the estimated coefficient it is deduced that Pakistan has not yet made sufficient investments in educational infrastructure to fully reap the long-term benefits. In Pakistan utilization of educational services has been increased overtime, and this increased associated with a moderate increase in productivity growth ranging from 0.19 to 0.28 percentage points during 1980-2022.

Table 9. 3 *LR Growth effect of social infrastructure (health): Estimated coefficients: a summary*

	$LnSghtq_t$	$LnSghtq_t$^2		$LnSghthu_t$	$LnSghtu_t$^2
5/3-1(1.1)	0.688		5/3-2(1.1)	0.128	
5/3-3(1.1)	0.188***		5/3-4(1.1)	0.012	
5/3-3(1.2)	-0.008		5/3-4(1.2)	0.008	
5/3-3(1.3)	7.273	-0.476	5/3-4(1.3)	-0.169**	0.027***
5-A	0.653*		5-B	0.087***	
5-C	0.291***		5-D	0.021*	

The impact of the investing in the health infrastructures in both physical and the infrastructural utilization terms both significantly contributed in the long-run (LR) productivity growth. The LR models 5-C and 5-D indicate highly significant coefficients for physical infrastructure. From these results, it can be concluded that a one percent increase in the number of hospital beds per million workers leads to a LR productivity growth increase of 0.291 percentage points. Additionally, a one percent increase in BCG immunization coverage contributes to a LR productivity growth increase of 0.087 percentage points.

Table 9. 4 *LR Growth effect of financial infrastructure: Estimated coefficients: a summary*

	$LnSgfq_t$	$LnSgfq_t$^2		$LnSgfu_t$	$LnSgfu_t$^2
5/4-1(1.1)	-0.033		5/4-2(1.1)	0.334*	
5/4-3(1.1)	0.006		5/4-4(1.1)	0.081***	
5/4-3(1.2)	0.002		5/4-4(1.2)	0.015	
5/4-3(1.3)	0.668**	0.060**	5/4-4(1.3)	0.176	-0.021
5-A	-0.085		5-B	0.139	
5-C	0.013		5-D	0.105***	

The estimated growth models for the financial infrastructure highlight the crucial role of financial services in achieving positive long-run (LR) productivity growth. Across all models, except for 7-3, the coefficients measuring the LR linear effects were found to be insignificant. However, model 7-3 supported the argument for the existence of a positive LR growth effect, which exhibits an increasing rate. Therefore, based on these findings, we can conclude that financial infrastructure plays a key role in enhancing productivity growth, and this effect increases at an accelerating rate over time. Moreover, the utilization of financial infrastructure directly contributes to LR productivity growth, as evidenced by the results from models 7-2(A), 7-4(A), and 8-D. These results emphasize the significance of effectively utilizing financial services to drive productivity growth in the long run.

In the short run (SR), all components of infrastructure, especially the physical stock of core infrastructure, play a crucial role in stimulating productivity growth in Pakistan. Additionally, the provision of infrastructural services in education and finance significantly contributes to generating productivity growth in the short run. A summary of the coefficients related to these effects is presented in table 9.5.

Table 9. 5 *SR Growth Effect of infrastructure: A summary*

$Sgcq_t$	5/1-1(1.1)	0.053**	$Sgcu_t$	5/1-2(1.1)	0.005
$lnSgedq_t$	5/2-1(1.1)	0.014	$lnSgedu_t$	5/2-2(1.1)	0.155***
$lnSghtq_t$	5/3-5(1.1)	0.14	$lnSghtu_t$	5/3-2(1.1)	-0.004
$LnSgfq_t$	5/4-1(1.1)	-0.033	$LnSgfu_t$	5/4-2(1.1)	0.082*

This study also examines whether the effects of increasing or decreasing physical infrastructure and infrastructural services on long-run (LR) and short-run (SR) productivity growth are significantly different. To investigate this, the NARDL bound testing approach has been employed, and a summary of the estimated coefficients can be found in table 9.6.

Table 9. 6 ***Asymmetric effects: Productivity growth and infrastructure's components***

	LR	SR		LR	SR
$Sgcq^{+}{}_{t}$	0.173***	0.115**	$Sgcu^{+}{}_{t}$	-0.014	0.020*
$Sgcq^{-}{}_{t}$	0.154	0.009	$Sgcu^{-}{}_{t}$	-0.320	-0.090
$lnSgedq^{+}{}_{t}$	0.432***	0.021	$lnSgedu^{+}{}_{t}$	0.135	0.131**
$lnSgedq^{-}{}_{t}$	-0.323	0.097	$lnSgedu^{-}{}_{t}$	-0.945	0.094
$lnSghtq^{+}{}_{t}$	0.852***	0.404***	$lnSghtu^{+}{}_{t}$	1.073	-0.008
$lnSghtq^{-}{}_{t}$	0.148	0.058	$lnSghtu^{-}{}_{t}$	10.056	0.098
$lnSgfq^{+}{}_{t}$	0.505***	-0.097	$lnSgfu^{+}{}_{t}$	0.435	0.089
$lnSgfq^{-}{}_{t}$	0.029	0.024	$lnSgfu^{-}{}_{t}$	0.832	0.116

The study revealed that investing in physical infrastructure for various components creates asymmetric effects on Pakistan's economy in the LR and SR. Based on the Wald coefficient test to analyse the symmetric effects to asymmetric effects, it can be inferred that investing in the physical infrastructure for each component of infrastructure leads to asymmetric effects in the long run (LR) on Pakistan's economy. Specifically, the coefficients of the $Sgcq^{+}{}_{t}$, $lnSgedq^{+}{}_{t}$, $lnSghtq^{+}{}_{t}, and\ lnSgfq^{+}{}_{t}$ are positive and statistically significant, while $Sgcq^{-}{}_{t}$, $lnSgedq^{-}{}_{t}, lnSghtq^{-}{}_{t} lnSgfq^{-}{}_{t}$ are found to be statistically insignificant. In the short run (SR), the asymmetric effect is significant for the core infrastructure and educational infrastructure. Notably, the utilization of infrastructural services does not create a significant asymmetric effect on the economy in the long run. However, in the short run, the asymmetric effect exist for the infrastructural services of core and education infrastructure. Therefore, we can conclude that in the short run, the utilization of infrastructural services from core and education infrastructure has a significant positive influence on productivity growth, while reductions in these services do not show significant statistical effects.

