Discussion Paper No. 16-038

Social Infrastructure and Productivity of Manufacturing Firms – Evidence from Pakistan

Riaz Ahmed

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 16-038

Social Infrastructure and Productivity of Manufacturing Firms – Evidence from Pakistan

Riaz Ahmed

SOCIAL INFRASTRUCTURE AND PRODUCTIVITY OF MANUFACTURING FIRMS - EVIDENCE FROM PAKISTAN

RIAZ AHMED*

South Asian Institute (SAI), Heidelberg University, Germany,

Centre for European Economic Research (ZEW), Mannheim, Germany, and

Department of Management Sciences, University of Turbat, Pakistan

April 20, 2016

Abstract

Does investment in social infrastructure affect the productivity of manufacturing firms in developing countries? To test this question, I empirically investigate the impact of social infrastructure indicators at district level on firm productivity using firm level data from Pakistan. I split my sample into rural and urban regions to capture the effect of regional disparities in investment in social goods while controlling for a potential selection bias from firms' decision to locate in regions with better infrastructure equipment. My findings reveal that indicators of health and education are positively and significantly related to firm level productivity in manufacturing industries in Pakistan. However, these results hold for urban districts only. For rural regions, both health and education show a negative impact on firm productivity.

Keywords: Firm Productivity, Social Infrastructure, Health and Education, Pakistan

JEL classification: D24, H51, H52, I15, I25

This research is part of a PhD project and has received funding from the Higher Education Commission of Pakistan via The German Academic Exchange Service (DAAD). I am thankful to my supervisors, Prof. Stefan Klonner (South Asian Institute (SAI), Heidelberg University) and Dr. Christian Rammer (ZEW) for their valuable comments and supervision. Furthermore, I thank ZEW for providing research facilities.

* Riaz Ahmed, ZEW, L 7, 1 68161 Mannheim, Germany.

Tel: +49(0)157-80328539; Email: riazmazar_786@hotmail.com

1. Introduction

The role of government spending for economic growth and development in general, and its impact on the performance and productivity of business firms in particular has been a main issue in economic analysis for a long time. Most of the works try to quantify this relationship based on cross country macroeconomic data. The problem with macroeconomic data is that they do not capture the true heterogeneity among regions and economic actors since investment in social infrastructure, e.g. education or health, tends to be concentrated on some regions which hence may grow faster than others. There are rather few studies which focus on micro-level (firm-level) data and most micro-level analyses focus on a limited number of government activities. For instance, Datta (2012), Ghani, Goswami, and Kerr (2013), Holl (2013), and Rothenberg (2012) study the relationship between government spending and firm performance by looking only at road and transportation infrastructure and Allcott, Collard-Wexler, and O'Connell (2014), Fisher-Vanden, Mansur, and Wang (2015), Moyo (2013), and Pasha, Ghaus, and Malik (1989) focus on energy provision to study this relationship. There are few studies which used cross country firm level data. One of the most relevant papers is by Dollar, Hallward-Driemeier, and Mengistae (2005) who analyzed why countries like China, India, Bangladesh, and Pakistan which started their journey at almost the same level of growth in 1990 but after a decade their growth rates depict a different picture with 7.2%, 4.2%, 2.7%, and 1.3% respectively.

To understand the relationship between spending on social infrastructure and productivity, one should consider infrastructure data at a more disaggregated level such as provincial, county or district level. In a recent study, Kneller and Misch (2014) use data from South Africa and estimate a micro-economic production function model in order to disentangle the effect of specific public spending on firm productivity. Their findings are limited for two main reasons, however. First, they use monetary values of public spending which rather poorly represents the physical output of social infrastructure in the context of developing counties. Second, they use provincial level data which do not accurately represent the infrastructure that is actually available for each firm owing to the heterogeneity of infrastructure distribution within a province (given the vast extent of provinces in South Africa).

A similar study was done by Sumarto and De Silva (2014) to test the effect of health and education indicators at district level on regional (provincial) growth and poverty. There are two difference between this work compared to the study by Kneller and Misch (2014). First, Sumarto and De Silva employ district level education and health indicators rather than monetary values as proxies for human capital (social infrastructure). Second, they use regional GDP as dependent variable rather than firm productivity or output. A serious concern with both papers is that they do not take into account a potential selection bias of more productive firms that may choose to locate in growing regions or regions with a better infrastructure provision.

This study builds upon the micro-level studies cited above and tries to extend these studies for the case of Pakistan along three lines: First, by considering different types of social infrastructure, secondly by using more disaggregated regional data at a district level, and thirdly by addressing potential selection bias. Since districts in Pakistan typically represent an area which often correlates with a firm's geographical spread of business activities (in terms of supply and sales networks), the availability and quality of the social infrastructure in a district should represent the infrastructure relevant to a firm's business activities. I go a step further in details to test the effect of rural and urban infrastructure on firm productivity at district level. In this research my main hypothesis is that *the variation of firm productivity is positively and significantly related to the variation of social infrastructure across districts*. By social infrastructure I mean the institutions and activities which provide skills, knowledge, health, and other benefits to the labor force of a country. Here I specifically take health and education indicators as social infrastructure.

Firm location decisions are one of the sources for a potential selection bias. In this study, I encounter this potential bias by including the Ellison-Glaeser (EG) index (see sections 4 and 6 for more details) in the regression analysis to capture geographical concentration assuming that firms choose their location based on advantages of external economies of scales due to industrial agglomeration. Firms from the same industry located near to each other (localization) or near to firms in related (supplier or user) industries (urbanization) in the same region get benefits from industrial agglomeration in the form of low transaction cost, sharing of knowledge and information, availability of skilled labor and cheap inputs, ease access to customer markets etc. But on the other hand, it might be the case that higher agglomeration might cause problems for industry in the form of higher land values, higher transportation and other input costs, tough competition in output markets etc. To control the latter case, I also include the squared term of the EG index which may capture negative agglomeration effects if industrial agglomerations exceed a certain size.

To investigate the impact of investment in social infrastructure on productivity of business enterprises I take Pakistan as a case study. I use the 2005/06 Census of Manufacturing Industries (CMI 2005-06) data set along with other data sources for social infrastructure at district level.

In short to medium run, my findings show that government itself or by the help of private partner can increase firm productivity in manufacturing industries by investing in health and education infrastructure. For instance, increasing a 1% in the net enrolment of primary schools in urban regions within a district seemingly increase firm output by 0.37%. In addition, the positive coefficient of health infrastructure indicates that a 1-percentage-point increase of urban population satisfied with services and facilities provided by the basic health units[1] (BHU) in a given district is apparently positively associated with an increase in firm output in manufacturing industries by 0.58%. Based on a variety of sensitivity checks, my findings remain robust.

[1] According to the Development Statistics books of Khyber Pakhtunkhwa, A Basic Health Unit (BHU) is provided to serve about 5,000 to 10,000 populations. A Basic Health Unit is responsible for comprehensive health care which, among other things, includes midwifery, child care, immunization, diarrhea diseases, malaria control, child spacing, mental and school health services with in its areas.

The rest of the paper is structured as follows. In the next section I present some of the growing literature based on theoretical as well as empirical works on the effect of social infrastructure on productivity and firm output. First I put a general view of this relationship followed by some macro level empirical evidence and a summary of the few empirical works conducted at the micro level. In section 3, I discuss the regional disparities of social infrastructure i.e. education and health indicators in Pakistan. In Section 4, I elaborate the modeling framework and estimation strategies of the empirical analysis. Section 5 describes the data I use, and Section 6 presents and discusses the empirical findings, followed by some concluding remarks.

2. Effect of Social Infrastructure on Productivity and Output of Firm

Literature: General Public Infrastructure Spending

The role of public infrastructure for economic growth has been studied since the book "Wealth of Nations" by Adam Smith. Research on the issue rejuvenated when in a series of empirical seminal papers, Aschauer (1989a, 1989b, 1989c), using time series data for the U.S and some other developed countries, found a very strong effect of public infrastructure capital on total factor productivity. There are many studies looking at the role of government spending for economic growth and productivity. In this literature review, the focus is on recent empirical works related to the effects of social infrastructure on firm productivity. Aschauer's works were validated by Munnell (1990, 1992). She found the same relation as Aschauer but the magnitude of the cofficients she calculated were slighlty lower than those of Aschauer. Barro (1991) decomposed public spending into productive and non-productive and found a negative relationship between non-productive government consumption and the growth rate.

More specifically focusing on developing countries, a worthwhile study by Devarajan, Swaroop, and Zou (1996) investigated the relationship of public expenditure and economic growth based on the composition of public expenditure came. They find that changing the composition of public spending, i.e. shifting resources from infrastructure to current expenditure surprisingly yields a positive effect on growth. They conclude that in developing country there is a misconception about investing more in infrastructure at the cost of current expenditure.

Whether public spending in monetary values or public infrastructure capital stocks in physical units are better proxies for the infrastructure investment has been a controversially debated issue. In this regard, Sanchez-Robles (1998) found no conclusive results for monetary values of public spending, however he found a positive effect of public infrastructure on growth per capita when considering capital stock in physical units.

Another issue when estimating the relationship between public spending and growth is the problem of endogeneity and reverse causality. To account for a potential endogeneity, Calderón and Servén (2004) apply GMM estimators and use lagged variables of some exogenous variables such as population density, urban population, and labor force as instrument variables and find statistically and economically significant effects of public infrastructure on economic growth.

Literature examined the impact of infrastructure on productivity by using different models and techniques. The Mexican economy demonstrated this impact very clearly from the early 1980s to the 1990s. Government reduced investment in public infrastructure from 12% to 5% of GDP which dramatically reduced the output growth rate and created macroeconomic imbalances (Mamatzakis, 2007). Mamatzakis also suggested that the Mexican economy can raise productivity in the private sector by prioritizing public expenditure in infrastructure rather than consumption expenditure.

Literature: Health

A number of studies investigated the role of health infrastructure in different ways, some of them relying on expenditure data while other stressing that the final outcome may be similar when using health expenditures or health indicators. For instance, Lorentzen, McMillan, and Wacziarg (2008) investigated the relationship of health indicators and economic growth in countries in southern Africa based on country level data from 1960 to 2000 and interestingly they found a very strong negative relationship between adult mortality and economic growth. Life expectancy is one of the health indicators which many scholars have been using. Among them, Bloom, Canning, and Sevilla (2004) examined the impact of life expectancy of population on growth of output. Their results suggest that government investment in health infrastructure resulting in a one percentage point increase in life expectancy raises output by 4%. In another paper, Bloom and Canning (2005) found that increasing adult survival rates by one percentage point could increase labor productivity by 2.8 percent.

In a similar study, Weil (2007) first estimated a proxy for health infrastructure at the macro level by taking the estimates of height, adult survival rates, and age at menarche at the micro level and then investigated the effect of health on GDP per worker through a cross country analysis. He found a statistically significant positive relationship between health infrastructure and output which was, however, smaller than other cross country regressions found.

Recently, Dube, Phiri, and Bahmani-Oskooee (2015) study the relationship between health and growth by taking nutrition as health indictor based on data for South Africa for the years 1961-2013. They find a positive relationship between nutrition and economic growth. They also find a causal relationship which flows from nutrition to economic growth much stronger than in the other direction.

Literature: Education

While the theoretical literature strongly supports the view of positive externalities from education expenditure on growth and development of a country, empirically there are not many consistent and robust findings (see Dastidar, Mohan, & Chatterji, 2013). Beside these inconclusive results, reverse causality was observed in many cases. For instance, in China, Feng-ying (2013) observe a huge impact of government spending on education[2] on both short

[2] Measured as per worker education expenditure

run and long run economic growth, however, he found a relatively larger growth in the long run probably because of the accretion of human capital. But at the same time he also observed a reverse causality between public spending on education and economic growth. A recent similar study by Mekdad, Dahmani, and Louaj (2014) tested this hypothesis in Algeria and they also find a significant positive impact of Algerian government expenditure on education on economic growth during 1974-2012.

In contrast to this supporting evidence, in some countries' expenditure on education shows either a negative or no effect on economic growth. For instance, Farzanegan (2011) in reference to the Iranian economy found no effect of public spending on education on economic growth during the period of 1959–2007. He identified that one main reason behind this unproductive spending is the absence of effective political and economic institutions.

Literature: Health and Education

The paper by Baldacci, Clements, Gupta, & Cui (2008) suggests that governments in developing countries need to raise their public spending on both education and health simultaneously in order to achieving Millennium Development Goals (MDGs). To quantify the relationship between government spending, human capital and economic growth they used a simulation method which unleashes some very interesting findings for policy making purpose. For instance, to attain a 0.5 percentage point yearly growth, on average governments of developing countries need to raise the country's net enrollment rate from 90% to 99% and to shrink its child mortality rate (under five years) from 76 to 70[3]. And to get that level of education and health capitals, a country would have to expand its public spending on both education and health by a percentage point of GDP for each sector. But this could be possible only if governments of developing countries successfully retain other complementary policies as well; providing better administrative structures and controlling inflation for example.

This view is supported by Popa (2012) who explains the effect of social factors (health education, poverty and unemployment rate) on economic growth taking a panel data set from 2005 to 2009 in the framework of European countries. Her findings reveal a positive and significant effect of '*the expected year of schooling'* and '*life expectancy*' and a negative effect of '*population at risk'* and '*unemployment rate'* on the economic growth (measured by GDP per capita).

Literature: Pakistan

There is a large number of published and unpublished studies explaining the importance of education and health (and human capital in general) on output and total factor productivity in the context of the Pakistan economy. In the following I discuss those studies that are most relevant to my own study.

Ali and Ramay (2014) highlighted the role of human capital measured in average years of schooling[4] for total factor productivity (TFP) and output in the context of Pakistan taking country time series data from 1961 to 2013. They concluded that human capital has an impact

[3] The unit of mortality rate is death of children(under five years) per thousand children
[4] For more details on how to calculate human capital see Ali and Ramay (2014).

on output and total factor productivity both directly and indirectly. They found a positive and significant impact of human capital on both output and total factor productivity when including human capital directly in the regression models. To test for an indirect influence, they include an interacted term of the human capital variable with labor force and physical capital, respectively, and found a similar effect on output and TFP. They claimed that labor and capital could be utilized more productively in the presence of higher human capital.

Afzal, Farooq, Ahmad, Begum, and Quddus (2010) applied the method of Pesaran, Shin, and Smith (2001) to observe both long and short run relationships between school education and economic growth and found that the school enrolment ratio is positively and significantly related to GDP growth both in the short as well as in the long run. In the short run they also observed a reverse causality. In another country level study, Afzal, Arshed, and Sarwar (2013) analyzed time series data from 1971 to 2011 to test the relationship between human capital (health and education), inflation in food prices and economic growth in Pakistan. With respect to health and education they found a positive and significant impact of human capital on economic growth both in the short as well as in the long run. Since they found a reverse causality between education and economic growth they suggested to maintain stable economic growth and to spend more on education because both create a virtuous circle.

Country level data are only able to depict a general picture. For policy making purposes more disaggregated findings on the effects of investment in human capital on output and productivity in different sectors and industries and at the firm level would be needed. In this respect, Amjad, Ghani, ud Din, and Mahmood (2012) analyzed a firm level perception survey[5] which asked exporting firms about key hurdles and limitations when accessing international markets. Among other factors, they found lack of skilled labor to be one of the main factors for Pakistan's lagging internationalization performance, particularly with respect to a lack of training institutions and the low standard of education. The findings of this study were limited, however, by the small sample size (40 firms), the unbalanced distribution of firms across industries and regions, and the neglect of other relevant factors that may explain firms' decisions and ability to engage in international activities.

