Essays on Infrastructure, Firm Productivity, Natural Disaster and Life Course Transition in South Asia

Inaugural-Dissertation

zur Erlangung des Doktorgrades

des Fachbereichs Wirtschaftswissenschaften

der Ruprecht Karls Universität Heidelberg

vorgelegt von

Riaz AHMED

Heidelberg

April 2017

****

This dissertation is dedicated to my late parents;
Mazar (father) and Ladi (mother),
who spent their lives in poverty and prioritized my
education, so I can live a prosperous life.

****

Acknowledgement

This piece of work has not been accomplished without the moral and financial supports, persistence encouragements, insightful comments and kind suggestions from many mentors, colleagues, friends, family members, and institutions both in Pakistan and in Germany.

First of all, I would like to thank Prof. Stefan Klonner for being my dissertation first supervisor, who intellectually furnished my level of understanding and thinking about the deeper issues related to my dissertation's key hypotheses by providing his very valuable comments and suggestions during each meeting we have had. I am also grateful to Christian Rammer (ZEW, Mannheim) for being a kind and gentle mentor throughout my stay in Germany and Florian Misch (ZEW, Mannheim). Both helped me designing the initial structure and the first two papers of this dissertation. I am really thankful to Prof. Melanie Arntz for being the second reviewer of my dissertation and defense committee member. Her valuable comments on the dissertation really help me in restructuring the essays for journal publication and my further future research works. I am also thankful to Prof. Rahul Mukherji for being the third defense committee member.

I am also obliged to thank Christian Oldiges and two discussants, Nick Barton and Rahul Mehrotra, of my paper presentation at the June 2016 Development Economics and Policy conference in Heidelberg for their valuable comments and suggestions on chapter 3 of this dissertation.

I am really thankful to the Higher Education Commission (HEC) of Pakistan, the German Academic Exchange Service (DAAD) and the Centre for European Economic Research (ZEW) for their financial and administrative support and the University of Turbat for providing study leaves. My 4-year PhD program has fully been funded by the HEC via DAAD and partially funded (a completion grant for three months) by DAAD. I am also obliged to thank the people at ZEW who provided me all types of equipment, technical facilities, and service support (particularly Thomas Eckert, Heidrun Förster and Heidi Halder) as well as research oriented guidance and suggestions since the beginning of my PhD program. At the ZEW, I also want to thank Prof. Francois Laisney with whom I shared an office and improved my econometrics skills.

I indebted to say thanks Christian Rammer, Mukhtar Hussain, Adeel Ahmed, Adnan Riaz and Waseem Barkat for reading the first draft versions of my dissertation.

I thank the following generous people who helped me in the collection and processing of data for the empirical part of my studies: Nohman Ishtiaq (Ministries of Finance, Pakistan); Jamal Sarwar (National Electric Power Regulatory Authority, Pakistan); Nasir Amin, Bilal Khalil and Adeel Zia (Academy of Educational Planning & Management, Pakistan); Syed Hassan Raza, Mirza Humayun and Naseer Ahmed Baig (Pakistan Bureau of Statistics); Dr. Muhammad Waheed (The World Bank, Islamabad); Ch. Sajid Rasul, Muhammad Arif Aziz and Sana Gull (Bureau of Statistics, Punjab); Waqas Alam (Additional Secretary, Communication & Work Department, Punjab); Mohammad Mehdi (National Transmission & Despatch Company, Pakistan); Tariq Mahmood, Abid Alam, Alamgir Khan, Qazi Fahad, Shahnawaz Khan, Saqib Afridi and Jahanzaib Khan (Bureau of Statistics, Khyber Pakhtunkhwa); Abdul Haleem Bangash (P&D Department, Khyber Pakhtunkhwa); Muneer Hussain Qureshi, Nadeem Khan, Ali Dino Gahoti, and Izat Ali Chachar (Sindh Bureau of Statistics); Khair Muhammad Kalwar (Additional Secretary, Finance Department, Sindh); Eng. Abdul Fattah Tunio (P&D Department, Sindh); Ghulam Jan Naseer and Imran Baloch (Government of Balochistan, Quetta); Raheem Bakhsh and Zai-ur-Rehman (Sate Bank of Pakistan, Karachi); Sehar Saeed (ASER, Idara-e-Taleem-o-Aagahi, Pakistan); Professor Dr. Abid Aman Burki (LUMS, Pakistan); Arvind K. Jain (World Bank Group, Washington). I am thankful to Sher Bugti, Nazir Ahmed, Wali Bugti, Jehand Nazeer, Muhammad Younis, Habib Ahmed and Ifran Ullah who provided me accommodation during my stay in Pakistan for data collection.

I once again say thanks to Christian Rammer and Agnes Korn who treated me as a family member and celebrated all my joys with me and console me in time of distress throughout my stay in Germany. They have invited me with my family at their home many times and their hospitability became the part of my lifetime memories. I am really thankful to Angelika Weber (STUBE-BW, Stuttgart) and DAAD-Neckar group for arranging weekend's seminars and other excursion events that helped me for socialization and integration.

I acknowledge the emotional and administrative supports of my colleagues and friends; Professor Rasool Bakhsh Raisani, Farooq Badiani, Dr. Ameer Bakhsh, Siraj Ahmed, Khalil

Ahmed, Aqeel Ahmed, Gohram Baloch, Irfan Mirwani, Ganguzar Baloch, Shahid Noor, Nadeem Ahmed Latif Baloch, Bahadin Nazeer, and Maqsood Karim.

Special thanks to those who made my life easy during the program: my new and ex-colleagues and friends at the University of Turbat and the University of Balochistan, Dr. Badal Khan, who has been supporting me from the beginning, my elder brothers and specially, Nasir Mazar, who has been supporting me, both financially and emotionally, throughout my career, my in-laws, who took care of my family during my stay in Germany. Last but not least, love for my wife, Zaiba Riaz and my son, Ghazzain Riaz, who bore all those tough time in Pakistan without me and kept patience during my 4-year stay in Germany.

Table of Contents

List of Figures

List of Tables

Abbreviations and Definitions

AEI	American Enterprise Institute
ARF	ASEAN Regional Forum
ASEAN	Association of Southeast Asian Nations
ASER	Annual Status of Education Report
BCG	Bacillus Calmette–Guérin, a vaccine used against tuberculosis
BHUs	Basic Health Units
CMI	Census of Manufacturing Industries
CMRA 1929	Child Marriage Restraint Act 1929
CTP	Critical Threats Project
DB Index	Doing Business Index
DHS	Demographic and Health Surveys
DiD	Difference-in-Difference
DPT	Diphtheria, Pertussis, and Tetanus (DPT): A combination of vaccines used against the three infectious diseases in human
EGI	Ellison and Glaeser Index
FD	Flooded Districts
FDI	Foreign Direct Investment
fGLS	Feasible Generalized Least Square
GCI	Global Competitiveness Index
GDP	Gross Demotic Product
GMM	Generalized Method of Moments
GoP	Government of Pakistan
HDI	Human Development Index
HEIF	High Electricity-Intensive Firms
HH	Household
HPF	High Performing Firms
ICFI	Inner City Fund International
ICT	Information and Communication Technology
IDP	Internally Displaced People/Person
IPCC	International Panel on Climate Change
IPT	Inverse Probability of Treatment
IPTWs	Inverse Probability of Treatment Weights
ISIC	International Standard Industrial Classification
KPK	Khyber Pakhtunkhwa, previously known as NWFP, North West Frontier Province
LEIF	Low Electricity-Intensive Firms
LPF	Low Performing Firms
MDGs	Millennium Development Goals
MFLO	Muslim Family Laws Ordinance (1961)

NDMA	National Disaster Management Authority
NER	Net Enrolment Rate
NFD	Non-flooded Districts
NGOs	Non-Governmental Organizations
NIPS	National Institute of Population Studies
NYC	New York City
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Square
PBS	Pakistan Bureau of Statistic
PSIC	Pakistan Standard Industrial Classification
PSLM	Pakistan Social and Living Standards Measurement
PSUs	Primary Sampling Units
RdRv	Road Paved per Registered Vehicles
RDS	Reduction of Standardized Difference
RHCs	Rural Health Centers
SA	South Asia
SA	Sample Area
SAARC	South Asian Association for Regional Cooperation
SAR	South Asia Region
SCMRA	Sindh Child Marriage Restraint Act (2014)
SD	Standardized Difference
SDGs	Sustainable Development Goals
SSA	Sub-Saharan African
SSUs	Secondary Sampling Units
TFP	Total Factor Productivity
TMT	Terror Management Theory
TT-Injection	Tetanus Toxoid Injection
UNDAC	United Nations Disaster Assessment and Coordination
UNDP	United Nations Development Program
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
WB	World Bank
WBES	World Bank Enterprise Survey
WEF	World Economic Forum
WHO	World Health Organization
WTC	World Trade Center

Chapter 1

Introduction and Background to the Studies

South Asia, the world's 'fastest growing emerging market and developing economy region' (World Bank Group 2017), accommodates one-fourth of the world's population (including India, Pakistan and Bangladesh, the second, sixth and seventh most populous countries of the world respectively) and 40% of the world's poor people. Most of them live in rural areas of the region. In the recent past, South Asia has acquired a momentum growth: since the 1990s, South Asia has obtained an unprecedented annual GDP growth rate of 6 percent (Khilji and Rowley 2013). Real GDP growth has increased from 6.7 percent in 2014 to 7.0 percent in 2015 and is projected to reach 7.2 percent in 2017 (World Bank 2016) and 7.4 percent during 2018-19 (World Bank Group 2017). Despite its high and steady economic growth, on average, the region has failed to achieve many social development goals—e.g. to provide their citizens with rudimentary necessities, minimum living standards, quality of health and education, clean drinking water, hygiene and sanitation facilities, etc.—under the Millennium Development Goals (MDGs) (Kumar and Pandey 2010). However, the level of their failure in obtaining the MDGs varies from country to country in the region depending upon the nature and structure of their economies. For instance, some countries have failed to make social inclusion while others have failed to alleviate inequality and regional infrastructure gaps (World Bank 2016).

In the 21st century's global economic growth and development of the world economy, South Asia has the capacity to play a huge role by utilizing its potential, such as its burgeoning population, rapid growing skilled labors, etc. Unlike many developed countries, the young population of South Asia is growing rapidly and if they properly acquire skills and expertise, these young skilled labors would be demanded in the world market very soon (Khilji 2012).

However, South Asia is facing a number of challenges that relate to several domains including economic and social issues, the environment and security (Palit and Spittel 2013). The main challenges include poverty, deprivation, political confrontation, unemployment, security, corruption, ill-health and illiteracy, child malnutrition, poor human development, social exclusion, inequalities, gender disparities, infrastructure gaps, and government failures (Khilji and Rowley 2013; Kumar and Pandey 2010; World Bank 2016). For instance, just by looking at the current Transparency International's Corruption Perceptions Index 2016, one can easily observe the extent to which the region suffers from corruption—all of South Asia except India is in the red zone (Transparency International 2016).

Unlike the MDGs, the United Nations have widened the horizon of the development goals in the Sustainable Development Goals (SDGs) Agenda, from social development to sustainable development including goals and targets from three dimensions of sustainable development—namely the economic, the social and the environmental. These SDGs are much more integrated and inseparable than the MDGs. Though some goals are well specified and defined while others are not, the agenda itself is above all else universal in nature, covering all countries and people. A brief version of the Agenda's vision which jointly fits to all of the studies of this dissertation reads as follows: "In these [Sustainable Development] Goals and targets, we envisage a world free of poverty, hunger, disease, […] fear and violence, [a] world with universal literacy, […] equitable and universal access to quality education at all levels, to health care and social protection, where physical, mental and social well-being are assured, […] where human habitats are safe, resilient and sustainable and where there is universal access to affordable, reliable and sustainable energy, […] in which humanity lives in harmony with nature and, […] wildlife and other living species are protected" (United Nations 2015, p. 3-4).

This dissertation aims to investigate the relationships among some of the key domains under the umbrella of the SDGs which could ensure sustainable development. In order to obtain sustainable development in South Asia Region (SAR), key areas such as health, education, energy, natural disasters etc., deserve much more attention than others, because SAR is a region where one-fourth of the world population and a huge number of illiterate people with low health facilities live as well being a region which is extremely vulnerable to climate change and natural disasters. For instance, the value of the Human Development Index (HDI) for South Asia shows an urgency of research attention in these areas—the HDI for the region in 2014 was 0.607, which was much lower than the world average, 0.711, and even lower than the average of developing countries, 0.660 (Jahan, et al. 2015). All these areas come in goals of the SDGs; 3, 4, 7, 8, 9, 11, and 13 (see Figure 1.1 on page 8). In order to ensure inclusive and sustainable economic growth and building a productive work force (Goal 1), economies of the developing world should invest and focus on the health sector which warrants healthy lives and well-being for all individuals irrespective of their ages (Goal 3) and on the education sector by providing quality education and lifelong learning opportunities to all individuals without any type of discrimination (Goal 4). Furthermore, a reliable, affordable and resilient infrastructure network such as energy is also a prerequisite for sustainable industrialization and economic growth (see Goals 7 and 11). Ensuring overall human well-being by making their settlement more safe and resilient (Goal 13), these countries should also take immediate actions for climate change and its impacts (Goal 11).

In order to achieve these sustainable development goals by 2030, South Asia will need to maintain its current economic growth above 7 percent per year (World Bank 2016). To uphold this growth momentum, South Asian countries need to cooperate and strengthen their infrastructural networks such as transportation, energy, information etc., within and across the region (Sahoo and Dash 2012). In ensuring an inclusive economic growth, governments of these economies should provide basic health care facilities, quality education, safe drinking water, better sanitation, and sufficient food. Because of the impressive economic growth that SAR has achieved during the last decade, the region has a great opportunity for the first time in history to reduce its poverty by providing basic health, education, and infrastructure facilities which help poor people to come out of the poverty trap. Though South Asian countries have a shared history, they failed to build and sustain trust-based regional cooperation and integrated networks. Because of its geographical location, South Asia is severely vulnerable to environmental hazards such as floods, heat waves, earthquakes, etc. For instance, between 1990 and 2008, the weather-produced disasters smashed about $45 billion amount of region's properties, killed 60,000 people and affected half of the regional population (World Bank 2009).

The purpose of this dissertation is to better understand the role of infrastructure and natural disasters for the socio-economic development of the region. The role of infrastructure in development outcomes and poverty reduction can be seen in both households and in business enterprises. Household uses physical infrastructure such as electricity as a consumption good while business enterprise use it as an intermediate production factor. Infrastructure affects firm productivity directly when it enters into the production process as an input factor but indirectly it also affects technological progress by improving other factors of productivity such as labor productivity (Straub and Terada-Hagiwara 2010). Furthermore, human capital in the form of health and education outcome also plays a huge role in economic growth and development. It directly enhances labor productivity at the firm level by increasing the level of employees' understanding and performance directly, and indirectly through positive externalities, it also creates a 'ripple effect' throughout the economy (Sumarto and De Silva 2014). In order to achieve sustainable growth and development, developing economies face several challenges to build and maintain a desired level of reliable infrastructure on the one hand while on the other hand, managing development outcomes such as mental, psychological and physical health that are vulnerable to the effects of climate change and natural disasters.

The importance of infrastructure development in economic growth has been widely documented in the literature (Aschauer 1989c; Calderón and Servén 2004; Canning and Pedroni 2004; Easterly and Rebelo 1993; Röller and Waverman 2001). Many studies found a positive association between investment for social and physical infrastructure and sustainable economic growth and poverty reduction (Estache 2007; Sahoo and Dash 2012). Though the ongoing debate on the role of infrastructure for economic growth and productivity is inconclusive, the importance of infrastructure capital such as roads, telecommunication, energy, water and sanitation, health and education, etc., for accomplishing development policies is often highlighted (Straub 2008). Straub (2008) also stressed that public infrastructure is essential for both households (they use infrastructure as consumption goods) and business enterprises (they use infrastructure as intermediate consumption goods).

Individual wellbeing in South Asia is not only vulnerable due to a lack of public infrastructure but also due to climate change, which generates natural hazards that affect human wellbeing. For several reasons poor people suffer severely from climate change. But the most frightening issue about climate change is that it will endanger sustainable development by creating a new poor population between now and 2100 in both developed and developing countries (Olsson, et al. 2014). In the fifth assessment report of the intergovernmental panel on climate change, the International Panel on Climate Change (IPCC) argued that natural hazards affect human health in both direct and indirect ways (see Smith, et al. 2014). For instance, variations in temperature, precipitation, heat waves, floods, droughts etc., affect human health directly by injuring the physical or/and mental health of human beings. At the same time, these hazards can also affect individuals' health indirectly through their impact on failure of crops or displacement caused by floods or drought, for example. Beside health issues, weather or climate conditions may also repress human freedom in practicing their culture in terms of 'livings, narratives, world views, identity, community, cohesion, sense of place' (Adger, et al. 2014), and life course transitions**.** Natural disasters may destabilize martial relationships or become hurdles for building new marriage unions.

Since the United Nations' SDGs are closely linked to enhancing human wellbeing, a holistic view on well-being is required that also considers various areas—such as life course transitions, beyond the mental and physical health issues. For instance, in order to acquire and maintain an individual's overall well-being, individuals need to experience sexual satisfaction, pleasure and positive self-esteem that improves not only their sexual health but also mental and physical health (Anderson 2013). According to the World Health Organization's (WHO), sexual health is defined as "a state of physical, emotional, mental and social well-being in relation to sexuality; it is not merely the absence of disease, dysfunction or infirmity. Sexual health requires a positive and respectful approach to sexuality and sexual relationships, as well as the possibility of having pleasurable and safe sexual experiences, free of coercion, discrimination and violence. For sexual health to be attained and maintained, the sexual rights of all persons must be respected, protected and fulfilled" (World Health Organization 2010, p. 3). In developing countries, beside the sexual health advantage of a marriage union, marriage also plays an important role for providing children a favorable environment for health and education. Based on the definition above, individuals in South Asia could be deprived by their basic right of sexual health if they could not afford a marriage union as a consequence of a disaster that hit their region. For example, a premarital sexual relationship is not generally accepted in the societies of the SAR and is considered as a taboo in many Asian countries.

The decline in marriage not only affect the overall human social wellbeing in terms of deteriorating sexual, mental and physical health, but also in terms of economic costs and benefits. There is also evidence in favor of higher income of married individuals than unmarried ones. For instance, Bardasi and Taylor (2008); Blanquicett and Duarte (2016); Cornaglia and Feldman (2011); de Hoon, Keizer, and Dykstra (2015); Lincoln (2008); Pollmann-Schult (2011); and Shtudiner (2015) have all found that married men are more productive and earn much higher income than unmarried men, even after controlling for other factors. There are many reasons that married men gain a higher income and a higher level of productivity, which is beyond the scope of this dissertation (for more detail about this, see Blanquicett and Duarte 2016; Jakobsson and Kotsadam 2016; Pollmann-Schult 2011). In a nutshell, marriage is positively associated with physical and mental health, raising children, wellbeing, financial and other benefits (Waite and Lehrer 2003; Wilson and Oswald 2005). In this regard, governments and NGOs should also consider sexual health along with other dimensions of wellbeing when they come forward to assist and rehabilitate disaster-affected adult individuals.

Due to very complex and dynamic challenges embedded in the South Asia region, it is very difficult to analyze all those issues under one study while giving all issues and countries equal importance. In a series of empirical case studies in the South Asia Region conducted as a part of this dissertation, I have therefore focused my analysis on certain domains within the SDGs: health, education, energy, human settlements and climate changes (see Figure 1.1). The dissertation consists of two main parts. Part I comprises Chapter 2 which discusses the effect of natural disasters on one of the life course transitions, marriage, in Pakistan. In part II, I analyze the association of infrastructure with firm productivity. This part comprises two chapters: Chapter 3 studies the conditional correlation social infrastructure (i.e. health and education) with firm productivity and Chapter 4 analyses the correlation between physical infrastructure (i.e. electricity) and firm productivity by taking Pakistan and South Asia respectively as case studies.

In Chapter 2, titled "***The Impact of Flooding on Marriage: Evidence from Pakistan"***, I examine the causal effect of flood on marriages in flood affected households compared to marriages in unaffected households by utilizing the 2010 Pakistani flood as a type of natural experiment and a difference-in-difference (DiD) analytical strategy. Pakistan experienced a colossal flooding in 2010 which affected almost 10 percent of the country's population and affected 78 out of 140 districts across the country. The flood affected about 18 million people, damaged 1.7 million houses, killed 1,984 and wounded 2,946 people (Government of Pakistan 2011a). The DiD research design allows me to estimate the effect of the flood in flood-stricken districts after the flood by comparing them with unaffected districts before and after the flood and affected districts before the flood. Furthermore, the availability of data on the level of individuals from the Pakistan Social and Living Standards Measurement (PSLM) surveys allows me to disentangle the impact of a natural disaster on marriage which otherwise could not be done in a small sample.

In order to formulate the model for this study, I theorize a model based on two perspectives. The first perspective includes attachment theory (Bowlby 1969) and terror management theory (Greenberg, Pyszczynski, and Solomon 1986; Pyszczynski, Greenberg, and Solomon 1999; Solomon, Greenberg, and Pyszczynski 1991) which predicts that stressful events may bring favorable impacts on life course transitions, such as causing an increase of marriage rate. The second perspective comprises research on economic circumstances which predicts the contrary results.

In line with the above mentioned perspectives, my study contributes to the literature in the following ways. First, to the best of my knowledge, it is the first study which investigates the impact of flood on marriage rates by using a type of natural experiment. Second, I examine the relationship between flood disaster and marriage based on geographic location (rural vs. urban) and gender (male vs. female). Third, I use a rich and updated household survey data and utilize a DiD approach to disentangle both the immediate as well as the long term causal impact of the 2010 Pakistani flood on marriage rate. Fourth, I also investigate the impact of natural disasters on child marriage which may be an interesting debate for policy making purpose.

As above theories and research on economic circumstances predicted, my findings show that the 2010 Pakistani flood caused a net decrease of marriages by 17 marriages per 1,000 individuals aged 15-50 years during the flood year (2010-11) and after that the effect of the flood disappeared. Geographical location matters most for determining the impact of the flood on marriage. For instance, the flood 2010 significantly reduced the number of marriages (25 marriages per 1,000 individuals) in rural areas of flooded districts compared to their counterparts, rural areas of non-flooded districts. Based on the Pakistan's Child Marriage Restraint Act 1929 (CMRA) the minimum marriage age limits—men (18 years) and women (16 years)—I find no evidence for any significant impact of the flood 2010 on child marriage in flooded areas after the flood.

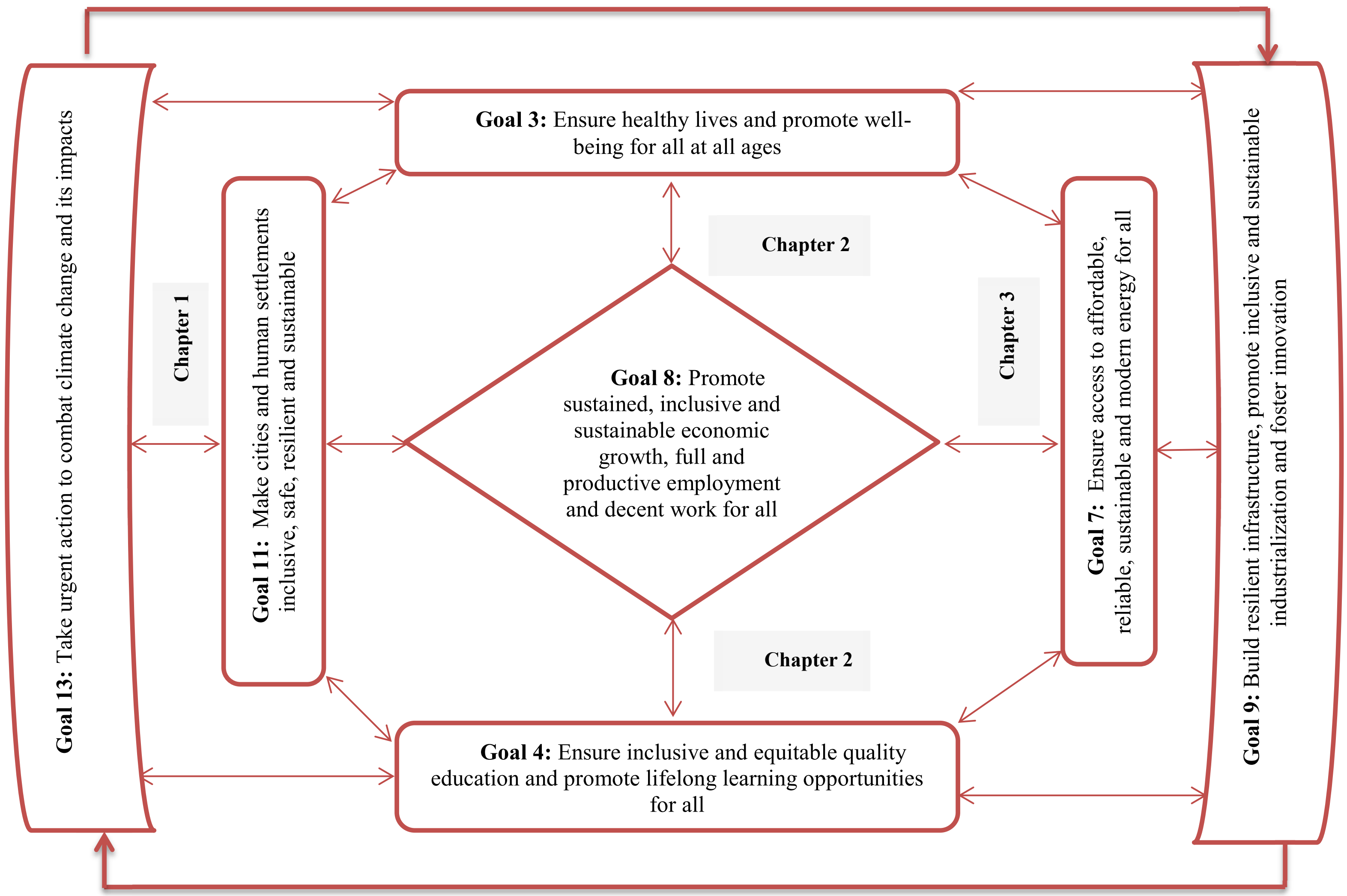


Figure 1.1: Dissertation's Chapters Alignment with the United Nations' Sustainable Development Goals (SDGs)

My findings suggest the following implications for policy. Flood disasters seemingly destabilize adults' romantic relationships and marital life through flood-related economic and financial losses in flooded districts. So in order to compensate or rehabilitate the disaster-affected people fully, policy makers or humanitarian aid-providers should also consider the life course transitions along with other physical or mental health issues, at least for the year directly following the disasters.

In the second essay, Chapter 3, titled ***"Social Infrastructure and Firm Productivity: Evidence from Pakistan"***, I empirically investigate the correlation between social infrastructure indicators at district level and firm productivity by using firm level data from Pakistan. Social infrastructure in this context comprises health and education. Through many channels and positive externalities, both health and education formulate human capital which is one of the most important factors for economic growth, poverty reduction, and human wellbeing and development in general (Sumarto and De Silva 2014).

In this study, I analyze relationship between investment in social infrastructure and firm productivity by utilizing a firm level census dataset on manufacturing firms in Pakistan, the 'Census of Manufacturing Industries 2005-06 (CMI 2005-06),' along with a district level dataset on health and education indicators from the reports of the Pakistan Social & Living Standards Measurement (PSLM) surveys—2004-05 and 2006-07. Following an augmented growth model by Sumarto and De Silva (2014) in which one can add human capital as an additional factor to the production function, I use a cross section firm level dataset (see Chapter 3, Sections 3.4 and 3.5). In this essay, I use the log of firm's output (in numbers) as dependent variable and a set of firm and district control variables along with health and education indicators at district level as independent variables. I measure health infrastructure by two indicators; 1) the percentage of population satisfied with services provided by the basic health units (BHU) at district level, and 2) the percentage of currently married women aged 15-49 years who had a birth in the last three years and received a tetanus toxoid injection during the last pregnancy. Similarly, education infrastructure is also measured by two indicators; 1) total net primary schools enrolment, and 2) the adult literacy rate (for people aged 15 or above). In order to control for a potential selection bias—which could happen due to the firms' location decision since more productive firms may deliberately choose to locate in a region with superior infrastructure supply—I also include an Ellison and Glaeser index (EGI) in the model. I estimate the model by ordinary least square (OLS) estimators.