Based on the analysis conducted in the preceding Chapter 9 at the sub-national provincial level, the significance of investing in infrastructure in shaping regional economies becomes evident. Our conclusion underscores that the pivotal determinants of regional economic development in Pakistan are investments in human capital, private capital, and infrastructure. By employing dynamic spatial panel data analysis, we derived estimations that shed light on these dynamics. In the short run, the cumulative impact of core infrastructure on regional economic development doesn't display substantial significance, as per our estimations. However, the long-term perspective reveals a positive net effect, underscoring the paramount importance of infrastructure in driving prolonged economic growth and development. The reason behind the limited immediate impact of infrastructure, encompassing road and energy initiatives, is the inherent protraction associated with infrastructure projects. Such endeavours necessitate considerable time for preliminary feasibility assessments, design validations, contractual solicitations, and the allocation of substantial fiscal resources. Additionally, the involvement of international funding entities like the World Bank and initiatives such as the China-Pakistan Economic Corridor (CPEC) further contributes to time delays. Consequently, our empirical findings align logically with these underlying factors.

Considering the positive direct and indirect outcomes, it becomes imperative for Pakistan to commit to core infrastructure investments, acknowledging the potential short-term setbacks in favour of accruing enduring advantages. A summarized presentation of the outcomes can be observed in Table 9.7.

Table 9. 7 *The Marginal effects of Core infrastructure (index)*

	Short-Run	**Long Run**
Direct Effect	-0.0507***	0.166**
Indirect Effect	-0.0736***	0.169
Total Effect	-0.124***	0.334

In addition to this, impact of road infrastructure investment was assessed in uplifting the quality of life of people by tapping the direct, indirect effects and spatial spillovers or externalities. The road infrastructure and social welfare does have the spatial dependence, therefore we developed model spatial models. The results of panel-data regression revealed the critical importance of road infrastructure in the poverty reduction and welfare improvement in Pakistan via significant direct positive effect of road infrastructure. The results show on average multidimensional poverty index is lowered by 0.03 points when high-type road length increases by one percentage point, keeping the spatial effects and intensity land-utilization constant during 2004/5 to 2014/15. At the same time, incidence of multidimensional poverty is reduced by 0.12 percentage points on average when the road network in each district is increased by one percent keeping the spatial effects and intensity land-utilization in constant. Therefore, these results reinforce the presence of the positive direct effect of the transport infrastructure in the welfare improvement through geographic connectivity and integration. Moreover, the results also depicting negative indirect effects, the reduction in welfare (increase in incidence of multidimensional poverty) as results of factor out mobility from the regions with high social deprivation to lower ones as results of the road infrastructure in the nearby regions. The results revealed that a percent increase in the road length in the neighbouring districts (on average) can increase the multidimensional poverty in a district by 0.022 points.

Moreover, the estimated local as well global spatial autocorrelation (spatial spillover effects) were positive in magnitude statistically significant therefore, the finding are resonating with the development policy of Pakistan for boosting the regional connectivity and economic cohesion by investing in road-networks. The estimated size of these global spillovers or network externalities are high and ranging from 0.37 to 0.82. Based on the rigorous estimation, we can therefore conclude that landscape of social development in Pakistan at district level is determined size of the transport infrastructure network, population density, urbanization, relative geographic covered area of district, and size network externalities. Moreover, the province Punjab and KPK has significantly higher level of social development at the district level compared to other provinces. It is therefore provisioning of road-access by investment in the road infrastructure would be promising to reduce the social deprivation directly in the most poverty-stricken districts of Pakistan.

Furthermore, an assessment was conducted to gauge the impact of road infrastructure investment on enhancing people's quality of life. This assessment encompassed direct and indirect effects, as well as spatial spillovers or externalities. Recognizing the interdependency between road infrastructure and social well-being, we formulated spatial models to explore these dynamics. The outcomes of panel-data regression underscore the paramount importance of road infrastructure in mitigating poverty and advancing welfare in Pakistan. A notable positive direct effect of road infrastructure was observed. Specifically, our results indicate that, on average, a one percentage point increase in high-quality road length leads to a 0.03-point reduction in the multidimensional poverty index score, while a one percent increase in the road network within a district corresponds to a 0.12 percentage point decline in incidence of multidimensional poverty. These findings reinforce the affirmative direct influence of transport infrastructure on welfare enhancement through geographical connectivity and integration. However, our findings also reveal negative indirect effects. The mobility of individuals from regions marked by significant social deprivation to less-deprived areas due to improved road infrastructure resulted in reduced welfare (increased incidence of multidimensional poverty) in the former regions. Notably, an average one percent increase in road length in neighbouring districts led to a 0.022-point increase in multidimensional poverty within a district. Moreover, the analysis of local and global spatial autocorrelation (spatial spillover effects) demonstrated their positive and statistically significant magnitudes. These findings resonate with Pakistan's development policy, emphasizing the importance of investing in road networks to bolster regional connectivity and economic cohesion. The substantial size of these global spillovers or network externalities, ranging from 0.37 to 0.82, underscores their significance.

Through rigorous estimation, we deduced that the social development landscape at the district level in Pakistan is shaped by the size of the transport infrastructure network, population density, urbanization, relative geographic coverage of districts, and the extent of network externalities. Remarkably, Punjab and Khyber Pakhtunkhwa exhibit significantly higher levels of social development at the district level compared to other provinces. Hence, investing in road infrastructure to ensure road access holds promise in directly alleviating social deprivation in the most impoverished districts of Pakistan. Based on these findings, it is evident that infrastructure, particularly transport infrastructure, plays a pivotal role in poverty reduction and mitigating social deprivation. This study delves further into the crucial role of public policy within the context of infrastructure development in Pakistan.