Based on the existing literature one may conclude that there is a evidence for a significant positive impact of investment in social infrastructure on economic output and productivity in Pakistan. To the best of my knowledge, no a study has been conducted yet that analyzes this impact at the firm level. The purpose of this study is to fill this gap and to provide more concrete findings for policy making issues in developing countries.

3. Manufacturing Industries, Social Infrastructure, and Fiscal Policy in Pakistan

Pakistan, the second largest country in south Asia and the sixth largest in the world in terms of population, is still in its infantry stage of development. The World Economic Forum (2014) defines a number of factors required for development of an economy depending upon the

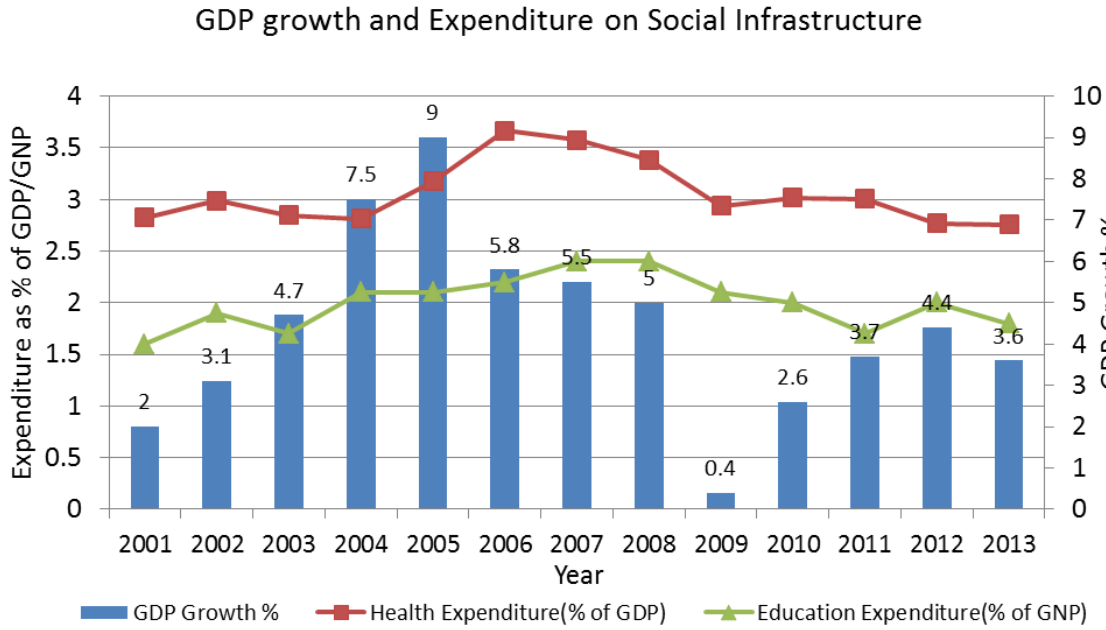
[5] The survey was conducted by the Lahore Chamber of Commerce and Industry and the Pakistan Institute of Development Economics.

existing stage of development. According to the WEF ranking, Pakistan finds itself in the lowest stage of development. Following the WEF, Pakistan should focus on developing the basic determinants, namely institutions, infrastructure, macroeconomic environment and education. Yet, based on several criteria, Pakistan is one of the poorest performers in the South Asia region. In this section, I briefly explain the role of the manufacturing sector in Pakistan and present a comparative descriptive analysis of south Asian nations to provide a broader picture of the Pakistani manufacturing in the region. In addition, I will briefly describe the conditions of social infrastructure (focusing on health and education) both at the South Asia regional level and at the district level within Pakistan. Finally, I will shed some light on the fiscal policy and budgetary constraints of the Pakistani economy that may help to understand how these constraints limit the country in investing on social infrastructure.

In a developing country, the manufacturing sector could be considered as an engine for economic growth and development. It is the sector that creates more opportunities for employment to accommodate urbanized population, boosts export of value added products, and generates more revenue for government in the form of corporate income tax and sales tax (Sanchez-Triana et al., 2014). This role of manufacturing for growth and development of an economy has been realized for a long time. But in Pakistan, partly because of government's fiscal policy which highly taxes manufacturing while exempting the agriculture sector (which contributes about 21% in GDP in 2014-15) and parts of the service sector (which contributes about 59% in GDP in 2014-15), investors have been discouraged to invest in manufacturing industries (Sanchez-Triana et al., 2014). As a consequence Pakistan's real annual GDP growth rate was around 4.4% during the last one and half decades which is rather low for a country with a young and rapidly increasing population. Figure 1 shows that the GDP growth rate in Pakistan fluctuated drastically during the 2000s starting at 2% in 2001 and reaching a peak point of 9% in 2005. The growth rate dramatically dropped to 0.4% in 2009 and recovered to an average of 3.6% in the following four years. This "boom-bust cycles of GDP growth" has been particularly pronounced for the manufacturing sector (Sanchez-Triana et al., 2014).

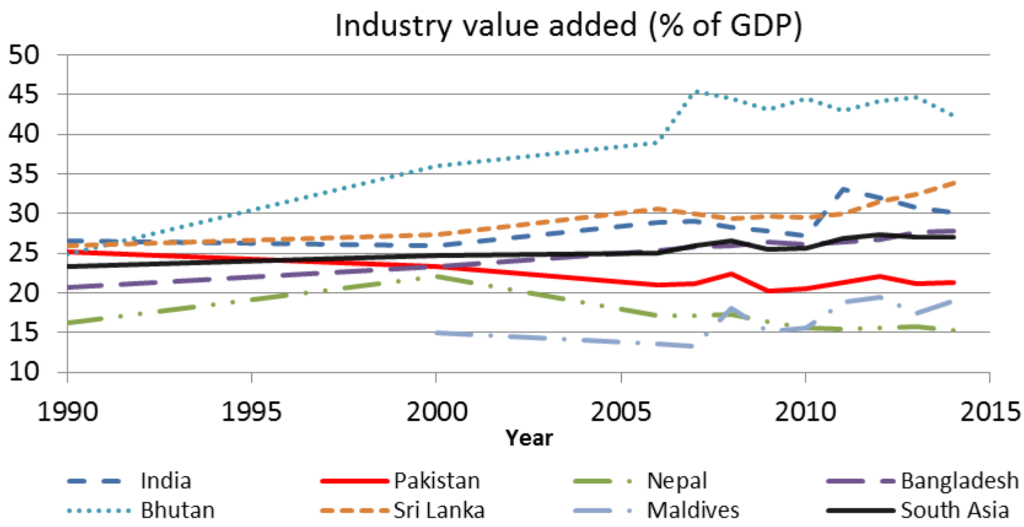
Over the past decade, the service sector in Pakistan grew massively and clearly gained in significance for the Pakistani economy. In 1999, the shares of agriculture, manufacturing and services in Pakistani GDP were about 25.4%, 25.7%, and 48.9%, respectively. Until 2015, the share of agriculture declined by 4 percentage points and that of manufacturing by 4.8 percentage points while the share of service sector increased dramatically and reached at 58.8% in 2015. While agriculture contributes 20 to 23% of GDP, its share in tax revenue is only about 1 to 1.5%. This peculiar sectoral shift is one of the motivations for this study Contrary to other South Asian countries, Pakistan is moving on an opposite track. According to Dutz and O'Connell (2013), sectoral reallocation in Pakistan remained far behind its neighboring countries over the periods 1980-2008. In Pakistan, sectoral reallocation contributed about 15 percent to the average annual growth in labor productivity, compared to 16 percent in Sri Lanka, 25 percent in India, 40 percent in Bangladesh, and 91 percent in Nepal.

Figure 1: Real GDP growth and expenditure on social infrastructure in Pakistan, 2001-2013



Source: Economic Survey of Pakistan (various issue)

Figure 2: Share of Industrial Value Added in GDP in South Asia, 1990-2014



Source: World Bank Development Data

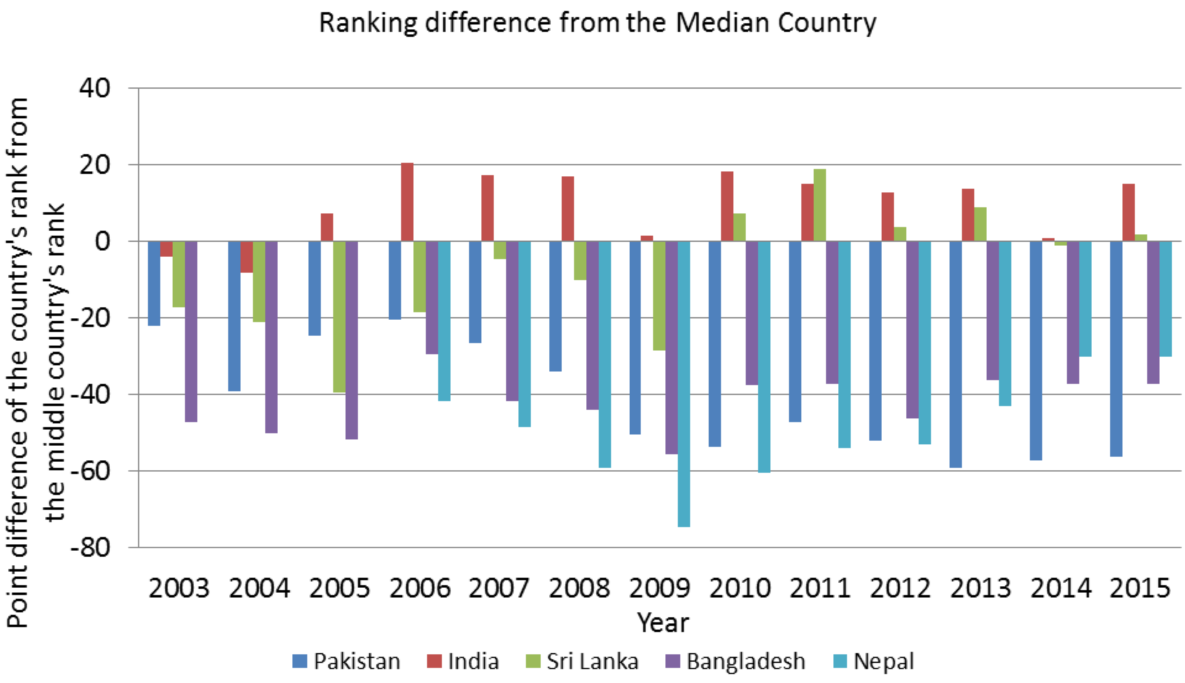
While a declining share of manufacturing in GDP is a phenomenon common to all developed countries, such a development may be seen as critical for a developing country since manufacturing industries have a high potential for creating productive jobs, employing skilled labor, and generate export income, while much of the growth in the service sector is on trade, low-productivity services and the informal economy. The share of manufacturing value added in total GDP is hence a relevant measure to evaluate the progress a developing economy is making towards higher productivity levels. Based on this measure the industrial sector in Pakistan has been performing very poorly. For instance, Pakistan's share of industrial value added in GDP is well below to the regional average and has been continuously decreasing over time since 2000 (see Figure 2). Furthermore, the value for Pakistan was same as in India, Sri Lanka, and Bhutan and above to the regional average by two percentage points in 1990. But in 2014, Pakistan remained far below to its competing countries to gain the share of industrial value added in GDP. During that period, Pakistan was the only country in the region that has lost 4 percentage points while other countries - Bhutan, Bangladesh, Sri Lanka, India, Maldives, and Nepal - manufacturing has gained in percentage points. Likewise, the regional average share of industrial value added in GDP that was below to the Pakistani value in 1990, the former surpassed the latter by six percentage points in 2014.

The weak performance of manufacturing in Pakistan was mainly due to irregular fluctuations of large-scale manufacturing while the growth rate of small-scale manufacturing has been increasing quite stable over time (Sanchez-Triana et al., 2014). Sanchez-Triana et al. (2014) also emphasize the importance of the textile industry and call it a backbone for economic growth and development of the Pakistani economy. Pakistan successfully positioned itself among the top five cotton producing countries in the world, contributing 9% to the world cotton production, 8.5% to Pakistani GDP, 38% to total manufacturing labor force in Pakistan, 46% of total manufacturing output in Pakistan, and 54% to Pakistan's total exports. More specifically Pakistan has comparative advantages in the garment industry. The future of the garment industry, if properly managed, is seen bright because global garment demand is continuously growing. This sector is not only able to absorb parts of the huge unemployed labor force in Pakistan but can also boost the economy through exports. A main comparative advantage of Pakistan in this specific industry are low labor costs, which are three times lower than in the Chinese garment industry and 50% lower than in India (Sanchez-Triana et al., 2014).

Sanchez-Triana et al. (2014) also identified some challenges of the manufacturing sector in Pakistan. First, manufacturing is highly concentrated in either low value-added consumer products such as food, beverages etc. or in products for which the world market is continuously declining. As a consequence, Pakistani manufacturing is less attractive to foreign direct investment. Second, the country is also facing the problem of a huge inter-sectoral investment shift from the industrial sector to the service sector. For example, in total investment, the portion of transport and communication investment doubled (12% to 24%) between 2000 and 2010, while the portion of industrial sector investment declined from 38% to 20%.

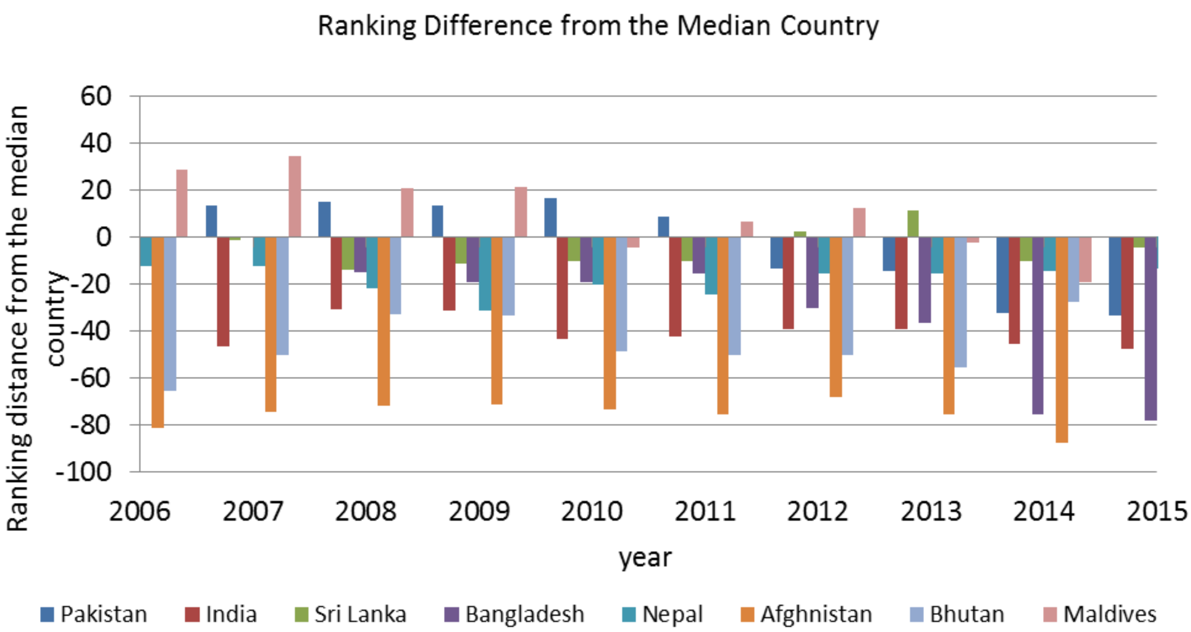
There may be other factors that cause a discouraging trend in FDI in manufacturing. One is related to the failure of the Pakistani government to create a conducive business environment. In order to bring foreign direct investment into the country, global comparative indices such as the Global Competitiveness Index (GCI) or the World Bank's Doing Business (DB) index provide important information for investors. Pakistan is gradually deteriorating its position in these indices. For instance in the GCI, Pakistan held 73rd position (out of 101 countries) of the 2003 and dropped to 126th position (out of 140 countries) in 2015. Its rank was also far below that of competing countries in South Asia. Figure 3 shows the point difference of each South Asian country's rank to the median country's rank in the GCI (2003 to 2015). Pakistan has been moving further away from the average country while India and Sri Lanka performed better than the average country over the past couple of years.