My results read as follows. For the overall sample, my findings show that the indicators on both types of social infrastructure are positively correlated with firm productivity, conditional on various control. According to my point estimates, if government invest in facilities of basic health units (BHUs) in urban region of a district, and as a result of that investment if the share of the population that is satisfied with BHUs' facilities increase by one percent (say from 40% to 41%), this is associated—ceteris paribus—with firm output by 0.41% which is significant at the conventional levels. The conditional positive correlation between health infrastructure in urban region and firm productivity holds based on various robustness check. The results which I find in my analysis for health infrastructure indicators are within the line of the results of a study by Sumarto and De Silva (2014), who studied the effect of health and education indicator at district level on the subnational growth and poverty in Indonesia.

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

In the final study, Chapter 4, titled ***"Electricity Infrastructure and Firm Productivity: Evidence from South Asia"***, I investigate the association of electricity infrastructure with firm productivity in both manufacturing and service sectors in six South Asian countries, including Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka. Electricity is one of the main sources of energy and a necessary input in the production process for both manufacturing and service enterprises in developing countries (Fisher-Vanden, Mansur, and Wang 2015). Shortage and unreliability of electricity supply may markedly affect the productivity of these enterprises.

In line with the theoretical framework postulated by Carlin, Schaffer, and Seabright (2006) and the model by Moyo (2013), I formulate an empirical model that allows me to embed the electricity infrastructure along with other firm's input variables in a firm production function. In this model, I measure firm output, the dependent variable, as the logarithm of sales volume and electricity infrastructure, my variable of interest, by two types of proxies—subjective-qualitative and subjective-quantitative—including electricity shortages, losses due to electricity shortages, electricity perceived as an obstacle for business operations, etc. I use a subjective-qualitative measure for electricity infrastructure in which managers or owner of a firm were asked to rate the electricity is as an obstacle for doing business operations on a 5-point Liket-type scale (where 0 shows no obstacle and 4 shows a very severe obstacle) as key variable. The data for this study comes from the World Bank Enterprise Surveys (WBES) for the six South Asian countries; Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka covering the years 2002 to 2014 (see Table 4.6 on page 144). For empirical analysis, I pool enterprise surveys' data across countries and obtain OLS regression estimates—clustered at the regional level. Regions, in this study, are the second tiers of administrative units (e.g. states in India and provinces in Pakistan).

In this study, I find the following results: Compared to the base category (in which electricity is not an obstacle), I find that the OLS estimates of all four categories are negative and significant at the 10% level except in one case. Interestingly, the magnitudes on the coefficients of all categories are equal, indicating that the electricity infrastructure based on managers' perception is closely related to firms output. On average, the percentage of sales volume of firms which complain that electricity is an obstacle, is 8-12% lower than for the firms in base category. In addition, I test the hypothesis whether the correlation between electricity infrastructure and firm productivity would be different for a firm if it had an internal capacity to substitute the publicly provided electrical power by its own power generation at the plant level. Compared with the base category, my findings indicate that both types of firms, with or without own/share generator facilities, seemingly suffer by power outages from public grid. But the level of this suffrage is much higher and robust for firms that own/share generators compared to those which do not. This may be because of a) in developing countries, electricity generation and distribution at government disposal costs enterprises much less than their own power production, b) firms that have their own power generation capacity suffer the most due to the high cost of fuels for generating power at their disposal. I also use subjective-quantitative measures as proxies for electricity infrastructure, one of which is the percentage of sales losses due to the power outages. I find that, on average, the percentage of total annual sales losses due to power outages and the annual sales of a firm are negatively correlated.

Based on different alternative measures and robustness checks, I conclude that there exists a negative conditional correlation between the poor quality of electricity infrastructure and firm productivity in South Asia and the main subjective-qualitative measure seems to estimate the correlation better than other measures. My overall findings suggest that even though there exists evidence for a negative conditional correlation between electricity as an obstacle and firm productivity, the relationship is not too strong economically and/or statistically. I conclude that the World Bank enterprise surveys—the only source of data for cross country analysis in South Asia which provides the basic insights about how business enterprises face infrastructure obstacle in their daily business operations—could not be used for generalizing the results for policy recommendations.

Chapter 2

The Impact of Flooding on Marriage: Evidence from Pakistan

Abstract

This study designs to examine the causal effect of the flood on marriages in flood affected households compared to marriages in unaffected households by utilizing the 2010 Pakistani flood as a type of natural experiment. A difference-in-difference approach based on a district and province-year fixed effects model is used to estimate the impact of the flood on individual marital life. The goal of this paper is twofold. First, it investigates an overall impact of the flood on marriage rate in the flooded districts compared to non-flooded districts, and second, it further scrutinizes the effect, if any, on individual marriage behavior based on geographic (rural vs. urban) and gender (male vs. female) sample distribution. Results show that the July 2010 Pakistani flood decreased marriage rate in flooded districts immediately during the disaster (2010-11) which discontinued from the second (2012-13) to the fourth (2014-15) years following the flood. Geographic location matters a lot for such type of natural disasters. In this study, the negative impact of the flood on marital life is significantly, both economically and statistically, higher for rural individuals than urban ones.

Keywords: Life Course Transitions, Marriage, Natural Disasters, Flood, Pakistan

JEL classification: H84, I3, J12, Q54

2.1 Introduction

Disasters in general and natural disasters in particular destroy public and private infrastructure and properties, cause fatalities and injuries, create traumatic mental and health issues in people lives and specially push developing countries further below in terms of economic growth and poverty reduction. Natural disasters shake human well-beings at individual, family and community level (Cohan and Cole 2002). To this point, researchers nowadays agreed that most of the severe natural disasters are caused due to the climate changes (Kousky 2014). While the most immediate impacts of natural disasters get a lot of attention from media, governments and NGOs, the focus of the aftermath effect of the events mostly remains calm, perhaps due to the difficulty in quantifying long-term impacts (Anttila-Hughes and Hsiang 2013). In order to understand the mechanism of economic growth and development of developing countries holistically and to devise long term policy goals for these economies such as MDGs and SDGs, as I have already explained in the first chapter, policy makers and international organizations should not only look at the immediate impacts on social and economic lives of people caused by natural disasters. It is also important to consider the long-term effects on mental health and life course events.

The issue of disasters' impacts on mental health has received much attention in research. But only until recently the impacts of disasters on life course transitions such as marriage, divorce, birth, etc., had not been studied. Though there exists a sound theoretical background for studying these issues (see Section 2.2), it seems that only after the 9/11 attack on WTC in New York City, researchers started to look at on this type of consequences of disasters. Most of the previous studies, either at micro level or country-specific, have been focused on some very specific type of events in the past: a) man-made disasters such as the 9/11 attack in NYC (Cohan, Cole, and Schoen 2009; Hansel, Nakonezny, and Rodgers 2011; Ruther 2010), the 1995 Oklahoma City Bombing (Nakonezny, Reddick, and Rodgers 2004), or b) natural disasters such as the 2011-East Japan Earthquake (Hamamatsu, et al. 2014; Hamamatsu, Watanabe, and Umezaki 2015; O'Donnell and Behie 2014, 2015), the 1997-Umbria-Marche earthquake, Italy (Prati and Pietrantoni 2014), the earthquakes in Turkey (1999), India (2001), and Pakistan (2005) (Finlay 2009), and the 1989 Hurricane Hugo (Cohan and Cole 2002). Perhaps due to the limited availability of data, research on the impacts of floods on life course transitions, to the best of my knowledge, have not received any attention at all. This study tries to fill this gap in the literature by utilizing a very rich and representative survey-based dataset from Pakistan and investigating the 2010 Pakistani flood as a natural experiment.

This study revisits and tests the research question posed by Cohan and Cole (2002) whether natural disasters have any significant impacts on 'life course transitions'. Cohan and Cole (2002) tested the impacts of the 1989 Hugo Hurricane, South Carolina on the life course transitions (marriage, divorce and birth rates) and found that the disaster increased marriage rate in the affected areas. This relationship is also validated by Xu and Feng (2016) who used Chinese province level panel data (31 provinces) between 2000 and 2011 and found that earthquake disasters increased marriage rate in China. By contrast, a number of studies find negative relationship between disasters and marriage rate. For instance, Prati and Pietrantoni (2014) found a decrease in marriage rate in 15 affected municipalities due to the 1997 Umbria-Marche earthquake in Italy. In a similar study, Hamamatsu, et al. (2014) also found the evidence of marriage reduction after the East Japan Earthquake 2011. A disaster may affect marital life either positively or negatively, depending upon the circumstances or the level of severity the disaster creates (Cohan, et al. 2009). Literature tells us that the exposure from a disaster may bring some psychological issues in individual life such as parenting practices, but the issue of the effect of natural disasters on marital life has rarely been explored (Lowe, Rhodes, and Scoglio 2012). So the focus of this study is to investigate the aftermath effect of the 2010 Pakistani flood on marriage rate by utilizing individual level data from Pakistan.

Pakistan experienced a colossal flooding in 2010 which affected almost 10 percent of the country's population and affected 78 out of 140 districts across the country. Even though the number of death tolls was not as high as caused by the 2005 earthquake in Kashmir, in terms of affected areas and people, it was the largest flood in the history of Pakistan. The flood affected about 18 million people, damaged 1.7 million houses, killed 1,984 and wounded 2,946 people (Government of Pakistan 2011a, p. 14). The main origin of the flood was the Indus River, and both sides of the Indus River Basin were affected by the flood. It destroyed public and private infrastructure stocks; namely, roads and bridges, water and sanitation networks, agriculture lands and crops, and live-stocks, which further affected the trembling economy of the country, including its human capital stock and households' social well-being. The most heavily affected region in Pakistan was the Sindh province.

This study investigates the impact of the July 2010 flood on marriage across all affected and unaffected districts of Pakistan by using a difference-in-difference (DiD) analytical strategy. The DiD research design allows me to estimate the effect of the flood in flood-stricken districts after the flood by comparing them with unaffected districts before and after the flood and affected districts before the flood. Furthermore, the availability of data on the level of individuals from the Pakistan Social and Living Standards Measurement (PSLM) surveys allows me to disentangle the impact of a natural disaster on marriage which otherwise could not be done in a small sample.

This study contributes to the literature in various ways: First, to the best of my knowledge, it is unique of its kind in the sense that there is no such work has been previously done which studies the impacts of floods on marriage rates by using a type of natural experiment. Second, it does not only examine the relationship between flood disasters and life course transitions at household level but it also digs down deeper to find out whether or not there exists any impacts of the flood disaster on individual marital lives in flood affected households based on geographic location (rural vs. urban) and gender (male vs. female). Third, this paper uses a rich and updated household surveys data and utilizes a DiD approach to disentangle both the immediate as well as the long run causal impact of the 2010 Pakistani flood on marriage rate. Fourth, it also briefly looks at the impact of natural disaster on child marriage which may be an interesting debate for policy making purpose.

Finally, based on the attachment and terror management theories (see Section 2.2), as suggested by Cohan and Cole (2002), this study also tests the given question in a different country context, Pakistan—an Islamic country where dating or courtship is not allowed and according to the Muslim Family Laws Ordinance, 1961, a man is legally allowed to marry more than one wife (National Legislative Bodies/National Authorities 1961). The general predictions of the above theories dictate that disasters may increase marriage rate in the disaster-stricken-regions. But in case of Pakistan, religious, social and legal constraints do not allow individuals' behavior—toward marriage—to vary. So the predictions could be explained based on the above mentioned theories as well as economic circumstances and according to these bases, natural disasters either increase or decrease marriage rate in the disaster affected areas (see Section 2.2 for explanation). The flood severely damaged and caused economic losses to individuals in flooded districts, so the marriage decision of these individuals could mainly be influenced by their current economic circumstances while the positive impact of the flood on individual marriages in both flooded and non-flooded districts also cannot be ruled out based on sociology and psychological factors—the positive effect in non-flooded districts could be happened due to the spillover effects. In time of the disaster, people in non-flooded districts could be also be exposed to the flood consequences such as NGOs' volunteers' works, etc. for collecting donations and other basic living stuffs for the affectees or via mass or electronic media that eventually could affect individuals' behavior and enhance the desire of attachment and marital solidarity.

In line with the results of prior research, my findings show that the 2010 Pakistani flood decreased marriages, 17 marriages per 1,000 individuals aged 15-50 years in the flooded districts during the flood year (2010-11). When comparing individuals from the most severely affected districts with individuals in the unaffected districts, 2 points (17 to 19 marriages per 1,000 individuals) further reduced in the most severely affected districts during the flood. The 2010 flood had had an immediate shock on marital life and there is no evidence further for the long term effect. Geographical location matters alot for determining the impact of the flood on marriage. For instance, the 2010 flood significantly reduced the number of marriages (22 marriages per 1,000 individuals) in rural areas of flooded districts compared to their counterparts, rural areas of non-flooded districts. Based on the Pakistan's Child Marriage Restraint Act 1929 (CMRA) the minimum age limits—men (18 years) and women (16 years)—I find no evidence for any significant impact of the 2010 flood on child marriage in flooded areas after the flood.

The rest of the chapter proceeds as follows: Section 2.2 explains the theoretical background for this study and based on those theories, this section also includes the main hypotheses to be tested in this study. Section 2.3 describes data sources and key descriptive statistics. Section 2.4 discusses empirical modeling framework, followed by a results and discussion section 2.5. Section 2.6 includes some sensitivity checks of the main findings and expands the analysis to child marriage and migration effect. Finally, Section 2.7 concludes and suggests policy implication.

2.2 Theoretical Background and Hypotheses Derivation

On December, 2004, an earthquake of 9.1–9.3 (on moment magnitude scale) struck beneath the Indian Ocean near Indonesia (Satake and Atwater 2007), which caused a giant tsunami across 14 countries and killed about 230,000 lives. Its impacts were too severe that were felt in the eastern parts of Africa. Though it was really a deadly disaster, it created a wonderful story of love, affection, and relationship between two entirely different species; hippopotamus and tortoise. The story began on the very day near Malindi town, Kenya when a heavy rainfall drowned a group of hippopotamuses down the Sabaki River and then to the sea. On the next day, only a baby hippo (one year old) was found, rescued, named 'Owen' and transported to the Hall Park. In Hall Park, when Owen was released, it immediately ran to a giant 130-years-old tortoise, Mzee, perhaps for the purpose of safety, love, affection or any other instinctual drives. The bond between Owen and Mzee remained for a very long time and then they were separated deliberately because now the grown-up hippopotamus began damaging the shell of the tortoise and needed other hippopotamuses for socialization. Though this true story scientifically proved nothing about what was going on in their minds during their touching, trolling, etc., it moved thousands of people, media, reporters, etc., around the world to watch this natural bondage with their own eyes. It became so popular that Hatkoff and Hatkoff (2016) wrote a book, titled, 'Owen and Mzee: The true story of a remarkable friendship'. This is relevant to the current study based on evolutionary perspective (explained below); when species including Homo sapiens face life threating events, such as disasters, then they look for caretakers.

The theoretical base for this study comes from sociological and psychological (Cohan and Cole 2002) as well as economic literature, from which the prediction of disasters' impacts on life course transitions and marriage, in this particular case, can be explained. From the known literature, these can be grouped into two categories: The first group includes theories which predict that stressful events may bring favorable impacts on life course transitions, such as causing an increase of marriage rate. This group includes two evolutionary perspective theories; the attachment theory (Bowlby 1969) and the terror management theory (Greenberg, et al. 1986; Pyszczynski, et al. 1999; Solomon, et al. 1991). Second group comprises research on economic circumstances which predict the contrary results.

Starting with the first group, the attachment theory, postulated by Bowlby (1969), is very popular in this arena which states that 'marital solidarity' may occur after a disaster. According to the attachment theory when a disaster happens, then people would like to be near to their closed or loved ones and in case of severe dangers, both adults and children run toward their proximate relationships for security and relaxation. One of the important assumptions of the theory is that the attachment system is active over the entire life span which can also explain the adults' behavior for seeking association and support in time of distress (Hazan and Shaver 1994). Furthermore, Cohan, et al. (2009) also suggest that marital bonds become more strong in case of extreme events which eventually may cause an increase of marriage and a reduction of divorce rate in disaster affected areas after a disaster because the role of being a romantic partner rather than stay attached to parents or guardian would change after a disaster and dating couple become life partners and get and maintain more intimacy to in partnership (Prati and Pietrantoni 2014). Based on this theory marriage rate in flood-affected districts compared to unaffected districts in Pakistan, due to the July 2010 flood, should increase.

The second theory is the terror management theory (TMT), which has been postulated by Greenberg, et al. (1986); Pyszczynski, et al. (1999); Solomon, et al. (1991) based on Ernest Becker's 1973 book, 'The Denial of Death'. The theory states that a psychological conflict arises in human beings as a result of having desire for being immortal but having awareness that death is not escapable and this conflict leads to terror. There are two fundamental realities hidden in the theory i.e. 1) human with other species have instinctual drives for reproduction and 2) only human beings have awareness of their death (Mikulincer, Florian, and Hirschberger 2003). In order to manage the terror and mitigate or remove the awareness of their own death from their consciousness, human beings create symbolic systems, which attach meanings and values to life; for instance believing in immortality or afterlife in religious sense. But reminders of death such as disasters may trigger again the sense of attachment (Cohan, et al. 2009), and these death reminders may enhance the inspiration of building and maintaining the close relationships which could provide a camouflage against the awareness of their own death while breaking the close associations may enhance the awareness of death (Mikulincer, et al. 2003). Again based on this theory, marriage rate should increase in flooded districts after the disaster.

In contrast to the above mentioned theories, a second base of my hypotheses formulation comes from the research on economic circumstances. Current economic circumstances of individuals such as real wages, employment opportunities, etc., affect individuals' marriage decisions. Poor economic circumstances may discourage them to be engaged in long term relationships because these circumstances influence their ability to build separate households (Oppenheimer 1988; White and Rogers 2000). Economic benefits such as high earnings, secure jobs, etc., of both genders are associated with life course transitions; higher economic advantages lead to more marriages and vice versa. So, based on research on economic circumstances, marriage rate in flooded district should increase after the flood.

Empirical evidence of disasters' impact on marriage rate is rare—if any—rather mixed. To the best of my knowledge there are only four papers available so far which study the impacts of disasters on the given life course transition, marriage. Three of them (Hamamatsu, et al. 2014; Prati and Pietrantoni 2014; Xu and Feng 2016) study earthquake impact and one of them (Cohan and Cole 2002) studies hurricane impact on marriage rate. Cohan and Cole (2002) and Xu and Feng (2016) find a positive while Hamamatsu, et al. (2014) and Prati and Pietrantoni (2014) find a negative relationship between natural disasters and marriage rate.

Based on disaster triggering agents, disasters can be grouped into two major categories; natural disasters and man-made disasters. Disasters' impact on life course events would vary depending on the nature and type of the disasters. Since comparatively natural disasters such as floods tend to cause more physical damages while man-made disasters such as terrorist attacks cause more human losses (Cohan, et al. 2009), marriage rate is likely to decrease in favor of economic loss perspective (former situation) while it is likely to increase in favor of attachment and terror management theories (latter situation).

In a nutshell, based on the above theoretical background and empirical literature, this study tries to answer the following questions:

Q1: Had the July 2010 Pakistani flood have any significant impact on individual marriage in the flooded households?

Q2: If yes, how long does it take these households to adapt well in the face of the hardships?

Q3: Did the July 2010 Pakistani flood affect households differently based on their geographic location they reside (rural vs. urban) and their gender (male vs. female)?

One of the novelties of this study is to investigate these hypotheses based on the attachment theory (Bowlby 1969) in the context of Pakistan—an Islamic country where certain constraints on life course transitions already exist. For instance, intimacy before marriage is not allowed while polygamy is allowed which may distort the direction of relationship between disaster and marriage. Even with similar type of disaster in different countries, its impacts on marital life would vary due to their cultural differences dealing with such disastrous events (Prati and Pietrantoni 2014). Thus the effect of the 2010 Pakistan flood on marriage rate could go in both directions.

2.3 Data Sources and Descriptive Statistics

2.3.1 *Marriage Data*

Marital status of individuals as well as other individual characteristics and household level data are taken from the Pakistan Social and Living Standards Measurement (PSLM) surveys. These are household level surveys conducted by Pakistan Bureau of Statistic (PBS) for each district in a biennial rhythm (Government of Pakistan 2005, 2008b, 2010, 2011b, 2014, 2016). The project was initiated in 2004 and has been completed in 2015. In this study, I use all six waves of surveys, i.e. 2004-05, 2006-07, 2008-09, 2010-11, 2012-13, and 2014-15. This dataset provides a very huge source of information about individuals' education, health, employment, household living standards, assets in possession, marital life, etc., in a household as well women and children health related information in Pakistan. These surveys captured almost all of the districts[1] from the four provinces of Pakistan; Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan, which provide about 3 million observations at the individual level for all six years (roughly speaking 0.5 million per year). Of course not all of the individual level observations are useable for investigating the research questions of this paper. For more detail about sampling distribution at the provincial as well as the district level and primary sampling units (PSUs) and secondary sampling units (SSUs) kindly refer to Table 2.3 and Table 2.4. The outcome variable of interest, marriage as an indicator variable, is derived from the marital status information in the surveys. In the surveys' questionnaire, marital status is coded as; never married (1), currently married (2), widow/widower (3), divorced (4), and nikkah solemnized but rukhsati not taken place[2] (5). I use only the first two categories for marriage variable, never-married (coded as 0) and currently-married (coded as 1) which represent 98.27% (60.15% married and 38.12% unmarried) of all individuals aged 15-50 years in the sample.

2.3.2 *Geospatial Data for Flooded and Non-flooded Districts*

Geospatial data for treated (flood-affected districts) and comparison (flood-unaffected districts) groups for the main analysis, comes from two maps: First map (see Figure 2.3 on page 59) is taken from the MapAction (MapAction 2010). The MapAction closely works with the humanitarian crisis teams of the United Nations Disaster Assessment and Coordination (UNDAC), the United Nations High Commissioner for Refugees (UNHCR), the Red Cross/Red Crescent Movement and international NGOs. Its emergency mapping services provide more immediate assistance and response to more than 60 types of humanitarian crisis including floods. The map used in this analysis, is created on 6 September 2010, which color-coded districts according to the level of severity the districts affected due to the 2010 flood—moderately affected districts if less than 100,000 people were affected and severely affected districts if more than 100,000 people were affected in a district. The second map (see Figure 2.4 on page 60) created by the Critical Threats Project (CTP), which was formed in 2009 by the American Enterprise Institute (AEI). The purpose of producing this map was to track the

[1]Few exemptions: Due to the law and order situation in Pakistan, some PSUs have been dropped from the surveys—in 2014-15 survey 13 PSUs from KPK, 7 from Sindh, and 82 from Balochistan and in 2012-13 survey, 1 PSU from Sindh and 26 PSUs from Balochistan.

[2]Nikah (نِکاح) in the Islamic legal system implies a marriage contract or official wedding ceremony, when solemnized, meaning that the marriage contract has been done, but Rukhsati (رُخصتی) is the departure (farewell) ceremony of the marriage, in which the groom's family will leave together with the couples to their own home. Actually, from Nikah to Rukhsati, it often takes several days.

flood affected areas, the level of damages and their potential consequences on the population and to guide international donors to response to the disaster by aids and other assistances. With very few exemptions[3] in which the severity levels of the floods differ in both maps, they both declared almost the same districts as flood affected and unaffected districts. For the purpose of this analysis, I utilize geospatial data provided by the above two maps to create a treated group of flood affected districts (62 districts) if the 2010 Pakistani flood affected the districts irrespective of the level of severity and a comparison group of districts (53 districts) which were not affected by the 2010 flood. From the maps, it is clear that these affected districts were tightly clustered across the Indus River.

For further robustness checks and sensitivity analysis, I use another source of geospatial data for the treatment assignment, the National Disaster Management Authority (Government of Pakistan 2011a), along with the aforementioned two maps to create a new treated group of flood exposed districts. These districts were severely affected by the 2010 flood and needed immediate aid response. For comparison group, I still use the maps' unaffected districts. Moderately affected districts mentioned in the maps are excluded from this analysis (see Table 2.5 on page 46 for severely affected districts).

2.3.3 *Descriptive Statistics*

Table 2.1 shows descriptive statistics. After cleaning the data, I obtained about 1.4 million individual level observations; 370,701 and 342,887 individuals respectively residing in non-flooded and flooded districts before the flood (2004-05 to 2008-09) and 349,002 and 363,017 individuals respectively residing in the non-flooded and flooded districts after the flood (2010-11 to 2014-15).

Only a slight difference between the number of marriages, on average, in flood affected and unaffected districts in Pakistan before the flood was observed. For instance, about 63% of adults, aged 15-50 years, had been married in flood-affected districts while about 59 % in unaffected districts before the 2010 flood. But during and after the flood, the percentage of marriages increased in both—flood affected and unaffected districts—but in affected districts, it increased by just 0.18% while in unaffected districts, it increased by 1.32% which is significantly greater than the former districts. A rural-urban comparison, marriage rate is significantly higher in rural areas, but rural marriage rate decreased in flood-affected districts (65.14% to 64.69%) while increased in unaffected districts (62.54% to 63.06%) after the 2010

[3]For instance, Mardan was declared as a severely affected district in the CTP while it was declared a moderately affected district in the MapAction. Only one district, Lasbela, was declared as flood-affected districts in MapAction while not in the CTP.

flood. There is also a huge difference in men and women marriage rates. For example, 52.08% of men and 66.39% of women aged 15-50 before the floods while 53.66% of men and 67.19% of women after the flood in unaffected districts got married. In flood-affected districts, the average percentage of married men was 54.60% before the flood, which increased to 54.98% after the flood while 70.94% of women married before the flood, which reduced to 70.64% after the flood. The difference in the percentage of marriages between women and men may be due to the Pakistani polygynous society.

Table 2.1: Descriptive Statistics of Outcome Variable of Interest (Marriage)

Outcome Variable: Marriage	Non-flooded Districts(a)		Flooded Districts(a)		Severely flooded Districts(b)	
	Obs.	%	Obs.	%	Obs.	%
Before Flood: 2004-05 to 2008-09						
Total	370,701	59.07	342,887	62.69	99,472	63.34
Rural	205,301	62.54	241,090	65.14	74,457	65.33
Urban	165,400	54.75	101,797	56.91	25,015	57.41
Male	189,669	52.08	173,006	54.60	50,437	55.09
Urban Male	85,353	48.36	52,187	50.06	12,757	50.51
Rural Male	104,316	55.12	120,819	56.56	37,680	56.65
Female	181,032	66.39	169,881	70.94	49,035	71.81
Urban Female	80,047	61.56	49,610	64.12	12,258	64.59
Rural Female	100,985	70.22	120,271	73.75	36,777	74.21
During and After Flood: 2010-11 to 2014-15						
Total	349,002	60.39	363,017	62.87	117,076	63.27
Rural	220,047	63.06	270,606	64.69	92,049	64.87
Urban	128,995	55.82	92,411	57.54	25,027	57.36
Male	175,455	53.66	180,085	54.98	58,813	55.25
Urban Male	65,630	49.53	47,129	50.85	12,835	50.62
Rural Male	109,825	56.12	132,956	56.44	45,978	56.55
Female	173,547	67.19	182,932	70.64	58,263	71.36
Urban Female	63,325	62.35	45,282	64.49	12,192	64.46
Rural Female	110,222	69.98	137,650	72.66	46,071	73.18

Notes: Outcome variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). (a) Districts according to the Maps; the MapAction and the Critical Threats and (b) districts according to the NDMA. Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan and author's own calculation.

According to the PSLM surveys data, the ratio between married women to men aged 15-50, on average, is about 1.12, which makes sense in the context of Pakistan because in Pakistan, according to the Muslim Family Laws Ordinance, (1961), a man is legally allowed to marry more than one wife conditionally based on obtaining a written consent (permission) from his existing wife/wives. However, the number of men, who currently married more than one wife, is very small. For instance, the DHS[4] (2012-13) observed about 3.4% of men in total, 4% in urban and 3% in rural areas of Pakistan were in polygynous unions (National Institute of Population Studies - NIPS/Pakistan and ICF International 2013).

From the PSLM surveys dataset, it is also observed that there exists a huge gap in the percentage of marriages between flood affected and unaffected districts to begin with (see Figure 2.1, A). This may be due to the covariates' differences in both groups of districts. In order to obtain biased-reduced estimates, I reweight the sample based on Inverse Probability of Treatment Weights (IPTWs), which I will explain in the next section.