9.3 Existing Policies of Infrastructure and the way-forward

Concerning transport infrastructure and transportation services, the National Transport Policy (NTP) was officially promulgated in 2018, establishing a comprehensive framework encompassing key objectives in the realm of transport policy. These core objectives aimed to enhance overall connectivity and accessibility through an integrated transport network, elevate trade competitiveness, facilitate sustainable urban development, foster equitable growth across Pakistan, optimize utilization of existing and emerging transport infrastructure, enhance user-friendliness and consistency of transport services, ensure the safety of all transport users and their environs, and uphold environmental preservation and conservation principles. Moreover, the NTP outlined specific implementation arrangements, setting the stage for the establishment of a Cabinet Committee on Transport. This committee, constituted by the Prime Minister or Cabinet within three months of the Policy's endorsement, is mandated to coordinate the NTP's execution. Anticipated to convene biannually, with the possibility of more frequent meetings as needed, the Cabinet Committee on Transport holds the primary responsibility of supervising a wide array of matters. This encompasses the development, monitoring, and enforcement of the NTP, as well as the formulation and execution of the National Transport Master Plan. It is important to note, however, that as of now, the cabinet committee has not yet to be formed under the auspices of the NTP 2018.

Furthermore, within the context of the prevailing energy shortages in Pakistan, formulating an energy infrastructure policy becomes of paramount importance. To address this challenge, Pakistan has established operational bodies such as the Cabinet Committee on Energy (CCoE) and the Ministry of Energy, encompassing divisions dedicated to power and petroleum. These entities undertake the crucial responsibilities of managing energy aspects, including pricing, demand-supply dynamics, and regulatory oversight. Both federal and provincial authorities exercise their respective jurisdiction in this domain. To ensure the seamless provision of energy, the Government of Pakistan (GoP) has introduced several policies. Notable among these are "The National Power Policy 2013," "The Power Generation Policy 2015," and the "Alternative and Renewable Energy Policy 2019" (GoP, 2022). The core focus of "The National Power Policy 2013" is to cultivate an efficient and consumer-oriented power generation, transmission, and distribution system. The ultimate objective is the complete eradication of load shedding, reduction of average electricity generation costs, mitigation of transmission and distribution losses, augmentation of revenue collection, and alleviation of the delays that incur costs. The "Power Generation Policy 2015" seeks to facilitate private sector investments within the power sector. It presents incentives to the private sector for the establishment of new power generation units, primarily emphasizing the concept of least-cost power generation capacity. Furthermore, the "Alternative and Renewable Energy Policy 2019" aims to stimulate and promote the growth of renewable resources in the country. Its fundamental goal is to foster a supportive environment for renewable energy projects, leading to a substantial increase in the proportion of green energy capacity. The policy sets ambitious targets, aiming for a 20 percent share of green energy capacity by the year 2025 and an even more substantial 30 percent by 2030. Therefore, these energy policies represent critical endeavours by the Government of Pakistan to tackle the energy deficit challenge, enhance energy infrastructure, and promote sustainable energy solutions for the nation's future.

The energy sector in Pakistan confronts a series of challenges, notably the persistent issue of mounting circular debt. For instance, in the fiscal year 2013, the circular debt stood at approximately Rs 450 billion, but by March 2022, it had surged to Rs 2467 billion. This escalation can be attributed to several factors, including the heightened reliance on imported liquefied natural gas (LNG) due to the depletion of domestic natural gas reserves. Furthermore, Pakistan's energy input mix is predominantly reliant on imported resources, while the potential of hydro-energy generation remains underutilized, with only 25% of the installed capacity being hydro-based. The depreciation of the Pakistani Rupee has further exacerbated the energy generation import expenditure (GoP, 2022). Presently, the China-Pakistan Economic Corridor (CPEC) is driving the development of energy-infrastructure projects in the country. For instance, the Private Power and Infrastructure Board (PPIB) is facilitating thirteen power generation projects with a total capacity of 11,648 MW under the ambit of CPEC. This portfolio comprises four hydropower projects with a capacity of 3,428 MW, five projects based on Thar coal with a capacity of 3,960 MW, four projects using imported coal with a capacity of 4,260 MW, and a 660 kV High-Voltage Direct Current (HVDC) Transmission Line Project. Among these, three imported coal-based power projects with a cumulative capacity of 3,960 MW and a Thar coal-based power project with a capacity of 660 MW have already been commissioned. Additionally, the ±660 kV Matiari-Lahore High-voltage direct current (HVDC) Transmission Line Project, developed by the private sector, has commenced commercial operations as of September 1st, 2021. This project not only marks Pakistan's first HVDC transmission line but also represents a significant stride toward enhancing transmission capacity. Furthermore, the full operationalization of the Thar coal-based project is anticipated to yield substantial benefits by curbing the need for imported coal in energy generation. This, in turn, would contribute to a noteworthy reduction in the import bill associated with energy generation.

In 2015, a comprehensive telecommunication policy was promulgated with the explicit aim of amplifying the diffusion of ICT services throughout the nation. The policy's overarching objective was to align the regulation of this sector with the national pursuit of economic growth (GoP, 2015). Over time, the realm of telecommunication and postal services has witnessed marked improvement through the implementation of multiple initiatives involving public-private partnerships and private sector investments. A noticeable augmentation in teledensity has been achieved, accompanied by a concurrent rise in financial transactions attributed to the synergistic impact of technological advancements in secure real-time banking services, encompassing e-banking and online banking. This study's findings further underscore this trend, revealing a positive long-term productivity growth effect due to increased utilization of both financial and core infrastructure components.

Nonetheless, the potential for growth effects could be substantially amplified through more coordinated endeavours, driven by the implementation of the National Transport Policy (NTP) and the establishment of a dedicated cabinet committee. This concerted effort would span federal, provincial, and district administrative levels, thereby enhancing the strategic road development's networking effect by concurrently uplifting the quality of district and provincial infrastructure. Energy, as a vital input, yields both short-term and long-term impacts on economic growth. Hence, the enhancement of demand-supply management and the reduction of inefficiencies, particularly in curbing line losses, emerge as imperatives. Addressing these challenges could lead to a reduction in reliance on imported fuels for electricity and petroleum product generation. Investments in the energy sector, especially within the ambit of the China-Pakistan Economic Corridor (CPEC) and the Thar-Coal projects, hold the potential to mitigate future energy prices. However, the fruition of these positive impacts hinges on the concurrent reduction of sectoral inefficiencies and adept price management. Moreover, a recent (policy-brief) study by PIDE advocates for the adoption of a flat unit rate of electricity for all users as a strategic move to bolster the revenues generated from electricity sales. Yet, considering the looming consequences of climate change, transitioning from thermal to hydro energy sources becomes an imperative in shaping Pakistan's future long-term economic growth trajectory.