Figure 3: Global Competitive Index Ranking in South Asia, 2003-2015



Source: World Bank Economic Forum: Global Competitiveness Report (various issues)

Figure 4: Ease of Doing Business Ranking in South Asia, 2006-2015



Source: World Bank: Doing Business (various issues)

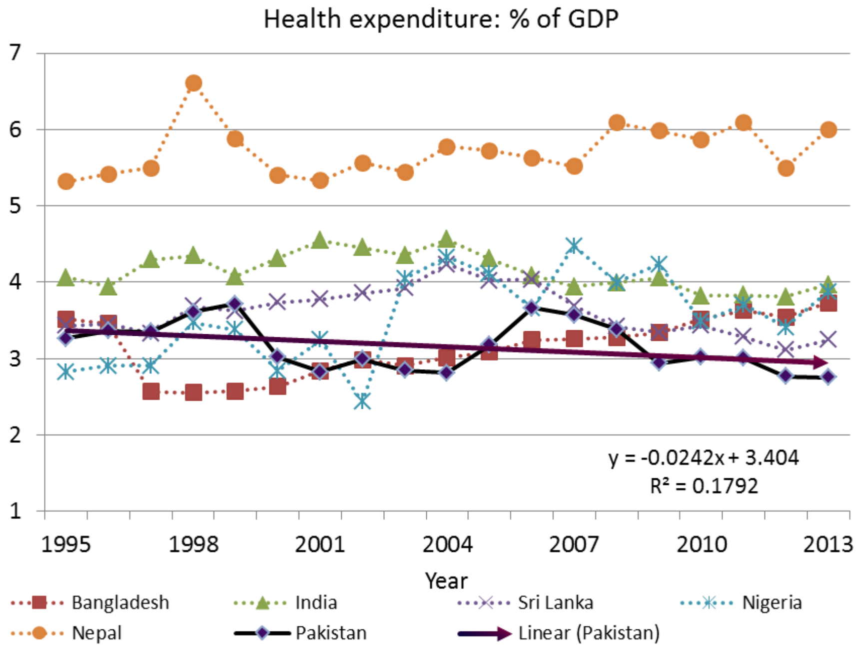
Pakistan is also continuously losing its ranking position in the *Doing Business* (DB) index. While Pakistan's rank (128th out of 189 countries) in the 2015 index is still above the ranks of some other South Asian countries like Afghanistan (183), Bangladesh (173) and India (142), its rank has been significantly declining over time (see in Figure 4). It started it journey from 11 points above the average rank in 2006 and ended up with 33 point below the rank of the average country in 2015. Though India's rank is below that of Pakistan, India's position was stable over the past decade (see in Figure 4).

Pakistan, the most rapidly urbanizing country in the South Asia, is facing some serious issues associated with the provision and quality of infrastructure (Ahmed et al., 2013). The population of Pakistan is heavily shifting from rural to urban regions. Unlike other South Asian countries, Pakistan is not reaping the benefits of this urbanization. Urbanization that creates agglomeration economies by increasing industrial productivity in urban areas can contribute to better social and infrastructure services, and can generate skilled labor force for more productive industries. Agglomeration economies, furthermore, can help employees in raising their skills by exchanging information and knowledge. From this perspective, urbanization could be an opportunity for developing countries to achieve productivity growth provided that the infrastructural and social problems associated with rapidly growing cities can be resolved (Fuller & Romer, 2014). However, urbanization in Pakistan rather contributes to regional inequality since some regions are growing much faster than others. It is essential to identify those social infrastructures which could become either contributing factors or hurdles for the performance of the business sector and hence for the development of the entire economy. With the present study, I want to shed some light on the impact of social infrastructure on productivity in manufacturing industries in Pakistan.

For achieving the desired level of growth and development and in order to switch from low productive to high productive economic activities, focusing on the development of the manufacturing is a promising strategy for Pakistan. To achieve that end, Pakistan will have to flourish its labor force (human resource) by investing in education and health. But to date, Pakistan widely failed to attain a level of social infrastructure which could boost the industrial sector of the economy. To test whether Pakistan is serious towards in providing social services to the population in general, I compare social indicators on health and education at both the South Asia Regional (SAR) level and at the district level within Pakistan.

At the SAR level, I take some stylized facts on education from the performance reports which were produced each year after the agreement of 164 nations on the goals of the "*Education for All*" at the World Education Forum in Dakar, Senegal in April 2000. Ironically, Pakistan's commitment towards the targets for each indicators established in the Forum was inadequate. For instance, Pakistan committed to achieve a literacy rate of 88% in 2015, but only achieved 58% by 2012 (Planning Comission GoP, 2013). One of the main reasons for this worse situation is the low tax collection. In the Monitoring Report 2012, UNESCO (2012) mentioned that the world second largest numbers of out-of-school children, after Nigeria, live in Pakistan. In other words, one out of every twelve out-of-school children in the world lives in Pakistan. Using the latest Monitoring Report of UNESCO (2015), I compare Pakistan with India in some dimensions of development targets and achievements. For instance, in 2000, the number of out-of-school children was about 16.9 million for India and 8.8 million for Pakistan. By 2012, India has successfully reduced its out-of-school children by 92% while Pakistan has only achieved 39% over the entire twelve-year period. The situation is even worse for female education: for every 10 out-of-primary school girls of the world, one lives in Pakistan. Furthermore, gender inequality in term of access to education is also a serious issue in Pakistan. Comparing with India, there were 32 fewer girls for every 100 boys enrolled in primary school in Pakistan in 2000 and by 2012, a 13-points gap in their enrollment still remained unfulfilled. On the other hand, India achieved the target and even crossed the line; started from 16 fewer girls enrollment in primary school in 2000 and ended up with 2 more girls' enrollment in 2012. With regard to student to teacher ratio, this ratio increased from 33 (in 2000) to 44 (in 2012) in Pakistan while it reduced from 40 to 35 in India during the same period, indicating that Pakistan hired fewer teachers to fulfill teaching requirements. Another way to test the seriousness of the Pakistani government towards education is to look at the amount of money that was spent on education. Despite the Pakistani government's promise on spending up to 6% of GDP on education in order to achieve the goals of the "*Dakar Framework for Action, Education for All"* (UNESCO, 2015), education spending as a percentage of GDP was only around 2% during 2001 to 2013 (see Figure 1).

Figure 5: Health Expenditure Comparison: Selected South Asian Countries and Nigeria, 1995-2013



Source: World Health Organization: Global Health Expenditure Database

Likewise, indicators associated to the health sector are equally depressing. For instance, data from the Global Health Expenditure Database (World Health Organization) shows that the situation of health care in Pakistan is not very impressive when comparing it with other South Asian countries (or Nigeria, a country of similar size to Pakistan). In fact, public spending on the health sector has been reduced over time (see Figure 5). Recently, health expenditure as a percentage of GDP even fell below 3% in Pakistan. Beside this low investment in social infrastructure, the unequal distribution of public spending across regions within Pakistan makes the situation even worse.

Regional discrepancies in term of accessing to social services, which is estimated by the Human Development Index[6] (HDI), reveals that there is unequal distribution of resources on public services across Pakistan (see Figure A1 in Appendix). The HDI values at district level indicate that most of the districts in Punjab are well above the average value while the districts in Balochistan, with a few exemptions, are well below the average. In other words, none of the Punjabi districts is among the 35 least developed districts of Pakistan.

Figure A2 maps the accessibility to education across Pakistani district and further highlights the discrepancies across regions. Beside accessibility to education, quality of education equally matters for economic growth and development. Using ASER[7] 2014 data, I map the districts according to the percentage of population (aged 6-16 years) that could solve a 2-digit arithmetic division question[8] (see Figure A3 for more details). Again, Balochistan is not only below in terms of access to education but also in terms of quality of education. More precisely, there are several districts in Balochistan where less than 10% of the pupils could do math 2-digit division problem in contrast to Punjab where even in rural areas the score is above 50%.

Regional inequality of social services is not only present in the education sector. The indicators for the health sector also depict similar stories. In fact most of the time education and health sectors go hand in hand. Districts having higher values for education indicators are more likely to provide better health facilities. Just to mention one example: In fiscal year 2012-13, less than 15% of children in Balochistan were born in government hospitals/RHCs/BHUs[9], whereas the figure for Punjab is much higher (see Figure A4).

In order to better understand the constraints and limitations of public spending on social infrastructure at the district level, I briefly summarize some main features of fiscal policy in Pakistan based on the study by Ahmed (2013, pp. 67-111). Federal, provinces and districts are the three tiers of the government system in Pakistan. The Federal government collects more than 80% of the national revenue and spends more than 65% of total state expenditure. The second tier is provincial governments which collect about 17-18% of national revenue and the third tier is local (district) governments which collect only 2-3%. In Pakistan there are some severe issues in fiscal policy. First, the tax to GDP ratio is very low, around 10 to 11%, and did not change much since 1976. One of the major reasons behind this low ratio is the tax exemption of agriculture and of parts of the service sector because most of the advantages from these sectors go to the ruling elites in the country. Second, spending on defense is one of the escalating expenditures, which limits policy makers to allocate resources in social sectors such as health, education, transport infrastructure etc. Almost one third of total budget is

[6] District HDI, calculated by Jamal, H., & Khan, A. J. (2007) for 2005.

[7] ASER stands for The Annual Status of Education Report, a household based survey that provides the schooling status of children aged 3-16 years residing in all rural and few urban districts of Pakistan.

[8] In theory there are other quality measures but owing to different education standards in Pakistan other measures are less comparable. For instance reading stories or learning level in different languages such as English, Urdu, Sindhi, Pashto etc. are not common in all Pakistan. But the language of mathematics is the same for all of them.

[9] RHCs stands for rural health centers and BHUs stands for Basic health units.

consumed on defense and military expenditure which grew on average by 10% every year since the 1950s. There is little doubt that higher investment in social sectors would contribute to tackle the country's main challenges, e.g. the rapidly growing population, grinding poverty, lack of education and health facilities, low literacy rate, and poor infrastructures. In addition, expenditure on social, economic and community services are not only low but were also reduced over time. Third, to finance the burgeoning defense expenditure, the country hugely relies on public debt on the expense of low tax revenue. As a result, the high amount of debt service is another serious fiscal challenge for Pakistani government. Fourth, high level of corruption is another major issue. Politicians and bureaucrats indulge in corruption and they support to one another and allocate resources on their own interest rather than delivering services to the general public. General administration expenditure is around 1% of GDP and not being utilized properly because of huge corruption. In a nutshell, investment in social services and infrastructure, and poverty alleviation projects are negatively affected by high expenditure for non-development, defense and debt service.

4. Modeling the Relationship between Social Infrastructure and Firm Productivity

For investigating the impact of social infrastructure on productivity and growth empirically, a number of previous studies employed the framework of the Cobb-Douglas production function (Aziz et al., 2008; Benhabib & Spiegel, 1994; Brandt et al., 2012; Mekdad et al., 2014; and Ping, 2005). In this study I will follow the most recent empirical paper by Sumarto and De Silva (2014), who used an augmented growth model in which one can add human capital as an additional factor to the production function. But I use a slightly different approach since I have cross section firm level data at hand. I use a firm level production function in which firm uses fixed capital (K), labor (L) and material inputs (M) to produce gross output (Y). In addition, a firm's output may also be affected by the available social infrastructure (S) which can provide spillovers to a firm's productivity through a better educated and healthier workforce. Using the familiar Cobb-Douglas production function, the basic model is given by

[1] $$Y = \alpha L^{\beta 1} K^{\beta 2} M^{\beta 3} S^{\lambda} e^{u}$$

In this model, α represents total factor of productivity, $\beta_{1,}$ β_2, β_3, and λ are output elasticities of labor, capital, materials and social infrastructure, respectively and u is the usual error term. Taking natural logarithms of the above equation one gets

[2] $$y= a + \beta_1 l + \beta_2 k + \beta_3 m +\lambda s + u$$

For empirically estimating model [2], further refinements are required. While *y, l, k* and *m* can be measured directly through accounting data, there is no single measure for *s*. Social infrastructure will instead be measured by different indicators ($Z \in n=1,...,N$) representing various aspects of the amount and quality of education and health investment by the government and private sector. These indicators are measured for each district *k* in which a firm *i* is located and also include district-specific control variables. Since a firm's location decision may have been influenced by the availability and quality of social infrastructure, I follow Ellison and Glaeser (1997) and include their index (EGI) to measure geographical concentration of industries. This index is measured at the level of the industry sector *j* a firm *i* belongs to. Finally, it is useful to add further control variables ($X \in m=1,...,M$) representing a firm's ownership and market orientation. The equation to be estimated reads as follows:

$$[3] \qquad y_{ijk} = a + \beta_1 l_i + \beta_2 k_i + \beta_3 m_i + \lambda s_k + \sum_n \delta_n Z_{nk} + \varphi EGI_j + \sum_m \gamma_m X_{mi} + \varepsilon_{ijk}$$

where *Z* represents district-specific variables and ε_{ijk} is an idiosyncratic error term. Model [3] will be estimated by using ordinary least square (OLS) and feasible generalized least square (fGLS) estimators.

Controlling for a potential selection bias resulting from the firms' location decision is essential since more productive firms may deliberately choose to locate in a region with superior infrastructure supply. This could be particularly the case for firms that require well trained staff, which is often the case for more productive firms. Including a measure of industry agglomeration can control for such a bias. The EGI used for this purpose is defined as follows:

$$[4] \qquad EGI_j = \frac{\sum_{i=1}^{M}\left(S_{jk} - X_k\right)^2 - \left(1 - \sum_k X_k^2\right) H_j}{\left(1 - \sum_k X_k^2\right)\left(1 - H_j\right)}$$

with S_{jk} being the *j*th share of employment in district *k* and X_k respresents total manufacturing employment in district *k*. H_j is the Herfindahl index which measures the *i*th plant level concentration of employment in industry *j*. It is defined by:

$$[5] \qquad H_j = \sum_{i=1}^{M} S_{ij}^2$$

where S_{ij} is the share of the *i*th plant in total employment of the industry *j*.

There are mainly two approaches to measuring health and education infrastructure. The first approach is to take the amount of money invested on health and education by governments, potentially also including private investment. Hong and Ahmed (2009) as well as Kneller and Misch (2014) follow this approach in their econometric models. In the context of developing countries, monetary values of government spending do have some drawbacks, however, since these values may not necessarily represent effective investment. For instance, one of the most obvious complications in public spending is corruption. But there is also another more compelling reason why monetary values of spending on infrastructure could not be good proxies because the impact of current spending will not immediately transfer into an output relevant for the productivity of businesses. In other words, business firms could get benefits from investment in health institutions or schools only as soon as these investments actually start rendering services. Pritchett (1996), for example, estimated that a dollar spent on social infrastructure in developing countries transfers into a capital stock worth less than 50 cents.