In order to answer this very specific kind of question in the context of Pakistan, PSLM, the only source of information on life course transitions, is available at individual level. There is another one, the Demographic and Health Surveys (DHS), 2006-07 and 2012-2013, which could provide better information on outcome variable, marriage rate, but still could not be used in this specific context for the following two main reasons: First, the time period between these two waves is too large to identify the immediate impact of the 2010 Pakistani flood on marriage rate. Because DHS (2006-07) was conducted four years before the 2010 flood and DHS (2012-13) was conducted two years after the flood. In addition, it is very challenging to isolate the time-varying factors' effect in estimation of the coefficients of interest given that there are only two waves of survey data. Second, though the DHS surveys data, based on provincial urban and rural areas, were claimed to be a nationally representative sample (National Institute of Population Studies - NIPS/Pakistan and ICF International 2013), these data could not be disaggregated at district level because the sample data at district level do not necessarily represent the district population. Thus in order to test the given questions in this paper, the PSLM surveys dataset, which is free from the above mentioned drawbacks, is used. Furthermore, to test the reliability and quality of data on marriage in the PSLM surveys, I make a comparative descriptive statistics analysis between the PSLM and the DHS datasets (see Table 2.6). These two data sets provide fairly exactly the same percentage distribution of men and women marital status. For instance, according to the PSLM survey (2012-13), 48% of men and 32 % of women and to the DHS survey (2012-13), 48% of men and 33 % of women have never been married while 52% of men from the both surveys and 66% (PSLM) and 64% (DHS) of women showed their current marital status—married. Thus these two datasets provide quite similar statistics and because of rich and updated data, I use the PSLM surveys dataset to test the given hypotheses in this study.

[4]Demographic and Health Surveys (DHS).

2.4 Empirical Analysis

2.4.1 *Modeling Framework*

In order to identify the areas that were affected by the 2010 flood, I use two maps; one was generated by the MapAction and the other was generated by the Critical Threats Project (CTP). Each map shows the districts declared as flood-stricken by the July 2010 flood in Pakistan (see maps in Appendix to Chapter 2). Based on these maps, treated areas (flood affected districts) and comparison areas (flood unaffected districts) have been separated by creating an indicator variable which holds a value of one if a district is affected by the 2010 flood and zero otherwise.

In order to build a dataset for this study, I combine the geocoded household level survey data, which come from the six waves of the district level household surveys, Pakistan Social and Living Standards Measurement (PSLM), with the above mentioned control and treated groups. The PLSM surveys cover a decade of household data (2004-05 to 2014-15), three waves (2004-05, 2006-07, and 2008-09) before and three waves (2010-11, 2012-13, and 2014-15) during and after the 2010 flood. There are 115 districts in total, out of which 62 districts are in the treated and 53 districts are in the comparison groups. In this study, I use district level variations in damages caused by the 2010 flood in order to test the effect of the flood on one of the life course transitions, marriage, by applying a difference in difference (DiD) approach.

The PSLM surveys dataset does not provide only a rich source of individual level information but these surveys have been collected data, independent of the natural disaster, which may reduce household biases towards exaggerating or misrepresenting of data in which they usually do after a disaster in anticipation of getting more reliefs or empathy from governments, NGOs, etc. To analyze the impact of the 2010 Pakistani flood on marriage, I use a province-year and district fixed effects regression model:

$$y_{idpt} = \alpha_{dp} + \gamma_{pt} + \beta_1 D11_t * flood_{dp} + \beta_2 D13_t * flood_{dp} + \beta_3 D15_t * flood_{dp} + \varepsilon_{idpt} \quad (2.1)$$

where y_{idpt} is an outcome variable, which represents whether an individual i, in d district of province p in time t is married. The variable $flood_{dp}$ is, whether a district is exposed to the 2010 Pakistani flood and *D11*, *D13* and *D15* are dummy variables representing years during and after (2010-11), second (2012-13), and fourth (2014-15) years following the flood respectively. The unit of observation in this study is individual in household who is whether married or unmarried within the age window 15-50 years. β_1, β_2 and β_3 are the difference in difference estimates of the coefficients of interest which represent the immediate (2010-11), second (2012-13) and fourth (2014-15) years' differential effect of the flood on individual marriages in the flood exposed districts compared to individual marriages in non-flooded districts. Regardless of 2010 flood, there might be some systematic trends in data over time, so by inclusion of province-year fixed effects, γ_{pt}, in the model specification may capture time-varying effects of marriage as well as other covariates. Similarly, there would also be other geographical, demographic or time-invariant differences among districts, thus including district fixed effects may capture these unobservable differences at district level. In order to estimate the βs (parameters) of the above model, the usual OLS estimation method is used. All regressions are clustered at the district level to obtain robust standard errors. In order to avoid obtaining unbiased/reduced-biased estimates due to the covariates' differences between flooded and non-flooded individuals, the sample is reweighted for balancing the flood exposed and un-exposed individuals' covariates. The weights are obtained by a method, inverse probability of treatment weighting, which I will explain in the next sub-section.

2.4.2 *Inverse Probability of Treatment Weighting: Balancing Flooded and Non-flooded Individuals' Observable Covariates*

In a randomized controlled trial setting, the effect of treatment on outcome can be measured directly by comparing the average outcome of treated group with the average outcome of control group because randomization by definition ensures the similarities between treated and control groups' observed and unobserved covariates. By contrast in an observational study setting, one cannot measure the impact of treatment on outcome directly because in such a setting, the effect of treatment on outcome may be subjected to treatment selection bias (Austin and Stuart 2015). In this study, however the case of the 2010 Pakistani flood is a type of exogenous shock for people who were residing in the country at the time of flood. The flood affected people in different districts systematically in different ways. For instance, people in Pakistan whose living sustenance come from the agriculture sector, compel to dwell near to rivers (which mostly caused flooding) have less access to education and health services, foods, water and sanitation facilities, etc., and these people have suffered the most in the flood in Pakistan (Deen 2015). Maps 3 and 4 also depict these flooding districts which were mostly agglomerated across the Indus and other major rivers in Pakistan. So in order to obtain unbiased or reduced-biased estimates, balancing of individuals' covariates in flood affected and unaffected districts, is essential. In this study, I obtain a reweighted-balanced-pseudo-sample by using the inverse probability of treatment weights (IPTWs) based on flooded and non-flooded individuals' observable covariates.

Estimating the exposure effect on outcome, measurable confounding and selection bias, based on time varying covariates, can be taken into account by using IPTWs (Cole and Hernán 2008). The objective of weighting individuals based on IPTWs is to create an artificial sample in which treatment assignment is independent of the measured baseline covariates (Austin and Stuart 2015). Following Lunt (2014) guide for propensity analysis, propensity scores for the IPTWs are estimated by using a logistic model in which treatment assignment (flooded district) is regressed on the 22 covariates[5] along with survey-year fixed effects. The predicted value of the outcome variable is the propensity scores of individual. Austin and Stuart (2015) define propensity score is the probability of an individual being treated conditional on observable baseline covariates. Next, I estimate the IPTWs by the following way; individuals in the flood stricken districts get a weight of 1/Propensity Score and individuals in the unaffected districts get a weight of 1/ (1- Propensity Score).

One of the benefits of using propensity score methods is that they almost mimic randomized controlled experiments in which research design is separated and independent from the main analysis of the treatment effect on outcome variable. The weights obtained by the inverse probability of treatment weighting method actually alter the distribution of covariates of individuals in flooded and non-flooded districts in such a way that reduces their distributional gap (Lunt 2014). According to the author "the IPT weighted analysis therefore compares what we would expect to see if everyone received treatment to what we would expect to see if no-one received treatment". Weighting, furthermore, allows me to use the entire sample without discarding any individual from the sample while reducing the bias by giving more weights to those individuals with closer propensity scores. In order to estimate the propensity scores for reweighting the sample, model misspecification may be one the concerns but it is difficult to validate this assumption. So, Austin and Stuart (2015) suggest that rather than focusing the model misspecification, we should focus on the balancing of the individuals' covariates between flooded districts and non-flooded districts. Baseline diagnostic checks—whether IPTWs have removed the systematic covariates differences between flooded and non-flooded districts—are essential. Lunt's user written program for checking the covariates' balance between treated and control groups is used. Table 2.7 shows that sample reweighting, on average, reduces the observable covariates gap between the flood

[5]A list of the covariates is given in Table 2.7.

affected and unaffected individuals. For instance, the standardized difference of age between flooded and non-flooded individuals was -0.037 before reweighting which reduced to -0.001 (a 97.30% reduction) after reweighting the sample. Reweighting also did well in balancing of other covariates—including assets possession of iron, sewing machine, table-chair, etc.,—by reducing the covariates' gap between flood affected and unaffected individuals up to more than 100% (for more detail, see Table 2.7 on page 48). This diagnostic assessment indicates that weighing the original sample created a pseudo-sample in which means values of the covariates between flooded and non-flooded individuals remain almost similar. Thus this study uses these IPTWs for obtaining biased-reduced estimates throughout the entire analysis, if not mentioned otherwise. Since the weights are estimated rather than known, so a more robust standard errors estimation method for inferences should be used (Joffe, et al. 2004). In this study, I use a more conservative standard errors correction method in which errors are clustered at the district level. This method is used throughout the entire analysis.

2.5 Results and Discussion

My outcome variable of interest in this study is marriage—an indicator variable which holds a value of 1 for individuals (aged 15-50) who represented their marital status is married and 0 if they represented their marital status is never-married.

The number of marriages at the national level is higher for individuals who resided in flood-affected districts than for those who resided in unaffected districts, both before and after the 2010 flood in Pakistan (see A, Figure 2.1). On average, 63% of individuals aged 15-50 in flood affected districts and 59% of individuals in flood unaffected districts showed their marital status as married before the 2010 flood while 63% of individuals in flooded districts and 61% of individuals in non-flooded districts represented their marital status as married after the flood. Running regressions of marriage on flooded districts (a dummy variable which holds a value of 1 if individuals resided in flooded districts and 0 otherwise), survey-year dummies (survey 2004-05 as base year) and province dummies (Punjab province as reference group) show that, on average, there were 41 more marriages per 1,000 individuals aged 15-50 in flooded districts compared to non-flooded districts (see Table 2.8 on page 49). Furthermore, seemingly there exist time trends and province specific differences in the data, so in order to capture these effects, I include province-year and district fixed effects in all of the regression analysis. In order to test the hypothesis—whether the 2010 flood in Pakistan have brought any significant impact, positive or negative, on marriage—I use the ordinary least square (OLS) method to estimates the coefficients of interest. Table 2.2 presents the main results. The standard errors are clustered at the district level (115 clusters for total and rural areas and 107 clusters for urban areas) and the IPTWs are also included for estimation.

2.5.1 *The Impact of the 2010 Pakistani Flood on Marriage*

The results of the province-year and district fixed effects model of the impact of the 2010 Pakistani flood on marriage of individuals aged 15-50 from the DiD analysis are given in Table 2.2 on page 33. The table includes only coefficients and standard errors of the variables of interest and excludes province-year and district dummies due to space limitation. Panel A and Panel B contain the estimated coefficients of βs without and with IPTWs respectively and columns 1 to 7 represent the coefficient estimates based on different sample and sub-samples of individuals including total, rural vs. urban, male vs. female, etc. Since the weighted estimates are assumed to be biased-reduced, so the interpretation of all of the regression results would be based on IPTWs, if not mentioned otherwise. However, only a very slight variation in magnitudes and significance levels of the estimates in both, weighted and un-weighted samples was observed (see panels A and B).

The estimated coefficients of βs from the baseline regression (1) indicate that the 2010 Pakistani flood affected marriages of individuals aged 15-50 negatively in flood affected districts compared with marriages of individuals in flood unaffected districts during the flood year (2010-2011), and after that the effect of the flood on marriage disappeared, at least for those years for which data are available, 2012-13 and 2014-15. In other words, 17 fewer individuals aged 15-50 per 1,000 married and unmarried individuals at the same age group have gotten married in flood affected districts compared with individuals in flood unaffected districts during the flood year 2010-11. The impact of the flood on marriage in flooded districts vanished immediately after the second (2012-13) and fourth (2014-15) years following the flood (2010-11). Since the PSLM surveys have been conducted for each alternative year (e.g. 2004-05, 2006-07, etc.) rather than for each consecutive year (e.g. 2004-05, 2005-06, etc.) and because data for the year 2011-12, one year after the flood, is not available to analyze the impact of the flood on individual marital life, so based on this data set, it is difficult to infer that whether the negative impact of the 2010 flood on marriage still persisted in year 2011-12 or it just disappeared immediately after the flood.

In order to investigate the geographical heterogeneous effect of the 2010 flood on marriage rate, I run two separate regressions; one for urban regions (2) and one for rural regions (3). Interestingly, my findings show that the 2010 flood in Pakistan did not bring any immediate or long-lasting significant impact on marriage preferences of urban affected individuals compared with their counterparts who dwelled in the urban regions of unaffected districts. On the other hand, the flood reduced the number of marriages in rural regions in affected districts by a number of 22 per 1,000 married and unmarried individuals aged 15-50. The findings coming from the geographically distributed samples (rural vs. urban) reveal that the impact of the flood on marriage using full sample is mainly driven by the rural regions.

Columns (4) and (5) show the results of the 2010 flood impact on gender wise (male vs. female) marriage rates. Negative impact of the flood on female marriages in the flood affected districts is slightly lower than the male marriages compared to their respective counterparts in the flood unaffected districts. Further analysis shows that the difference between male and female estimated coefficients is due to the urban areas because the 2010 flood affected male and female marriage rates roughly equally in the rural areas (see Columns 6 and 7).

Columns (8 and 9) show that the impact of the flood on marriage in flooded districts compared with marriages in adjacent non-flooded districts is not significant while the effect of the flood on marriage in flooded districts (far away districts from the flooded areas) compared with marriages in non-adjacent non-flooded districts actually contributed the overall impact in the total sample. Adding the information from Figure 2.1, B with these regression findings reveal that the marriage rate in Pakistan was affected by two ways due to the floods—the flood increased of marriage rate in rural unaffected areas while it decreased marriage rate in the affected districts. Thus the overall negative impact of the flood on marriage rate is the DiD estimates by comparing individuals in flooded districts with individuals in non-flooded districts.

Further analysis of the flood impact on marriage rate based on different age groups (15-24, 25-34, and 35-50) shows that the flood affected the first two age groups negatively (see Table 2.9). Out of 1,000 married and unmarried individuals in flooded districts compared with the non-flooded districts, the disaster reduced, on average, 25 and 17 individual marriages in the first and second age groups respectively. The reduction of marriages due to the flood happened only in the rural flooded areas with a reduction of 37, 22, and 6 marriages per 1,000 individuals in the first, second and third age groups respectively. A comparison between men and women marriages, the disaster reduced male marriage by 17 and 25 marriages in the first and second age groups respectively while it reduced female marriage by 32 marriages in the first age group in flooded districts. The first age group which also includes child marriage

Table 2.2: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Flood on Marriage (Main Analysis)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome Variable: Marriage		Location		Gender		Rural - Gender		Rural Comparison - Group	
Panel A: Without Weight	Overall Sample	Urban	Rural	Male	Female	Rural Male	Rural Female	Adjacent District	Non-adjacent Districts
Flooded District x Year 2010-11	-0.018**	-0.005	-0.025**	-0.016**	-0.020**	-0.023**	-0.026**	-0.022	-0.026**
	(0.008)	(0.006)	(0.011)	(0007)	(0.009)	(0.011)	(0.013)	(0.016)	(0.012)
Flooded District x Year 2012-13	-0.003	-0.007	0.006	-0.000	-0.006	0.009	0.003	0.005	0.009
	(0.007)	(0.006)	(0.010)	(0.008)	(0.007)	(0.011)	(0.010)	(0.015)	(0.010)
Flooded District x Year 2014-15	0.001	-0.002	-0.001	0.002	-0.001	0.003	-0.009*	-0.003	0.003
	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)	(0.010)	(0.009)	(0.012)	(0.010)
Observations	1,425,607	488,563	937,044	718,215	707,392	467,916	469,128	659,247	789,493
Panel B: With IPTWs									
Flooded District x Year 2010-11	-0.017**	-0.004	-0.022**	-0.019***	-0.016*	-0.023**	-0.024*	-0.019	-0.024**
	(0.007)	(0.005)	(0.010)	(0.007)	(0.009)	(0.011)	(0.013)	(0.010)	(0.011)
Flooded District x Year 2012-13	0.003	-0.003	0.010	0.001	0.003	0.011	0.004	0.008	0.013
	(0.007)	(0.007)	(0.010)	(0.008)	(0.007)	(0.011)	(0.010)	(0.016)	(0.010)
Flooded District x Year 2014-15	0.007	0.006	0.001	0.005	0.007	0.003	-0.005	-0.002	0.006
	(0.007)	(0.009)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)	(0.012)	(0.009)
Observations	1,425,607	488,563	937,044	718,215	707,392	467,916	469,128	659,247	789,493
Province-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	115	107	115	115	115	115	115	81	96

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in Section 2.4 are used. Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

may be interesting to policy makers: the results for this age group require further exploration, which will be done in the next section.

2.5.2 *Discussion*

This study utilizes a DiD approach to investigate the effect of the 2010 Pakistani flood on individual marriages in flood affected districts compared with individual marriages in flood unaffected districts. Based on the theoretical background, on the one hand, I predicted that the number of marriages in flood-stricken districts of Pakistan after the 2010 flood would increase due to the attachment (Bowlby 1969) and terror management (Solomon, et al. 1991) theories, but on the other hand, the prediction would be opposite due to the research on economic circumstances (Cohan and Cole 2002).

My findings support both perspectives. Due to the 2010 flood in Pakistan, the number of marriages in flood-stricken districts compared with unaffected districts before and after the flood and affected districts before the flood decreased immediately during the flood year (2010-11). This may be mainly because of economic circumstances. These findings accord with the findings of a number of studies that include different types of disasters such as man-made or natural disasters, etc. For instances, Hamamatsu, et al. (2014), observed a decrease in the number of marriages after the East Japan Earthquake in Japan, and Prati and Pietrantoni (2014), with reference to the 1997 Umbria-Marche Earthquake in Italy also found a reduction in marriage rates in 15 disaster-stricken municipalities. In contrast to the above studies' findings, the Figure 2.1, B shows that marriage rate in non-flooded districts increased during the flood-year (this may be due to due to the above mentioned theories predict) which are in the line of the studies by Cohan and Cole (2002) and Xu and Feng (2016), who found a positive impact of natural disasters on marriage rate. These regression findings along with Figure 2.1, B reveal that the overall impact of the flood on marriage was negative however, the effect would be considered as the residual impact of the flood on marriage rate in flooded districts (negative) and non-flooded districts (positive). Given the limited information in dataset, it is difficult to disentangle the effect of both type of factors, psychological as well economic, separately.

Why did the 2010 Pakistani flood not increase marriage rate in the flooded districts— the terror management theory's main assumption would predict? This theory proposes natural disasters or other life threating events trigger the individual's feelings of anxiety or fear about his/her own death. In order to camouflage this fear/anxiety, individuals usually run towards their closed or loved ones. The sign of the relationship between disasters and the life course transition (marriage) tends to depend on the nature of disasters or the nature of consequences disasters generate. For instance, manmade disasters or disasters which cause more causalities are more likely tend to trigger death anxiety that could bring positive impact of disasters on marriage while on the other hand, disasters which cause more economic losses than human losses, the effect of the disasters would be negative.

Unlike above mentioned two studies which show a positive relationship between disasters and marriage, the 2010 Pakistani flood caused more economic losses than human casualties in flooded areas. The flood severely damaged infrastructure, businesses, houses, agriculture and livestock (Deen 2015) and the case of 2010 Pakistani flood was the biggest disaster in the history of Pakistan in terms of economic losses and number of affected people which was ranked the first, among the top 10 natural disasters since 1900 (Asgary, Anjum, and Azimi 2012). As economic losses of the flood were severe in rural Pakistan, so did the impact on the life course transition (marriage). Almost all of the cases, the effect of the flood on marriage rate of the rural population compared to their counterparts were high and significant at the conventional levels. For example, 22 fewer individuals per 1,000 individuals (married and unmarried aged 15-50 years) married in rural regions of the flooded districts compared with rural regions of the non-flooded districts, and that relationship is statistically significant at the 5% level.

Natural disasters may also affect life course transitions differently than other type of disasters such as terrorist attacks, bombings, etc., because in case of natural disasters such as flood or earthquake, they not only generate psychological and traumatic issues in individual life, but have also had a considerable economic and social impact on the society as a whole. There is evidence that economic stability leads to marital life stability in the context of Pakistan. For instance, Khaleek and Hussain (2015) found a strong connection between financial instability and marriage instability in Lahore, the capital city of Punjab province (Pakistan), based on a qualitative research—interviewing from lawyers, married and divorced people.

Beside short term and immediate impact of natural disaster on life course transition, this paper also investigated the long-lasting effect of the natural disaster on marriage rate. Insignificant coefficients on the 'Flooded District x Year 2012-13' and 'Flooded District x Year 2014-15' in Table 2.2 show that there exists no long term effect of the flood on individual marriages in flooded districts. These findings accord with Rubonis and Bickman (1991), who reviewed 52 empirical papers intensively and concluded that, despite methodological differences in estimation, there exists evidence that disasters affect victims' psychopathology and the disaster related psychopathological issues are timely dependent, with the passage of time these issues would be reduced or disappeared. Usually disaster related mental health problems disappear after the third year of disasters, but in order to follow up the impact of disaster on mental health issues after several years following the disaster, no such studies have done so far (Cohan and Cole 2002).

Despite the facts that this study finds interesting and significant results, we should be careful about interpreting these results especially in the context of Pakistan—a country where family is the most influential institution in settling of marriage decision. For instance, Pakistani society endows men the patriarchal and prominent power and in case of any dispute in marriage, wife suffers her husband abused acts, both physically and psychologically. Usually a reduction in marriage rate is considered a bad sign, but based on moral ground, in a country like Pakistan due to lack of family support, education and earning power, and emotional feelings about raising children and where divorce is a shame (Bhutta, et al. 2015), a reduction of marriages and particularly women marriages, if caused by disasters, not necessarily a bad sign. Furthermore, if disasters reduce child marriage, which I will explain in the next section, it is even a good sign because there are ample evidence in literature on the negative impact of early marriage—especially girl child marriage—on education (Field and Ambrus 2008; Malhotra, et al. 2011; Vogelstein 2013; Wodon, Nguyen, and Tsimpo 2016), health (Nour 2009; Raj 2010), labor force participation (Parsons, et al. 2015), earning potential and productive use of their earnings (Duflo 2012; Parsons, et al. 2015).

2.6 Robustness Checks and Further Extensions

2.6.1 *Placebo Experiment*

One of the potential problems in difference in difference estimation technique is that, in presence of time trends in data, irrespective of the flood, the estimates obtained by the natural experiment setting may be driven by other confounding factors. In order to check the sensitivity of my findings based on this assumption, I did a placebo experiment, in which I assumed that the devastating flood was not actually happened in year 2010-11 but rather in year 2006-07 (first pseudo flood) or in year 2008-09 (second pseudo flood). The availability of data from three surveys (i.e. 2004-05, 2006-07, 2008-09) before the actual flood (2010-11) allows me to estimate the effect of the pseudo floods (2006-07 or 2008-09) on number of marriages in flooded districts. Results from this analysis (see Table 2.10 on page 51), with one exemption of the rural female sub-sample in first pseudo flood, show that all the coefficients of interest in the placebo flood experiments are insignificant (see panels A and B in Table 2.10). Evidence from this experiment further suggests that the negative impact of the flood on individual marriages in the main analysis is due to the flood.

2.6.2 *Alternative Approach for Flooded Districts Selection*

In this subsection, I encounter the assumption related to the selection of flooded districts (treated group). In the main analysis, it was assumed that the 2010 flood affected—if any districts—equally and the severity level of damages caused by the flood do not contribute in anyhow determining the flood impact on marriage rate, which may not be a realistic assumption in practice. Previously the selection of flooded district was based on the criteria that, whether a district was affected by the flood by utilizing the maps, produced by the MapAction and the CTP, may not be helpful in identifying the effect of the flood on marriage rate. In order to test the sensitiveness of my findings based on this assumption, I use a different alternative approach for selecting the treated group. In this regard, information provided by the National Disaster Management Authority (NDMA), is utilized for generating a new treated group (flood affected districts). The NDMA, immediately after the flood, declared 29 out of 78 affected districts as the most severely affected districts for providing immediate relief and response for recovery. I use 28 out of 29 these affected districts, for which data is available, as a treated group and previously the unaffected districts as comparison group. Last columns of Table 2.1 on page 25 show that, on average, 63.34% of the population aged 15-50 among married and unmarried individuals in the severely flooded districts showed their marital status as married in the surveys before the flood, which reduced slightly to become 63.27%. For rural and female subsamples, it reduced to 0.46 and 0.45 percentage points respectively and for rural-female sample, it reduced to 1.03% (see last column, Table 2.1 on page 25). However in non-flooded districts, it increased from 59.07% to 60.39% on average and except rural-female subsample, it increased in all of the subsamples by approximately 1% (see column 1, Table 2.1). Coming to the regression results for this exercise, Table 2.11 on page 52 is a replication of the main results in Panel B, Table 2.2 on page 33. Estimated magnitudes and the directions of all of the coefficients of interest, with one exemption[6], remained fairly similar to those of the main analysis, ranging between -0.016 and -0.022. In addition, I did

[6]The 2010 flood affected marriages in the flooded urban areas against the main analysis.

another exercise by taking the severely affected districts as treated group and unaffected districts as comparison group from the given maps and excluded moderately affected districts this time. My findings still remained similar[7]. These findings further strengthen the hypothesis that the 2010 flood in Pakistan brought an immediate negative impact on the life course transition (marriage). The level of severity does not matter for identifying the flood impact on marriage, however the fact—whether or not a district affected by the flood—matters.

2.6.3 *Restricted Analysis Based on Common support and Trimming data*

Common support is one the main objective by utilizing propensity score approaches. Common support actually limits the sample within a restriction which "implies that the test of the balancing property is performed only on the observations whose propensity score belongs to the intersection of the supports of the propensity score of treated and controls" (Becker and Ichino 2002). Including individuals outside the common support can produce biased estimates and "when doing propensity score analysis we should restrict our analysis in between the range of propensity scores at which we observe both treated and untreated individuals" (Lunt 2014). Table 2.13 shows common support restriction limits in which the lower bound is restricted to the higher value of minimum propensity scores while the upper bound is restricted to the lower value of maximum propensity scores among the flooded and non-flooded districts. Based on these restrictions, the main weighted estimates from Panel B in Table 2.2 are replicated in Panel A, Table 2.14 on page 54. Only a small number of observations[8] under the common support restriction limits are dropped out from the analysis. That is why the reduction of the individuals—having extreme propensity score values—does not change the magnitudes and directions of all of the variables of interest at all.

In addition, I also use a more conservative approach, trimming, suggested by Lunt (2014), to include all those individuals in the flood affected and unaffected districts under a certain propensity scores limits. The limits can be obtained by calculating "the x^{th} centile of the propensity score in the treated and the $100 - x^{th}$ centile in the untreated, and remove all subjects outside these limits" (Lunt 2014). Based on basic command in stata, 'proprtrim treatment_variable propensity_score', I created dummy variables, 'keep_0', 'keep_1', and 'keep_5', that holds a value of 1 for individuals to be included in the analysis and 0

[7]Results will be shared upon request.
[8]Less than 50 individuals in all other samples and 1,731 individuals in rural sample (698 male and 1,071 female from rural sample) are dropped out from the analysis.

otherwise for their each respective centile categories. Since trimming at the 0^{th} centile is similar to the common support, so I limit this analysis to the first centile (Panel B, Table 2.14) and the fifth centile (Panel C, Table 2.14). Despite of excluding about 6% - 10% and 24% - 27% of observations in the first and fifth centiles analysis respectively, an efficiency gain is observed only in two cases of the first centile analysis; in female and rural-female samples. However, the magnitudes of the 1^{st} centile cases increased a little bit, up to a 2 to 5 percentage points. In the fifth centile trimming scheme, neither magnitudes of the variables of interest varied nor an efficiency gain was observed. These findings further strengthen the evidence that there exists a significant immediate and negative impact of the 2010 flood on individuals' marital lives in flood affected districts in Pakistan.