Moreover, the cost of communication services holds paramount importance in expanding reach and ensuring access to modern infrastructural services. Enhanced communicational services hold the potential to bridge gaps, contingent upon pricing mechanisms. According to experts, Pakistan's internet bundle costs significantly exceed those of neighbouring countries like India. Therefore, an adept pricing mechanism emerges as a critical policy tool to stimulate demand, utilization, and penetration of core infrastructural services within the country.

The social sectors of healthcare and education play an instrumental role as drivers of economic growth. In Pakistan, the decentralization of the health and education sectors to the provinces occurred under the 18th Amendment to the 1973 Constitution. Recently, the National Health Vision 2025 has been formulated, aligning with the World Health Organization's health system framework. This vision is structured around six thematic pillars encompassing health financing, health service delivery, human resources for health, health information systems, governance, essential medicines and technology, and cross-sectoral linkages. This policy framework is poised to guide the trajectory of national health planning and health sector governance in the future.

Furthermore, the study findings underscore the pivotal role of both the physical infrastructure for healthcare and the utilization of healthcare services in driving long-term productivity growth in Pakistan. Consequently, an extensive array of coordinated policy measures is warranted at both the national and provincial levels. These measures are envisaged to enhance the quality of healthcare services delivery, especially within the realm of public sector healthcare organizations. Equally critical is the expansion of the outreach of healthcare and vaccination services, achieved by ensuring financially feasible and physically accessible quality healthcare infrastructure and services.

Similarly, the establishment of infrastructure for educational services assumes a paramount role in nurturing the human capital development process. In Pakistan, although there has been an overall increase in the utilization of educational services, this upswing remains suboptimal in the face of a burgeoning population. Thus, ensuring the presence of quality infrastructure and accessible educational services is imperative for Pakistan's economic advancement. Notably, the challenge of deprivation and limited access to educational services, particularly among the female population, underscores the gravity of the issue. In this context, integrated efforts employing modern information and communication technologies (ICT) and other communication tools are essential to amplify access to productive resources, thereby uplifting the overall educational landscape.

Furthermore, following the post-revolutionary era of banking and telecommunication industry deregulation, the utilization of financial infrastructure has witnessed substantial growth. This upswing has made a notable contribution to the economy through enhanced integration and the facilitation of transactions. The integration of embedded e-commerce and telecommunication services, along with the convenience of e-banking, has further augmented the economy's efficiency and streamlined the distribution of goods and services. However, it remains essential to underscore that achieving financial inclusion is a pivotal factor in nurturing Pakistan's long-term economic growth trajectory. The nation still faces the imperative of intensifying efforts to broaden financial inclusivity and elevate economic participation, particularly among the marginalized segments of society.

In conclusion, this study accentuates the profound importance of diverse forms of infrastructure at the national level in fostering sustained economic growth over the long term. Both the tangible stock and effective utilization of infrastructural services emerge as pivotal drivers of productivity growth. To ensure the continuity of productive expansion in Pakistan, a dual focus on demand-side and supply-side policies becomes indispensable. Additionally, the attainment of balanced regional development necessitates strategic investments in the construction of both physical and human capital, alongside infrastructure augmentation. Moreover, the findings underscore the substantial impact of direct investments in transportation infrastructure in mitigating social and economic deprivation across Pakistani districts. As such, cohesive endeavours are warranted to ensure the seamless connection of local populations with the national highways. Through a strategic emphasis on and prioritization of infrastructure investment, Pakistan can effectively foster an all-encompassing and thriving economic landscape.

9.4 Limitations of the Study

This research endeavour aims to evaluate the aggregated impact of major components of infrastructure on economic growth and development. However, it is essential to acknowledge that the study possesses inherent limitations. The foremost limitation pertains to the absence of control over the influence exerted by the quality of diverse facets of infrastructural services. Regrettably, data pertaining to quality remains unavailable, underscoring the necessity for a more comprehensive investigation into this particular aspect. Moreover, existing scholarly discourse underscores the significance of institutional quality in conjunction with its impact on sustained economic growth. Remarkably, the analysis of drivers behind long-term economic expansion in Pakistan lacks adequate control over the quality of institutions, optimal practices, and implementation strategies associated with demand-side policies, relative to each infrastructure component. Another noteworthy constraint arises from the unavailability of gross domestic product data at the regional and district levels. Consequently, the assessment of growth effects in terms of productivity remains unquantifiable within the Pakistani context. In addressing this concern, the study employs GNI per capita (USD) indicators to facilitate a relative comparison of developmental states among provinces in Pakistan. Furthermore, it is imperative to emphasize the demand for additional research regarding the impact of public infrastructure provisioning at the household level. Further inquiry into this subject is deemed necessary to glean a more comprehensive understanding of its ramifications.

References

Adil, F., & Jalil, A. (2020). Determining the financial inclusion output of banking sector of pakistan—supply-side analysis. *Economies*, *8*(2), 1–20. https://doi.org/10.3390/ECONOMIES8020042

Aghion, P., Howitt, P., Brant-Collett, M., & García-Peñalosa, C. (1998). *Endogenous growth theory*. MIT press.

Aghion, P., & Howitt, P. (2009). The Economics of Growth. *Economic Record* (Vol. 86, Issue 272). https://doi.org/10.1111/j.1475-4932.2010.00622.x

Agénor, P.-R., & Moreno-Dodson, B. (2012). Public Infrastructure and Growth: New Channels and Policy Implications. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2005043

Ahmed, K., Bhattacharya, M., Qasim, A., & Ahmed, N. (2021). Research in Transportation Economics Transport infrastructure and industrial output in Pakistan : An empirical investigation. *Research in Transportation Economics*, *August 2020*, 101040. https://doi.org/10.1016/j.retrec.2021.101040

Ahmed, R. (2016). *social infrastructure and productivity of manudfacturing firms-evidence from Pakistan* (No. 16–038).

Ali, I., & Pernia, E. M. (2003). Infrastructure and Poverty Reduction — What is the Connection ? *Asian Development Bank .ERD Policy Brief No. 13*, *13*.

Alkire, S., Foster, J., Seth, S., Santos, M. E., Roche, J. M., & Ballon, P. (2014). Multidimensional Poverty Measurement and Analysis. In *OPHI Working Paper* (Vol. 82). https://doi.org/10.2139/ssrn.2564702

Alvarez-Áyuso, I. C., & Delgado-Rodriguez, M. J. (2012). High-capacity road networks and spatial spillovers in spanish regions. *Journal of Transport Economics and Policy*, *46*(2), 281–292.