The second approach is to measure investment in social infrastructure by output indicators of social infrastructure. Sumarto and De Silva (2014), for example, used four health-related indicators (prevalence of waterborne disease, skill birth attendance, immunization rate, and incidences of self-medication as proxies to measure health system) and three education indicators (gross secondary school enrolment ratio, share of population with secondary education, and years of schooling). Years of schooling (Bils & Klenow, 2000; Krueger & Lindahl, 2000), student teacher ratio, and literacy rate are some of among others widely used as proxies for education in literature. Weil (2007) used three indicators as proxies for health: height, adult survival rates, and age at menarche. McDonald and Roberts (2002) used the mean years of total education as a proxy for education capital and infant mortality as well as life expectancy at birth as proxies for health capital.

There are a number of mechanisms which justify the main arguments of spending in infrastructure in monetary values could not provide services to the private firms and even sometimes they reduce firm productivity. For example, if a government spends a huge amount of money on a big highway project which is in construction over a long period of time will rather become a hurdle for businesses to access markets or for plant sites to deliver goods during the construction period. Obviously, in the long run the situation will entirely change. The same argument can be made for infrastructure investment in health and education. It takes several years to build a government hospital in developing countries. During the time funds are allocated to build hospitals, there are no effects of this investment in terms of an improved health situation as long as the hospital has not started operation. Once it would be completed the ex-ante probability of getting benefits from the investment will be very high. For this reason I use indicators on the existing physical social infrastructure in a given year rather than the amount of spending on that specific infrastructure.

Another serious concern relates to the issue that government spending on social infrastructure is an incomplete proxy for the existing social capital stock since it excludes the role of the private sector in the formation of social capital. In order to study the effects of social infrastructure on firm productivity holistically would require including private investment along with public investment in social infrastructure. For instance, a very big portion of investment in immunization programs in developing countries comes from the private sector and from foreign donors. From the viewpoint of a manufacturing firm it does not matter whether better social services that translate into better conditions for increasing productivity are financed from public or private funds. At the end of the day, it is the social infrastructure outcomes which matter for economic growth and development. By focusing on public spending in monetary values only, one cannot adequately capture the effects of the social infrastructure on firm productivity. Using output indicators related to social infrastructure will hence provide better measures for investigating the research question.

As demonstrated above, there are various indicators to measure the latent variables, health (*H*) and education (*E*). When choosing among these indicators, availability, reliability and comparability of data are important dimensions. Given the data situation at the district level in Pakistan, I employ two proxies for measuring health capital, the percentage of satisfied population with basic health units in 2005-06, and the percentage of children (aged 12 to 23 months) who were fully immunized.[10] For measuring education capital, I use the net

[10] Fully immunization means children who received vaccines of BCG, DPT1, DPT2, DPT3, Polio1, Polio2, Polio3 and measles.

enrolment in primary school and the literacy rate of people aged 10 and above. The second indicator for health is the share of the immunized population. The immunization rate is also reflecting private investment in health services since immunization programs are mostly funded by international donators and other agencies.[11] As the aim of immunization programs is to save children from diseases that can be prevented by vaccine, I use this indicator as proxy for health facilities rather than a direct input indicator. Intuitively, regions which are marginalized and are having lack of other basic facilities to get less attention for vaccination programs and poor families who are already suffering from health facilities, their children are less likely to get fully immunized.

For education infrastructure, the first proxy is the net enrolment in primary schools which is calculated by the number of students of the age group 5 to 9 who are enrolled in primary school divided by the total number of children of the same age in each district. The second indicator for education infrastructure is the literacy rate which has widely been used in literature as a proxy for education capital stock or human capital. I use the literacy rate for the population 10 years or older.

In order to investigate likely differences in the effects of health and education infrastructure in urban and rural areas, health and education indicators are not only measured at the district level, but also for urban and rural areas within each district. This specification is used to analyze whether the effects of investment in social infrastructure differ if these are made in urban or in rural areas.

To take into account district specific factors which may influence firm productivity across regions, I include a set of district level control variables. The first control variable is the log of district estimated population in number of people (*Pop*) in 2006 [12] and the percentage of urban population in districts in 1998. I assume that firms which are located in highly populated districts would be more productive than firms in districts with a low population. This is because in higher populated districts firm may have better opportunities to hire skilled labor force and more importantly it can serve a bigger customer market. As the population in log and the percentage population at district level are highly correlated, I will use them alternatively. The second district control variable is the number of kilometers of high-type/paved roads in a district per registered vehicle in that district (*RdRv*). This control variable is used as a proxy for the quality of other infrastructure available at the district level.

Another potential source for productivity differences across firms are factors such as geography, the political situation, demography etc. For example, provincial government policies towards its manufacturing industries, e.g. subsidies to output process, may alter productivity. To control for such effects I include province dummies (*Prov*) in all regressions. I also control for industry specific effects by including two-digit industry dummies (*Ind*).

[11] Financially Government of Pakistan is only contributed approximately 20% of the program and the program is being managed by the Expended Program on Immunization(EPI), the GAVI Alliance, the WHO, and UNICEF (Research and Development Solutions, 2012).
[12] The only latest census available is the census conducted in 1998.

A further regional control variable is the distance to Karachi as the main international port of Pakistan (Dis). This variable is intended to capture a likely advantage a firm might get from its location in or near to the port. Ahrend, Farchy, Kaplanis, and Lembcke (2014) in the context of cities in five OECD[13] countries found port cities are 2 to 4% more productive than other cities. Distance is measured between the capital city of a district and Karachi using longitude and latitude data and applying the Haversine formula[14]. Since all districts in the Sindh Province are located nearer to Karachi than districts in other provinces, the distance variable and dummy for the Sindh province are highly correlated; I show the distance effects only in the robustness section rather when excluding the Sindh dummy from the model.

I also include the Human Development Index (*HDI)* to control for the level of human development at district level. Since the HDI is highly correlated with the main variables of interest, health and education, I do not include the HDI in the main regressions but rather in the robustness check.

5. Data and Descriptive Statistics

In this paper I attempt to analyze the effect of investment in social infrastructure on firm productivity. To that end, I need both firm data on productivity and the main drivers of productivity within a firm, as well as district level data on social infrastructure. Firm level data are taken from the Census of Manufacturing Industries 2005-06 (CMI 2005-06). The CMI 2005-06 was conducted by the Statistical Office of Pakistan. The data set consists of 6,417 manufacturing establishments which are either registered or qualify for registration under the Factories Act 1934[15]. Manufacturing activities of establishments are classified at the 5-digit level of the Pakistan Standard Industrial Classification (PSIC) 2007, which is derived from the UN International Standard Industrial Classification ISIC Rev-3.1. In the CMI 2005-06 both financial and non-financial information such as production output at producer prices, employment, working hours, capital stock, raw material inputs etc. are covered in detail.

Infrastructure indicators are taken from the Pakistan Social & Living Standard Measurement (PSLM) surveys conducted by the Pakistan Bureau of Statistics every year since 2004-2005, and each alternative year it is being conducted at district level and province level respectively. Since the firm level data are available for the fiscal year 2005-06, I use district level indicators of health and education from the PSLM surveys 2004-05 and 2006-07[16] which are available on the official website of the Pakistan Bureau of Statistics. From these surveys I calculate the averages in order to represent information for the year 2005-06.

I use a human development index based on the estimates calculated by Jamal and Khan (2007) at district level in Pakistan for the year 2005. The estimated population in year 2006 comes from the provincial development statistic books and to calculate the percentage of

[13] OECD countries were Germany, Mexico, Spain, United Kingdom, and United States.

[14] The Haversine formula for distance between two points on earth in kilometers in is given by: Distance=ACOS(SIN(Lat_city1)*SIN(Lat_city2)+COS(Lat_city1)*COS(Lat_city1)*COS(Long_city1-Long_city2))*6371

[15] The survey was conducted by the Pakistan Bureau of Statistics (PBS) with the help of the Provincial Directorates of Industries and Provincial Bureaus of Statistics (BOS) under sections 9 & 10 of the General Statistics Act 1975 and section 5 & 6 of the Industrial Statistics Act 1942.

[16] PSLM data at district level are not available for year 2005-06

urban population in district I take the urban and total population of district from the population census 1998. In addition, I use the number of registered vehicles by road kilometer to control for transport infrastructure provision. Data on these variables are taken from the provincial development statistics books.

Table 1: Firm and District Level Variable: Description and Sources

Variable	Description	Sources
Output (Y)	Log of total output at producer price in Pakistani Rupees	CMI 2005-2006
Labor (L)	Log of average number of employees (both production and non-production) during the fiscal year 2005-2006.	CMI 2005-2006
Capital (K)	Log of net book value of capital (land, building machinery, vehicles, and other equipment etc.) in Pakistani Rupees. It is calculated by taking fixed assets on July 1st 2005 plus purchases of fixed assets during the year plus fixed assets produced for own use minus sales of fixed assets during the year.	CMI 2005-2006
Materials (M)	Log of cost of materials in Pakistani Rupees consumed during the fiscal year2005-2006. It includes fuels, electricity, raw materials, chemicals & dyes, packing materials, spare parts, lubricants, and others both locally purchased and imported.	CMI 2005-2006
Government	State ownership dummy. Equals 1 if the firm owned by Pakistani state.	CMI 2005-2006
Foreign	Ownership fully or partially owned by foreigners. Equal 1 if either government or private firms fully or partially owned by foreigners.	CMI 2005-2006
Import	Firms involved in importing materials abroad. Equal 1 if the value of firm imported material is greater than zero.	CMI 2005-2006
EG Index	It is calculated based on data available in CMI 2005-2006 data set	CMI 2005-2006
Estimated Population 2006	It is log of estimated population (inhabitants) of districts in fiscal year 2005-2006 based on the annual growth rate of previous population censuses.	Provincial Development Statistics
% of Urban Population 1998	It is the percentage share of population distribution in urban region in the Population Census 1998.	Population Census 1998
Number of Km per Registered Vehicle [a)]	It is the ratio of number of kilometers of high type roads in district to the number of registered vehicles in the district.	Provincial Development Statistics
HDI 2005[b)]	HDI is been estimated according to the definition of UNDP which consists three components of HDI; health, education, and income. Life expectancy at birth was used to measure health, enrolment rate and literacy rate were used to estimate education and for income since there was no district level data so they apply a method by taking the provincial level data to estimate district level income. For more detail please refer to original source.	Jamal and Khan (2007)
Distance to Karachi	It is the distance between the given districts to Karachi in kilometers. For its calculation see section 4.	www.wemakemaps.com www.distancesfrom.com
Population Satisfied with BHU*	Percentage of people who recorded that they are satisfied with services provided by the basic health units (BHU) at district level.	PSLMS 2004-05 and 2006-2007
Immunized Population*	Percentage of children aged 12-23 months that have been fully immunized based on record.	PSLMS 2004-05 and 2006-2007
Primary Net Enrolment (5-9 years)*	Total net primary schools enrolment is the number of children aged 5-9 years enrolled in primary schools divided by the total number of children aged 5-9 (PSLM definition)	PSLMS 2004-05 and 2006-2007
Literacy Rate (age 10 or above)*	Population aged 10 years and older who can read a newspaper and write a simple letter.	PSLMS 2004-05 and 2006-2007

* These variables are differentiated by urban and rural regions.

a) Road Data for districts of Punjab in the year 2006 are missing, so I estimated the average values by taking data of years 2005 and 2007.

b) The value of HDI for Islamabad is not given in the given source, to compensate I took the value of 2012 from the following source. https://en.wikipedia.org/wiki/List_of_districts_of_Pakistan_by_Human_Development_Index.

Firm and district level variables and their brief description are given in Table 1 and for definition of each variable please refer to the appendix.

Table 2: Firm and District Level Variables: Descriptive Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Firm level variables					
Output (Y), log	6,416	10.910	2.085	4.828	18.387
Capital (K), log	6,212	9.459	2.152	3.219	17.686
Labor (L), log	6,417	3.686	1.323	0	9.725
Materials (M), log	6,414	10.409	2.228	1.946	18.016
Government Ownership	6,417	0.086	0.280	0	1
Foreign Ownership	6,417	0.044	0.206	0	1
Importing Firm	6,413	0.136	0.342	0	1
EG Index at Industry Level	6,416	0.058	0.119	-0.823	3.677
EG Index Square at Industry Level	6,416	0.018	0.252	0	13.524
Gini Coefficient	6,416	0.095	0.102	0.025	0.976
District level variables					
Estimated Population 2006, log	73	7.466	1.057	5.730	14.369
% of Urban Population 1998	73	24.055	17.818	0	94.752
Number of Kilometers per Registered Vehicle	69	0.3805	2.4281	0.0001	20.222
HDI 2005	73	0.6486	0.0736	0.3137	0.892
Primary Net Enrolment (5-9) Total	73	53.993	12.483	29	85.5
Primary Net Enrolment (5-9) Urban	73	62.137	14.478	0	86.5
Primary Net Enrolment (5-9) Rural	73	51.137	13.393	23.5	85.5
Literacy Rate (Age 10 or Above) Total	73	50.377	12.040	32	85.5
Literacy Rate (Age 10 or Above) Urban	73	64.219	14.028	0	92
Literacy Rate (Age 10 or Above) Rural	73	44.945	11.413	25	74.5
Population Satisfied with BHU-Total	73	39.345	13.892	7.395	76.02
Population Satisfied with BHU - Urban	73	37.835	18.456	0	84.29
Population Satisfied with BHU-Rural	73	40.109	13.182	16.3	73.55
Immunized Total Population	73	49.267	18.974	7.5	90.5
Immunized Urban Population	73	57.082	20.802	0	89.5
Immunized Rural Population	73	45.747	20.090	6	91

Note: For details description and definition of variables refer to Table 1

I did some minor adjustments to both firm level and district level data. The CMI 2005-06 dataset contains a total of 6,417 firms in 22 2-digit industries covering 73 out of 120 Pakistani districts existing at the time of the survey. After cleaning the data, observations which were missing for two variables; capital and number of kilometer road paved per registered vehicles (*RdRv*), have been adjusted by including two dummies for each variable and setting missing values to zero. Missing observations may have some systematic bias because almost 40% of missing observations on capital are from the food and beverage industry (Division 15) and 75% missing observations of *RdRv* are from the Islamabad district as I could not find an authentic source of the number of kilometers of paved roads in 2006 for this district. But I also do an exercise in the section 6 to check whether or not these missing observation are sensitive to our findings. Excluding missing observation leaves 6,059 observations in the sample. So the final sample varies between 6,059 and 6,413 observations depending on the availability of data at the district level. While I have regional data for all 120 district of Pakistan, the CMI 2005-06 was only conducted in 73 districts. Table 2 shows the descriptive statistics of dependent and independent variables. The list of districts and 2-digit industries are given in Table A2 and A3 in Appendix respectively.

6. Estimation Results and Discussion

Estimation results of OLS regressions for different model variants are shown in columns (1) to (7) of Table 3. In all regressions, I use a set of controls at the firm level to capture firm specific effects and a set of controls at the district level to control for district specific effects along with industry and province dummies. To capture industry specific fixed effects I include two-digit industry dummies in all models. The coefficients of industry dummies are not reported in the table while the coefficients of province dummies are included (using the Punjab province as a reference).