2.6.4 *The Disaster Effect on Child Marriages (Age 10-17 Years)*

Natural disasters may compel the affected parents to marry their children, especially girl children, at a very younger age in order to reduce the economic burden of raising them further. Child marriage is defined as "any legal or customary union involving a boy or girl below the age of 18" (Parsons, et al. 2015). But the age limits for child marriage differ from country to country. For instance, the minimum marriage age limits for both genders in South Asia are given as follows: for male it is 18 years in Afghanistan, Bhutan, Maldives, Pakistan and Sri Lanka, and 21 years in Bangladesh, India and Nepal while for female, it is 18 years in Bangladesh, Bhutan, India, Maldives, Nepal and Sri Lanka, and 16 years in Afghanistan and Pakistan (Khanna, Verma, and Weiss 2013). Child marriage is not only a form of human rights violation, but it also becomes one of the barriers for social and economic development (Parsons, et al. 2015).

It is estimated that 21% of children before the age of 18 married in Pakistan which is lower than in Afghanistan (33%), Bangladesh (52%), Bhutan (26%), India (47%), and Nepal (37%) in 2016 (see Figure 2.2) however, child marriage is one of the major concerns in Pakistan because child marriage is more vulnerable to the spouse controlling behavior and violence than adult marriage (Nasrullah, Zakar, and Zakar 2014). Furthermore, in reference to the 2010 flood in Pakistan, Bhatti (July 31, 2011) also makes an argument that the floods increased child marriage in flood stricken areas after the flood. Due to limited family resources and lack of financial supports from governments or NGOs after the disaster, parents actually sold their daughters in the name of marriage to get some financial assistance from the child husband (Bhatti July 31, 2011). Furthermore he says:

> Child marriage in India, Afghanistan and Pakistan is a centuries-old traditional way of arranging marriages. But if parents marry their children in childhood, they would be sent back to live with their parents until the end of immaturity. But after the flash floods, people in Pakistan started giving away their minor girls to their husbands on the day of marriage.

Thus based on the above stylized facts, I hypothesized that the children in flooded districts would experience more marriages compared to those in non-flooded districts due to economic circumstances. The sample of married and unmarried individuals with narrow age window 10-17 allows me to test the given hypothesis. This sample represents only those individuals who married recently under the age window 10-17 and excluded all those (18-50) who may be married at the age of 18 or below but were older at the time the PSLM surveys were conducted. Table 2.12 on page 53 depicts descriptive statistics of child marriage. Only 2.6% of children aged 10-17 and 0.63% of children aged 10-15 showed their marital status as married in the surveys. The Pakistan's Child Marriage Restraint Act 1929 (CMRA) defines the minimum marriage age limits for male (18 years) and for female (16 years). Based on the Act, 0.88% of male and 1.61% of female children in flooded districts while 0.43% of male and 0.67% of female children in non-flooded districts married below their legal age limits. In order to test the hypothesis whether the 2010 flood in Pakistan may have any impact on child marriage, I run several regressions. Results are shown in Table 2.15. Regression results based on age window 10-17 shows a slight decrease in child marriage (column 1) but running regressions based on one year below the legal marriage age limits for both gender, male (column 3), and female (column 4), the effect of the flood on child marriage vanished. Based on the legal marriage age restriction, there is no strong evidence that child marriages in flooded districts increased after the 2010 flood. In addition, the Sindh provincial assembly approved the Sindh Child Marriage Restraint Act in April 2014 that increased the age limit for girl marriage in Sindh province. Running regressions without the last PSLM survey data (2014-15) show no effect of the flood on child marriage[9]. In a nutshell, this study finds no evidence of the flood impact on child marriage in flooded districts of Pakistan.

2.6.5 *The Disaster Effect on Marriages: Consideration of Migration*

Migration and recovery process could play a huge role in household resilience to the disasters. In—or—out migration/internally displaced people (IDP) could distort my findings. One of the limitations of this study is that there is no information for

[9]Results will be shared upon request.

migration/IDP in the surveys. However, there is evidence that Pakistani households recovered from the 2010 flood very rapidly. For instance, as on 24 January 2011, 95% of the flood affectees have returned to their areas of residence, reported in a report on an inter-agency assessment of the humanitarian response to the 2010 flood in Pakistan (Polastro, et al. 2011, p. 92). Similarly, Weiss, et al. (2014) compared recovery process of the affectees from the 2010 Haiti earthquake and the 2010 Pakistani flood and found that the Pakistani households recovered much faster in term of dwelling, i.e. 61% of Pakistani households had returned to a permanent residential structure such as a house or apartment within six months compared to only half of the Haitian households, who recovered after one year following the disasters.

Random Migration does not influence my findings, but it matters when there happened a systematically different types of migration/displacement between married and unmarried individuals aged 15-50. Due to data limitation on migration, permanent residential status information is used to capture the migration/IDP impact on marriage, though it is a week proxy for migration/IDP. The notion behind using this proxy is that individuals aged 15-50 who were dwelling in their own permanent residences at the time when the surveys were conducted were less likely to be migrated/displaced from other places than the people who were not living in their personal residences. Intuitively, at the time when surveys were conducted, if the flood affectees migrated/displaced internally from affected area to non-affected ones, then they would be living in temporary places on rent rather than their own personal residences.

Residential status is a dummy variable, which holds a value of one if an individual residential status at present is personal and zero if an individual residential status at present is on rent, on subsidized rent, and without rent. Table 2.16 on page 56 depicts the descriptive statistics of residential status. There is very little variations in the average percentage of individuals aged 15-50, who were residing in their personal residences between before and after the 2010 flood in both flooded and non-flooded districts. For instance, the percentage of individuals having their own personal residences increased by 0.37% (from 89.54% to 89.87%) in flood affected districts and by 0.08% (from 86.94% to 87.01%) in flood unaffected districts.

This dummy variable is interacted with my main variable of interest. Regression results are given in Table 2.17 on 57. The coefficient of the residential status indicates that, at the time of the survey interviews, an additional of 38 individuals per 1,000 individuals aged 15-50, who were living in their personal residences, showed their marital status as married than those individuals who were living in other types of residences. Except in the overall sample (column 1) in which the coefficients of the interaction terms are positive but not highly statistically significant, in all other cases, the coefficients of interaction terms are insignificant indicating that there is no evidence for temporarily migration impact on marriage rate due to the flood 2010 in Pakistan. Coming to the variables of interest, the magnitudes of the coefficients increased, in terms of direction, my findings still remain robust however.

In addition to the above analysis, I also did another exercise in which I assume that the people who were affected by the flood, were more likely to migrate to the nearer adjacent non-flooded districts (19 districts) than to far-flung, non-adjacent non-flooded, districts (34 districts). Furthermore, in this analysis I compare marriage rate in the most severely affected districts to the marriage rates in non-adjacent non-flooded and far-flung districts. The DiD Estimates in Columns 8 and 9 of Table 2.11 show that the impact of the flood on marriages in rural severely affected areas compared to marriages in rural unaffected non-adjacent areas in both type of districts is negative but is significant only for the rural areas in non-flooded far-flung districts. Which further strengthen the findings that the flood 2010 affected individual marriages, on average, negatively in Pakistan.

2.7 Conclusion

The present study designs to examine the aftermath impact of a natural disaster (flood) on one of the life course transitions, marriage, in developing countries with a special reference to the July 2010 Pakistani flood. Based on the terror management and attachment theories, I predicted that marriages in flooded districts after the flood would increase while research on economic circumstances provides a base for negative association between natural disaster and marriage.

My findings, based on research on economic circumstances and above theories, reveal that the 2010 flood in Pakistan decreased individual marriages in 62 flooded districts compared with 53 non-flooded districts during the flood year 2010-11. However, the impact of the disaster on marriage disappeared immediately after the flood year. The flood impact on marriage rate in rural parts of the affected areas of Pakistan were much worse than in urban parts compared with their counterpart areas in unaffected districts. This reduction in the number of marriages in affected districts seems to be mostly related to financial and economic constraints rather than psychological factors because the disaster impact vanished immediately after the disaster year 2010-11. One of the explanations for the disappearance of any longer term impact of the 2010 flood on marriage is that it took only a short period to recover from the flood. There exists some evidence that most of the Pakistani flood-affected households returned to their home/recovered within one year following the flood (Pastor and Morello-Frosch 2014; Weiss, et al. 2014). But the flood impact on individual marital life psychologically also cannot be ruled out as the non-flooded districts depict a positive response of marriage.

My findings have the following implication. Flood disasters seemingly destabilize adults' marital life through flood related economic and financial losses in flooded districts, so in order to compensate or rehabilitate the disaster-affected people fully, policy makers or humanitarian aid-providers should also consider life course transitions along with other physical or mental health issues, at least for the year directly following the disasters.

This study has some limitations which should be mentioned. In order to identify the exact mechanism through which the natural disaster may alter individual behavior towards life course events in a post-disaster circumstance, I realized that much more factors beyond economic circumstances—such as post-traumatic stress disorder, anxiety, depression, or any other physical, psychological, or mental health problems that affect marriage decision—needed to be explored. Second, the outcome variable used in this study is derived from the individual's current status of marital life and the corresponding age window (15-50) and does not represent the individual's age at marriage. Thus, a dataset with individual's age at marriage could provide a better or more reliable answer for the present research question.

2.8 Appendix to Chapter 2

2.8.1 *PSLM Surveys: 2004-2005 to 2014-2015*

Individual or household level data are taken from the nationally representative surveys at district level, namely the Pakistan Social and Living Standards Measurement (PSLM). These surveys have been conducted by the Pakistan Bureau of Statistics (PBS) since 2004 which provide data on individuals in households at the district level for each alternate year. The project was initiated in 2004 and has finished in 2015. In this study, I use all six waves of the surveys, i.e. 2004-05, 2006-07, 2008-09, 2010-11, 2012-13, and 2014-15. The PBS used its own sampling frame for these surveys which consists of all urban and rural areas of Pakistan except some military restricted areas. The sampling frame for urban areas was defined and updated by the PBS in 2003 and 2013 while for rural areas; the 1998 population census was used for constructing the sampling frame. Urban areas consist of cities and towns which were further divided into smaller units, called enumeration blocks (E.Bs). Each enumeration block consists of 200-250 households. In the updated urban-area sampling frame, there are 26,698 enumeration blocks while rural area sampling frame consists of 50,590 villages and mouzas/deh[10]. The PBS created three types of stratum within a district; big cities[11], urban, and rural stratum. The big cities stratum were further (sub) stratified based on low, middle and high income groups. The final sample for each surveys were obtained by a two stage stratified sampling design. In the first stage, the survey sample of the primary sampling units (PSUs) were selected from the population PSUs—enumeration blocks/villages and mouzas—by using a probability proportional to size sampling method. In the second stage, the sample secondary sampling units (SSUs)—which are in fact households—were drawn from the first stage selected PSUs, using a systematic sampling technique with a random start, including 16 households from rural PSUs and 12 households from urban PSUs. Although, the sampling method was same for all surveys, the samples did not necessarily include the same households over time. So obtaining a balanced panel dataset is not possible. Table 2.3 and Table 2.4 describe the sampling distribution further in details in terms coverage, PSUs and SSUs.

[10] In Pakistan, the farming communities inhabit in compact and or in scattered units called 'Mouza or Deh'. The areas of Mouzas/Dehs are demarcated, measured, and recorded properly in the documents of the Revenue Department. Source: Pakistan Mouza Census 2008, PBS, Islamabad.

[11] Islamabad, Lahore, Gujranwala, Faisalabad, Rawalpindi, Multan, Bahawalpur, Sargodha, Sialkot, Karachi, Hyderabad, Sukkur, Peshawar and Quetta

2.8.2 *Severely Flood Affected Districts*

The National Disaster Management Authority (NDMA) declared 29 out of 78 affected districts as severely affected districts for the purpose of emergency aids and rehabilitation. In this study, I use 28 districts, for which data are available, as treated group and the list of these affected districts are given in Table 2.5.

2.8.3 *Tables*

Table 2.3: Province Wise Sample Distribution

Region	No. of District /Cities	Type of Stratum	No. of Stratum
Big Cities[(a)]	14 Cities	Low, Medium, High	3
Punjab	34 Districts	Rural, Urban	2
Sindh	16 Districts	Rural, Urban	2
KPK	24 Districts	Rural, Urban	2
Balochistan	26 Districts	Rural, Urban	2

Notes: (a) Islamabad, Lahore, Gujranwala, Faisalabad, Rawalpindi, Multan, Bahawalpur, Sargodha, Sialkot, Karachi, Hyderabad, Sukkur, Peshawar and Quetta.
Source: PLSM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.4: Primary and Secondary Sampling Units: 2004-05 to 2014-15

PSLM Survey	2004-2005		2006-2007		2008-2009		2010-2011		2012-2013		2014-2015	
Region	SA (PSUs)	HH (SSUs)	SA (PSUs)	HH (SSUs)	SA (PSUs)	HH (SSUs)	SA (PSUs)	HH (SSUs)	SA (PSUs)	HH (SSUs)	SA (PSUs)	HH (SSUs)
Punjab[(a)]	2,313	32,544	2,313	32,242	2,313	32,816	2,344	32,972	2,344	31,916	2,500	36,571
Sindh	1,326	18,648	1,320	18,532	1,376	19,300	1,407	19,728	1,407	19,480	1,276	18,735
KPK	849	12,552	849	12,525	831	12,264	849	12,552	866	12,473	868	13,082
Balochistan	716	10,676	716	10,654	778	11,668	813	12,236	794	11,647	682	10,247
Overall	5,204	74,420	5,198	73,953	5,298	76,048	5,413	77,488	5,411	75,516	5,326	78,635

Notes: SA = Sample Area, HH = Household, PSUs = Primary Sampling Units and SSUs = Secondary Sampling Units. (a) Islamabad is included in Punjab Province.
Data Source: PLSM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.5: List of Severely Flood Affected Districts Declared by the NDMA

Province	Severely Affected Districts
Punjab	Muzzafargarh, Rajanpur, Mianwali, R. Y. Khan, Layyah, D.G. Khan, Bhakkar
Sindh	Kashmore, Shikarpur, Jaccobabad, Larkana, Shahdad Kot, Thatta, Dadu, Jamshoro, Ghotki
KPK	Tank, D.I. Khan, Kohistan, Peshawar, Charsada, Nowshera, Dir Lower, Dir Upper, Shangla, Swat
Balochistan	Naseerabad, Jaffarabad

Notes: Source: National Disaster Management Authority (Government of Pakistan 2011a).

Table 2.6: Percentage Distribution of Individuals' (aged 15-49) Current Marital Status: 2004-05 to 2014-15

Survey Year	Unmarried	Married	Divorced	Widow/Widower	Nikha solemnized but Rukhsati not taken place[(a)]
			Men		
2004-05	47.95	51.01	0.25	0.53	0.26
2006-07	47.45	51.65	0.19	0.48	0.24
2008-09	48.29	51.00	0.16	0.42	0.14
2010-11	48.28	50.90	0.17	0.48	0.17
2012-13	47.69	51.52	0.22	0.41	0.17
2014-15	43.91	55.09	0.25	0.50	0.26
2012-13[(b)]	47.60	51.30	0.20	0.50	-
			Women		
2004-05	31.20	66.51	0.40	1.69	0.20
2006-07	31.33	66.61	0.32	1.51	0.22
2008-09	32.49	65.67	0.31	1.40	0.13
2010-11	32.86	65.02	0.30	1.65	0.17
2012-13	32.15	65.70	0.36	1.65	0.15
2014-15	28.59	69.22	0.37	1.57	0.25
2006-07[(c)]	34.80	62.20	0.30	2.10	-
2012-13[(b)]	33.30	63.70	0.60	1.90	-

Notes: (a) Nikah (نِکاح) in the Islamic legal system implies a marriage contract or official wedding ceremony, when solemnized, meaning that the marriage contract has been done, but Rukhsati (رُخصتی) is the departure (farewell) ceremony of the marriage, in which the groom's family will leave together with the couples to their own home. Actually, from Nikah to Rukhsati, it often takes several days.
Data Sources: PSLM Surveys (various issues) and author's own calculation. (b) National Institute of Population Studies - NIPS/Pakistan and ICF International (2013). (c) National Institute of Population Studies - NIPS/Pakistan and Macro International (2008).

Table 2.7: Balancing of Covariates between Flood Affected and Unaffected Districts

	Un-weighted Means			Weighted Means			% RSD
	FD	NFD	SD	FD	NFD	SD	
Age	29.89	30.52	-0.037	29.04	29.06	-0.001	97.30
Female	0.49	0.48	0.010	0.5	0.5	0.003	70.00
Reading literacy	0.49	0.59	-0.199	0.54	0.54	-0.003	98.50
Math literacy	0.72	0.8	-0.178	0.77	0.78	-0.016	91.01
Illness	0.06	0.06	0.022	0.04	0.04	0.001	95.45
Number individuals in a household	8.31	7.44	0.244	7.8	7.85	-0.013	105.33
Rural	0.73	0.6	0.282	0.66	0.65	0.016	94.33
Household possession of:							
Iron	0.69	0.77	-0.185	0.73	0.73	0.003	101.62
Sewing machine	0.56	0.68	-0.252	0.62	0.62	0.001	100.40
Video or Cassette Player	0.31	0.31	-0.010	0.31	0.31	0.008	180.00
Table, Chair	0.63	0.65	-0.036	0.63	0.62	0.016	144.44
Watch, Clock	0.79	0.85	-0.144	0.82	0.81	0.015	110.42
TV	0.46	0.61	-0.296	0.54	0.55	-0.012	95.95
VCR, VCP, VCD	0.07	0.11	-0.113	0.09	0.09	-0.004	96.46
Refrigerator	0.33	0.41	-0.166	0.37	0.37	-0.010	93.98
Air collar	0.10	0.09	0.037	0.10	0.10	-0.002	105.41
Air conditioner	0.04	0.06	-0.090	0.05	0.05	-0.007	92.22
Computer/Laptop	0.07	0.09	-0.067	0.08	0.08	-0.005	92.54
Bicycle	0.30	0.37	-0.150	0.34	0.34	-0.002	98.67
Motor/Scoter	0.24	0.33	-0.215	0.29	0.30	-0.015	93.02
Car/Truck	0.05	0.07	-0.083	0.06	0.06	-0.002	97.59
Tractor	0.03	0.04	-0.028	0.04	0.04	0.002	107.14
Agriculture land	0.39	0.27	0.262	0.32	0.31	0.019	92.75
Non-agriculture land	0.03	0.03	-0.006	0.03	0.03	0.006	200.00
Residential building	0.90	0.87	0.086	0.88	0.88	0.002	97.67
Commercial building	0.04	0.05	-0.022	0.04	0.04	0.003	113.64
Livestock	0.44	0.28	0.335	0.35	0.34	0.019	94.33
Sheep/Goat	0.32	0.24	0.179	0.28	0.27	0.008	95.53
Animal	0.10	0.11	-0.021	0.10	0.10	0.005	123.81
Poultry	0.27	0.11	0.439	0.18	0.17	0.030	93.17
Roof with high materials	0.21	0.31	-0.234	0.26	0.26	-0.006	97.44
Wall with high materials	0.56	0.72	-0.329	0.65	0.65	-0.008	97.57
Tap drinking water	0.29	0.35	-0.144	0.31	0.31	0.006	104.17
Flush toilet	0.55	0.64	-0.175	0.60	0.60	0.004	102.29
Cooking by gas and electricity	0.22	0.37	-0.326	0.30	0.31	-0.013	96.01
Lighting by electricity	0.89	0.89	0.014	0.89	0.89	-0.003	121.42
Home Telephone	0.65	0.67	-0.026	0.66	0.66	-0.007	73.077

Notes: FD = Flooded districts, NFD = Non-flooded districts, SD = Standardized difference, and % RDS = % reduction of standardized difference.
Data Source: PSLM Surveys (various issues) and author's own calculation.

Table 2.8: Flooded District, Province and Survey-Year Fixed Effects on Marriage

	(1)	(2)	(3)
Outcome Variable: Marriage		Location	
	Overall	Urban	Rural
Flooded District	0.041***	0.023***	0.0033***
Year Fixed effects (Ref. Year: 2004-05)			
Year 2006-07	0.001	0.001	0.002
Year 2008-09	-0.010***	-0.020	-0.016***
Year 2010-11	-0.011***	-0.007**	-0.013***
Year 2012-13	-0.010	0.013***	-0.018***
Year 2014-15	0.030***	0.033***	0.012***
Province Fixed effects (Ref. Province: Punjab)			
Sindh	0.025	-0.001	0.048***
KPK	-0.015	-0.008	-0.021
Balochistan	0.059***	0.041***	0.050***
Observations	1,425,607	488,563	937,044

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS.

Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.9: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Flood on Marriage (Age Groups Analysis)

	(1)	(2)	(3)	(4)	(5)
Outcome Variable: Marriage		Location		Gender	
Panel A: Age Group 1: 15-24 Years	Overall Sample	Urban	Rural	Male	Female
Flooded District x Year 2010-11	-0.025**	-0.003	-0.037**	-0.017**	-0.032**
	(0.010)	(0.008)	(0.014)	(0007)	(0.013)
Flooded District x Year 2012-13	-0.009	-0.000	-0.012	-0.008	-0.013
	(0.009)	(0.009)	(0.012)	(0.007)	(0.011)
Flooded District x Year 2014-15	0.005	0.004	-0.005	0.001	0.004
	(0.010)	(0.012)	(0.012)	(0.008)	(0.013)
Observations	597,799	209,508	388,291	310,595	287,204
Panel B: Age Group 2: 25-34 Years					
Flooded District x Year 2010-11	-0.017*	-0.003	-0.022**	-0.025**	-0.010
	(0.009)	(0.010)	(0.010)	(0.012)	(0.008)
Flooded District x Year 2012-13	0.001	0.005	0.002	-0.005	0.006
	(0.010)	(0.010)	(0.013)	(0.013)	(0.009)
Flooded District x Year 2014-15	-0.001	0.000	-0.009	-0.008	0.005
	(0.007)	(0.012)	(0.008)	(0.010)	(0.007)
Observations	371,345	125,656	245,689	176,541	194,804
Panel C: Age Group 3: 35-50 Years					
Flooded District x Year 2010-11	-0.002	0.004	-0.006***	-0.004	-0.001
	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Flooded District x Year 2012-13	-0.002	-0.004	-0.000	-0.003	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Flooded District x Year 2014-15	0.000	0.008**	-0.003	0.001	-0.001
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)
Observations	456,463	153,399	303,064	231,079	225,384
Province-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes
Clusters	115	107	115	115	115

Notes: Dependent variable: Marriage - whether an individual (aged 15-24 or 25-34 or 35-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in section 2.4 are used.

Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.10: District and Province-Year Fixed Effects Estimates of the Pseudo Floods (2006-07 or 2008-09) on Marriage (Robustness Check-I)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome Variable: Marriage		Location		Gender		Rural - Gender	
Panel A: Pseudo Flood in 2006-07	Overall Sample	Urban	Rural	Male	Female	Rural Male	Rural Female
Flooded District x Year 2006-07	0.0001	0.008	-0.004	0.000	0.000	-0.001	-0.009
	(0.006)	(0.008)	(0.008)	(0007)	(0.007)	(0.010)	(0.010)
Flooded District x Year 2008-09	-0.004	0.001	-0.002	-0.000	-0.010	0.008	-0.019*
	(0.007)	(0.008)	(0.009)	(0.007)	(0.008)	(0.010)	(0.010)
Observations	713,588	267,197	446,391	362,675	350,913	225,135	221,256
Clusters	110	103	110	110	110	110	110
Panel B: Pseudo Flood in 2008-09							
Flooded District x Year 2008-09	-0.004	-0.004	-0.000	-0.001	-0.010	0.009	-0.014
	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	(0.010)	(0.009)
Observations	713,588	267,197	446,391	362,675	350,913	225,135	221,256
Clusters	110	103	110	110	110	110	110
Province-Year Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in the section 2.4 are used. Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.11: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Flood on Marriage (Robustness Check-II)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome Variable: Marriage		Location		Gender		Rural - Gender		Comparison - Group	
Severely Affected District[(a)]	Overall Sample	Urban	Rural	Male	Female	Rural Male	Rural Female	Adjacent District	Non-adjacent Districts
Flooded District x Year 2010-11	-0.019***	-0.016**	-0.020**	-0.020**	-0.018**	-0.022*	-0.019*	-0.016	-0.022*
	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.013)	(0.010)	(0.012)	(0.011)
Flooded District x Year 2012-13	-0.007	-0.015	0.002	-0.010	-0.006	0.002	-0.004	-0.004	0.006
	(0.009)	(0.010)	(0.010)	(0.011)	(0.009)	(0.015)	(0.009)	(0.018)	(0.010)
Flooded District x Year 2014-15	0.003	-0.011	-0.001	0.003	-0.002	0.005	-0.012	-0.012	0.004
	(0.009)	(0.013)	(0.011)	(0.012)	(0.010)	(0.015)	(0.009)	(0.012)	(0.012)
Observations	936,251	344,397	591,854	474,374	461,877	297,799	294,055	314,057	444,303
Cluster	73	69	73	73	73	73	73	39	54
Province-Year Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in the section 2.4 are used. Data Sources: PSLM surveys (various issues), PBS, Islamabad, Pakistan. (a) NDMA (Government of Pakistan 2011a).

Table 2.12: Descriptive Statistics of Child Marriage (Age 10-17 Years)

Outcome Variable: Child Marriage	(Age 10-17 years)		(Age 10-15 years)	
	Observations	%	Observations	%
Overall Sample				
Total	206,143	2.6%	72,675	0.63%
Male	110,674	0.66%	-	-
Female	-	-	33,315	1.14%
Flooded Districts				
Male	56,990	0.88%	-	-
Female	-	-	16,836	1.61%
Non-flooded Districts				
Male	53,684	0.43%	-	-
Female	-	-	16,479	0.67%

Notes: Outcome variable: Child Marriage - whether an individual (aged 10-15 or 10-17 years) is married (Child Marriage=1) or unmarried (Child Marriage=0).
Data Source: PSLM surveys (various issues) and author's own calculation.

Table 2.13: Common Support Restriction Based on Higher Minimum and Lower Maximum Propensity Score Values

Sample	Districts	Observation	Mean	Min	Max
Total	Not-flooded	724,779	0.430741	0.044982	0.993332
	Flooded	725,985	0.569975	0.059942*	0.991682**
Urban	Not-flooded	297,829	0.362154	0.044982	0.993332
	Flooded	199,533	0.459439	0.066776*	0.972976**
Rural	Not-flooded	426,950	0.478585	0.056909	0.956064**
	Flooded	526,452	0.611870	0.059942*	0.991682
Male	Not-flooded	364,682	0.431049	0.048344	0.992402
	Flooded	360,544	0.564004	0.065077*	0.991682**
Female	Not-flooded	360,097	0.430429	0.044982	0.993332
	Flooded	365,441	0.575866	0.059942*	0.990981**
Rural Male	Not-flooded	213,346	0.478026	0.056909	0.956064**
	Flooded	259,540	0.604897	0.065077*	0.991682
Rural Female	Not-flooded	213,604	0.479144	0.061445*	0.955730**
	Flooded	266,912	0.618650	0.059942	0.990981

Notes: The values with asterisk show common support restriction limits. * and ** show higher minimum and lower maximum propensity score values respectively.
Data Source: PSLM surveys (various issues) and author's own calculation.