Álvarez, I. C., Barbero, J., & Zofío, J. L. (2016). A spatial autoregressive panel model to analyze road network spillovers on production. *Transportation Research Part A: Policy and Practice*, *93*, 83–92. https://doi.org/10.1016/j.tra.2016.08.018

Andersson, M., Urban, G., & Karlsson, C. (2007). Spatial Economie System and Productivity in a Human Capital. *Annales d'Économie et de Statistique*, *87/88*, 125–143. https://doi.org/10.2307/27650045

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin, L. (2007). Spatial econometrics in RSUE: Retrospect and prospect. *Regional Science and Urban Economics*, *37*(4), 450–456. https://doi.org/10.1016/j.regsciurbeco.2006.11.009

Arrow, K. J., & Kurz, M. (1970). Optimal Growth with Irreversible Investment in a Ramsey Model. *Econometrica*, *38*(2), 331–344.

Aschauer, D. A. (1989a). Does public capital crowd out private capital? *Journal of Monetary Economics*, *24*(2), 171–188. https://doi.org/10.1016/0304-3932(89)90002-0

Aschauer, D. A. (1989b). Is public expenditure productive? *Journal of Monetary Economics*, *23*(2), 177–200. https://doi.org/10.1016/0304-3932(89)90047-0

Bajar, S. (2013). The Impact of Infrastructure Provisioning on Inequality. *United Nations*, *December*, 1–12. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2018/07/1-2.pdf

Bajar, S., & Rajeev, M. (2016). The Impact of Infrastructure Provisioning on Inequality in India: Does the Level of Development Matter? *Journal of Comparative Asian Development*, *15*(1), 122–155. https://doi.org/10.1080/15339114.2016.1150188

Barrett, P., Treves, A., Shmis, T., Ambasz, D., & Ustinova, M. (2019). *The Impact of School Infrastructure on Learning: A Synthesis of the Evidence*. International Bank for Reconstruction and Development, The World Bank.

Barro, R. J. (1990). Government Spending in a Simple Model of Endogeneous Growth. *Journal of Political Economy*, *98*(5, Part 2), S103–S125. https://doi.org/10.1086/261726

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. In *The Quarterly Journal of Economics* (Vol. 106, Issue 2).

Barro, R. J., & Sala-i-martin, X. (1992). Convergence. *Journal of Political Economy*, *100*(2), 223–251. https://www.jstor.org/stable/2138606

Belotti, F., Hughes, G., & Piano Mortari, A. (2017). Spatial Panel Data Models Using Stata. *SSRN Electronic Journal*, *17*(1). https://doi.org/10.2139/ssrn.2754703

Bluyssen, P. M. (2017). Health, comfort and performance of children in classrooms – New directions for research. *Indoor and Built Environment*, *26*(8), 1040–1050. https://doi.org/10.1177/1420326X16661866

Boarnet, M. G. (1998). Spillovers and the locational effects of public infrastructure. *Journal of Regional Science*, *38*(3), 381–400. https://doi.org/10.1111/0022-4146.00099

Bom, P. R. D., & Ligthart, J. E. (2009). *How Productive is Public Capital? A Meta-Regression Analysis*. *August*.

Bom, P. R. D., & Ligthart, J. E. (2014). What have we learned from three decades of research on the productivity of public capital ? *Journal of Economic Survey*, *28*(5), 889–916. https://doi.org/10.1111/joes.12037

Bridge, G., Özkaynak, B., & Turhan, E. (2018). Energy Research & Social Science Energy infrastructure and the fate of the nation : Introduction to special issue. *Energy Research & Social Science*, *41*(March), 1–11. https://doi.org/10.1016/j.erss.2018.04.029

Calderón, C., & Servén, L. (2004). *The Effects of Infrastructure Development on Growth and Income Distribution*.

Canning, D. (1999).. *The World Bank Policy Research Working Paper, 2246*, *November*, 20. http://documents.worldbank.org/curated/en/563251468739159840/Infrastructures-contribution-to-aggregate-output

Canning, D., & Pedroni, P. (2008). *I*nfrastructure , long-run economic growth and causality tests for cointegrated panels. The Manchester School, *76*(5), 504–527.

Cantos, P., Gumbau-Albert, M., & Maudos, J. (2005). Transport infrastructures, spillover effects and regional growth: Evidence of the Spanish case. *Transport Reviews*, *25*(1), 25–50. https://doi.org/10.1080/014416410001676852

Cellini, R. (1997). Implications of Solow's Growth Model in the Presence of a Stochastic Steady State. *University of Bologna Bologna, Itlay*, *19*(1), 135–153.

Chatterjee, S., & Turnovsky, S. J. (2012). Infrastructure and inequality. *European Economic Review*, *56*(8), 1730–1745. https://doi.org/10.1016/j.euroecorev.2012.08.003

Chen, Z., & Haynes, K. E. (2015). Spatial Impact of Transportation Infrastructure: A Spatial Econometric CGE Approach. *SSRN Electronic Journal*, *April*, 1–460. https://doi.org/10.1007/978-3-319-07305-7

Chin, M. S., & Chou, Y. K. (2004). Modelling Social Infrastructure and Economic Growth. *Australian Economic Papers*, *43*(2), 136–157. https://doi.org/10.1111/j.1467-8454.2004.00221.x

Chudik, A., Pesaran, M. . H., & Elisa, T. (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal*, *14*(1), C45–C90. https://doi.org/10.1111/j.l

Cohen, J. P. (2010). The broader effects of transportation infrastructure: Spatial econometrics and productivity approaches. *Transportation Research Part E: Logistics and Transportation Review*, *46*(3), 317–326. https://doi.org/10.1016/j.tre.2009.11.003

Cohen, J. P., & Paul, C. J. M. (2004). Public infrastructure investment, interstate spatial spillovers, and manufacturing costs. *Review of Economics and Statistics*, *86*(2), 551–560. https://doi.org/10.1162/003465304323031102

Czernich, N., Falck, O., Kretschmer, T., & Woessmann, L. (2011). Broadband Infrastructure and Economic Growth. *Economic Journal*, *121*(552), 505–532. https://doi.org/10.1111/j.1468-0297.2011.02420.x