I start the discussion with the base line model shown in column 1 which represents the estimation results of equation 3 in section 4. The elasticity of firms' conventional inputs; labor, capital, and materials are in their expected range; 0.15, 0.057, and 0.81 respectively at the 99% confidence interval which are consistent with the findings of Kneller and Misch (2014).

Among other firm control variables, the coefficient of the foreign firm dummy is positive and highly statistically significant. There are many reasons why foreign firms outperform local ones in developing countries. It is not just the foreign ownership, it could be firm characteristics, firm assets, or even home/host countries' government policies (Bellak, 2004). Foreign firms may use better, more modern and sophisticated technology and equipment in their production processes. They tend to be more innovative (Taymaz & Lenger, 2004) and more respondent to market demand. They may have better management and marketing strategies, and they may boost their employees' skills by training and development etc. The coefficients of other firm level controls are insignificant whereas the coefficients of import oriented firms and the EG Index are negative while the coefficient of government-owned firms is positive. The interpretation of the EG index is that a firm produces more output if it is located in a highly industrial concentrated region than a firm with the same amount of input and technology located in a widely dispersed industrial region. This relation only holds up to a certain level of concentration, however. After that limit, firms happen to face diseconomies of scales, as shown by a model variant that included the squared term of the EG index (see the robustness check section). When excluding the squared term, the estimation results imply that agglomeration does not matter in manufacturing industries in Pakistan[17].

[17] I also check the impact of industrial concentration on firm productivity by putting EG Gini coefficients which also show an insignificant impact on firm productivity. The results are not included in the table but available upon request.

Table 3: Effect of Social Infrastructure on Firm Output: Results from OLS Regressions

Dependent Variable: Total Output at Producer Price				**Large Firm**	**Small Firm**	**High Capital Intensive Firm**	**Low Capital Intensive Firm**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Capital	0.0568***	0.0559***	0.0585***	0.0660***	0.0654***	0.0891***	0.0330***
	(0.0068)	(0.0069)	(0.0070)	(0.0096)	(0.0108)	(0.0146)	(0.0123)
Labor	0.1513***	0.1537***	0.1553***	0.1643***	0.1832***	0.1079***	0.2205***
	(0.0102)	(0.0103)	(0.0102)	(0.0167)	(0.0282)	(0.0186)	(0.0206)
Materials	0.8138***	0.8132***	0.8085***	0.7797***	0.8136***	0.8057***	0.7938***
	(0.0075)	(0.0075)	(0.0076)	(0.0122)	(0.0107)	(0.0137)	(0.0096)
Government	0.0186	0.0184	0.002	-0.0315	0.058	-0.0436	0.1146**
	(0.0300)	(0.0300)	(0.0299)	(0.0362)	(0.0517)	(0.0386)	(0.0448)
Foreign	0.1348***	0.1405***	0.1496***	0.1284***	0.1379*	0.2605***	-0.0523
	(0.0389)	(0.0387)	(0.0384)	(0.0411)	(0.0771)	(0.0505)	(0.0663)
Import	-0.0077	-0.0051	-0.0046	-0.0057	0.008	-0.0063	0.0154
	(0.0186)	(0.0186)	(0.0184)	(0.0214)	(0.0373)	(0.0239)	(0.0279)
EG Index	-0.0793	-0.0927	-0.0643	-0.1602*	-0.016	-0.1059**	0.0321
	(0.0689)	(0.0674)	(0.0702)	(0.0959)	(0.0614)	(0.0528)	(0.1624)
% of Urban Population 1998[a)]	0.0167**	0.0138*	0.0003	0.0009	-0.001**	-0.0006	0.0006
	(0.0076)	(0.0079)	(0.0003)	(0.0006)	(0.0005)	(0.0005)	(0.0005)
Number of km per Registered Vehicle	-0.0038	0.0031	-0.0011	-0.0639	-0.0015	0.006	-0.1626*
	(0.0102)	(0.0079)	(0.0099)	(0.1635)	(0.0080)	(0.0065)	(0.0986)
Population Satisfied with BHU - Total		0.0019***					
		(0.0006)					
Primary Net Enrol-ment (5-9 years), total		0.0020***					
		(0.0008)					
Population Satisfied with BHU - Urban			0.0058***	0.0049***	0.0050***	0.0032**	0.0086***
			(0.0008)	(0.0010)	(0.0012)	(0.0013)	(0.0009)
Population Satisfied with BHU - Rural			-0.0045***	-.0025**	-0.0043***	-0.0036***	-0.0046***
			(0.0008)	(0.0012)	(0.0012)	(0.0011)	(0.0014)
Primary Net Enrolment (5-9 years), urban			0.0037***	0.0064***	0.0027**	0.0053**	0.0077***
			(0.0011)	(0.0025)	(0.0013)	(0.0022)	(0.0023)
Primary Net Enrol-ment (5-9 years), rural			-0.0014	0.0011	-0.0029***	-0.0036**	-0.0021
			(0.0009)	(0.0017)	(0.0011)	(0.0016)	(0.0014)
Sindh	0.1117***	0.1054***	-0.0141	0.0458	-0.0191	0.0960**	-0.1049**
	(0.0146)	(0.0154)	(0.0288)	(0.0444)	(0.0444)	(0.0455)	(0.0435)
KPK	0.0184	-0.0109	-0.0770**	-0.0248	-0.100***	-0.0200	-0.1753***
	(0.0224)	(0.0283)	(0.0303)	(0.0423)	(0.0427)	(0.0411)	(0.0478)
Balochistan	0.5741***	0.6015***	0.4929***	0.7898***	0.2134**	0.3578***	0.9988***
	(0.0713)	(0.0720)	(0.0736)	(0.0102)	(0.0985)	(0.0747)	(0.1961)
Constant	1.1780***	1.0379***	1.1734***	0.9846***	1.2070***	1.1245***	1.0209***
	(0.0794)	(0.0806)	(0.0766)	(0.1312)	(0.1281)	(0.1563)	(0.1527)
Observations	6,413	6,413	6,411	3,205	3,206	3,092	3,082
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.9474	0.9475	0.9482	0.9279	0.9347	0.9402	0.9415

Note: Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1. Estimation based on OLS.
a) in regressions (1) and (2):log of Estimated Population in 2006.

At the district level I use two control variables; estimated population in 2006 and the number of paved roads in kilometers available per registered vehicles at district level. The effect of population in the baseline regression shows a positive sign which is statically significant while the number of kilometers per registered vehicle at district level shows a negative sign which is statistically insignificant at conventional level.

Now I turn my attention to the main question, does social infrastructure affect firm total factor productivity? In column (2), I test the effects of social capital (health and education) on firm total factor productivity. Indicators for both social infrastructure i.e. health and education show positive effect on firm productivity. More precisely, if government and/or non-government agents invest in facilities of basic health units (BHUs) at district level, and as a result of that investment the share of the population that is satisfied with BHUs' facilities increase by one percent (say from 40% to 41%), this can raise - ceteris paribus - firm level output by 0.19% which is quite a substantial increase which is significant at the 1% level. Beside this, spending on education can also have a substantial impact on firm output in manufacturing in Pakistan. My sample translated this relationship in the way that if the net enrolment in primary school increases by 1% of population at the age of 5 to 9 years or primary classes from 1 to 5, as a result, on average firm output in manufacturing industry likely increases by 0.20% holding all other factors constant. The results which I find in my analysis for both social infrastructure indicators are within the line of the study by Sumarto and De Silva (2014) who study the effect of health and education indicator at district level on the subnational growth and poverty in Indonesia.

Now I check this relationship by including separate indicators for rural and urban regions within a district. The findings of health and education indicators are shown in column (3). I replace the estimated district population in 2006 by the share of urban population in 1998[18] to capture the urban and rural proportion in district population. In the reminder of the paper, I will use this share as an indicator for separating urban and rural districts, except stated otherwise. In the urban sub-sample, the effects of both indicators on firm productivity are positive and significant. However, a negative and significant impact of health and a negative but insignificant impact of education are recorded.

Investment in education infrastructure is positively (and statistically significant at the 1% level) correlated with firm productivity in urban regions whereas in rural regions it is negatively (but statistically insignificant) associated. More precisely, if the net effects of spending on education increase the net enrolment in primary school by 1% in urban region, firm output may likely increase by 0.37% while education indicators in rural region show a negative but insignificant impact on industrial productivity. Here my findings are not consistent with the findings of Fan, Hazell, and Thorat (2000) who investigate the relationship of government spending on rural education with other rural spending on productivity taking a country level time series data set in India. They found a positive effect.

The health sector depicts a different picture compared to education because health indicators show a positive effect in urban regions but a negative effect in rural regions, and both of them

[18] The latest available population census is the Population Census 1998.

are highly statistically significant. From this exercise, I interpret the results that if investment (whether by government or private agents) in health facilities caused an increase of the share of population satisfied with BHUs by 1% in urban regions, this comes together with an average increase of firms' output by 0.58%. However, I find a negative effect of increasing satisfied population with BHU on manufacturing productivity in rural region and this might be the case that manufacturing industries which are the major tax contributing entities are mostly located in urban regions and government investment to the health sector in urban region directly benefits to the employees of these firms while on the other hand I believe that government investment in rural health infrastructure could not translate this effect on the firm productivity immediately, at least not in the short run. Moreover, in the short run, rural population is highly engaged in agricultural activities, which is a tax-exempted sector in Pakistan. The rural population can benefit from government investment in rural region which are largely funded from taxes collected in urban regions. In the long run, this might bring some impact because healthy labor force move on to cities and may get engaged in industrial production. Here my results are again different for rural investment in the health sector as compared to the results of Fan et al. (2000).

In all of the regressions so far, I did not control for larger and strongly capital intensive firms which might affect my findings as firms behave differently to the available social infrastructure at district level. In this regard, I first check for firm size effects by splitting my sample into two parts; small firms and big firms based, using the median value of the number of employees per firm in the industry. The results are shown in the column (4) and (5) in Table 3. Both small and large firms get similar benefits from urban investment in health infrastructure while smaller firms get much more benefits than large firms from education investment in urban regions. Except for primary net enrolment in rural region which is positive though insignificant, all other indicators in rural region show a negative effect.

In order to test the impact of both rural and urban social indicators on productivity of high vs. low capital intensive firms, again I spilt the sample into two parts based on the median capital intensity. Here I find that low capital-intensive firms get much more benefits from investment in human capital (both on health and education) in urban region than high capital-intensive firms. This may be because health and education matter more for firms when they opt for a labor-intensive mode of production. Investment in rural social infrastructure portrays a similar picture as rural health indicators show a negative and significant effect while education shows a negative and insignificant effect on firm productivity.

Robustness Checks

For investigating the robustness of the findings presented above, I run a number of alternative model specifications to test my hypotheses. First, I apply the feasible generalized least square (fGLS) estimator instead of OLS. I show these findings for the effects of health and education indicators in rural and urban region on firm productivity at district level in column (1) of Table A4 which is the replication of column (3) of Table 3. Despite some variation in the results for control variables, my variables of interests remain unchanged in term of their sign. The effect of education indicators in urban region on firm productivity remains exactly the same while for rural education the effect is still negative but this time it is statistically significant at the 10 % level.

Another source of concern on the findings presented above relates to potentially erroneous assumptions about the structure of error terms. Firms within the same district might be affected by some similar unobservable shocks, e.g. natural disasters. If this is the case then error terms of firms from the same district would be systematically correlated. Heteroskedastic-robust standard errors usually do not take into account this issue. The result would be too high t-statistics and too narrow confidence intervals and hence wrong conclusions about rejecting null hypotheses (Cameron & Miller, 2015). Based on the assumption that errors are correlated within districts but uncorrelated across districts, I test my hypotheses while clustering the errors at the district level. Here, I also aggregate 22 industries to six larger sector groups to get more observations per group. The estimated coefficients are given in column 8 of Table A4. Once again I found the same effects of the health indicators on firm productivity at the 99% confidence level in urban and rural regions. For the education indicators, the effects at the 90% confidence level only hold for urban regions.

From the main findings I observe that firms located in Balochistan are more productive than firms in other provinces. To check the impact of this peculiar result, I run a regression excluding firms from the Balochistan province. The results are shown in the column 2 of Table A4. The findings remain robust.

In regression (3), I investigate whether the level of available social infrastructure at the district level matters in explaining firm-level productivity in the first place. To check this relationship I include the Human Development Index 2005 (HDI[19]) in my regression analysis but here I exclude population since both the HDI and the log of population are rather highly correlated. From this exercise I find that the effect of HDI on firm level output is positive and statistically significant. Turning to my variable of concern, I find similar and robust results in term of sign and magnitude against the specification (3) of Table 3.

In column 4 of Table A4, I alter the infrastructure indicators as some time indicators may mislead the results. For this analysis I take other available proxies for health and education. For health I take the share of children aged 12-23 months that have been fully immunized based on record. Full immunization means that the child has received: BCG, DPT1, DPT2, DPT3, Polio1, Polio2, Polio3 and measles. For education I take the literacy rate of the population at age 10 or above. The magnitudes of the estimated coefficients vary slightly while the sign for both rural and urban social infrastructure remains the same.

In the main findings I did include two dummy variables to capture missing observations of two variables; capital[20], and number of paved road available per registered vehicles[21]. When running the regression excluding all missing observations only very little and negligible variations in my findings occur.

[19] District values of HDI_2005 are taken from the paper by Jamal, H., & Khan, A. J. (2007).

[20] About 200 firms have no capital records and among them 40% firms are from the food and beverage industries.

[21] Since this variable derived by dividing the number of paved road in kilometer to the registered vehicles at district level and I could not find the number of kilometers for Islamabad district from any available sources so I put a dummy variable for including the missing observation.

It is well known in the literature that concentration of industries creates problems such as congestion and high transport cost due to mismanagement of urban cities. Agglomeration can hence lead to diseconomies of scales. In the presence of such negative agglomeration effects, the EG Index would not remain a linear function. In order to test non-linearity, I include the squared term of the EG Index to capture diseconomies of scale. While doing this exercise, EG index shows a positive sign while the squared term shows a negative sign but both coefficients are remain insignificant at conventional significance levels. Here my findings are not consistent with the findings of Burki and Khan (2013) who observed a negative relationship of industrial agglomeration and technical inefficiency of manufacturing firm in Pakistan[22]. Despite the insignificance results my findings are consistent with the findings of Lin, Li, and Yang (2011) who find a positive effect of industrial agglomeration on labor productivity in the textile industry of China to a certain threshold and after that threshold industrial agglomeration create diseconomies of scales. My main findings still remain robust.

Next I checked the effect of human capital supply at the district level on firm productivity when controlling for a district's proximity to export markets (in this case Karachi which is the business hub and the main port city in Pakistan). The results are shown in the column 7 of Table A4. The coefficient for the distance between a firm's location district and Karachi is negative and statistically significant. This result suggests that firms located closer to Karachi get better access to a large consumer market and to international markets for exporting their products and importing inputs. The magnitude of the distance term shows that the output of firms on average is reduced by 0.0004 percent as they located their plants one kilometer away from Karachi[23]. Doing this analysis does not change the sign of my variables of interest. The magnitudes of the coefficients are reduced moderately. This might be because the distance variable and the dummy for the Sindh province are highly correlated.