Table 2.14: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Flood on Marriage (Common Support and Trimming)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome Variable: Marriage		Location		Gender		Rural - Gender	
Panel A: Common Support	Overall Sample	Urban	Rural	Male	Female	Rural Male	Rural Female
Flooded District x Year 2010-11	-0.018**	-0.004	-0.023**	-0.019**	-0.016*	-0.024**	-0.024*
	(0.007)	(0.005)	(0.010)	(0007)	(0.008)	(0.010)	(0.012)
Flooded District x Year 2012-13	0.002	-0.001	0.011	0.002	0.003	0.012	0.004
	(0.007)	(0.007)	(0.009)	(0.008)	(0.007)	(0.011)	(0.009)
Flooded District x Year 2014-15	0.005	0.004	-0.001	0.002	0.006	0.001	-0.007
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
Observations	1,425,590	488,526	935,313	718,203	707,375	467,218	468,057
Panel B: Trimming (Based on 1st Centile)							
Flooded District x Year 2010-11	-0.020**	-0.005	-0.028**	-0.021***	-0.018**	-0.027**	-0.027**
	(0.007)	(0.006)	(0.010)	(0.007)	(0.009)	(0.010)	(0.012)
Flooded District x Year 2012-13	0.005	0.002	0.009	0.005	0.002	0.011	0.002
	(0.007)	(0.007)	(0.009)	(0.008)	(0.007)	(0.010)	(0.009)
Flooded District x Year 2014-15	0.004	0.005	-0.003	0.001	0.004	-0.001	-0.009
	(0.007)	(0.009)	(0.009)	(0.008)	(0.008)	(0.011)	(0.009)
Observations	1,304,383	459,044	845,339	660,960	643,423	424,683	420,656
Panel C: Trimming (Based on 5th Centile)							
Flooded District x Year 2010-11	-0.017**	-0.004	-0.024**	-0.017***	-0.016*	-0.023**	-0.025*
	(0.007)	(0.006)	(0.011)	(0.007)	(0.009)	(0.011)	(0.013)
Flooded District x Year 2012-13	0.006	0.003	0.010	0.006	0.004	0.010	0.006
	(0.007)	(0.007)	(0.010)	(0.008)	(0.007)	(0.011)	(0.010)
Flooded District x Year 2014-15	0.000	0.006	-0.007	-0.002	0.002	-0.008	-0.009
	(0.007)	(0.009)	(0.009)	(0.008)	(0.008)	(0.011)	(0.009)
Observations	1,064,824	372,298	692,526	541,346	523,478	348,574	343,952
Province-Year Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	115	107	115	115	115	115	115

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in the section 2.4 are used. Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

Table 2.15: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Floods on Child Marriage (Age 10-17 Years)

	(1)	(2)	(3)	(4)
	Overall Sample		Age Limits According to CMRA 1929[a]	
Outcome Variable: Child Marriage	(Age 10-17 years)	(Age 10-15 years)	Male (Age 10-17 years)	Female (Age 10-15 years)
Flooded District x Year 2010-11	-0.008**	0.000	-0.002	0.000
	(0.004)	(0.002)	(0.002)	(0.004)
Flooded District x Year 2012-13	-0.006	0.002	-0.002	0.002
	(0.004)	(0.002)	(0.002)	(0.004)
Flooded District x Year 2014-15	-0.002	0.010	0.000	0.004
	(0.004)	(0.003)	(0.003)	(0.005)
Observations	206,143	72,675	110,674	33,315
Province-Year Fixed effects	Yes	Yes	Yes	Yes
District Fixed effects	Yes	Yes	Yes	Yes
Clusters	115	115	115	115

Notes: Dependent variable: Child Marriage - whether an individual (aged 10-15 or 10-17 years) is married (Child Marriage=1) or unmarried (Child Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in the section 2.4 are used. (a) CMRA 1929 stands for Child Marriage Restraint Act 1929 in Pakistan

Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

.

Table 2.16: Descriptive Statistics of Residential Status

Outcome Variable: Residential Status	Non-flooded Districts		Flooded Districts	
	Obs.	%	Obs.	%
Before Flood: 2004-05 to 2008-09				
Total	366,866	86.94	346,719	89.54
Rural	202,457	92.65	243,931	93.39
Urban	164,409	79.91	102,788	80.40
Male	187,610	86.54	175,064	89.28
Urban Male	84,833	79.36	52,707	79.85
Rural Male	102,777	92.46	122,357	93.34
Female	179,256	87.36	171,655	89.80
Urban Female	79,576	80.49	50,081	80.97
Rural Female	99,680	92.85	121,574	93.44
During and After Flood : 2010-11 to 2014-15				
Total	345,311	87.01	366,708	89.87
Rural	217,021	91.98	273,632	92.94
Urban	128,290	78.61	93,076	80.83
Male	173,505	86.78	182,035	89.62
Urban Male	65,289	78.34	47,470	80.39
Rural Male	108,216	91.88	134,565	92.88
Female	171,806	87.24	184,673	90.11
Urban Female	63,001	78.88	45,606	81.29
Rural Female	108,805	92.08	139,067	93.00

Notes: Outcome Variable: Residential Status - whether an individual (aged 15-50 years) residential status at present is personal (Residential Status =1) or on rent, on subsidized rent, and without rent (Residential Status =0).
Data Sources: PSLM surveys (various issues) and author's own calculation. Districts according to the Maps; the MapAction and the Critical Threats.

Table 2.17: District and Province-Year Fixed Effects Estimates of the 2010 Pakistani Flood on Marriage (Interaction with Residential Status)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Location		Gender		Rural - Gender	
Outcome Variable: Marriage	Overall Sample	Urban	Rural	Overall Male	Overall Female	Rural Male	Rural Female
Flooded District x Year 2010-11	-0.031***	-0.012	-0.036**	-0.030***	-0.032***	-0.032*	-0.041**
	(0.009)	(0.008)	(0.015)	(0.010)	(0.012)	(0.017)	(0.017)
Flooded District x Year 2012-13	-0.007	-0.010	0.001	-0.008	-0.006	0.013	-0.014
	(0.010)	(0.010)	(0.015)	(0.012)	(0.011)	(0.018)	(0.016)
Flooded District x Year 2014-15	-0.007	-0.011	-0.010	-0.010	-0.006	-0.005	-0.019
	(0.010)	(0.014)	(0.012)	(0.012)	(0.011)	(0.015)	(0.012)
Residential Status (=1 if own ownership)	-0.038***	-0.048***	-0.050***	-0.044***	-0.035***	-0.056***	-0.047***
	(0.004)	(0.0005)	(0.005)	(0.008)	(0.005)	(0.006)	(0.005)
Residential Status Interaction with:							
Flooded District x Year 2010-11	0.015*	0.011	0.014	0.012	0.018	0.008	0.018
	(0.009)	(0.010)	(0.009)	(0.008)	(0.011)	(0.014)	(0.015)
Flooded District x Year 2012-13	0.012	0.010	0.010	0.012	0.010	-0.001	0.020
	(0.008)	(0.008)	(0.012)	(0.009)	(0.008)	(0.015)	(0.013)
Flooded District x Year 2014-15	0.014*	0.021	0.010	0.014	0.013	0.006	0.013
	(0.008)	(0.013)	(0.009)	(0.010)	(0.008)	(0.011)	(0.009)
Observations	1,425,604	488,563	937,041	718,214	707,390	467,915	469,126
Clusters	115	107	115	115	115	115	115
District Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Marriage - whether an individual (aged 15-50 years) is married (Marriage=1) or unmarried (Marriage=0). Robust standard errors in parentheses, clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. Models estimated by OLS. District and province-year dummies are included. Weights: IPTWs estimated in the section 2.4 are used. Data Source: PSLM surveys (various issues), PBS, Islamabad, Pakistan.

2.8.4 *Figures*

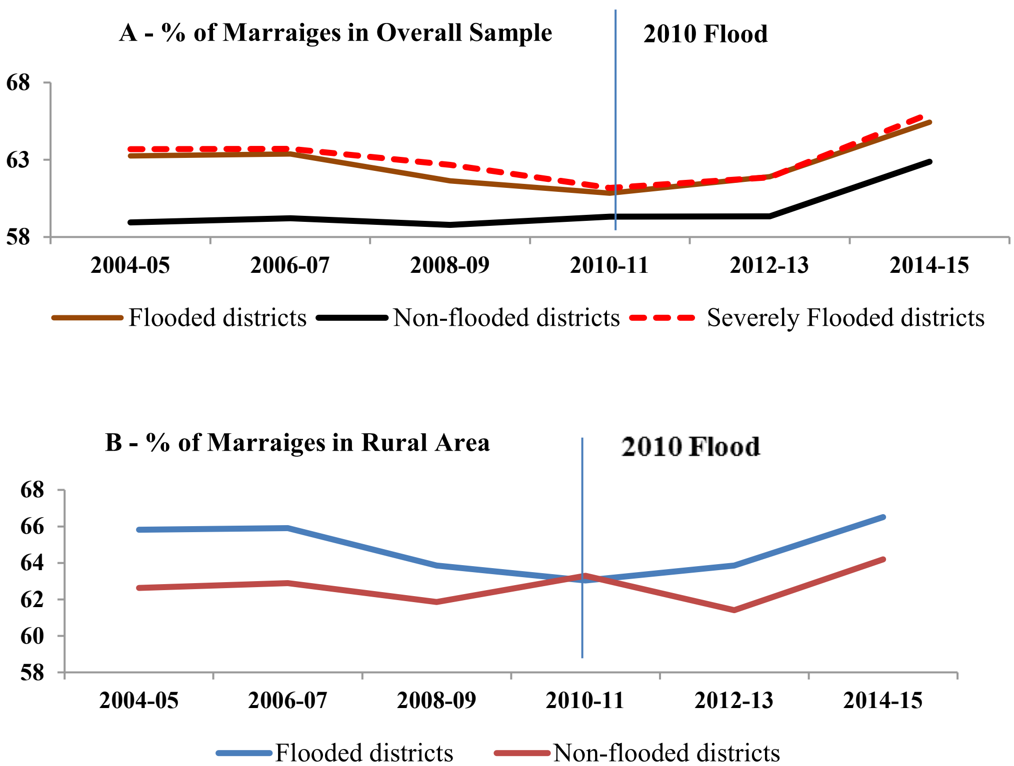


Figure 2.1: Percentage of Married Individuals Aged 15-50: 2004-05 to 2014-15
Data Source: PSLM surveys data (various issues), the Critical Threats, MapAction, National Disaster Management Authority (NDMA), and author own calculation.

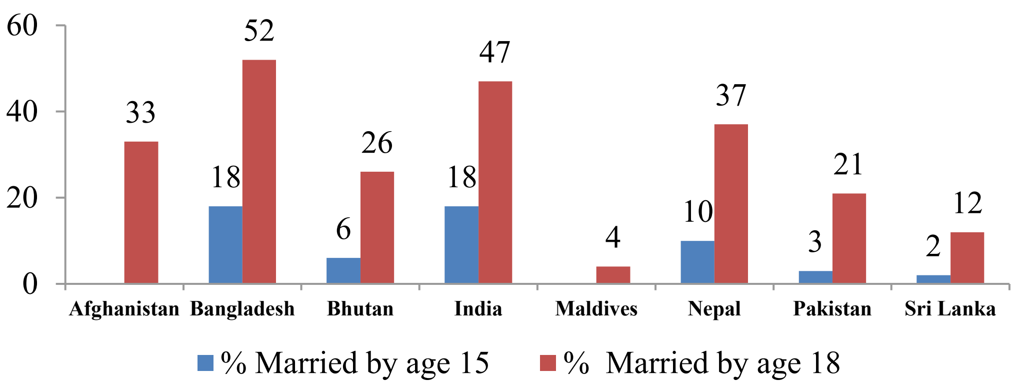


Figure 2.2: Child Marriage Rates (UNICEF, 2016)
Data Source: http://www.girlsnotbrides.org/about-child-Marriage/

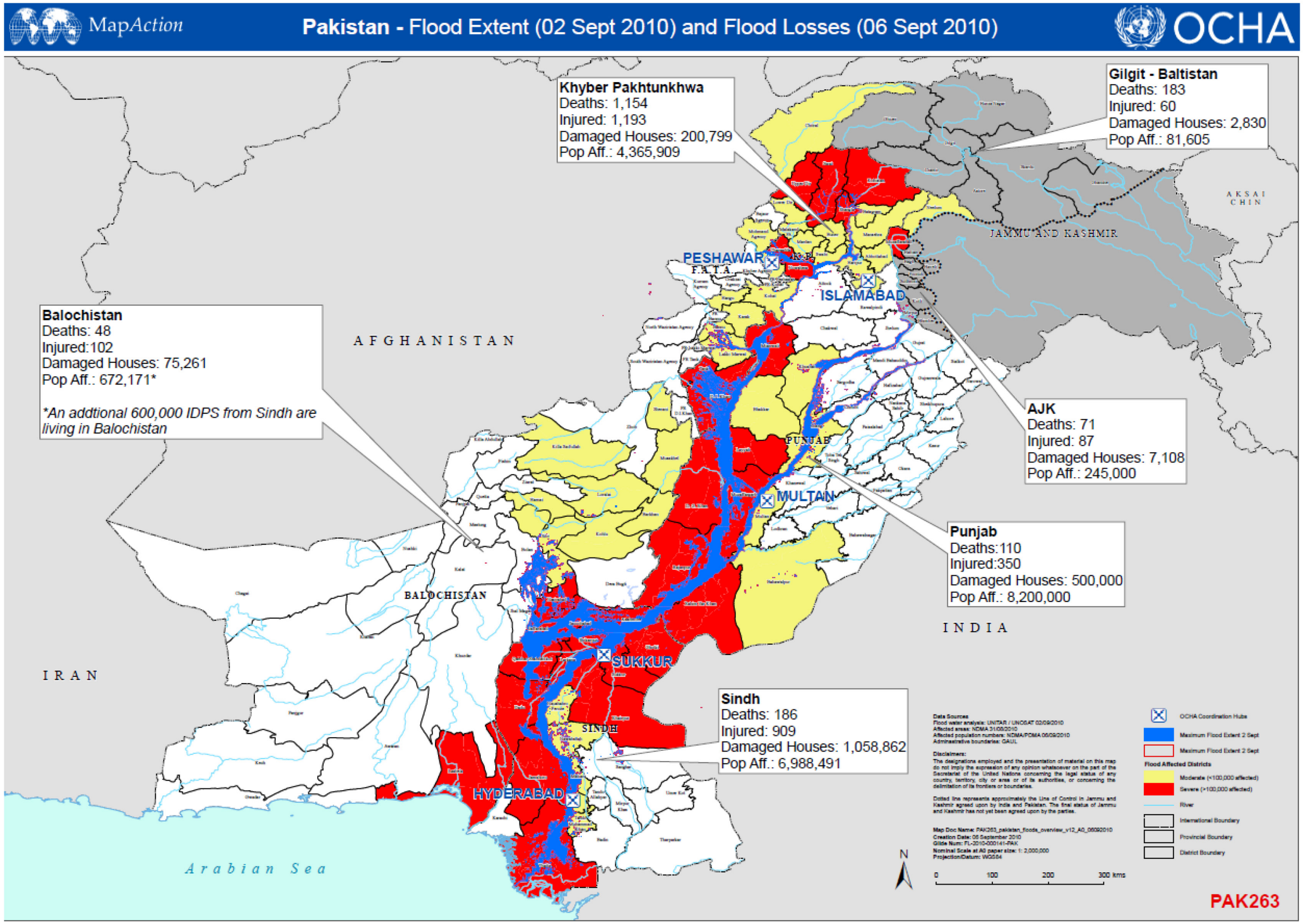
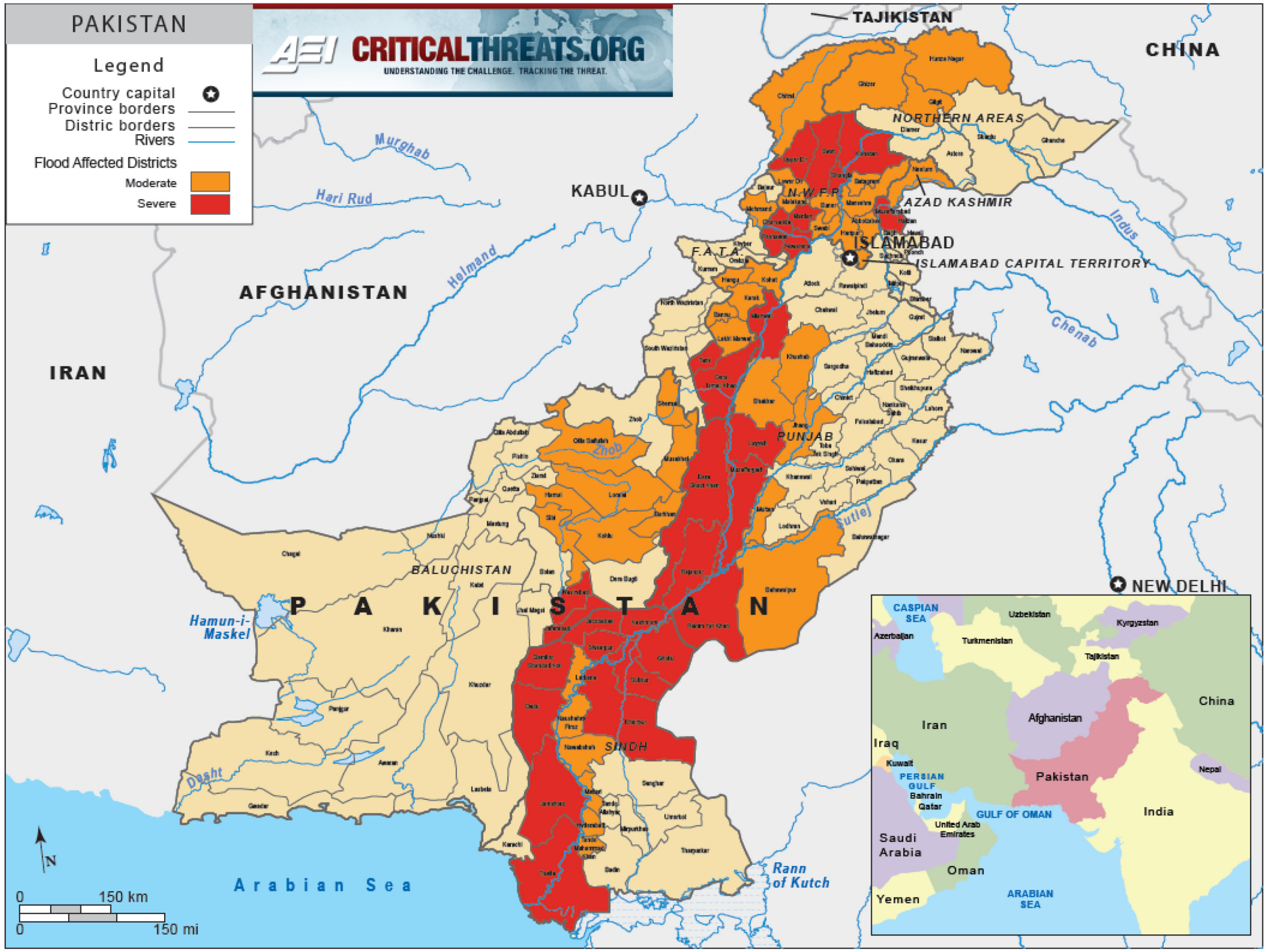


Figure 2.3: Map 1 for the 2010 Flood-affected Districts in Pakistan
Source: MapAction (2010)



Flood affected districts in Pakistan as of 26 August 2010 (Source: UN OCHA, Dawn News)

Figure 2.4: Map 2 for the 2010 Flood-affected Districts in Pakistan

Source: Critical Threats Project (2010)

Chapter 3

Social Infrastructure and Firm Productivity: Evidence From Pakistan*

Abstract

Does investment in social infrastructure affect the productivity of manufacturing firms in developing countries? To test this question, I empirically investigate the conditional correlation between social infrastructure indicators at district level and firm productivity by using a firm level dataset from Pakistan. Further, 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 infrastructure, conditional on various control, are positively and correlated with firm output in manufacturing industries in Pakistan. However, these results hold only for health infrastructure in urban regions of a district. For rural regions, both health and education infrastructure show insignificant relationship with firm productivity.

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

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

* Published as a discussion paper at ZEW, Mannheim; Discussion Paper No. 16-038.

3.1 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 in 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 (2016); Fisher-Vanden, et al. (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 firm firm productivity, one should consider infrastructure data at a more disaggregated level such as at 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) level growth and poverty.

There are two differences 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 correlated to the variation of social infrastructure across districts, conditional on various controls at both district and firm level*. 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 Section 3.5 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 the negative agglomeration effect if industrial agglomerations exceed a certain size.

To investigate the relationship of investment in social infrastructure with productivity of business enterprises, I take Pakistan as a case study. I use the 2005-06 Census of Manufacturing Industries (CMI 2005-06) dataset 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, 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[12] (BHU) in a given district is apparently positively associated with an increase in firm output in manufacturing industries by 0.57%. Based on a variety of sensitivity checks, my findings remain robust.

The rest of the chapter 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 output of firm. In this section, 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 3.3, I discuss the regional disparities of social infrastructure i.e. education and health indicators in Pakistan. Section 3.4 describes the data I use in this study. In Section 3.5, I elaborate the modeling framework and estimation strategies of the empirical analysis. Section 3.6 presents and discusses the empirical findings, followed by a concluding remarks section, Section 3.7.

3.2 Effect of Social Infrastructure on Productivity and Output of Firm

3.2.1 *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

[12]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.

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 relationship 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 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. 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) applied GMM estimators and used lagged variables of some exogenous variables such as population density, urban population, and labor force as instrument variables and found 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.

3.2.2 *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) studied the relationship between health and growth by taking nutrition as health indictor based on data for South Africa for the years 1961-2013. They found a positive relationship between nutrition and economic growth. They also found a causal relationship which flows from nutrition to economic growth much stronger than in the other direction.

3.2.3 *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, and 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[13] on

[13]Measured as per worker education expenditure.

both short 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), who tested this hypothesis in Algeria and they also found 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.

3.2.4 *Literature: Health and Education*

The paper by Baldacci, et al. (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 implication. 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[14]. 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 economic growth (measured by GDP per capita).

[14]The unit of mortality rate is death of children (under five years) per thousand children.

3.2.5 *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 Pakistani 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 for total factor productivity (TFP) and output in the context of Pakistan by taking the country's time series data from 1961 to 2013 (see Ali and Ramay 2014 for more details on how to calculate human capital). They concluded that human capital has an impact 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 regression models. To test for an indirect influence, they also included an interacted term of human capital variable with labor force and physical capital, respectively, and found a similar effect on output and TFP. They argued that labor and capital could be utilized more productively in the presence of higher human capital.

Afzal, et al. (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 only depict a general picture. But for policy implications, more disaggregated findings on the effects of investment in human capital on output and productivity of firms in different sectors and industries and at the firm level would be needed. In this respect, Amjad, et al. (2012) analyzed a firm level perception survey[15] which asked exporting firms about key hurdles and limitations when accessing

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

international markets. Among other factors, they found that the 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 they neglected 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 are 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 such 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.

3.3 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 Schwab and Sala-i-Martín (2014) defines a number of factors required for development of an economy depending upon the 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 South Asia Region (SAR). 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 sector 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 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). The role of manufacturing sector for growth and development of an economy has been realized for a long time. But in Pakistan, partly because of the government's fiscal policy which has been highly taxing manufacturing sector 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 3.1 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).

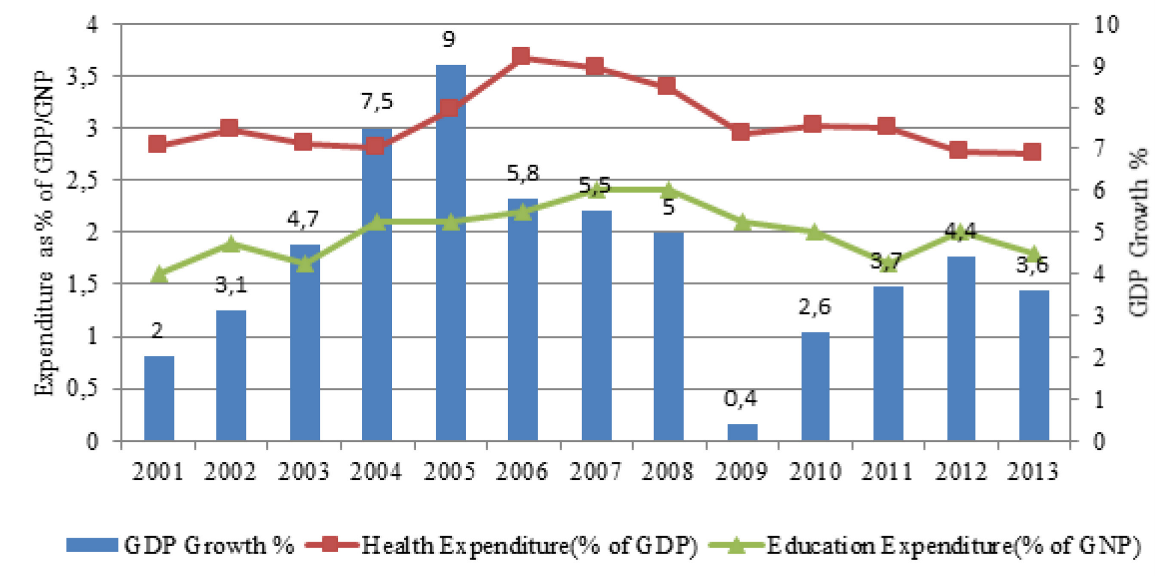


Figure 3.1: Real GDP Growth and Expenditure on Social Infrastructure in Pakistan, 2001-2013
Data Source: Economic Survey of Pakistan (various issue)

Over the past decade, the service sector in Pakistan grew massively and clearly gained in significance for 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 countries in SAR, 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.

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 3.2). 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.

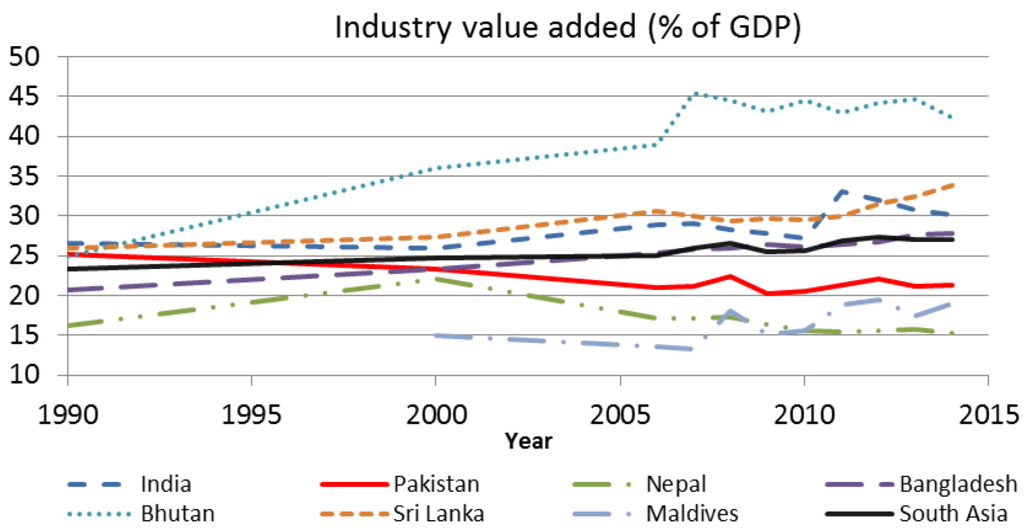


Figure 3.2: Share of Industrial Value Added in GDP in South Asia, 1990-2014
Data Source: World Bank Development Data

The weak performance of manufacturing sector 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 positioned itself among the top five cotton producing countries in the world successfully, 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 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 industrial sector to 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 of FDI in manufacturing sector. One is related to the failure of 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 to investors. Pakistan is gradually deteriorating its position in these indices. For instance in the GCI, Pakistan held 73rd position (out of 101 countries) in 2003 and dropped to 126th position (out of 140 countries) in 2015. Its rank was also far below than that of the competing countries in South Asia. Figure 3.3 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.

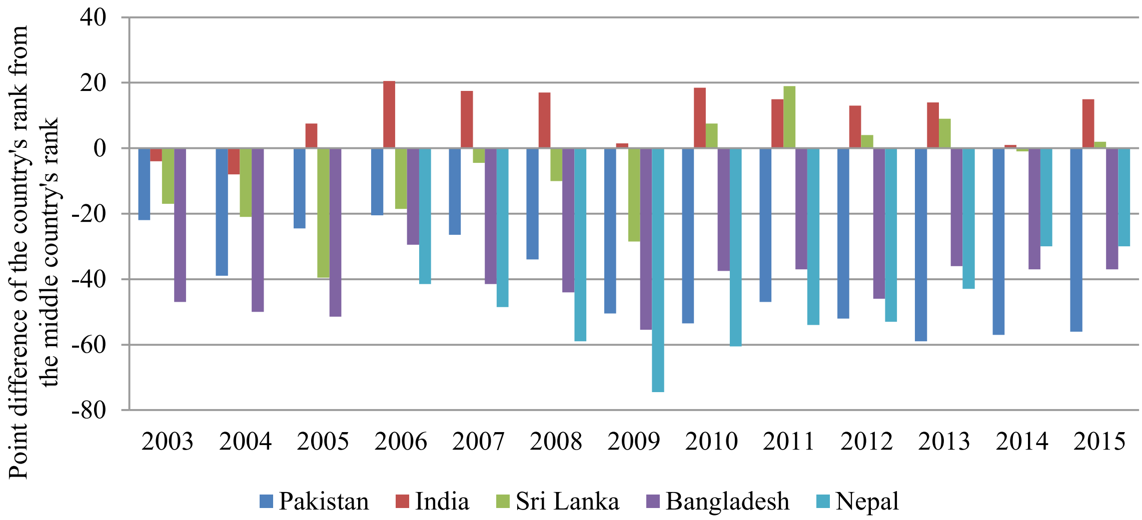


Figure 3.3: Global Competitive Index Ranking in South Asia, 2003-2015
Data Source: World Bank Economic Forum: Global Competitiveness Report (various issues)

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 Figure 3.4). It started its 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 Figure 3.4).