Debarsy, N., Ertur, C., & LeSage, J. P. (2012). Interpreting dynamic space-time panel data models. *Statistical Methodology*, *9*(1–2), 158–171. https://doi.org/10.1016/j.stamet.2011.02.002

Démurger, S. (2001). Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China? *Journal of Comparative Economics*, *29*(1), 95–117. https://doi.org/10.1006/jcec.2000.1693

Devarajan, S., Swaroop, V., & Zou, H. F. (1996). The composition of public expenditure and economic growth. *Journal of Monetary Economics*, *37*(2), 313–344. https://doi.org/10.1016/S0304-3932(96)90039-2

Elburz, Z., Nijkamp, P., & Pels, E. (2017). Public infrastructure and regional growth: Lessons from meta-analysis. *Journal of Transport Geography*, *58*, 1–8. https://doi.org/10.1016/j.jtrangeo.2016.10.013

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction : Representation , Estimation , and Testing. *Econometrica*, *55*(2), 251–276.

Ertur, C., & Koch, W. (2007). Growth , Technological Interdependence And Spatial Externalities: Theory And Evidence. *Journal of Applied Econometrics*, *1062*(August), 1033–1062. https://doi.org/10.1002/jae

Esfahani, H. S., & Ramírez, M. T. (2003). Institutions, infrastructure, and economic growth. *Journal of Development Economics*, *70*(2), 443–477. https://doi.org/10.1016/S0304-3878(02)00105-0

Fan, S., & Chan-Kang, C. (2005). Road Development, Economic Growth and Poverty Reduction in China. In *International Food policy Research Institute*.

Fan, S., Hazell, P., & Thorat, S. (1999). *Linkages Between Government Spending, Growth and*

Poverty in Rural India, IFPRI research Report 110.

Felipe, J. (2001). Aggregate Production Functions and the Measurement of Infrastructure Productivity : A Reassessment. *Eastern Economic Journal*, *27*(3), 323–344.

Ferreira, F. H. G. (1995). *Wealth Distribution Dynamics with Public-Private Capital Complimentarity*.

Finance, M. of. (2020). *Pakistan's Economic Survey 2019-20*.

Fourie, J. (2006). Economic infracstructure: A review of definitions, theory and empirics. *South African Journal of Economics*, *74*(3), 530–556. https://doi.org/10.1111/j.1813-6982.2006.00086.x

Frank, A. G. (1959). Industrial capital stocks and energy consumption. *The Economic Journal*, *69*(273), 170–174.

Garcia-Milà, T., & McGuire, T. J. (1992). The contribution of publicly provided inputs to states' economies. *Regional Science and Urban Economics*, *22*(2), 229–241. https://doi.org/10.1016/0166-0462(92)90013-Q

Getachew, Y. Y. (2010). Public capital and distributional dynamics in a two-sector growth model. *Journal of Macroeconomics*, *32*(2), 606–616. https://doi.org/10.1016/j.jmacro.2009.12.009

Getis, A. (2010). Spatial Autocorrelation. In *Handbook of Applied Spatial Analysis*. https://doi.org/10.1007/978-3-642-03647-7

Ghani, E., & Din, M. U. (2006). The impact of public investment on economic growth in Pakistan. *Pakistan Development Review*, *45*(1), 87–98. https://doi.org/10.30541/v45i1pp.87-98

Gianpiero, T. (2010). Public infrastructure : definition , classification and measurement issues. *Munich Personal RePEc Archive*, *25850*.

Glomm, G., & Ravikumar, B. (1994). Public investment in infrastructure in a simple growth model. *Journal of Economic Dynamics and Control*, *18*(6), 1173–1187. https://doi.org/10.1016/0165-1889(94)90052-3

Goldsmith, R. W. (1969). *Financial Structure and Development*. Yale Uni. Press, New Haven.

GoP. (2018). *National Transport Policy of Pakistan 2018, Planning Commission, Ministry of Planning & Development & Reform*.

Hall, R. E., & Jones, C. I. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others. *The Quarterly Journal of Economics*, *114*(1), 83–116.

Hansen, B. E. (1992). Tests for parameter instability in regreessions with I(1) processes. *Journal* *of Business and Economic Statistics*, *10*(3), 45–59. https://doi.org/10.1198/073500102753410381

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*(3), 607–668. https://doi.org/10.1257/jel.46.3.607

Haughwout, A. (2005). Infrastructure and Social Welfare in Metropolitian America. *Economic Policy Review*, *7*(3), 1–16.

Hina, H., & Qayyum, A. (2019). Effect of financial crisis on sustainable growth: Empirical evidence from Pakistan. *Journal of the Asia Pacific Economy*, *24*(1), 143–164. https://doi.org/10.1080/13547860.2019.1573453

Holtz-Eakin, D., & Schwartz, A. E. (1994). Infrastructure in a Structural Model of Economic Growth. *National Bureau of Economic Research*, *51*(1), 51.

Holtz-Eakin, D., & Schwartz, A. E. (1995). Spatial productivity spillovers from public infrastructure: Evidence from state highways. *International Tax and Public Finance*, *2*(3), 459–468. https://doi.org/10.1007/BF00872777

Hu, A., & Liu, S. (2010). Transportation, economic growth and spillover effects: The conclusion based on the spatial econometric model. *Frontiers of Economics in China*, *5*(2), 169–186. https://doi.org/10.1007/s11459-010-0009-0

Hulten, C. R., Bennathan, E., & Srinivasan, S. (2006). Infrastructure, externalities, and economic development: A study of the Indian manufacturing industry. *World Bank Economic Review*, *20*(2), 291–308. https://doi.org/10.1093/wber/lhj007

Hussain, A. (1996). An Alternative Planning Perspective for Regional Growth and Development in Pakistan. *The Lahore Journal of Economics*, *1*(1), 1–14. https://doi.org/10.35536/lje.1996.v1.i1.a1

Iimi, A. (2003). Efficiency in the Pakistani Banking Industry : Empirical Evidence after the Structural Reform in the Late 1990s. *The Pakistan Development Review , Spring 2003 , Vol . 42 , No . 1*, *42*(1), 41–57.