7. Conclusion

In this paper I investigated whether investment in social infrastructure (health and education) increase firm level productivity in manufacturing industries in Pakistan. I found significant positive impacts for both dimensions, health and education, though the health indicator shows a relatively smaller impact on firm output which is consistent from the literature. However, the positive impact of social infrastructure investment depends on whether the focus of spending is towards urban or rural regions. Investment in urban infrastructure has a substantial impact on firm level productivity while investment in education and health in rural regions tend to have negative impacts on firm output. This might be because rural health infrastructure mainly targets the rural population engaged in agriculture activities and has little spillovers to manufacturing.

[22] Perhaps this may be because Burki and Khan (2013) took the data from all three available manufacturing censuses, CMI 1995-96 to 2005-06, and used three-digit industrial classification according to PSIC 1970, but I used three-digit classification of PSIC 2007.

[23] I also performed this exercise for the distance between the capital city of district and the capital city of the province where the firm operated its business and in this case the signs of the distance coefficient for all provinces remain positive.

I also control for a likely firm self-selection bias by including the EG index which measures industrial agglomeration at the district level. The coefficient of the EG index tells that firms are more productive if they are located in a highly-concentrated industrial district.

The importance of this study is that it contributes in the growing microeconomic literature by quantifying the relationship of spending on social infrastructure on firm productivity in developing countries at a disaggregated firm level. This study also counters the quality of infrastructure data by taking social infrastructure indicators recorded by the Pakistan Bureau of Statistic in its PSLM surveys which are better predictors than the actual spending in documents.

A number of caveats should be considered in the interpretation of the results. The first cautious while interpreting the results is that I was unable to investigate any long term relationship between social infrastructure and firm productivity, because no longitudinal data at the firm level is available in Pakistan up till now. Another caveat which I am more concerned about is that the exact geographic location of firm is not known. I am not able to determine whether a firm in the data set is located in an urban or a rural region since the most disaggregated geographical information available in the data is the district in which a firm operates. Knowing the location of a firm exactly could yield much better results for policy making purposes. Finally, there might be some reverse causality among dependent and independent variables. Unfortunately, single cross section data set does not allow controlling for this issue.

Above all, my findings propose that the positive effect of urban investment in health and education infrastructure on firm level productivity is very robust as it prevails in a number of robustness checks.

APPENDIX

DATA SOURCES

Firm level data set: the CMI 2005-06

Firm level data come from the Census of Manufacturing Industries 2005-06 (CMI 2005-06). The CMI 2005-06 dataset consists of 6,417 manufacturing establishments which are either registered or qualify for registration under the Factories Act 1934[24]. The census is distrusted in the following ways; Punjab (55.9%), Sindh (28.4%), KPK[25] (10.5%), Balochistan (3.3%), and Islamabad (1.8%). Manufacturing activities of establishments are classified at the 5-digit level of the Pakistan Standard Industrial Classification (PSIC) 2007, which is derived from the UN International Standard Industrial Classification ISIC Rev-3.1. Government workshops and defense establishments, even though registered under Factories Act 1934, were not included in the CMI 2005-06. Data were collected for the fiscal year[26] 2005-06. The CMI covered large scale manufacturing industries comprising establishments having 10 or more employees at 5-digit industry. The 4-digits industries of the CMI are comparable with the International Standards Industrial Classification ISIC Rev3.1. It is conducted for every five years using the frame provided by the Provincial Labor Departments.

Table A1 Sampling and Response Rate

Region	CMI 2000-01	No. of Establishments on mailing list	% of filled Questionnaire received	No. of non-responding Factories (defaulters)	Closed Establishments	Establishments reported in the final tabulation CMI 2005-06	% coverage compared to mailing list
Punjab	2,357	8,288	49	2,431	1,403	3,590	43.32
Sindh	1,768	3,288	64	423	770	1,825	55.502
NWFP	236	972	75	76	165	673	69.24
Balochistan	93	309	72	74	14	212	68.61
Islamabad	74	338	35	62	12	117	34.62
Pakistan	4,528	13,145	55	3,213	2,364	6,417	48.82

Source: the CMI 2005-06

The CMI 2005-06 frame was updated by using industrial directories provided by the Provincial Directorates of Industries and results of economic census conducted by the Federal Bureau of Statistics. Information was gathered by a mailed questionnaire followed a field visit by the provincial Directorates of industries. And finally, annual reports of the stock exchange were also concerned to augment the coverage. The number of establishments coverage increased compared with the previous census 2000-01. Table A1 shows the sampling frame, a comparison with previous census 2000-01 and the response rate.

Limitation of Census Data

There are number of limitations of the cenus data representing the overall establishments in Pakistan

- The frame of the survey included only the factory which was registered or qulaified for registration under the factories Act 1934 but some factories may be eligible to suvery not registered themselves so the CMI 2005-06 do not represent those fiirms.

[24] The survey was conducted by the Pakistan Bureau of Statistics (PBS) with the help of the Provincial Directorates of Industries and Provincial Bureaus of Statistics (BOS) under sections 9 & 10 of the General Statistics Act 1975 and section 5 & 6 of the Industrial Statistics Act 1942.
[25] KPK, Khyber Pakhtoonkhwa, previously known as NWFP, North West Frontier Province.
[26] In Pakistan the fiscal year starts on July 1st and ends on June 30th.

- A considerable number of establishments (2,364) closed their businesses and a considerable number of establishments(3,213) are defaulters.
- 333 duplicated firms were recorded in the frame. All firms were from Punjab Province.
- 819 firms were found involved other activities than manufacturing

I used all of the 6,417 firms of the CMI 2005-06 in my sample. Among 205 firms have either missing capital values or their values are less than or equal to zero. So to keep more observation in the sample I use a strategy for including the missing observation in the regression by including a dummy variable which represents the missing observation.

Definitions of Firm level data

Value of Production (Output)

It includes the value of sales from own production (finished and semi-finished products), values of fixed assets produced for own use, receipts for work, value of electricity sold, value of sales of goods purchased for resale, receipts for contract, commission, repairs and maintenance work done for others, receipts from industrial waste, and the net increase in the value of work in process.

Employees (Labor)

All persons whether part time or full time who work in an establishment and receive renumeration in cash or in kind. Working properiters, unpaid family workers and home workers are excluded. More specifically it includes all ***Production workers*** (who work directly associated with production like manufacturing, assembling, packing, repairing etc. Working supervisors and persons engaged for repairs and maintenance are also included) and ***Non-Production workers*** (administrative and professional employees, white-collar office employees, drivers, watchmen/guards, peons, sweepers, etc.).

Material Consumed

Material consumed is the materials which establishments purchased locally or imported from aboard. It includes raw materials, fuels, electricity, chemicals & dyes, packing materials, spare parts, lubricates and others.

Capital Stock

Capital Stock or Value of Fixed Assets is calculated by the following way:

- **Capital Stock** = Fixed Assets at the beginning of the fiscal year 2005-06 plus Investment.
- **Investment** = Purchase of Fixed Assets plus Fixed Assets produced for own use minus Sales of fixed Assets.

Government Ownership

All those establishments which are either solely owned by state/public sector or are owned by state/public sector with foreign collaboration.

Foreign Ownership

All those private or state owned establishments which are either solely owned by foreigners or are partly owned by foreigners.

Import Oriented Firm

All those establishments which involved importing materials from abroad are categorized as importing oriented firms. Materials they imported from abroad included fuels, electricity, raw materials, chemicals & dyes, packing materials, spare parts, lubricants, and others.

Definitions of District level variables (variables of interest)

Population Satisfied with BHU

A perception based question was asked to households to give their opinion about whether or not they are satisfied with the basic health units (BHU). It is the percentage of satisfied population with the services provided by BHU at district level.

Immunized Total Population

Percentage of children aged 12-23 months that have been fully immunized based on record. Fully immunization means that the child has received: BCG, DPT1, DPT2, DPT3, Polio1, Polio2, Polio3 and measles vaccination. (For more details see PSLM 2004-05 reports)

Net Enrolment Rate (age 5-9)

"The NER at primary level refers to the number of students enrolled in primary school of primary school age divided by the number of children in the age group for that level of education. In other words, for Pakistan, the official primary NER is the number of children aged 5 to 9 years attending primary level divided by the number of children aged 5 to 9 years" (PSLM definition).

Literacy-Population 10 Years and Older

In the PSLM 2004-05 report, literacy population 10 years and older is defined in the following way:

- *Population aged 10 years and older that is literate expressed as a percentage of the population aged 10 years and older.*
- *Literacy is taken as the ability to read a newspaper and to write a simple letter.*

Table A2: Distribution of Establishments at the District Level

No.	District Name	# Firms	Percent	Province	No.	District Name	# Firms	Percent	Province
1	Attock	36	0.56	Punjab	38	Sukkur	54	0.84	Sindh
2	Rawalpindi	67	1.04	Punjab	39	Khairpur	15	0.23	Sindh
3	Jhelum	17	0.26	Punjab	40	Ghotki	41	0.64	Sindh
4	Chakwal	13	0.2	Punjab	41	Nawab Shah	24	0.37	Sindh
5	Gujranwala	440	6.86	Punjab	42	Naushero Feroze	18	0.28	Sindh
6	Gujrat	190	2.96	Punjab	43	Hyderabad	98	1.53	Sindh
7	Mandibahaudin	18	0.28	Punjab	44	Dadu	82	1.28	Sindh
8	Hafizabad	27	0.42	Punjab	45	Badin	6	0.09	Sindh
9	Sialkot	215	3.35	Punjab	46	Thatha	17	0.26	Sindh
10	Narowal	18	0.28	Punjab	47	Mirpur Khas	21	0.33	Sindh
11	Sargodha	34	0.53	Punjab	48	Thar par Khar	1	0.02	Sindh
12	Khushab	15	0.23	Punjab	49	sanghar	42	0.65	Sindh
13	Mainwali	7	0.11	Punjab	50	Karachi	1198	18.67	Sindh
14	Bhakkar	9	0.14	Punjab	51	Peshawar	245	3.82	KPP
15	Faisalabad	419	6.53	Punjab	52	Charsada	18	0.28	KPP
16	Jhang	79	1.23	Punjab	53	Naushera	37	0.58	KPP
17	Toba Tek Sing	49	0.76	Punjab	54	Mardan	23	0.36	KPP
18	Lahore	774	12.06	Punjab	55	Swabi	70	1.09	KPP
19	Shekupura	201	3.13	Punjab	56	Kohat	6	0.09	KPP
20	Kasur	158	2.46	Punjab	57	Hangu	7	0.11	KPP
21	Okara	51	0.79	Punjab	58	Kakar	2	0.03	KPP
22	Multan	116	1.81	Punjab	59	Haripur	95	1.48	KPP
23	Khanewal	79	1.23	Punjab	60	Batagram	3	0.1	KPP
24	Lodhran	47	0.73	Punjab	61	D.I. Khan	23	0.36	KPP
25	Vehari	25	0.39	Punjab	62	Tank	9	0.14	KPP
26	Saiwal	71	1.11	Punjab	63	Bannu	25	0.39	KPP
27	Pakpattan	18	0.28	Punjab	64	Lakki Marwat	9	0.14	KPP
28	D.G. Khan	64	1	Punjab	65	Swat	60	0.94	KPP
29	Muzaffar Garh	61	0.95	Punjab	66	Lower Dir	2	0.03	KPP
30	Rajanpur	31	0.48	Punjab	67	Upper Dir	1	0.02	KPP
31	Layyah	6	0.09	Punjab	68	MalaKand	3	0.05	KPP
32	Bahawalpur	114	1.78	Punjab	69	Bunair	35	0.55	KPP
33	Bahawal Nagar	17	0.26	Punjab	70	Quetta	41	0.64	Balochistan
34	R.Y Khan	106	1.65	Punjab	71	Pishin	2	0.03	Balochistan
35	Jaccobabad	97	1.51	Sindh	72	Lasbella	169	2.63	Balochistan
36	Larkana	84	1.31	Sindh	73	Islamabad	117	1.82	Capital
37	Shikarpur	25	0.39	Sindh		**Total**	**6,417**	**100**	

Note: District Chinot is included in Jhang, and all districts of Karachi are combined to form a bigger district Karachi

Table A3: Distribution of Establishments at 2-Digit Industry Level

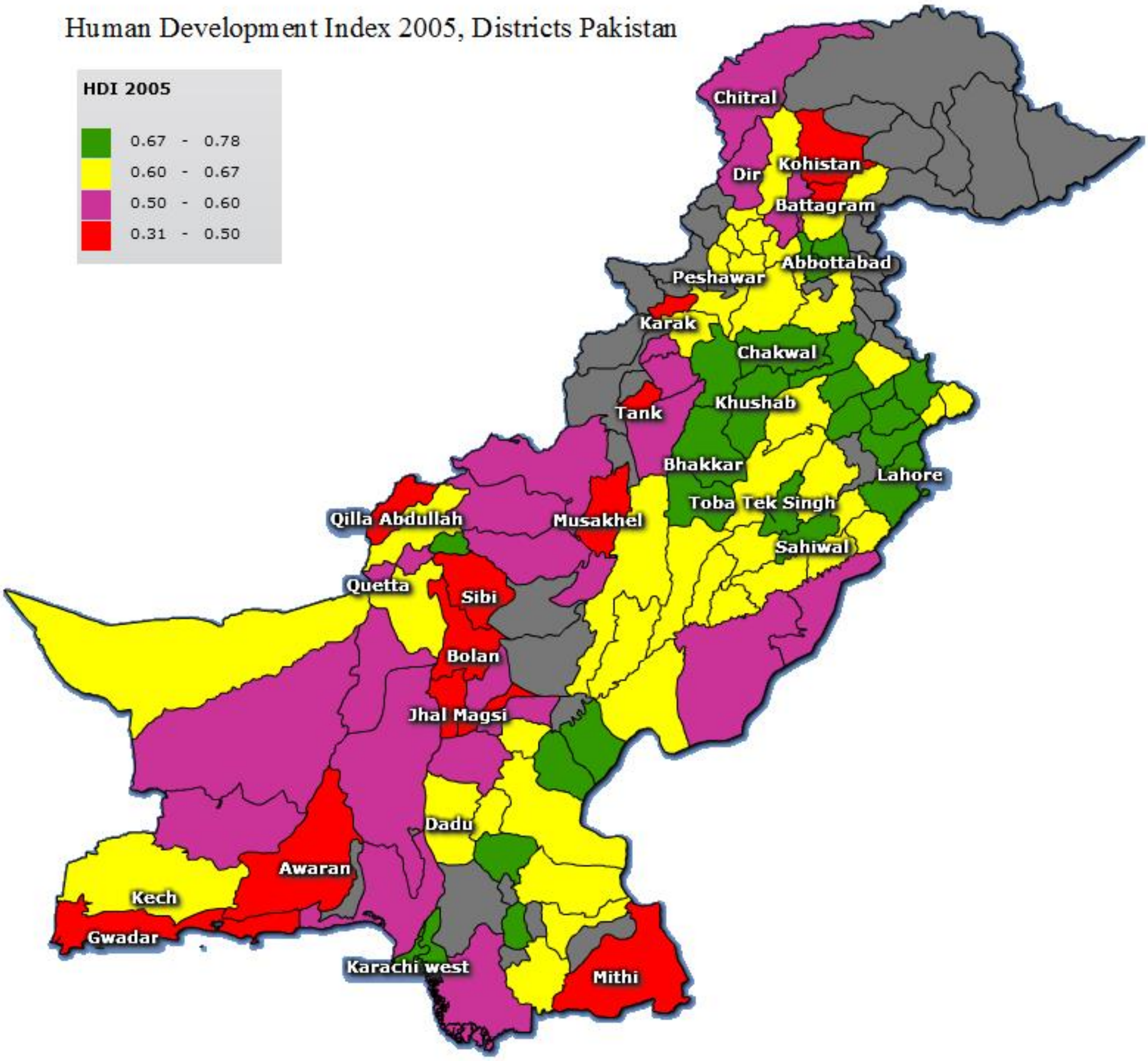
Division	Name of Industry	No. of Firms	Percent
15	Manufacture of food products and beverages	1,860	28.99
16	Manufacture of tobacco products	13	0.2
17	Manufacture of textiles	1,329	20.71
18	Manufacture of wearing apparel; dressing and dyeing of fur	326	5.08
19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	142	2.21
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	62	0.97
21	Manufacture of paper and paper products	133	2.07
22	Publishing, printing and reproduction of recorded media	47	0.73
23	Manufacture of coke, refined petroleum products and nuclear fuel	30	0.47
24	Manufacture of chemicals and chemical products	494	7.7
25	Manufacture of rubber and plastics products	170	2.65
26	Manufacture of other non-metallic mineral products	482	7.51
27	Manufacture of basic metals	291	4.53
28	Manufacture of fabricated metal products, except machinery and equipment	144	2.24
29	Manufacture of machinery and equipment n.e.c.	372	5.8
31	Manufacture of electrical machinery and apparatus n.e.c.	67	1.04
32	Manufacture of radio, television and communication equipment and apparatus	14	0.22
33	Manufacture of medical, precision and optical instruments, watches and clocks	95	1.48
34	Manufacture of motor vehicles, trailers and semi-trailers	139	2.17
35	Manufacture of other transport equipment	47	0.73
36	Manufacture of furniture; manufacturing n.e.c.	130	2.03
37	Recycling	30	0.47
	Total	**6,417**	**100**