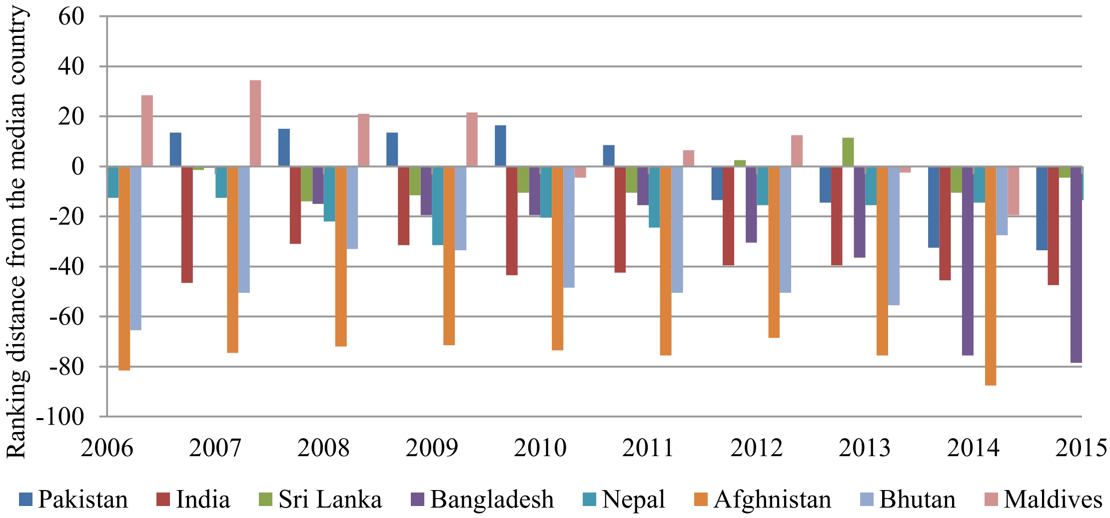


Figure 3.4: Ease of Doing Business Ranking in South Asia, 2006-2015
Data Source: World Bank: Doing Business (various issues)

Pakistan, the most rapidly urbanizing country in South Asia, is facing some serious issues associated with provision and quality of infrastructure (Cockburn, et al. 2013). The population of Pakistan has heavily been shifting from rural to urban regions. Unlike other South Asian countries, Pakistan is not reaping the benefits of 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 and 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 business sector and hence for the development of the entire economy. With present study, I want to shed some light on the association of social infrastructure with productivity of manufacturing firms in Pakistan.

For achieving the desired level of economic growth and development and switching from low productive to high productive economic activities, the focus on development of manufacturing industries 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 has widely been failed to attain a level of social infrastructure which could boost 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 Region (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 "*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 (Government of Pakistan 2013). One of the main reasons for this worse situation is the low tax collection. In 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, the 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 Pakistani government towards education is to look at the amount of money that was spent on education. Despite the government promise on spending up to 6% of GDP on education for achieving the goals of "*Dakar Framework for Action, Education for All"* (UNESCO 2015), education spending as a percentage of GDP was only around 2% during 2001-2013 (see Figure 3.1).

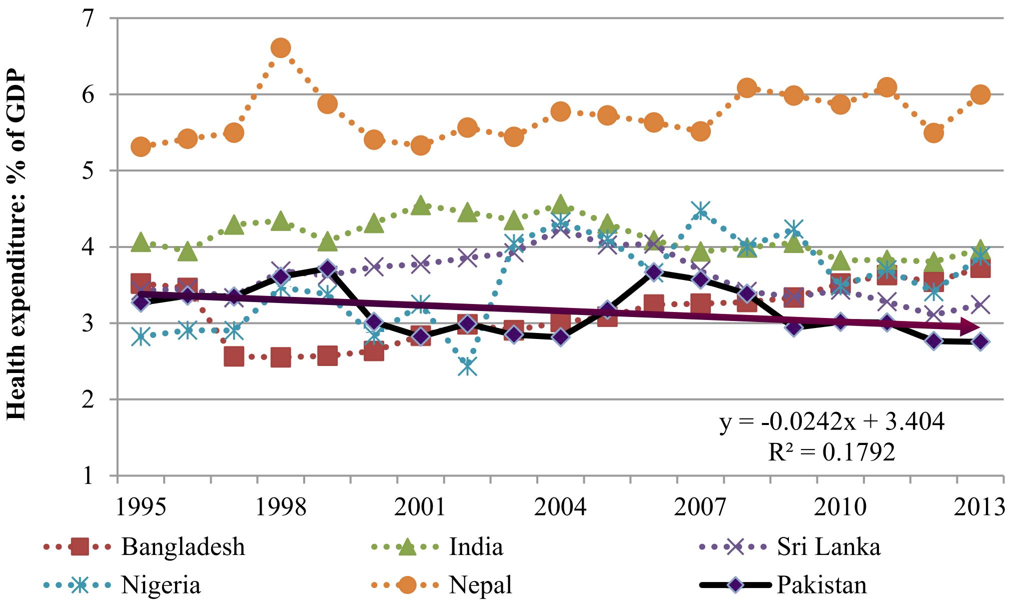


Figure 3.5: Health Expenditure: Selected South Asian Countries and Nigeria, 1995-2013
Data Source: World Health Organization: Global Health Expenditure Database

Likewise, indicators associated to 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 health sector has been reduced over time (see Figure 3.5). 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[16] (HDI), reveals that there is unequal distribution of resources on public services across Pakistan (see Figure 3.6 on page 104). The HDI values at district

[16]District HDI, calculated by Jamal, Haroon, and Amir Jahan Khan, "Trends in Regional Human Development Indices," Social Policy and Development Centre Research Report No. 73, 2007. for 2005.

level indicate that most of the districts in Punjab are well above the average value while the districts in Balochistan, with 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 3.7 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[17] 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[18] (see Figure 3.8 on page 106 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 pupils could do math 2-digit division problems in contrast to Punjab where even in rural areas the score is above 50%.

Regional inequality of social services is not only present in education sector; indicators for health sector also depict similar stories. In fact education and health sectors often go hand in hand. Districts with higher values of 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[19], whereas the percentage for Punjab is much higher (see Figure 3.9 on page 107).

In order to better understand the constraints and limitations of public spending on social infrastructure at district level, I briefly summarize some of the main features of fiscal policy of Pakistan based on a study done by Ahmed (2013, pp. 67-111). Federal, provinces and districts are 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 a part of service sectors because most of the advantages from these sectors are directly benefiting to the ruling-elites in the country. Second, spending on defense is one of the escalating expenditures,

[17]ASER stands for 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.
[18]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 across all districts in Pakistan. But the language of mathematics is the same for all of them.
[19]RHCs and BHUs stand for rural health centers and basic health units respectively.

which limits policy makers to allocate resources in social sectors such as health, education, transport infrastructure, etc. Almost one third of the total budget is 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. 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 has been reducing 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 each other by allocating 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 due to huge corruption. In a nutshell, investment in social services and infrastructure, and poverty alleviation projects are negatively affected by the high expenditure for non-development, defense and debt service.

3.4 Data and Descriptive Statistics

In this paper, I analyze the association of investment on social infrastructure with firm productivity in manufacturing industries. To that end, I need both firm level 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 (CMI) 2005-06—conducted by the Pakistan Bureau of Statistics (PBS). The dataset consists of 6,417 manufacturing establishments which are either registered or qualify for registration under the Factories Act 1934[20]. 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's 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.

[20]The survey was conducted by Pakistan Bureau of Statistics (PBS) with the help of 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.

Infrastructure indicators are taken from the Pakistan Social & Living Standards Measurement (PSLM) surveys—also conducted by the PBS for each alternative year at the district level since 2004-05. Since data at the firm level are available for fiscal year 2005-06, I use district level indicators of health and education from the PSLM 2004-05 and 2006-07 surveys' reports[21] which are available on the official website of PBS, Islamabad, Pakistan. From these surveys, I calculate the averages in order to represent information for the year 2005-06.

I use a district level control, human development index (HDI) in Pakistan for the year 2005-06—calculated by Jamal and Khan (2007). The estimated population in year 2006 is taken from the books of provincial development statistics, and for calculating the percentage of urban population in a 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 books of provincial development statistics.

Firm and district level variables and their brief description are given in Table 3.1 and for definition of each variable, please refer to the Appendix to Chapter 3.

I did some minor adjustments to both firm and district level data. The CMI 2005-06 dataset contains a total of 6,417 firms in twenty-two 2-digit industries, covering 73 Pakistani districts, existing at the time of the survey. After cleaning the data, observations which were missing for two variables; capital and number of kilometers road paved per registered vehicle (*RdRv*) have been adjusted by including two dummies for each variable and setting their missing values to zero. Missing observations may have some systematic bias because almost 40% of the missing observations on capital variable are from the food and beverage industry (Division 15) and 75% of the missing observations of *RdRv* are from Islamabad district—as I could not find an authentic source of data on the number of kilometers of paved roads in 2006 for Islamabad district.

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

Table 3.1: Description and Sources of Firm and District Level Variables

Variable	Description	Sources
Output (Y), log	Log of total output at producer price in Pakistani Rupees.	CMI 2005-06
Labor (L), log	Log of average number of employees (both production and non-production) during the fiscal year 2005-06.	CMI 2005-06
Capital (K), log	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 2005-06
Materials (M), log	Log of cost of materials in Pakistani Rupees consumed during the fiscal year 2005-06. It includes fuels, electricity, raw materials, chemicals & dyes, packing materials, spare parts, lubricants, and others both locally purchased and imported.	CMI 2005-06
Government	State ownership dummy. Equals 1 if the firm owned by Pakistani state.	CMI 2005-06
Foreign	Ownership fully or partially owned by foreigners. Equal 1 if either government or private firms fully or partially owned by foreigners.	CMI 2005-06
Import	Firms involved in importing materials abroad. Equal 1 if the value of firm imported material is greater than zero.	CMI 2005-06
EG Index	It is calculated based on data available in CMI 2005-06 data set.	CMI 2005-06
Estimated Population 2006	It is log of estimated population (inhabitants) of districts in fiscal year 2005-06 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 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 the authors applied a method by taking the provincial level data to estimate district level income.	Jamal and Khan (2007)
Distance to Karachi	It is the distance between the given districts to Karachi in kilometers— calculated by the formula in footnote 24 on page 88.	www.wemakemaps.com www.distancesfrom.com
Population Satisfied with BHU[c]	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 2006-07
TT- Injection[c]	Percentage of all currently married women aged 15-49 years who had a birth in the last three years and received a tetanus toxoid injection during the last pregnancy.	PSLMS 2004-05 and 2006-07
Primary Net-Enrolment (5-9 years)[c]	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 2006-07
Adult Literacy Rate (age 15 or above)[c]	Population aged 15 years and older who can read a newspaper and write a simple letter.	PSLMS 2004-05 and 2006-07

Notes: (a) Road data for districts of Punjab in year 2006 are missing, so the average of years, 2005 and 2007, are used. (b) The only available value of the HDI for Islamabad district for year 2012-13—used in the analysis and the value is retrieved from: Pasha, et al. (2016). (c) Averages between surveys 2004-05 and 2006-07.

Table 3.2: **Descriptive Statistics of Firm and District Level Variables**

Variable	Observations	Mean	Standard Deviation	Min.	Max.
Firm level variables					
Output (Y), log	6,413	10.909	2.084	4.83	18.39
Capital (K), log	6,213	9.458	2.153	3.22	17.69
Labor (L), log	6,413	3.685	1.322	0.70	9.73
Materials (M), log	6,413	10.408	2.227	1.95	18.02
Government Ownership	6,413	0.086	0.280	0	1
Foreign Ownership	6,413	0.044	0.206	0	1
Importing Firm	6,413	0.136	0.342	0	1
EG Index at Industry Level	6,413	0.058	0.119	-0.82	3.68
EG Index Square at Industry Level	6,413	0.018	0.252	0	13.52
Gini Coefficient	6,413	0.095	0.102	0.03	0.98
District level variables					
Estimated Population 2006, log	73	7.466	1.057	5.73	14.37
% of Urban Population 1998	73	24.055	17.818	0	94.75
Number of Kilometers per Registered Vehicle	69	0.3805	2.4281	0.00	20.22
HDI 2005	73	0.6486	0.0736	0.31	0.89
Primary Net Enrolment, Total	73	75.970	14.200	31.00	97.50
Primary Net Enrolment, Urban	73	57.170	19.660	0	88.50
Primary Net Enrolment, Rural	73	82.250	11.920	44.00	99.50
Literacy Rate, Total	73	46.550	11.880	26.50	83.50
Literacy Rate, Urban	73	60.840	14.440	0	90.50
Literacy Rate, Rural	73	40.850	11.170	21.50	71.50
Population Satisfied with BHU, Total	73	39.345	13.892	7.40	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
TT-Injection, Total	73	52.726	17.069	19.50	84.00
TT-Injection, Urban	73	65.342	16.893	0	93.50
TT-Injection, Rural	73	47.151	17.896	14.50	82.50

Note: For detail description and definition of each variable see Table 3.1.

In order to get robust findings, I therefore also do an exercise in section 3.6 to check whether or not these missing observations are sensitive to the findings by excluding those in the analysis. Excluding missing observations leaves a sample with 6,059 firms; so the final samples vary 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 3.2 shows the descriptive statistics of dependent and independent variables. The list of districts and 2-digit industries are given in Table 3.5 and Table 3.6 respectively.

3.5 Modeling the Relationship between Social Infrastructure and Firm Productivity

For investigating the relationship between social infrastructure and productivity or economic growth at firm level empirically, a number of previous studies employed the framework of the Cobb-Douglas production function (Aziz, Khan, and Aziz 2010; Benhabib and Spiegel 1994; Brandt, Van Biesebroeck, and Zhang 2012; Mekdad, et al. 2014; Ping 2005). In this study, I follow a 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 a cross-section firm-level dataset 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:

$$Y = \alpha L^{\beta 1} K^{\beta 2} M^{\beta 3} S^{\lambda} e^{u}. \tag{3.1}$$

In equation (3.1) above, α represents total factor of productivity, β_1, β_2, β_3, and λ are output elasticity of labor, capital, materials and social infrastructure respectively and u is the usual error term. Taking natural logarithms of the above equation one gets:

$$y = \alpha + \beta_1 l + \beta_2 k + \beta_3 m + \lambda s + u. \tag{3.2}$$

For empirically estimating the model (3.2), 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 ($Z \in q=1,...,Q$) 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 *d* in province *p* in which a firm *i* of industry *j* 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 the geographical concentration of industries. This index is measured at the level of industry sector *j* a firm *i* belongs to. Finally, it is useful to add further control variables ($X \in r=1,...,R$) representing a firm's ownership and market orientation. The equation to be estimated reads as follows:

$$y_{ijd} = \alpha + \beta_1 l_{ijd} + \beta_2 k_{ijd} + \beta_3 m_{ijd} + \sum_q \delta_q Z_{qd} + \varphi EGI_j + \sum_r \gamma_r X_{rijd} + D_j + D_p + \varepsilon_{ijdp}, \quad (3.3)$$

where Z represents district-specific variables including social infrastructure and ε_{ijdp} is an idiosyncratic error term. Industry (D_j) and province (D_p) dummies are included in the estimation. Equation (3.3) will be estimated by using ordinary least square (OLS) 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:

$$EGI_j = \frac{\sum_{i=1}^{M}(EI_{jd} - ED_d)^2 - (1 - \sum_d ED_d^2)\, H_j}{(1 - \sum_d ED_d^2)\,(1 - H_j)}, \quad (3.4)$$

with EI_{jd} being the jth share of employment in district d and ED_d respresents total manufacturing employment in district d. H_j is the Herfindahl index which measures the ith plant level concentration of employment in industry j. It is defined by:

$$H_j = \sum_{i=1}^{M} EP_{ij}^2, \quad (3.5)$$

where EP_{ij} is the share of the ith plant level employees 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 and Klenow 2000; Krueger and Lindahl 2000), student teacher ratio, and literacy rate are some of among others has widely been 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 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 be entirely changed. 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 when funds are allocated to build hospitals, there would be 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 infrastructure stock since it excludes the role of private sector in provision of these facilities. 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 infrastructure 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 married women (aged 15 to 49 years) who received TT-Injection during the last pregnancy. For measuring education capital, I use the net enrolment in primary school and the adult literacy rate of people aged 15 and above. 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 women are less likely to get TT injection.

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 group 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 adult literacy rate for the population of 15 years or older.

In order to investigate likely differences of the association 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.

In order 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[22] 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 district population in log and the percentage of urban 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.

Other potential sources 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. In order 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*).

A further regional control variable is the distance to Karachi (Dis); as Karachi is the main international port city of Pakistan. This variable is intended to capture a likely advantage a firm might get from its location in or near to the port city. Ahrend, et al. (2014) in the context of cities in five OECD[23] countries found port cities are 2 to 4 percent more productive than other cities. The distance variable (Dis) is measured the distance between the capital city of a district and Karachi using longitude and latitude data—by applying the Haversine formula[24]. Since the districts in Sindh Province are located nearer to Karachi than the districts in other provinces, the distance variable and Sindh province dummy are highly correlated; so the effect of distance on firm productivity is reported rather in the robustness checks section.

Finally, I also include the Human Development Index (*HDI)* to control for the level of human development at the 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 checks section.

[22]The only latest census data available is the fifth census conducted in 1998 and data collection for the sixth census, Census 2017 has been started in March 2017.

[23]OECD countries include Germany, Mexico, Spain, United Kingdom, and United States.

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

Another source of concern on the OLS estimates 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 lead towards wrong conclusions about rejecting null hypotheses (Cameron and Miller 2015). Based on the assumption that errors are correlated within districts but uncorrelated across districts, I test the hypotheses while clustering the errors at the district level.

3.6 Estimation Results and Discussion

3.6.1 *Main Results*

In econometric approach, I used a log-log production function in the cross-section analysis. Firm and district level control variables along with industry and province dummies are included in the estimation. Errors are clustered at the district level. And models are estimated by OLS. Estimation results of the variables of interest are shown in columns (1) to (5) of Table 3.3 on page 92. 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 all control variables are not reported in the table.

I start discussion with the base line model which represents the estimation results of equation (3.3) in section 3.5. The elasticity of firms' conventional inputs; labor, capital, and materials are in their expected range; 0.15, 0.053, 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 and 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 relationship 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 Sub-section 3.6.2). When excluding the squared term, the estimation results imply that agglomeration does not matter in manufacturing industries in Pakistan[25].

At the district level, I use two control variables; estimated population in 2006 and the number of paved roads in kilometers available for 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 the conventional levels.

Now I turn my attention to the main question, do social infrastructure and firm total factor productivity in a district go hand in hand? In column 1, panel A of Table 3.3, I test the association of social infrastructure (health and education) and firm total factor productivity. Indicators for both social infrastructure i.e. health and education are positively correlated with firm output but not significant at the district level. However, when I ran regressions at a more disaggregated levels i.e. rural (column 1, Panel B) vs. urban (column 1, Panel C), then interestingly I found a positive and significant conditional correlation between health infrastructure and firm output in urban areas only while in all other cases, the conditional correlation between social infrastructure and firm output is statistically insignificant. More precisely, if government and/or non-government agents invest in facilities of basic health units (BHUs) in urban areas of a district, and as a result of that investment if the share of the population that is satisfied with BHUs' facilities increase by one percent (say from 40% to 41%), this is positively associated with—ceteris paribus—firm level output by 0.41%, which is significant at the 10% level. The results which I find in my analysis for health infrastructure indicators are within the line of results found by the study of Sumarto and De Silva (2014), who studied the effect of health and education indicator at district level on the subnational growth and poverty in Indonesia.

[25] 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.

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 the district level. In this regard, I first check for firm size effects by splitting my sample into two parts; small firms (column 2) and large firms (column 3) by using the median value of the number of employees per firm in the industry. The output of both types of firms, small and large, is positively associated with health infrastructure in urban areas while output of small firm is positively associated with education infrastructure in rural areas which indicates that conditional on various control, manufacturing industries could get benefit from the investment in urban health infrastructure while small firms seemingly get benefit from the rural education infrastructure. In all other cases, the correlations are insignificant.

In order to test the association of both rural and urban social indicators with productivity of high vs. low capital intensive firms; again, I spilt the sample into two parts based on the median capital intensity. Results are shown in columns (4) and (5) respectively. Here I find that there exists no conditional correlation between social infrastructure and firm productivity of high capital intensive firms. On the other hand, for low-capital-intensive firms, investment in health infrastructure and facilities in urban areas while investment in education infrastructure in rural areas seemingly matter for their firm productivity. Surprisingly a negative and significant conditional correlation between primary net enrolment in urban areas and output of low capital intensive firms is also observed.

Table 3.3: Effect of Social Infrastructure on Firm Output: Results from OLS Regressions

Dependent Variable: Log(Total Output at Producer Price)	Overall Sample	Large Firm	Small Firm	High Capital Intensive Firm	Low Capital Intensive Firm
	(1)	(2)	(3)	(4)	(5)
Panel A: Social infrastructure at District Level					
Population Satisfied with BHU	0.0012	0.0020	0.0008	-0.0019	0.0038
	(0.0019)	(0.0023)	(0.0015)	(0.0013)	(0.0024)
Primary Net Enrolment	0.0005	-0.0037	0.0030**	0.0038**	-0.0016
	(0.0019)	(0.0030)	(0.0013)	(0.0015)	(0.0021)
Observations	6,413	3,205	3,208	3,104	3,105
Clusters	73	65	69	70	66
Panel B: Social Infrastructure at Urban Level					
Population Satisfied with BHU	0.0041*	0.0041**	0.0028*	0.0009	0.0069***
	(0.0021)	(0.020)	(0.0016)	(0.0014)	(0.0021)
Primary Net Enrolment	-0.0013	-0.0040	0.0008	0.0018	-0.0045**
	(0.0018)	(0.0024)	(0.0014)	(0.0014)	(0.0017)
Observations	6,413	3,205	3,208	3,104	3,105
Clusters	73	65	69	70	66
Panel C: Social Infrastructure at Rural Level					
Population Satisfied with BHU	-0.0002	-0.0003	0.0001	-0.0011	0.0005
	(0.0010)	(0.0012)	(0.0013)	(0.0012)	(0.0016)
Primary Net Enrolment	0.0030	-0.0025	0.0053*	0.0014	0.0060*
	(0.0028)	(0.0019)	(0.0054)	(0.0027)	(0.0035)
Observations	6,413	3,205	3,208	3,104	3,105
Clusters	73	65	69	70	66

Notes: Dependent Variable: Log (Total output at producer price). Other controls: Firm level variables (logs of capital, labor, and materials, EG index, government ownership, foreign ownership, and importing firm), district level variables (log district population for panel A/% of urban district population, 1998 for Panel B, and C and number of km per registered vehicle). Industry and Province dummies are also included. Models estimated by OLS. Two dummies for missing observations of log (capital) and number of km per registered vehicle are included, except in (5) and (6). Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1. Data Sources: See Table 3.1 on page 82.

3.6.2 *Robustness Checks*

In order to investigate the robustness of my findings presented above, I run a number of alternative model specifications to test my hypothesis. From the main findings I observed that firms located in Balochistan province 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 1 of Table 3.7 on page 103. The findings remain robust.

In column 2 of Table 3.7, I alter the social infrastructure indicators; as sometimes such proxies may mislead the results. For this analysis, I take two other available proxies for health and education. For health, I take the percentage of women received a tetanus toxoid injection during the last pregnancy. For education, I take the literacy rate of the population at age 15 or above. Results in column 2 show that the magnitude of the coefficient on health infrastructure, the variable of interest, decreased a little bit, still remain significant at the 10% level.

In the main findings, I did include two dummy variables to capture missing observations of two variables; capital[26], and number of kilometers of paved road available per registered vehicle[27]. When running the regression excluding all missing observations; this time, only very small and negligible variations in my findings occur (see column 3 Table 3.7).

In regression (4), 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 a district level human development index (HDI)[28] 2005 in my regression analysis. From this exercise, I find that the conditional correlation between HDI and firm level output is positive and statistically significant. Turning to my variables of concern, I find similar and robust results in term of their signs and magnitudes against the specification (1) of Table 3.3, panel B and C.

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, EG Index would not remain a linear function. In order to test the

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

[27]Since this variable derived by dividing the number of paved road in kilometers 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.

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

non-linearity of EG Index, I also 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 remain insignificant at the conventional significance levels (see column 5 Table 3.7). 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 firms in Pakistan[29]. However, my findings are consistent with the findings of Lin, Li, and Yang (2011), who found 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 estimates of the variables of interest 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 main port city in Pakistan). The results are shown in the column 6 of Table 3.7 on page 103. The coefficient for the distance between a firm's location district and Karachi is negative and statistically significant in both cases, urban vs. rural. These result suggest that firms located closer to Karachi get better access to a larger local consumer market and to international markets for exporting their products and importing inputs. The magnitude of the distance-variable shows that the output of a firm, on average, is lower by 0.04-0.06 percent; as if it located its plant(s) one kilometer away from Karachi. Doing this analysis does not change the signs of my variables of interest. The magnitudes and significant levels of the coefficients of interest decreased this time which might be because the distance variable and the dummy for Sindh province are highly correlated.

3.7 Conclusion

In this study, I investigated whether investment in social infrastructure (health and education) is correlated with firm productivity in manufacturing industries in Pakistan. At district level I found no significant relationship between social infrastructure and firm productivity. However, I found a positive and significant conditional correlation between health infrastructure in urban region of a district and firm output at firm level in that

[29]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.

district which is consistent from the literature. The positive conditional correlation between health infrastructure in urban regions and firm output remained robust in almost all of the main and further extension analysis. Furthermore, investment in urban education infrastructure seemingly matters for small and high capital intensive firms. This might be because both type of firms to be in market required skilled labor force. The association of investment in both social infrastructure in rural areas with firm output is a district is ambiguous this may be because the infrastructure which mainly targets to the rural population engaged in agriculture activities and has little spillovers to manufacturing sector or there may be reverse causality. Given that by using a single cross section dataset, it is hard to disentangle the causal relationship.

In addition, I also controlled for a likely firm self-selection bias by including the EG index which could measure 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.

This study contributes in the growing microeconomic literature by utilizing a more disaggregated firm level dataset to quantify conditional correlation between spending on social infrastructure and firm productivity in developing countries. This study also counters the quality of infrastructure data by taking social indicators as proxies for social infrastructure—recorded by Pakistan Bureau of Statistic in its PSLM surveys' reports, which I assume—are better predictors for actual social infrastructure supply than the recorded spending amount on the infrastructure projects in documents.

A number of caveats should be considered in the interpretation of the results. First, using the census of manufacturing industries data has several advantages, however: does not allow me to investigate any long term relationship between social infrastructure and firm productivity. The second caveat, which I am more concerned to, is that the geographic location information of a firm within a district in terms of rural vs. urban is not given precisely in the census. Therefore, this limitation also does not allow me 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 dataset 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. Thus, these findings should be considered as descriptive rather than causal relationship.

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

3.8 Appendix to Chapter 3

3.8.1 *Firm Level Dataset: 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[30]. The census is distributed in the following ways: Punjab (55.9%), Sindh (28.4%), KPK[31] (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[32] 2005-06. The CMI covered large scale manufacturing industries comprising establishments having 10 or more employees at 5-digit industries. The 4-digit industries of the CMI are comparable with the International Standards Industrial Classification (ISIC) Rev 3.1. In theory it is supposed to be conducted for every five years using the frame provided by the provincial labor departments. Though the survey for the CMI 2010-11 was done, the dissemination of data was stopped due to the lower response rate of the firms from Sindh province.[33]

The CMI (2005-06) frame was updated by using the industrial directories provided by the provincial directorates of industries and the results of economic census conducted by the Federal Bureau of Statistics. Information was gathered by a mailed questionnaire followed by a field visit by the provincial directorates of industries. And finally, annual reports of the listed companies in the stock exchange were also concerned to augment the coverage. Compared with the previous census, CMI 2000-01, the number of establishments' coverage in the CMI 2005-06 increased significantly (see Table 3.4).