Isaev, A. (2015). Transport Infrastructure and Economic Growth: Spatial Effects. *Spatial Economics*, *3*(43), 57–73. https://doi.org/10.14530/se.2015.3.057-073

J. B. Ramsey. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Society, Royal Statistical*, *31*(2), 350–371. https://www.jstor.org/stable/2984219

Jarque, C. M., & Bera, A. K. (1987). *A Test for Normality of Observations and Regression*

Residuals. *55*(2), 163–172.

Javid, M. (2019). Public and private infrastructure investment and economic growth in Pakistan: An aggregate and disaggregate analysis. *Sustainability (Switzerland)*, *11*(12). https://doi.org/10.3390/SU11123359

Jiang, X., Zhang, L., Xiong, C., & Wang, R. (2016). Transportation and Regional Economic Development: Analysis of Spatial Spillovers in China Provincial Regions. *Networks and Spatial Economics*, *16*(3), 769–790. https://doi.org/10.1007/s11067-015-9298-2

Knight, M. D., Loayza, N., & Villanueva, D. (1993). Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach. *IMF Working Papers*, *40*(3), i. https://doi.org/10.5089/9781451947052.001

Krueger, A. B., & Lindahl, M. (2001). Education for Growth : Why and For Whom ? *Journal OfEconomic Literature*, *XXXIX*(December), 1101–1136.

Kumbhakar, S. C., & Mavrotas, G. (2005). Financial Sector Development and Productivity Growth. In G. (eds) Guha-Khasnobis, B., Mavrotas (Ed.), *Financial Development, Institutions, Growth and Poverty Reduction.* Palgrave Macmillan, London. https://doi.org/https://doi.org/10.1057/9780230594029_11

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Lau, L.-S., Ng, C.-F., Cheah, S.-P., & Choong, C.-K. (2019). Panel Data Analysis (Stationarity, Cointegration, and Causality). In *Environmental Kuznets Curve (EKC)*. Elsevier. https://doi.org/10.1016/b978-0-12-816797-7.00009-6

Lesage, J. P. (1999). "Theory and Practice of Spatial Econometrics." In *https://www.researchgate.net/publication/266218273* (Vol. 10, Issue 3). https://doi.org/10.1080/17421772.2015.1062285

Levine, R. (1997). Financial Development and Economic Growth : Views and Agenda. *Journal of Economic Literature*, *35*(2), 688–726.

Mankiw, G. N., Romer, D., & Weil, D. N. (1992). A CONTRIBUTION TO THE EMPIRICS OF ECONOMIC GROWTH. *Journal of Policy Modeling*, *36*(2), 257–272. https://doi.org/10.1016/j.jpolmod.2014.01.001

Marinho, E., Campelo, G., França, J., & Araujo, J. (2017). Impact of infrastructure expenses in strategic sectors for Brazilian poverty. *EconomiA*, *18*(2), 244–259. https://doi.org/10.1016/j.econ.2017.01.002

McKinnon, R. I. (1973). *Money and Capital in Economic Development*. Brookings Institution,Washington, DC.

McMahon, W. W. (1979). Externalities in Education. *In Economics of Education: Research and Studies*. Pergamon Books Ltd. https://doi.org/10.1016/B978-0-08-033379-3.50028-0

Mehmood, K., Baig, S., Li, X., Ghanem, O., & Hanif, S. (2020). Research in Transportation Economics Causality between transportation infrastructure and economic development in Pakistan : An ARDL analysis. *Research in Transportation Economics*, *August*, 100974. https://doi.org/10.1016/j.retrec.2020.100974

Ministry of Finance. (2021). *Pakistan's Economic Survey 2020-21*.

Ministry of Finance. (2020). *Pakistan's Economic Survey 2019-20*.

Ministry of Finance. (2019). *Pakistan's Economic Survey 2019-20*.

Ministry of Finance. (2018). *Pakistan's Economic Survey 2018-19*.

Munnell, A. H. (1990). Why has productivity growth declined? Productivity and Public Investment. *New England Economic Review*, *January/February*, 3–22. http://ideas.repec.org/a/fip/fedbne/y1990ijanp3-22.html

Nawaz, S., & Mangla, I. U. (2021). The economic geography of infrastructure in Asia: The role of institutions and regional integration. *Research in Transportation Economics*, *88*(January), 101061. https://doi.org/10.1016/j.retrec.2021.101061

Nijkamp, P. (1986). Infrastructure and regional development: A multidimensional policy analysis. *Empirical Economics*, *11*(1), 1–21. https://doi.org/10.1007/BF01978142

Nijkamp, P., & Capello, R. (2009). Introduction: regional growth and development theories in the twenty-first century – recent theoretical advances and future challenges. *Handbook of Regiona Growth and Development (Edited by Roberta Capello, and Peter Nijkamp)*.

No Title. (1956). *70*(1), 1884513. https://doi.org/10.2307/1884513

NTDC. (2021). *Power System Statistics 45th Edition* (Issue March).

NTRC. (2020). *Digitalization of Roads Directory in the Country*.

Ottaviano, G. (2008). Infrastructure Investment, Growth and Cohesion. In *Infrastructure and economic geography: An overview of theory and evidence in Infrastructure Investment,* *Growth and Cohesion: The Economics of Regional Transport Investment (edited by) Hubert Strauss: Vol. 13 No 2*.

Pakistan, G. of. (2018). National Financial Inclusion Strategy (NFIS). In *Economic Survey of Pakistan* (Issue October).

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, *57*(6), 1361. https://doi.org/10.2307/1913712

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/jae.616

Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. In *Journal of Econometrics* (Vol. 68, Issue 1). https://doi.org/10.1016/0304-4076(94)01644-F

Phillips, A. P. C. B., & Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, *58*(1), 165–193.

Phillips, P. C. B. (1993). Fully Modified Least Squares and Vector Autoregression. In *Cowles Foundation Discussion Paper No. 1047* (Vol. 63, Issue 5). http://www.jstor.org/stable/2171721

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335–346. https://doi.org/10.1093/biomet/75.2.335

Ploberger, W., & Krämer, W. (1992). The Cusum Test with Ols Residuals Author. *Econometrica*, *60*(2), 271–285.