Table A4: Robustness Check

	fGLS	Regression Excluding Balochistani Firms	HDI_2005 included in the Place of Population	Alternative Proxies are used	Regression Excluding Missing Observations	Regression With EG Index Square term	Proximity to Karachi Included	Errors cluster at District-Level
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Capital	0.0396***	0.0621***	0.0578***	0.0591***	0.0623***	0.0590***	0.0583***	0.0606***
	(0.0034)	(0.0068)	(0.0070)	(0.0069)	(0.0073)	(0.0070)	(0.0069)	(0.0104)
Labor	0.0964***	0.1499***	0.1566***	0.1522***	0.1516***	0.1560***	0.1620***	0.1631 ***
	(.0061)	(0.0101)	(0.0104)	(0.0103)	(0.0107)	(0.0103)	(0.0106)	(0.0165)
Materials	0.8850***	0.8078***	0.8073***	0.8105***	0.8066***	0.8078***	0.8023***	0.8000***
	(0.0038)	(0.0078)	(0.0077)	(0.0079)	(0.0080)	(0.0077)	(0.0080)	(0.0166)
Government	0.0578***	0.0201	0.0015	0.0017	-0.0223	0.0024	-0.0081	-0.0269
	(0.0199)	(0.0305)	(0.0297)	(0.0301)	(0.0301)	(0.0299)	(0.0297)	(0.0448)
Foreign	0.1157***	0.1489***	0.1516***	0.1497***	0.1678***	0.1533***	0.1529***	0.1585***
	(0.0301)	(0.0398)	(0.0384)	(0.0391)	(0.0392)	(0.0387)	(0.0382)	(0.0434)
Import	-0.0169	0.0265	-0.0046	-0.0044	-0.0134	-0.0023	0.0056	-0.0070
	(.0136)	(0.0183)	(0.0184)	(0.0186)	(0.0181)	(.0184)	(0.0186)	(0.0314)
EG Index	-0.0324	0.0064	-0.0605	-0.0766	-0.0727	0.0503	-0.0724	0.6884***
	(0 .0694)	(0.0678)	(0.0712)	(0.0713)	(0.0686)	(0.1349)	(0.0718)	(0.2624)
EG Index Square						-0.0445		
						(0.0385)		
% of Urban Population 1998	0.00008	0.0007**	0.3363**	-0.0010**	0.0002	0.0003	-0.0001	0.0002
	(0.0002)	(.0003)	(0.1587)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0008)
# of KMs per Registered Vehicle	-0.0117	0.0066	0.0005	0.0035	0.0038	0.0049	0.0198	0.0022
	(0.0210)	(0.0065)	(0.0097)	(0.0063)	(0.0077)	(0.0074)	(0.0093)	(0.0113)
Distance to Karachi							-0.0004***	
							(0.0001)	
Population Satisfied with BHU - Urban	0.0035***	0.0063***	0.0054***	0.0031***	0.0064***	0.0063***	0.0033***	0.0056***
	(0.0005)	(0.0008)	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0021)
Population Satisfied with BHU - Rural	-0.0023***	-0.0042***	-0.0039***	-0.0008	-0.0044***	-0.0044***	-0.0016*	-0.0050***
	(0.0006)	(0.0008)	(0.0009)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0016)
Primary Net Enrolment - Urban	0.0037***	0.0075***	0.0031***	0.0051***	0.0064***	0.0064***	0.0028***	0.0040**
	(0.0009)	(0.0016)	(0.0011)	(0.0016)	(0.0017)	(0.0016)	(0.0011)	(0.0020)
Primary Net Enrolment - Rural	-0.0011*	-0.0024**	-0.0018**	-0.0060***	-0.0020*	-0.0024***	0.0021**	-0.0031
	(0.0006)	(0.0010)	(0.0009)	(0.0013)	(0.0011)	(0.0010)	(0.0010)	(0.0021)
Sindh	0.0171	-0.023	-0.0225	0.0607***	-0.0184	-0.0223	-0.2556***	-0.0380
	(0.0198)	(0.0288)	(0.0267)	(0.0230)	(0.0293)	(0.0289)	(0 .0476	(0.0729)
KPK	-0.0083	-0.0732***	-0.0850***	0.0348	-0.0820***	-0.0849***	0.0297***	-0.0969*
	(0.0218)	(0.0304)	(0.0307)	(0.0246)	(0.0313)	(0.0307)	(0.0323)	(0.0581)
Balochistan	0.4443***		0.5115***	0.5873***	0.5273***	0.5193***	0.3204***	0.6949***
	(0.0705)		(0.0735)	(0.0701)	(0.0757)	(0.0741)	(0.0728	(0.2294)
Constant	0.6541***	0.9292***	1.0307***	1.1412***	1.0105***	1.0257***	1.4398***	1.2516***
	(0.0616)	(0.0878)	(0 .0920)	(0.0803)	(0.0927)	(0.0885)	(0.0942)	(0.1933)
Observations	6,378	6,166	6,413	6,378	6,058	6,378	6,413	6,413
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.9634	0.9521	0.9482	0.9475	0.9493	0.9478	0.9486	0.9443

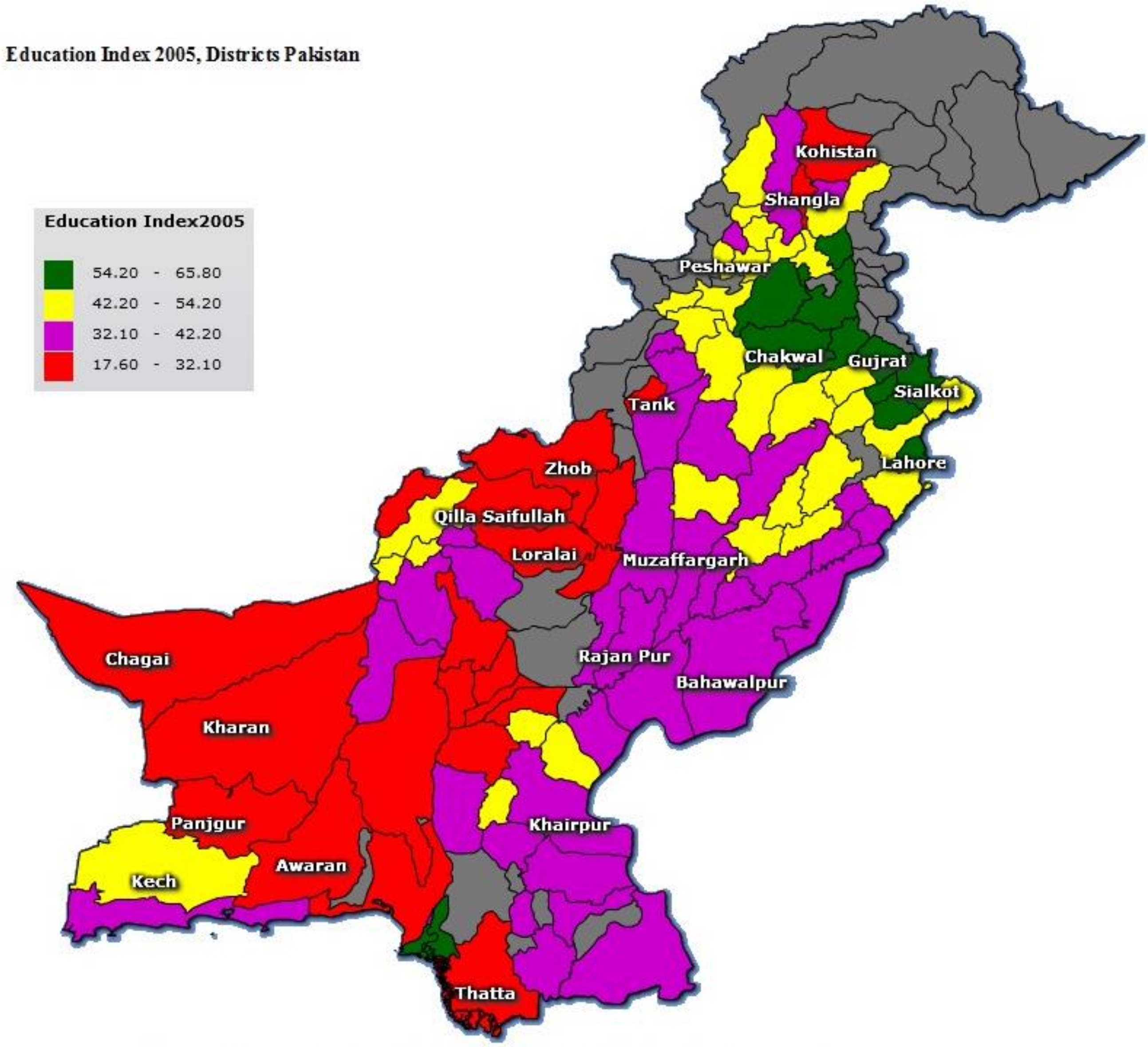
Note: Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1. Total Output at producer Price is the Dependent Variable. Except regression (1) all other models are estimated by OLS. In regression (2) firms from Balochistan are excluded. In regressions (3) % of Urban Population 1998 is replaced with HDI_2005, however, the variable name remains the same. In regression (4) Variables Population satisfied with BHU and Primary Net Enrolment are replaced with Immunized Total Population and Literacy Rate of Population at the Age of 10 or Above respectively . In column (5), results of the model without adjusting missing values.

Figure A1: Districts' Human Development Index 2005, Pakistan



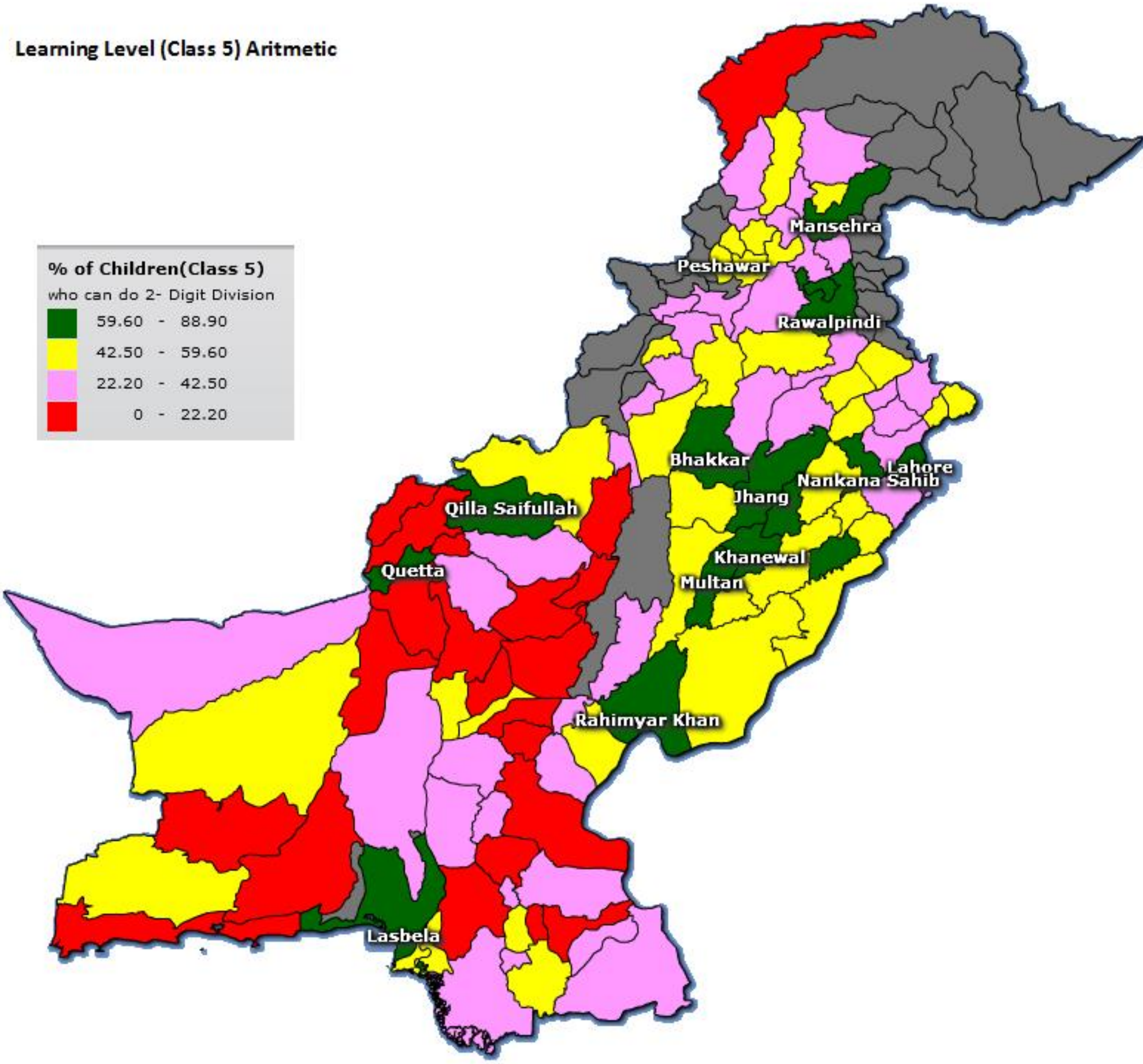
Sources: Data are taken from Jamal, H., & Khan, A. J. (2007) and map is prepared on the online website: https://www.targetmap.com.