---

[30]The survey was conducted by Pakistan Bureau of Statistics (PBS) with the help of 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.

[31]KPK, Khyber Pakhtunkhwa, previously known as NWFP, North West Frontier Province.

[32]In Pakistan the fiscal year starts on July 1st and ends on June 30th.

[33]According to concerned officials at the CMI section of Pakistan Bureau of Statistics, Islamabad, Pakistan—when I visited them for data collection in March, 2014

3.8.2 *Limitation of the CMI Data*

There are number of limitations by using the CMI data in terms of representative population of establishments in Pakistan. First, the survey frame included only those factories which were either registered or qlaified for registration under the Factories Act 1934. Factories, which may be eligible for the suvery, but not registered themselves: not included in the CMI 2005-06. Second, a considerable number of establishments—however were in the survey frame—were either closed their businesses (2,364 establishments) or defaulters (3,213 establishments). Third, there were 333 firms from Punjab province recorded double in the frame. Finally, 819 firms were found involved in activities other than the activities in manufacturing industries.

3.8.3 *Definitions of Firm Level Variables*

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 those *production workers*—whose works are directly associated with production such as manufacturing, assembling, packing, repairing, etc., including working supervisors and persons engaged for repairs and maintenance works—and *non-production workers*—administrative and professional employees, white-collar office employees, drivers, watchmen/guards, peons, sweepers, etc.

Material Consumed (Material)

Material consumed includes 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)

Capital stock or value of fixed assets is calculated by the following way: *capital stock* is the fixed assets at the beginning of the fiscal year 2005-06 plus *investment* and *investment* is the purchase of fixed assets plus fixed assets produced for own use minus sales of fixed assets.

Government Ownership (Government)

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

Foreign Ownership (Foregin)

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

Import Oriented Firm (Import)

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

3.8.4 *Definitions of District Level Variables (Variables of Interest)*

Population Satisfied with BHU

A perception based question was asked to households giving their opinions about whether or not they are satisfied with the basic health units (BHU) in their home town. The variable is used in this study is the percentage of population, who is satisfied with the services provided by BHUs at the district level.

Tetanus Toxoid Injections (TT Injections)

Percentage of all currently married women aged 15-49 years who had a birth in the last three years and received a tetanus toxoid injection during the last pregnancy.

Primary's Net Enrolment Rate (Aged 5-9)

The NER at primary level defines as the number of children aged 5-9 enrolled in primary school at district level divided by the total children of the same age group at the district level.

Adult Literacy Rate (Population 15 Years and Older)

Literacy rate of population 15 years and older is defined as the percentage of the population aged 15 years and older, who can read a newspaper and write a simple letter.

3.8.5 *Tables*

Table 3.4: 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: CMI 2005-06.

Table 3.5: Distribution of Establishments at the District Level

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

Notes: District Chinot is included in Jhang, and districts in Karachi are combined to form a district, Karachi. KPK stands for Khyber Pakhtunkhwa. Source: CMI 2005-06.

Table 3.6: Distribution of Establishments at the 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	0.20
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	7.70
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	5.80
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	100.00

Source: CMI 2005-06.

Table 3.7: Robustness Checks

Dependent Variable: Log (Total Output at Producer Price)	Sample Without Balochistani Firms	Alternative Proxies Used	Sample Without Missing Obs.	Overall Sample with an Additional Control (HDI)	Overall Sample with an Additional Control (EG-Index^2)	Overall Sample with an Additional Control (Proximity to Karachi)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Social in Urban Area Infrastructure						
Population Satisfied with BHU	0.0041**	0.0037*	0.0043**	0.0037**	0.0041*	0.0028
	(0.0017)	(0.0021)	(0.0021)	(0.0017)	(0.0021)	(0.0018)
Primary Net Enrolment	-0.0010	-0.0003	-0.0023	-0.0004	-0.0012	-0.0003
	(0.0013)	(0.0023)	(0.0021)	(0.0014)	(0.0019)	(0.0017)
Observations	6,201	6,413	6,059	6,413	6,413	6,413
Clusters	70	73	69	73	73	73
Panel B: Social in Rural Area Infrastructure						
Population Satisfied with BHU	0.0008	0.0016	-0.0004	0.0009	-0.0002	0.0007
	(0.0014)	(0.0020)	(0.0011)	(0.0014)	(0.0010)	(0.0009)
Primary Net Enrolment	0.0012	-0.0027	0.0032	0.0017	0.0031	-0.0062**
	(0.0025)	(0.0022)	(0.0029)	(0.0026)	(0.0028)	(0.0030)
Observations	6,201	6,413	6,059	6,413	6,413	6,413
Clusters	70	73	69	73	73	73

Notes: Dependent Variable: Log (Total output at producer price). Other control variables: log (capital), log (labor), log (materials), government, foreign, import, EG index, number of kilometers per registered vehicle, % of urban population and provincial fixed effects. Two dummies for missing observations of log (capital) and number of km per registered vehicle are included, except in (4). ***p<0.01; **p<0.05; *p<0.1. Models are estimated by OLS. Robust standard errors clustered at the district level in parentheses. In regression (1) firms from Balochistan are excluded. In regression (2) variables, population satisfied with BHU and primary net enrolment are replaced with TT Injection and adult literacy rate of population at the age of 15 or above respectively . In column (3), results of the model without adjusting missing values. In regression (4) HDI_2005 is included. In column (5), regression results with EG Index Square term. In regression (6), firm's distance to Karachi is included. Data Sources: See Table 3.1.

3.8.6 *Figures*

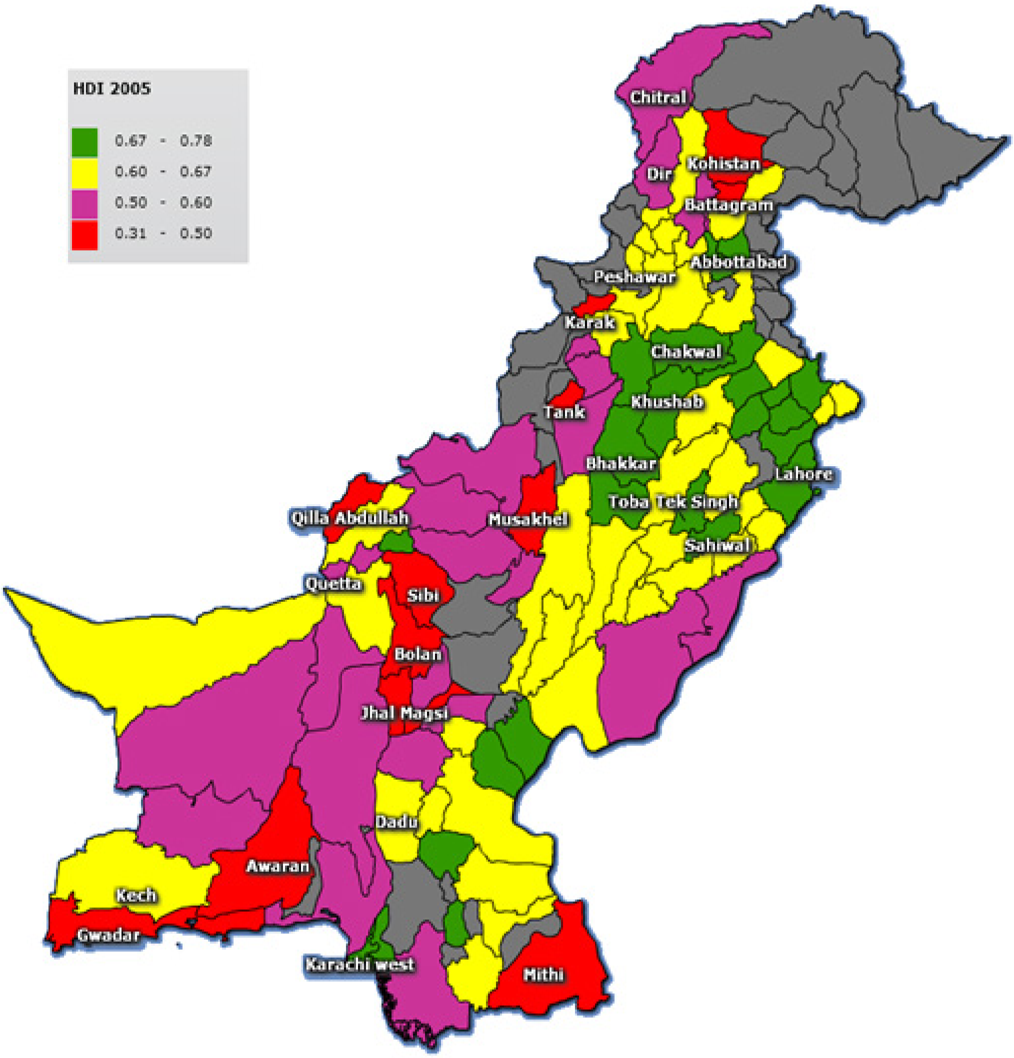


Figure 3.6: Districts' Human Development Index 2005, Pakistan
Data Sources: District level HDI data are taken from the paper by Jamal, H., & Khan, A. J. (2007) and map is prepared on the online website: http://www.targetmap.com/.

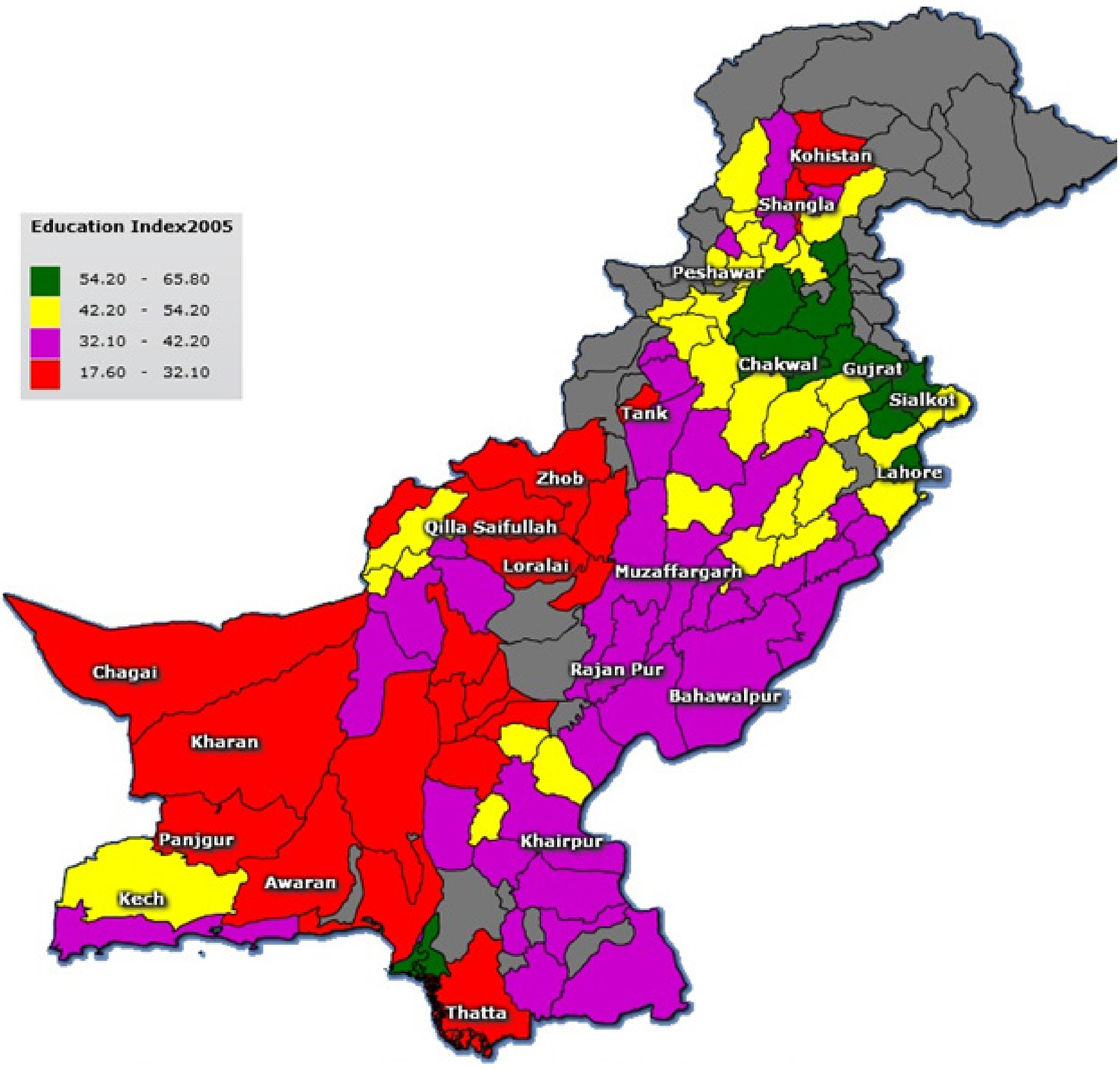


Figure 3.7: Districts' Education Index 2005, Pakistan
Data Sources: District level education index data are taken from the paper by Jamal, H., & Khan, A. J. (2007) and map is prepared on the online website:
http://www.targetmap.com/.

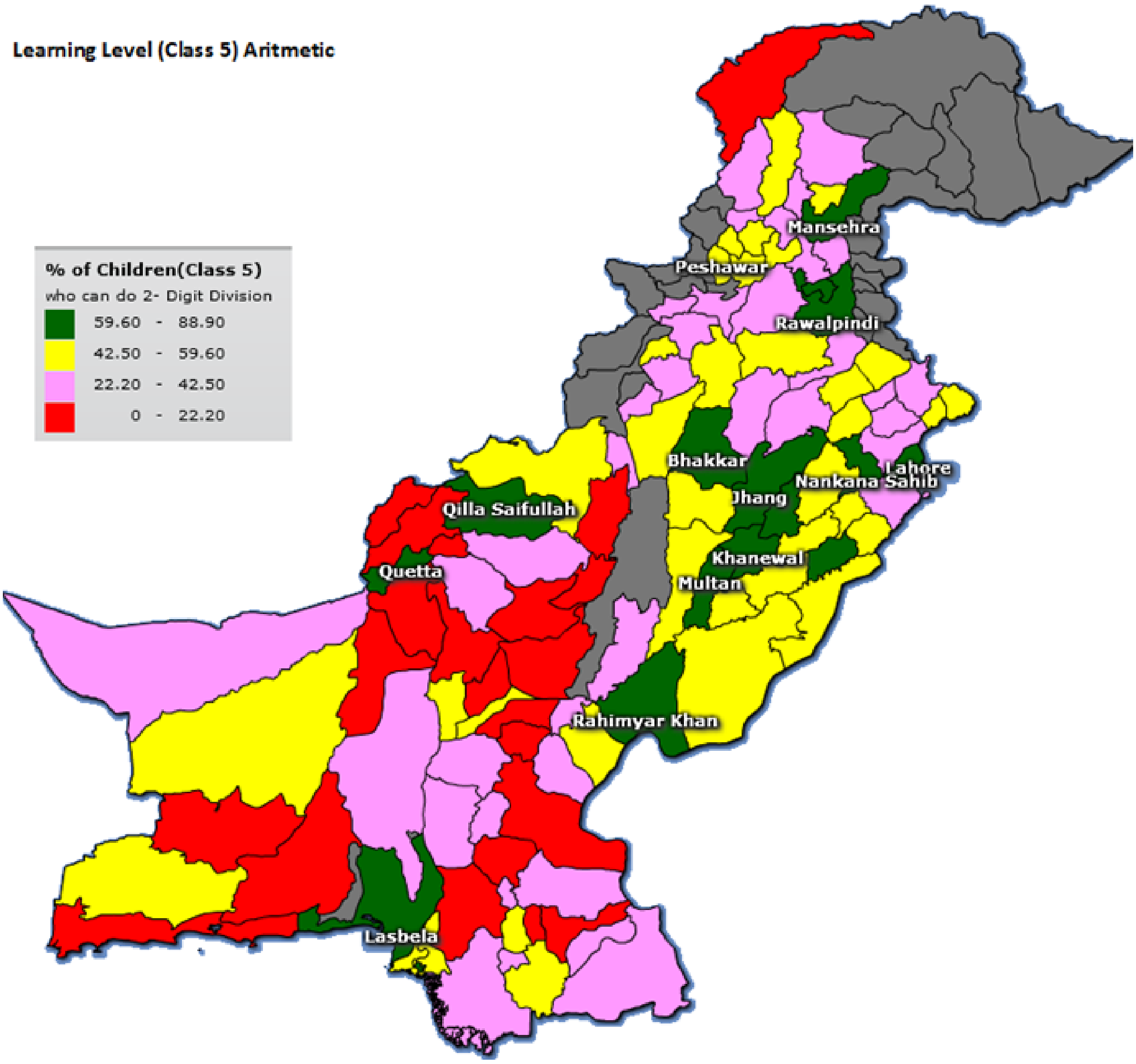


Figure 3.8: Districts' Quality of Education 2014, Pakistan
Data Sources: Learning level arithmetic data are taken from the annual status of education report (ASER) and map is prepared on the online website: http://www.targetmap.com/.

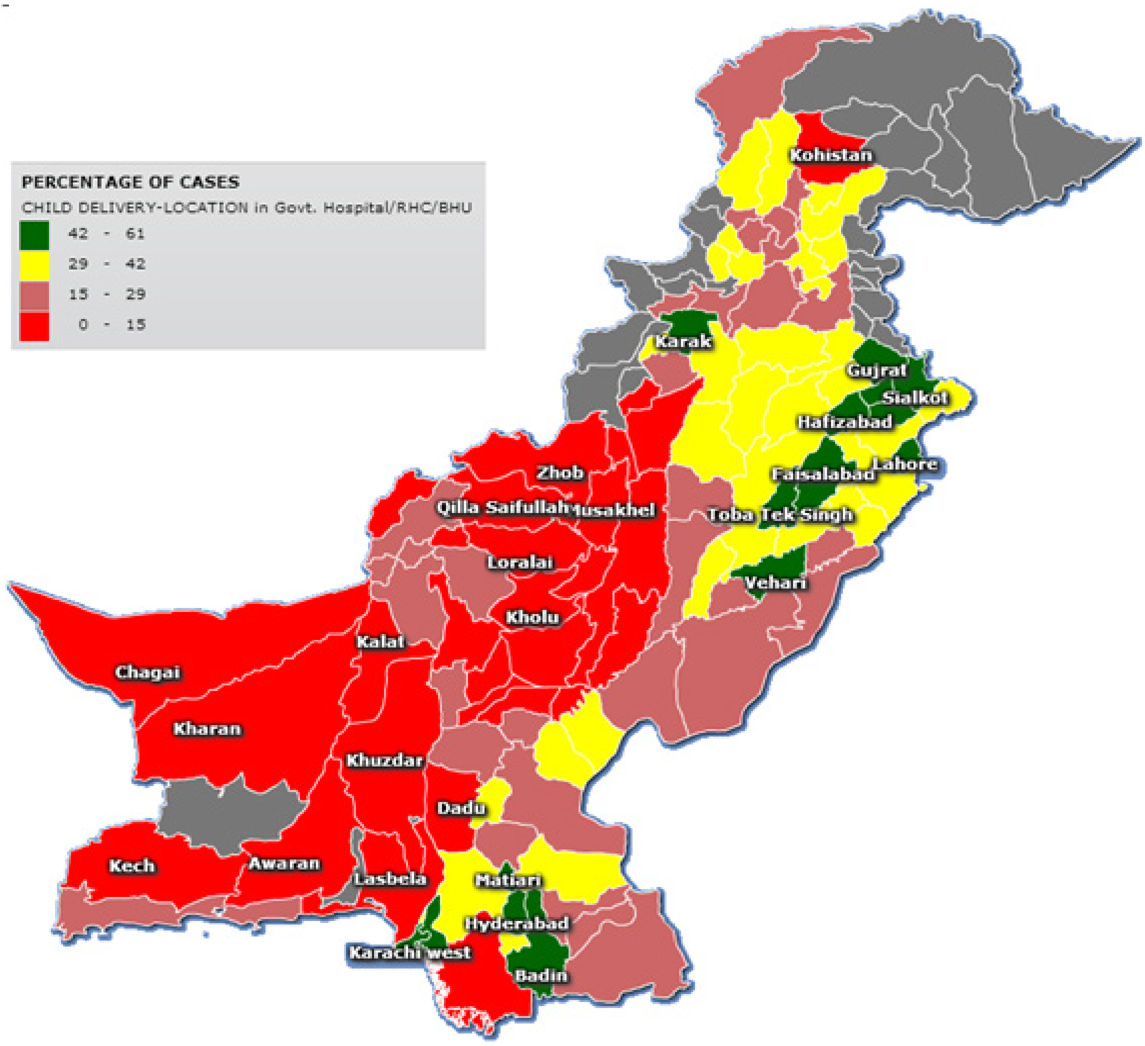


Figure 3.9: Districts' Child Delivery in Government Health Institutions 2012-013, Pakistan
Data Sources: Child delivery data are taken from the PSLM 2012-13 and Map is prepared on the website: http://www.targetmap.com/.

Chapter 4

Electricity Infrastructure and Firm productivity: Evidence from South Asia

Abstract

The purpose of this empirical chapter is to investigate the association of electricity infrastructure with firm productivity in South Asian countries based on a firm level dataset from the World Bank enterprise surveys. In the study both subjective-qualitative and subjective-quantitative measures are utilized for measuring the electricity infrastructure. Either measures, electricity as an obstacle (qualitative measure) or sales losses due to power outages (quantitative measure), are negatively correlated with firm productivity. However, the qualitative measures captured the effect more than the quantitative ones. The relationship between variable of interest and outcome variable, based on different robustness checks, appears to be robust.

Keywords: Firm Productivity, Electricity Infrastructure, South Asia

JEL classification: D24, H41, H54, O12, O13, O53, Q48

4.1 Introduction

The intriguing question of how industrial economies surpassed the developing status bounds in the policies of governments of these nations towards infrastructural development within and/or outside of their countries. Infrastructure—either in the form of physical structures or networks such as road, telecommunication, electricity, water and sanitation, etc., or in the form of social services or facilities such as health and education—are the most valuable capital stocks for an economy to generate economic growth and development. Being all other things constant, an economy grows much faster than other comparable ones if it has huge and reliable infrastructure networks because the existence of such networks makes other economic resources of the economy, human capital for instance, more productive.

Electricity, inter alia, is a key component of modern infrastructure and a prerequisite for applying advanced technologies such as ICT-based equipment. Sufficient and reliable energy supply is crucial for all types of enterprises, in both manufacturing and service sectors. Shortage or unreliable electricity supply may substantially impact the productivity of business enterprises.

In this chapter, I focus my investigation on the role of electricity infrastructure, which is to be considered a type of luxury good in most developing countries, in the production process of business enterprises in both manufacturing and service sectors in six South Asian countries[34], including Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka. In most of the developing countries, electricity still is one of the major sources of energy that both manufacturing and service firms use in their production process (Fisher-Vanden, et al. 2015). In production process, it is a type of complementary good (a necessary input in the production process) and if a firm, privately, financed the electricity by its own means (resources) this may crowd out other private investment (Reinikka and Svensson 2002). Insufficient or irregular electricity supply is one of the biggest obstacles for doing business operations in these countries because most firm operations depend, directly and indirectly, on electricity while these countries have been failed to acquire and maintain a reliable electricity infrastructure network so far. One of the reasons behind this failure is the failure

[34]Because of unrest and political instability of Afghanistan and non-availability of survey data for the Maldives, these two countries are not included in the study.

of governments of these nations in tackling inefficiencies and mismanagement of the power industry.

Although private market is considered to be a more efficient mean in supplying electricity to business community, government's monopoly over the power industry in terms of electricity generation, transmission, and distribution still exists in most of the developing countries. Even though private companies in some countries such as in India, Pakistan, etc., are allowed to do business of power generation but on the distribution side, it is the government only that can buy electricity from power-generating companies. For distribution of electricity to final users, business enterprises and household, only the government distribution companies are allowed to do business. In the context of South Asian countries, the business enterprises of transmission and distribution of electricity—with some exemption—are still mainly owned by the states (Singh, et al. 2015). As a result private electricity generation companies have very limited bargaining power over price. Furthermore, these private companies sometimes face huge debt crisis due to the delayed payments or rejection of payments by governments and they, therefore, either cut their power production units or shut down their power plants completely. Despite there exists a huge demand for electricity in developing countries, electricity markets in these countries are not growing impressively.

The purpose of this study is to add the body of literature (e.g. Carlin and Schaffer 2012; Dollar, et al. 2005)[35] in the context of South Asia by using a rich and updated dataset from the World Bank Enterprise Surveys (WBES) including all available countries, except Afghanistan, and years, 2002-2014. Unlike, Dollar, et al. (2005) and Carlin and Schaffer (2012), who studied the relationship between business environment (in general) and firm performance and growth in four Asian countries[36] and south Asian countries respectively, this study investigates, in depth, the association of electricity infrastructure only with firm output in South Asia Region (SAR)

In this study, my main hypothesis reads as follows: The inefficient and poor quality of electricity infrastructure is negatively and significantly associated to firm productivity of business enterprises in South Asian countries. To test the hypothesis, I use both subjective-qualitative and subjective-quantitative[37] measures as proxy for electricity infrastructure.

[35] For other countries or regions, see section 4.2 for more detail.

[36] Countries are: Bangladesh, China, India, and Pakistan.

[37] Contrary to the literature, I refer the subjective measures as subjective-qualitative and objective measures as subjective-quantitative measures. My argument of using them as subjective qualitative and quantitative is this: since both are subjective in nature and one is in qualitative and others are in quantitative values and there could not be objectively the records of losses due to the power outages for instance.

These measures were evaluated and provided by the top managers/owners of business enterprises, from the World Bank enterprise surveys in six South Asian countries. The nature of subjective-qualitative questions is like this: firm managers/owners were asked to evaluate the level of severity of electricity as an obstacle or rank electricity as an obstacle with reference to other obstacles. On the other hand, there are also questions that quantitatively measure the inefficiency of electricity infrastructure network such as losses due to power outages, numbers or hours of power outages, etc. One of the novelty of this study against most of the literature (see next section) is its usage of subjective-qualitative measure as a categorical variable rather than a continuous one in the regression analysis based on the following assumption: using it as a continuous variable requires a very strong assumption about the equality of distance between each interval on a Likert-type of scale which does not hold in practice (see sub-section 4.5.4 for more detail). In addition, contrary to most of the cross-country literature (see section 4.2 for literature), in which they included country dummies in cross-country regression to control for country specific shocks such as geographic, demographic, social or climate changes, etc., I argue that country dummies cannot control for regional variations within a country. Model estimations without taking into account both levels of variations, country and region[38], would be biased. In order to handle this issue, I estimate the model by including the country dummies to capture country-level variations and cluster the standard errors at the regional-level to correct the estimates for inferences.

The case of South Asia is important for the following reasons: first, the structural transformation of resources (from agriculture to manufacturing sector) in South Asia is much slower than in other regions of the world: South Asian economies are hugely dependent on agriculture and service sectors while manufacturing sector, except Bangladesh, nearly stagnated since 1980 (Nabi, et al. 2010; Sanchez-Triana, et al. 2014). Second, South Asian countries are far better than the Sub-Saharan African (SSA) countries in many socio-economic indicators (see Table 4.14), but when it comes to electricity provision to business enterprises, the South Asia region is far behind the rest of the world including the SSA (see Table 4.5).

Third, South Asia is the region in the world—according to the World Bank enterprise surveys' data (see section 4.3)—in which business enterprises are severely suffering from poor and unreliable electric supply. This finding is also reinforced by other sources such as

[38]Region, here, mean the second administrative tier of the government such as provinces in Pakistan or states in India.

media. For instance, Azad, Abdul Rasheed (September 13, 2012) in his article on the website[39] of Pakistan Defence, claimed that the 2/5 of the Pakistani textile industry moved to Bangladesh during the last five years because of the severe electricity problem they face; the industry experienced an average of six hours of power outages in a day. However, Saquib Saeed[40] (March 16, 2015) doubted that the claim of shrinking of garment industry in Pakistan and boosting of it in Bangladesh does not provide sufficient evidence for migrating the same companies from Pakistan to Bangladesh. Whatever the claim it might be, it would definitely discourage investment in these countries. So, a serious analysis is required to investigate, to what extent, if any, the quality of electricity infrastructure affects firm productivity in the South Asian context.