Pritchett, L. (2000). The tyranny of concepts: CUDIE (cumulated, depreciated, investment effort) is not capital. *Journal of Economic Growth*, *5*(4), 361–384. https://doi.org/10.1023/A:1026551519329

Ralls, D. (2019). “Becoming co-operative” – challenges and insights: repositioning school engagement as a collective endeavour repositioning school engagement as a collective endeavour. *International Journal of Inclusive Education*, *0*(0), 1–15. https://doi.org/10.1080/13603116.2019.1629159

Redding, J., & Turner, M. A. (2015). Transportation Costs and the Spatial Organization of Economic Activity. In *Handbook of Regional and Urban Economics, vol. 5* (1st ed., Vol. 5, pp. 1339–1398). Elsevier B.V. https://doi.org/10.1016/B978-0-444-59531-7.00020-X

Rietveld, P. (1994). Spatial economic impacts of transport infrastructure supply. *Transportation Research Part A*, *28*(4), 329–341. https://doi.org/10.1016/0965-8564(94)90007-8

Rives, J. M., & Heaney, M. T. (1995). Infrastructure and local economic development. *Regional Science Perspectives*, *25*(1).

Rizvi, K. A. S., Naqvi, B., & Tanveer, F. (2017). Mobile Banking: A Potential Catalyst for Financial Inclusion and Growth in Pakistan. *The Lahore Journal of Economics*, *22*(Special Edition), 251–281. https://doi.org/10.35536/lje.2017.v22.isp.a11

Sahoo, P., & Dash, R. K. (2012). Economic growth in South Asia: Role of infrastructure. In *Journal of International Trade and Economic Development* (Vol. 21, Issue 2). https://doi.org/10.1080/09638191003596994

Sanchez-Robles, B. (1998). Infrastructure Investment and Growth : Some Empirical Evidence. *Contemporary Economic Policy*, *Vol. XVI*, 98–108.

SBP. (2007). *The Pakistan Infrastructure Report*. *44*(0), 1–68. http://www.sbp.org.pk/departments/ihfd/InfrastructureTaskForceReport.pdf

Schneider, M. (2002). *Do School facilities Affect Academic Outcomes?* National Clearing house for educational facilities,.

Sen, A. (2003). On unit-root tests when the alternative is a trend-break stationary process. *Journal of Business and Economic Statistics*, *21*(1), 174–184. https://doi.org/10.1198/073500102288618874

Shaw, E. S. (1973). *Financial Deepening in Economic Development*. Oxford University Press.

Shi, Y., Guo, S., & Sun, P. (2017). The role of infrastructure in China's regional economic growth. *Journal of Asian Economics*, *49*, 26–41. https://doi.org/10.1016/j.asieco.2017.02.004

Shin, Y., Yu, B., & Greenwood-Nimmo, M. (2014). Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework. In *Festschrift in honor of Peter Schmidt: Econometric methods and applications* (pp. 281-314.).

Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *Journal of Finance and Data Science*, *4*(2), 71–89. https://doi.org/10.1016/j.jfds.2017.11.001

Siddiqui, Rehana. (2004). Energy and Economic Growth in Pakistan. *The Pakistan Development Review*, *43*(2), 175–200.

Siddiqui, Rizwana, & Pant, K. P. (2007). Quantifying the impact of development of the transport sector in Pakistan. *Pakistan Development Review*, *46*(4). https://doi.org/10.30541/v46i4iipp.779-802

Snieska, V., & Simkunaite, I. (2009). Socio-economic impact of infrastructure investments. *Engineering Economics*, *3*(63), 16–25. https://doi.org/10.5755/j01.ee.63.4.11638

Straub, S. (2008). Infrastructure and Growth in Developing Countries : Recent Advances and Research Challenges. *World Bank POLICY RESEARCH WORKING PAPER*, *January*, 54. https://doi.org/10.1596/1813-9450-4460

Straub, S., Charles, V., & Michael, W. (2006). Infrastructure and Economic Growth in East Asia. *Multinationals and Economic Growth in East Asia*, *April*. https://doi.org/10.4324/9780203966693

Stulz, R. (2001). Does Financial Structure Matter for Economic Growth? A Corporate Finance Perspective. In *Financial Structure and Economic Growth edited by Asli Demigruc-Kunt and Levine Ross*.

Timilsina, G., Hochman, G., & Song, Z. (2020). nfrastructure, Economic Growth, and Poverty: A Review. In *World Bank POLICY RESEARCH WORKING PAPER* (Issue May).

UNDP. (2016). Multidimensional Poverty in Pakistan. In *Pk.Undp.Org* (Vol. 9, Issue 3, p. 67). https://doi.org/10.1021/ja964229j

Valickova, P., Havranek, T., & Horvath, R. (2014). FINANCIAL DEVELOPMENT AND ECONOMIC GROWTH: A META-ANALYSIS. *Journal of Economic Surveys*, *00*(0), 1–21. https://doi.org/10.1111/joes.12068

Venables, A. J. (2010). New economic geography. *Written for Palgrave Dictionary of Economics*, 109–137. https://doi.org/10.4324/9780203310694_chapter_5

Weitzman, M. L. (1970). Optimal Economies Growth in the with Scale of Creation Overhead. *The Review of Economic Studies*, *37*(4), 555–570. https://doi.org/https://doi.org/10.2307/2296485

Xueliang, Z. (2013). *Social Sciences in China Has Transport Infrastructure Promoted Regional Economic Growth ?— With an Analysis of the Spatial Spillover Effects of Transport Infrastructure*. *December 2014*, 37–41. https://doi.org/10.1080/02529203.2013.787222

Yu, N., de Jong, M., Storm, S., & Mi, J. (2013). Spatial spillover effects of transport infrastructure: Evidence from Chinese regions. *Journal of Transport Geography*, *28*, 56–66. https://doi.org/10.1016/j.jtrangeo.2012.10.009

Yu, N., De Jong, M., Storm, S., & Mi, J. (2012). The growth impact of transport infrastructure investment: A regional analysis for China (1978-2008). *Policy and Society*, *31*(1), 25–38. https://doi.org/10.1016/j.polsoc.2012.01.004

Zhang, Y. F., & Ji, S. (2019). Infrastructure, externalities and regional industrial productivity in China: a spatial econometric approach. *Regional Studies*, *53*(8), 1112–1124. https://doi.org/10.1080/00343404.2018.1563678

Zulfiqar, K., Chaudhary, M. A., Aslam, A., & Aslam, M. (2016). Financial Inclusion and Its Implications for Inclusive Growth in Pakistan. *Social Review*, *54*(2), 297–325.