Figure A2: Districts Education Index 2005, Pakistan



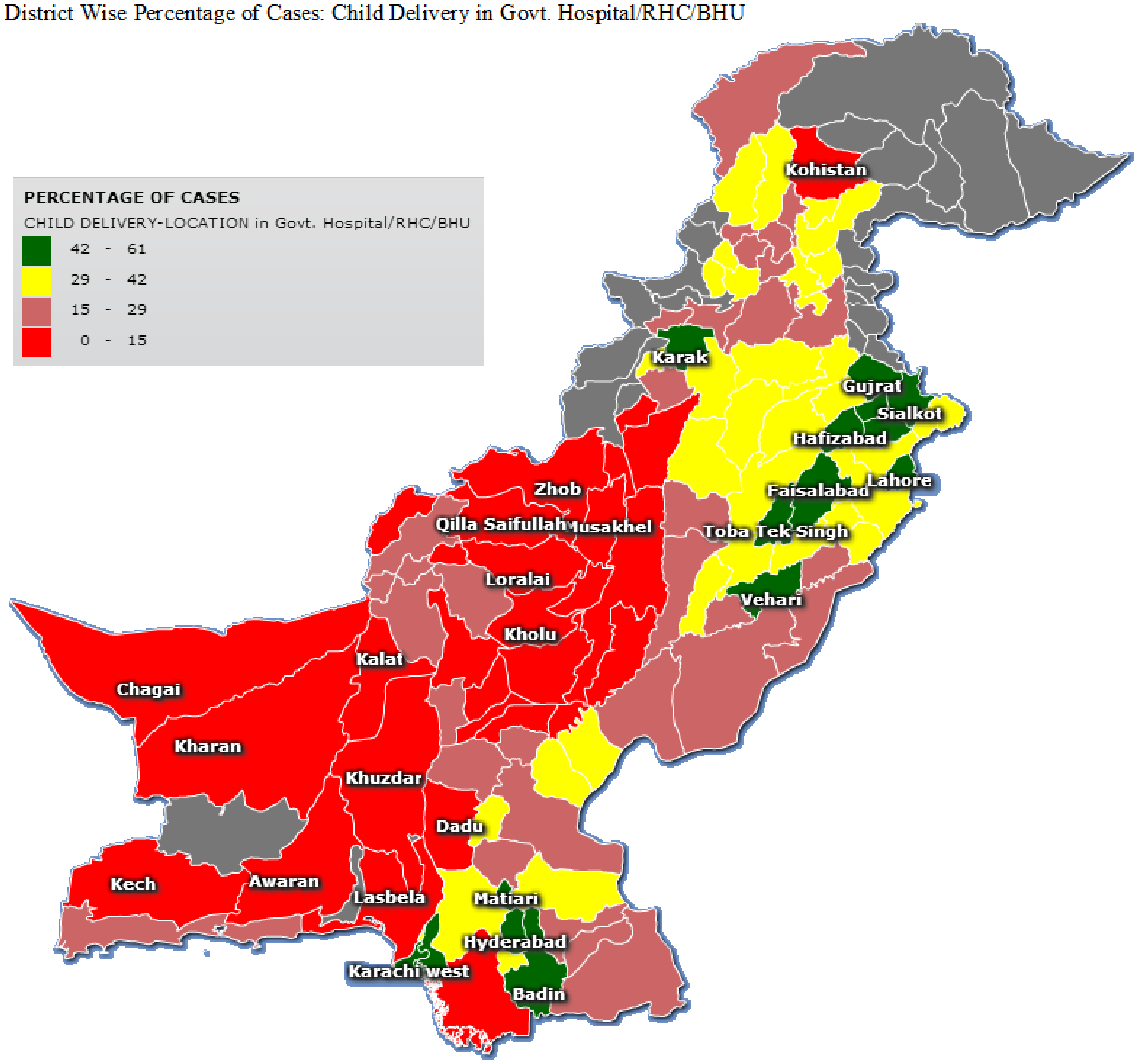
Sources: Data are taken from Jamal, H., & Khan, A. J. (2007) and map is prepared on the online website: https://www.targetmap.com.

Figure A3: Districts Quality of Education 2014, Pakistan



Source: annual status of education report: the largest citizen led household based initiative. Data are accessed as on 10-07-2015 at http://www.aserpakistan.org.Map is made on the online website: https://www.targetmap.com.

Figure A4: Districts Child Delivery in Government Health Institutions 2012-013, Pakistan



Reference: Data are taken from PSLM 2012-13 and map is made on the website: https://www.targetmap.com/

References

Afzal, M., Arshed, M. G., & Sarwar, K. (2013). Education, health, food inflation and economic growth in Pakistan. *Pakistan Economic and Social Review, 51*(2), 109-138.

Afzal, M., Farooq, M. S., Ahmad, H. K., Begum, I., & Quddus, M. A. (2010). Relationship Between School Education and Economic Growth in Pakistan: Ardls Bounds Testing Approach to Cointegration. *Pakistan Economic and Social Review, 48*(1), 39-60.

Ahmed, M. (2013). *Fiscal decenteralisation and politcal econmy of poverty reduction: Theory and evidence from Pakistan.* (Doctor of Philosophy), Durham University, Durham Theses. Retrieved from http://etheses.dur.ac.uk/7288/

Ahmed, V., Abbas, A., & Ahmed, S. (2013). Public infrastructure and economic growth in Pakistan: A dynamic CGE-microsimulation analysis. *Infrastructure and Economic Growth in Asia*, 117.

Ahrend, R., Farchy, E., Kaplanis, I., & Lembcke, A. C. (2014). *What makes cities more productive? Evidence on the role of urban governance from five oecd countries*. OECD Regional Development Working Papers, 2014/03. Organisation for Economic Cooperation and Development. Paris. Retrieved from http://dx.doi.org/10.1787/5jz432cf2d8p-en

Ali, L., & Ramay, M. I. (2014). Effects of human capital and trade orientation on output and total factor productivity in Pakistan. *Research Journal of Applied Sciences, Engineering and Technology, 8*(13), 1594-1606.

Allcott, H., Collard-Wexler, A., & O'Connell, S. D. (2014). *How do electricity shortages affect productivity? Evidence from India*. NBER Working Paper no. 19977. National Bureau of Economic Research. Retrieved from http://ssrn.com/abstract=2411280

Amjad, R., Ghani, E., ud Din, M., & Mahmood, T. (2012). Export Barriers in Pakistan: Results of a Firm Level Survey. *The Lahore Journal of Economics, 17*, 103–134.

Aschauer, D. A. (1989a). Does public capital crowd out private capital? *Journal of monetary economics, 24*(2), 171-188.

Aschauer, D. A. (1989b). Is Public Expenditure Productive? *Journal of monetary economics, 23*(2), 177-200.

Aschauer, D. A. (1989c). Public investment and productivity growth in the Group of Seven. *Economic Perspectives, 13*(5), 17-25.

Aziz, B., Khan, T., & Aziz, S. (2008). *Impact of higher education on economic growth of Pakistan*. Munich Personal RePEc Archive Paper no. 22912. Retrieved from http://mpra.ub.uni-muenchen.de/22912/

Baldacci, E., Clements, B., Gupta, S., & Cui, Q. (2008). Social spending, human capital, and growth in developing countries. *World Development, 36*(8), 1317-1341.

Barro, R. J. (1991). Government spending in a simple model of endogenous growth. *Journal of Political Economy, 98*(5), S103-S125.

Bellak, C. (2004). How performance gaps between domestic firms and foreign affiliates matter for economic policy. *Transnational Corporations, 13*(2), 29-56.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of monetary economics, 34*(2), 143-173.

Bils, M., & Klenow, P. J. (2000). Does schooling cause growth? *American Economic Review, 90*(5), 1160-1183.

Bloom, D. E., & Canning, D. (2005). *Schooling, health, and economic growth: Reconciling the micro and macro evidence*. unpublished, Harvard School of Public Health.

Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: A production function approach. *World Development, 32*(1), 1-13.

Brandt, L., Van Biesebroeck, J., & Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics, 97*(2), 339-351.

Burki, A. A., & Khan, M. A. (2013). *Agglomeration economies and their effects on technical inefficiency of manufacturing firms: Evidence from Pakistan*. unpublished Working Paper March 2013. International Growth Centre. Retrieved from http://igc.soapboxserver.co.uk/wp-content/uploads/2014/09/Burki-Khan-2013-Working-paper.pdf

Calderón, C., & Servén, L. (2004). *The effects of infrastructure development on growth and income distribution*. Policy Research Working Paper no. 3400. World Bank. Washington, D.C.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372.

Dastidar, S. G., Mohan, S., & Chatterji, M. (2013). *The relationship between public education expenditure and economic growth: The case of India*. SIRE-DP. Scottish Institute for Research in Economics Retrieved from http://repo.sire.ac.uk/handle/10943/435

Datta, S. (2012). The impact of improved highways on Indian firms. *Journal of Development Economics, 99*(1), 46-57. Retrieved from http://www.csae.ox.ac.uk/conferences/2011-EdiA/papers/519-Datta.pdf

Devarajan, S., Swaroop, V., & Zou, H.-F. (1996). The composition of public expenditure and economic performance. *Journal of monetary economics, 37*(2), 313-344.

Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2005). Investment climate and firm performance in developing economies. *Economic Development and Cultural Change, 54*(1), 1-31.

Dube, W. G., Phiri, A., & Bahmani-Oskooee, M. (2015). Nutrition and economic growth in South Africa: A threshold Co-Integration approach. *Journal of Economic Studies, 42*(1), 138-156.

Dutz, M. A., & O'Connell, S. D. (2013). *Productivity, innovation and growth in Sri Lanka: An empirical investigation*. World Bank Policy Research Working Paper, 6354.

Ellison, G., & Glaeser, Edward L. (1997). Geographic concentration in U.S. manufacturing industries: A Dartboard approach. *Journal of Political Economy, 105*(5), 889-927. Retrieved from http://www.jstor.org/stable/10.1086/262098

Fan, S., Hazell, P., & Thorat, S. (2000). Government spending, growth and poverty in rural India. *American Journal of Agricultural Economics, 82*(4), 1038-1051.

Farzanegan, M. R. (2011). *Education spending and productivity in Iran: Where have all the education expenditures gone?* Paper presented at the Politica and Economic Development: ERF 17th Annaul Conference, Renaissance Antalya Beach Resort & Spa, Turkey.

Feng-ying, G. (2013). The impact of government education expenditures on economic growth in China - The analysis based on the data from 1980 to 2011. *Technoeconomics & Management Research, 12*, 019.

Fisher-Vanden, K., Mansur, E. T., & Wang, Q. J. (2015). Electricity shortages and firm productivity: Evidence from China's industrial firms. *Journal of Development Economics, 114*(May 2015), 172-188.

Fuller, B., & Romer, P. (2014). Urbanization as opportunity. *World Bank Policy Research Working Paper, No. 6874*.

Ghani, E., Goswami, A. G., & Kerr, W. R. (2013). *The golden quadrilateral highway project and urban/rural manufacturing in India*. Policy Research Working Paper no. 6620. World Bank. Washington, DC.

Holl, A. (2013). *Highways and productivity in urban and rural locations*. Universitat de Barcelona-Institut d'Economia de Barcelona Working Paper. Retrieved from http://www.iedcyt.csic.es/sites/default/files/doc/cv/highways_Feb.pdf

Hong, H., & Ahmed, S. (2009). Government spending on public goods: evidence on growth and poverty. *Economic and Political Weekly, 44*(31), 102-108.

Jamal, H., & Khan, A. J. (2007). *Trends in regional human development indices*. Research Report no. 73. Social Policy and Development Centre.

Kneller, R., & Misch, F. (2014). The effects of public spending composition on firm productivity. *Economic Inquiry, 52*(4), 1525-1542.

Krueger, A. B., & Lindahl, M. (2000). *Education for growth: why and for whom?* NBER Working Paper no. 7591. National Bureau of Economic Research.

Lin, H.-L., Li, H.-Y., & Yang, C.-H. (2011). Agglomeration and productivity: Firm-level evidence from China's textile industry. *China Economic Review, 22*(3), 313-329. Retrieved from http://www.sciencedirect.com/science/article/pii/S1043951X11000253

Lorentzen, P., McMillan, J., & Wacziarg, R. (2008). Death and development. *Journal of Economic Growth, 13*(2), 81-124.

Mamatzakis, E. C. (2007). *An analysis of the impact of public infrastructure on productivity performance of Mexican industry*. CESifo Working Paper no. 2099. Centre for Economic Studies and Information and Forschung (Research) CESIFO.

McDonald, S., & Roberts, J. (2002). Growth and multiple forms of human capital in an augmented Solow model: a panel data investigation. *Economics Letters, 74*(2), 271-276. Retrieved from http://www.sciencedirect.com/science/article/pii/S0165176501005390

Mekdad, Y., Dahmani, A., & Louaj, M. (2014). Public spending on education and economic growth in Algeria: Causality test. *International Journal of Business and Management, 2*(3), 55-70.

Moyo, B. (2013). Power infrastructure quality and manufacturing productivity in Africa: A firm level analysis. *Energy Policy, 61*(October 2013), 1063-1070.

Munnell, A. H. (1990). Why has productivity growth declined? Productivity and public investment. *New England Economic Review*(Jan), 3-22.

Munnell, A. H. (1992). Policy watch: Infrastructure investment and economic growth. *The Journal of Economic Perspectives, 6*(4), 189-198.

Pasha, H. A., Ghaus, A., & Malik, S. (1989). The economic cost of power outages in the industrial sector of Pakistan. *Energy Economics, 11*(4), 301-318.

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.

Ping, H. (2005). *How does education at all levels influence productivity growth? Evidence from the Chinese provinces*. CERDI, Etudes et Documents, Ec 2005.15.

Planning Comission GoP. (2013). *Pakistan Millennium Development Goals Report 2013* (5). Retrieved from http://www.pk.undp.org/content/pakistan/en/home/library/mdg/pakistan-mdgs-report-2013.html

Popa, A.-M. (2012). The impact of social factors on economic growth: Empirical evidence for Romania and European Union Countries. *Romanian Journal of Fiscal Policy, 3*(2), 1-16.

Pritchett, L. (1996). *Mind your P's and Q's*. Policy Research Working Paper no. 1660. World Bank Washington, DC.

Research and Development Solutions. (2012). *Childhood Immunization in Pakistan*. Policy Briefs Series No. 3, February 2012. Retrieved from http://resdev.org/files/policy_brief/3/3.pdf.

Rothenberg, A. D. (2012). *Essays on Spatial Development.* (Doctor of Philosophy), University of California, Berkeley. Retrieved from http://escholarship.org/uc/item/5h82t2px

Sanchez-Triana, E., Biller, D., Nabi, I., Ortolano, L., Dezfuli, G., Afzal, J., & Enriquez, S. (2014). *Revitalizing industrial growth in Pakistan: Trade, infrastructure, and environmental performance*: World Bank Publications.

Sanchez-Robles, B. (1998). Infrastructure investment and growth: Some empirical evidence. *Contemporary economic policy, 16*(1), 98-108.

Sumarto, S., & De Silva, I. (2014). *Being healthy, wealthy, and wise: Dynamics of Indonesian subnational growth and poverty*. TNP2K Working Paper 09-2014. Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K). Jakarta, Indonesia.

Taymaz, E., & Lenger, A. (2004). *Multinational corporations as a vehicle for productivity spillovers in Turkey*. DRUID Working Paper no. 04-09. Danish Research Unit for Industrial Dynamics.

UNESCO. (2012). *EFA Global Monitoring Report 2012:* . Retrieved from Paris:

UNESCO. (2015). *EFA Global Monitoring Report 2015:* . Retrieved from Paris:

Weil, D. N. (2007). Accounting for the effect of health on economic growth. *The quarterly journal of economics, 122*(3), 1265-1306.

World Economic Forum. (2014). *Global Competitiveness Report 2014–2015*. Geneva.