My findings reveal that, on average, the productivity of a firm, which is severely affected by electricity irregularities, is lower than the firm which faces no obstacle for electricity at all. However, when taking electricity obstacle as a continuous variable, the correlation remains no more significant at the conventional level. On the other hand, based on subjective-quantitative measures, the conditional correlation, if any, is very significant both statistically and economically. In this study, I also investigate the association of electricity infrastructure with firms' sales growth and employment growth as an extension to verify the absence of selection bias of firms' managers/owners in choosing of the subjective-qualitative measure on a Likert- scale. Doing a number of other robustness-check exercises, my overall findings are inconclusive. While the subjective-qualitative measures are significant in most of the cases, subjective-quantitative measures show significant correlation—though only at a small magnitude—in few cases. Therefore, policy advisers or researchers should be cautious when interpreting the results based on the World Bank enterprise surveys' subjective, qualitative and quantitative, measures only.

The structure of this chapter reads as follows: In the next section, I give a brief overview of the empirical literature related to the effects of electricity infrastructure on economic growth and development, and discuss some of the key econometric issues related to the estimation. In section 4.3, some stylized facts about electricity infrastructure in the SAR are presented. Section 4.4 describes the data used in this study, followed by theoretical and empirical models in Section 4.5. Results are scrutinized in Section 4.6 and finally some concluding remarks are included in Section 4.7.

[39]http://defence.pk/threads/over-40-percent-of-pakistans-textile-industry-has-shifted-to-bangladesh.207565/#ixzz41vvL2q81. [Accessed on April 7, 2016].

[40]http://tribune.com.pk/story/853677/myth-vs-reality-pakistani-entities-not-migrating-to-bangladesh/. [Accessed on April 7, 2016].

4.2 Empirical Literature Review

A large body of literature has been produced that discusses the relationship between public infrastructure, and firm productivity and economic growth in both developed and developing countries. The earlier works mostly focused on cross-country studies using time series macro data. The issue has been capturing a lot of attention among academia and policy making circles since the seminal work of Aschauer (1989b) validated by Munnell (1990). Many other studies such as Easterly and Rebelo (1993), Gramlich (1994), Röller and Waverman (2001), Esfahani and Ramírez (2003), and Hulten, Bennathan, and Srinivasan (2006), and Sharma and Sehgal (2010) also investigate the relationship empirically by using cross-section and time-series data.

The previous works on the effect of infrastructure on firm performance and economic growth were mainly based on either country-level indicator data or firm-level financial data. The recent availability of data from enterprise surveys in general and the World Bank enterprise surveys in particular provide better sources of information related to public infrastructure that help out in disentangling the effects of infrastructure on firm productivity (Ayyagari, Demirgüç-Kunt, and Maksimovic 2008). Studies based on firm level data have some advantages over studies using cross-countries aggregated data however, they still failed to encounter some serious econometric issues; including endogeneity, omitted variable bias and measurement errors. While studying the effects of business environment on firm performance, factors in the business environment such as infrastructure, crime, corruption, access to finance, etc., were assumed as exogenous to firm and the problem of endogeneity was neglected in many studies (e.g. Dollar, et al. 2005; Pissarides, Singer, and Svejnar 2003). In contrast, other studies tried to handle the problem of endogeneity by taking the average of country-industry (Ayyagari, et al. 2008) and city-industry (Hallward-Driemeier, Wallsten, and Xu 2006) values of business climate variables as an instrument in their studies. In addition, Aterido and Hallward-Driemeier (2010) also address the endogeneity issue by using the average value of investment climate variables at the location-sector-size response (minus firm own response) rather than absolute values of firm responses.

In order to disentangle the casual effects of general conditions of doing business, such as the availability of electrical infrastructure on firm performance and to avoid endogeneity issue, Commander and Svejnar (2011) claim that the average value of a constraint to business operations at country-industry level acts like an external shock to firm performance and individual firm response and hence cannot distort the findings when taken as country-industry average. The problem to disentangle such type of relationship based on firm perception data is that the variables may work in reverse direction. Poorly performing firms might blame external constraints for their failures while high performing firms might require more public infrastructure to cope with their burgeoning future demand and therefore complain about lack in infrastructure provision—although their current need of infrastructure is fulfilled. While using instrumental variables for the variables of interest, the average value of the obstacle at the country-industry level in this case, it is less likely that firm performance could significantly influence country-industry average (Ayyagari, et al. 2008).

Beside endogeneity issues, studies using subjective measures[41] as proxies for the business environment may be subject to measurement errors. Subjective measures usually come either as rankings or as ratings. In ranking questions managers/owners are asked to rank the top obstacles among a set of obstacles (business environment) while in rating questions they are asked to rate the particular obstacles on a Likert-type scale. On the other hand, firm managers/owners are also asked to provide objective measures[42] on their business environment; for instance, the percentage of sales lost due to the power outages in order to measure the quality of electricity infrastructure. This may also be subject to endogeneity: if electricity supply is very bad, firms will avoid using devices that rely on electricity but switch to fuel-based engines or run their own generators, etc. I test this issue in the results and discussion section.

Many studies were concerned about the measurement issue and tried different ways to mitigate it. For instance, Dinh, Mavridis, and Nguyen (2010) use ranking measures to find out the most binding constraints for business operation employing both subjective and objective measures and investigate the extent to which those binding constraints affect firm performance. In a similar fashion, Commander and Svejnar (2011) doubt that there would be a systematic bias if mangers/owners had had optimistic or pessimistic attitudes towards the binding constraints in their business environment. For instance, it might be possible that some mangers/owners choose the inner options (e.g. 2 and 3 of a 4-point Likert scale) while others choose the outer most options (e.g. 1 and 4). In order to avoid such bias, the authors estimate individual firm's value by averaging all obstacles the firms rated, then subtracted it from the firm's specific constraint and finally divided the result by the

[41] In this study, I refer it as subjective-qualitative measures
[42] In this study, I refer it as subjective-quantitative measures.

standard deviation of all obstacles reported by the firm. Another way to mitigate the measurement error is by converting ratings into relative scores and then subtracting those from the mean complaint across investment climate dimensions that isolate a likely idiosyncratic perception bias of managers/owners (Aterido and Hallward-Driemeier 2010; Hallward-Driemeier and Aterido 2009). Hallward-Driemeier and Aterido (2009), furthermore, suggest that the objective measures for infrastructure services such as losses due to power outages (for electricity) or delays in transportation are more convincing than the subjective measures.

Based on subjective as well as objectives measures, the findings related to the effects of electricity infrastructure on firm performance are mixed in literature. For instance, Ayyagari, et al. (2008) use subjective measures for overall business environment and find no significant impact of infrastructure on firm performance. In a similar study, Commander and Svejnar (2011) also do not find any significant results in the South-Eastern Europe and Common Wealth of Independent States. On the other hand, Dinh, et al. (2010) find two binding constraints on firm performance; access to finance and competition from informal sectors. Electricity, which was observed the top most severe obstacle however, shows a positive sign (and significant) with association to firm performance which is counterintuitive.

Using power losses and losses in transit as proxies for infrastructure quality, Aterido and Hallward-Driemeier (2010) find that poor infrastructure is negatively associated with firm performance in Sub-Saharan African. Their findings were further validated by Stefan (2015), who also finds that the infrastructure is a binding constraint for the African firms, but this binging constraint is significantly robust for electricity and weak for transport infrastructure. Such findings are also observed by other studies. For instance, Dinh and Clarke (2012) find that electricity as an obstacle is negatively related to firm productivity in manufacturing industries and Lacovone, Ramachandran, and Schmidt (2014), by using subjective measures for power outages, find negative and significant impact of power outages on employment growth.

This study contributes to the literature in four regards. First, it focuses on a specific but crucial part of the business environment of firms in developing countries—the supply with electricity—and uses both subjective-qualitative (most of the literature refers it as subjective) and subjective-quantitative (most of the literature refers it as objective) measures as proxies for electricity infrastructure as potential obstacle for doing business. Second, unlike other studies, this research applies a different strategy for putting the subjective-qualitative measure of electricity in the regression analysis. In many previous studies subjective-qualitative Likert-scale measures of the business environment were treated as continuous variables rather than categorical or discrete variables. This requires a strong assumption about equal distance between each interval on the rating scale, which rarely hold in practice. So in order to avoid this issue, I put the main subjective-qualitative measure of electricity infrastructure as a categorical variable rather than a continuous one. Third, this research updates the analysis of the effects of electricity infrastructure on firm performance in the South Asia Region (SAR) by using a rich and updated dataset from the World Bank enterprise surveys. However, in the Asian context, there are papers which studied the relationship but in different perspectives. For instance, Dollar, et al. (2005) used a subjective-quantitative measure, the percentage of sales lost due to power outages, as a proxy for infrastructure and found a negative relationship between infrastructure and firm performance. In a similar study, Carlin and Schaffer (2012) investigated this relationship by using subjective-qualitative measures.

Fourth, this study provides some further evidence to the policy recommendations proposed by Andrés, Biller, and Dappe (2013, 2014); Jones (2006); Singh, et al. (2015); Timilsina, et al. (2015); and Wijayatunga, Chattopadhyay, and Fernando (2015) in the context of infrastructure gaps in general and electricity infrastructure in particular in South Asian countries and the benefits these countries get as a results of cooperation and coordination for building regional infrastructure networks. My research study is different from those in the Asian context in term of methodology, regional context, using a bigger and updated sample, and in-depth analysis focusing only on electricity infrastructure.

4.3 Stylized Facts about the SAR's Electricity Infrastructure

The South Asia Region (SAR)[43] could play a huge role in growth and development of the world economy in this century if it successfully utilized its resources by building and maintaining strong cooperation among and with their neighboring nations in order to accommodate its bourgeoning population and to cash its strategic geographic location with the help of different forums, such as the South Asian Association for Regional Cooperation (SAARC) and the ASEAN[44] Regional Forum (ARF). The latter could provide guidance and recommendations for regional cooperation (Palit and Spittel 2013). But currently, the

[43]The South Asia Region includes Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan, Sri Lanka
[44]ASEAN stands for The Association of Southeast Asian Nations.

SAR is seriously facing a huge gap between the demand for and the supply of infrastructure. In order to make inclusive growth and sustainable development, the countries in the SAR need to cooperate with each other not only to solve the infrastructure problem but also cope with other related challenges such as war and terrors, environment and climate changes, and social security. One way to get rid of such type of problems is to provide health and education facilities and other amenities to the general population. Andrés, et al. (2014) suggest that the SAR can fill this huge infrastructural gap in the region, if it will invest between US$1.7 trillion and US$2.5 trillion in mixed types of infrastructure and if this investment distributes evenly over each year, the SAR will need to invest between 6.6% - 9.9% of its GDP per year based on the 2010 price till 2020 to fill the gap. In order to accommodate the bourgeoning population of the South Asian countries, the governments of these nations should also include and encourage private partners' investment on infrastructure projects, especially in electricity (generation and distribution) projects to satisfy the escalating demand for infrastructure in the region.

One of the biggest hurdles which the South Asian business community faces is the poor quality of electricity infrastructure. The condition of power outages in the SAR, according to the World Bank Enterprise Surveys (WBES) data, is worse than in any other regions of the world (see Table 4.5 on page 143). In a typical month, the average number of power outages is 25.4 in the SAR, which is much higher than the world average (6.4), followed by Middle East & North Africa (17.6), and Sub-Saharan Africa (8.3). However, the number of power outages is unevenly distributed across the countries within SAR with 75.2, 64.5, 13.8, 11.5, 8.7, 4.1, and 0.4 in Pakistan, Bangladesh, India, Afghanistan, Nepal, Sri Lanka, and Bhutan respectively. Furthermore, South Asian firms, on average, loss 10.9% of their sales due to the poor quality of electricity infrastructure that is twofold higher than the world average; Pakistan, Nepal, and Afghanistan are the most affected economies in the region with 33.8%, 17%, and 9.6% respectively.

A very big portion of public utilities in the South Asian countries is still in the government domain. And, because of incompetency and superfluous bureaucratic rules and regulation, the governments of these economies failed to provide sufficient and reliable electricity supply to the business community. A firm in the South Asia, for instance, gets an electric connection on average 55 days after when it applies for the connection whereas the world average is 32 days (see Table 4.5). However, the delays in number of days within the region vary with lowest in Bhutan (21 days) and highest in Afghanistan (111 days). In theory, all types of infrastructure are important for businesses in developing countries, but when it comes to South Asia, the importance of the electricity infrastructure is strongly stressed by enterprises (from the World Bank Enterprise Surveys' data). The power generation capacity in the region is growing overtime (see Figure 4.1 on page 151), however, the slow growth rate of investment in the power sector does not accommodate the increasing growth of electricity demand by the current and future consumers (Singh, et al. 2015). One way to tackle the problem of electricity shortages in the region is to establish a cross-border electricity trade among the South Asian and their neighboring countries. This will not only reduce the economic costs (by utilizing regional energy endowments and economies of scales due to large project investment,) but also the environmental costs (low carbon dioxide emission) of electricity production (Timilsina, et al. 2015).

4.4 Data and Descriptive Statistics

This study uses the World Bank Enterprise Surveys (WBES) data for the six South Asian countries[45]; Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka covering the years 2002 to 2014 (see Table 4.6 on page 144). The surveys cover a wider range of topics in both manufacturing and service sectors including access to finance, corruption, infrastructure, crime, competition, and performance measures. A random stratified sampling methodology, based on three criteria (firm size, business sector, and geographic location), is used for sample selection. Firm size includes small (5-19 employees), medium (20-99 employees) and large (more than 99 employees). Business sector includes manufacturing, trade, and other services. Geographic region includes cities/towns and regions within a country. Due to the uniform methodology the surveys used in all countries, the surveys' data allow to do a cross-country analysis. The mode for data collection was a face-to-face questionnaire-based interview with managers or owners of business enterprises.

In this study, I use the second tier of administrative units (e.g. states in India and provinces in Pakistan) to capture demographic, geographic and political differences within each country. The two-digit ISIC Rev.2/ ISIC Rev.3 codes[46] are used to make larger sector groups ensuring that each group consists of more than 5 percent of all observations (for groups see Table 4.7 on page 144). The monetary values of all output/inputs (sales, capital,

[45]Because of unrest and political instability of Afghanistan and unavailability of survey data for the Maldives, these two countries are not included in the study.
[46]ISIC stands for International Standard Industrial Classification.

and raw materials) variables, which were in their respective local currencies, are adjusted for inflation and converted them into a common currency (US dollar) by dividing each value with the average real effective exchange rate[47] (index based on 2005 price levels) of each country and survey year.

Some other adjustments to the data are also made. The first adjustment is related to missing values of independent variables, capital and cost of raw materials, which are not of the primary interest of this study. As many observations have missing values which would result in a reduction of the number of observations for model estimations (implying a potential bias with respect to the representativeness of the data), missing values are set to zero and two dummy variables for missing values of each variable are included in the regressions to capture the effects of this data manipulation. Since data on the size of cities (population), in which firms were operating, are not given in the early surveys (Pakistan; 2002 and 2007, Bangladesh; 2007, and India; 2005), the population values of these cities were taken from other sources. Definitions and further details about all variables are given in Table 4.4 on page 142.

Table 4.1, Table 4.8 and Table 4.9 contain the descriptive statistics of all outcome variables, control variables, and the alternative measures of the variable of interest. The number of observations per variable varies between 17,331 and 10,267 depending on the usage of different alternative measures of the electricity infrastructure. In order to get more normally distributed data, I use the logarithmic values of sales, capital, number employees, and cost of raw materials data after adding 1 to their absolute values. Mean values and standard deviations of all key variables are in their reasonably conventional range.

The average age of a firm in the sample is about 18 years. Age is measured as the difference between the year in which the firm was established and the year in which the survey was conducted. There are only 2% of firms in the sample that were owned (fully or partially) by foreign investors. Only a very small fraction of firms (1 % of firms) are partly government-owned while there are no completely state-owned enterprises in the sample.

As can be seen from Table 4.1, the mean values and variations of the alternative measures for the electricity infrastructure are in their expected range. In a typical month, a South Asian firm experiences power outage, on average, about 55 times which last, on average, 2.5 hours. The variations in both variables are huge. All those who are dealing with the findings based on these measures should be aware that there might be some very

[47]The data for the real effective exchange rates are taken from the United Nation Conference on Trade and Development (UNCTAD) statistics.

extreme outliers in the data. The descriptive statistics also reveal that South Asian firms, on average, have lost about 8% of their sales due to electricity outages during the periods 2002 to 2014.

Table 4.1: Descriptive Statistics

Variables	Observations	Mean	Standard Deviation	Min.	Max.
Log (Sales)	17,331	11.60	3.22	0.00	21.30
Sales Growth	16,033	0.075	0.206	-4.549	3.809
Employment Growth	17,076	0.032	0.119	-1.714	2.839
Log (Labor)	17,331	3.58	1.34	0.00	9.31
Log (Capital)	10,137	9.88	4.00	0.00	20.04
Log (Materials)	10,137	10.44	3.74	0.00	20.33
Foreign Share	17,331	0.02	0.13	0.00	1.00
Government Share	17,331	0.01	0.07	0.00	1.00
Large Firm	17,331	0.22	0.41	0.00	1.00
Firm Age (years)	17,331	18.24	13.74	0.00	150.00
Exporting Firm	17,331	0.19	0.40	0.00	1.00
Average Power Outage-Duration (hours)	11,518	2.48	7.77	0.00	576.00
Average Power Outage-Number (numbers)	12,062	55.45	113.24	0.00	3500.00
Log of Average Power Outage-Duration (hours)	11,518	1.01	0.50	0.00	6.36
Log of Average Power Outage-Number (numbers)	12,062	3.26	1.35	0.00	8.16
Losses due to Power Outage (% of sales)	10,263	7.91	10.76	0.00	100.00
Electricity is the biggest obstacle among a set of other obstacles[(a)]	16,388	0.23	0.42	0.00	1.00
Generator Owned	17,037	0.55	0.50	0.00	1.00

Notes: (a) Other elements in the business environment are: access to finance; access to land; business licensing and permits; corruption; courts; crime, theft and disorder; customs and trade regulations; inadequately educated workforce; labor regulations; political instability; practices of competitors in the informal sector; tax administration; tax rates; and transport. Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation i.e. log(x+1). Data Sources: World Bank enterprise surveys (various issues).

Table 4.8 and Table 4.9 contain descriptive statistics of the main variable of interest. On average, the sample is almost equally distributed among the choices on the 5-point Likert-type scale. Each category consists of about 1/5 of the sample, however, for each individual country the distribution of firms varies. For instance, 29% of firms in Pakistan report that electricity supply is a very severe obstacle, followed by Nepal (35%) and Bangladesh (22%) whereas only a very small fraction of the sample, for instance 2% in Bhutan, 7% in Sri Lanka and 10% in Indian, lie in the category. Firms that own/share a generator are less likely to face electricity is an obstacle compared to those firms that do not own/share a generator (see Table 4.8).

These statistics are also consistent with the statistics on firms' ranking of the biggest obstacles. Among all business environment variables, electricity is the biggest obstacle, mentioned by about 51% of Pakistani firms, followed by 34%, 24%, 18%, 10%, and 1% of the firms in Bangladesh, Nepal, India, Sri Lanka and Bhutan respectively (see Figure 4.2 on page 152).

4.5 Methodology

4.5.1 *Theoretical Modeling Framework*[48]

In this section, I explain the theoretical framework—formulated and applied by Carlin, et al. (2006) and also applied by Stefan (2015)—in order to postulate the relationship between physical public capital (electricity infrastructure in this case) and firm productivity. The model begins with a profit maximization function for one year period in which labor L and public capital (electricity infrastructure) Z enter into a simple Cobb-Douglas production function to produce output Y with productivity A is as follow[49]:

$$Y = F(L, Z) = AL^{\alpha}Z^{1-\alpha}, \tag{4.1}$$

where α is less than 1 and marginal productivity of labor and infrastructure are respectively given below:

$$\frac{\partial Y}{\partial L} = \alpha AL^{\alpha-1}Z^{1-\alpha},$$

and

$$\frac{\partial Y}{\partial Z} = (1-\alpha)AL^{\alpha}Z^{-\alpha}.$$

In order to finance the public infrastructure, in theory, government needs to collect taxes. Now let's assume that there is a tax on firm output (t_Y) and a tax on employed labor (t_L), the profit function at firm level would be:

$$\pi = (p - t_Y)Y - (w + t_L)L - cZ, \tag{4.2}$$

[48]The following section is taken from Carlin, Schaffer, and Seabright (2006) and Stefan (2015).
[49]Private capital, as Carlin et al. (2006) suggests, does not directly enter into the model but rather it reflects its impact via "a cash-in-advance constraint on the payment of labor". For more detail about this concept kindly refer to the original papers by Carlin et al. (2006) and Stefan (2015).

where p is the price of output and c is the cost of infrastructure. A cash-in-advance constraint[50], $\overline{M}$, limits labor cost not to exceed the initial balance amount, M. So we have:

$$M = (w + t_L)L \leq \overline{M}. \tag{4.3}$$

Similarly, firms also face a public infrastructural constraint, as the supply of infrastructure as a public good at the country level is fixed:

$$Z \leq \bar{Z}. \tag{4.4}$$

Under the constraints of cash-in-advance and infrastructure, each firm wants to maximize its profits and faces the following maximization problem:

$$\max \quad (p - t_Y)Y - (w + t_L)L - cZ,$$

$$\text{s.t.} \quad \begin{aligned} \overline{M} - (w + t_L)L &\geq 0 \\ \bar{Z} - Z &\geq 0, \end{aligned}$$

so the Lagrange function can be obtained based on the firm's profit maximization problems:

$$\psi = (p - t_Y)Y - (w + t_L)L - cZ + \delta_L(\overline{M} - (w + t_L)L) + \delta_Z(\bar{Z} - Z). \tag{4.5}$$

The optimal values of binding inputs (labor and infrastructure) are decided by the values of both constraints so in this case each firm is at a corner solution because it cannot alter the input combination even in a condition of small changes in input prices which limits us to apply a comparative static analysis. An interesting feature of Stefan's (2015) and Carlin, et al. (2006) approach is to use 'the cost of the constraints in term of forgone profit' since we know that individual firm at a corner solution does not limit the empirical analysis in an environment where the costs of constraints usually vary between countries, between regions within the same countries and between firms within the same regions. When these binding constraints hold, and by replacing π, L, and Z with π^*, L^*, and $\bar{Z}$ respectively in equation (4.2) we can rewrite the profit maximization function as follow:

$$\pi^* = (p - t_Y)(L^*, \bar{Z}) - (w + t_L)L^* - c\bar{Z}. \tag{4.6}$$

From the above equation (4.3) we know that at the maximum utilization of initial money, $M = \overline{M}$, we get:

[50]Here capital indirectly enters in the model

$$L^* = \frac{\bar{M}}{w\ + t_L}. \tag{4.7}$$

Substituting the values of L^* in the above equation (4.6), we get the final constraint profit function:

$$\pi^* = (p - t_Y)A(\frac{\bar{M}}{w\ + t_L})^{\alpha}\bar{Z}^{-1-\alpha} - \bar{M} - c\bar{Z}. \tag{4.8}$$

By derivatives of the above profit function we get the costs of constraints, labor and infrastructure respectively in the following ways:

$$\lambda_M = \frac{\partial \pi^*}{\partial \bar{M}} = \ \alpha A \frac{(p - t_Y)}{(w\ + t_L)^{\alpha}}(\frac{\bar{Z}}{\bar{M}})^{1-\alpha} - 1, \tag{4.9}$$

and

$$\lambda_Z = \frac{\partial \pi^*}{\partial \bar{Z}} = (1 - \alpha)A \frac{(p - t_Y)}{(w\ + t_L)^{\alpha}}\left(\frac{\bar{M}}{\bar{Z}}\right)^{\alpha} - c. \tag{4.10}$$

From the above equation (4.10), I can further draw two derivatives which are the main interests of this study:

$$\frac{\partial \lambda_Z}{\partial A} = (1 - \alpha)\frac{(p - t_Y)}{(w\ + t_L)^{\alpha}}\left(\frac{\bar{M}}{\bar{Z}}\right)^{\alpha} > 0, \tag{4.11}$$

and

$$\frac{\partial \lambda_Z}{\partial \bar{M}} = \alpha(1 - \alpha)A \frac{(p - t_Y)}{(w\ + t_L)^{\alpha}}\bar{M}^{\alpha-1}\ \bar{Z}^{-\alpha}\ > 0. \tag{4.12}$$

The above equations show that both productivity and initial cash bind to the costs of infrastructure. That means that as firm productivity increases, it further raises the demand for infrastructure. In this empirical study, I investigate the relationship of electricity infrastructure with firm productivity in the six South Asian countries which will be discussed in a greater detail below.

4.5.2 *Empirical Modeling Framework*

In this study, I use a standard Cobb-Douglas production function in which a firm uses its private capital, labor, and materials to produce output:

$$y_i = \alpha + \beta_1 l_i + \beta_2 k_i + \beta_3 m_i\ + u_i, \tag{4.13}$$

where *y*, *l, k,* and *m* are logs of output, labor, capital and material respectively at the firm level. A model with public capital and without private capital produces biased estimates due to omitted variable, because public capital may increase private capital productivity and output (Erden & Holcombe, 2006). In order to avoid the omitted variable bias issue, hereafter, I include private capital. To estimate TFP at the firm level, the most common approach is borrowed from the macro (country) level literature in which the residual, which could not be explained by the necessary inputs in growth model, is considered TFP. This method, if one is rather interested to estimate the determinants of TFP other than the firm conventional inputs, requires a two-stage model. In the first-stage, residuals would be obtained by running a regression on Y using all conventional inputs in equation (4.13). In the second-stage regression, the estimated residual is treated as dependent variable while all possible determinants of TFP—such as public infrastructure (e.g. electricity, roads, etc.)—are part of the independent variables (Moyo 2013; and Wang and Schmidt 2002).

There are several other firm-specific characteristics, which will be discussed later in this section, that need to be included in the model before making any inferences based on the second-stage estimates. This paper follows the model proposed by Moyo (2013) in which my variable of interest, electricity infrastructure, could be included in the model to explain the part of the residuals (TFP) which is contributed by the electricity infrastructure. Instead of estimating firm productivity (TFP) in the first-stage regression and then running a regression of the estimated TFP on electricity infrastructure and firm's other characteristic variables that produces biased results (Wang and Schmidt 2002)[51], I embed all control variables along with electricity infrastructure in one equation. So the new augmented production becomes:

$$y_{ijc} = \alpha + \beta_1 l_{ijc} + \beta_2 k_{ijc} + \beta_3 m_{ijc} + \sum_n \lambda_n X_{nijc} + \gamma ELECT_{ijc} + D_j + D_c + \varepsilon_{ijc}, \quad (4.14)$$

where lowercases of the letters, stated above, represent the logs of their respective uppercases and the subscripts i, j, and c, represent firm, industry, and country respectively. Furthermore, X is a vector that represents firm's characteristic variables such as types of ownership, age of firm, exporting firm and large firm. $ELECT$ is the electricity infrastructure and D_j and D_c are dummies for industries and countries respectively. ε_{ijc} is the idiosyncratic error term.

[51]Based on simulation experiments, the authors suggest a one-step model instead of two-step model, because in a two-step model, the least square estimates would be biased, if any of the omitted variables, the variables of interest (say electricity infrastructure in my case) for the second-step, is correlated with both the output and regressors in the first step regression. For more technical details please refer to Wang & Schmidt (2002) and Moyo (2013) original papers.

4.5.3 *Measurement of Output, Input and Control Variables*

The typical variables in production or input/output functions, at firm level, are firm's output, labor, capital, and raw materials. This study uses firm's annual sales, total number of permanent full time production and non-production workers, net book values of capital (machinery, vehicles, and equipment), and cost of raw materials as proxies for output, labor, capital and materials respectively[52]. In order to control firm specific heterogeneity firm characteristics variables are also included.

In firm characteristic variables, firm age is the first control, which is one of the important determinants of total factor productivity. However, firm age is a positive function to TFP, it is nonlinear which means that in an already established industry, a new entrant, if it survives, grows much faster than the older ones in the industry but up to a limit (frontier of knowledge) and after that limit the growth rate of the former starts declining over time. This nonlinearity of age could be captured by including the square of age term too in the model. Since the coefficient of age square term is very small (about zero), so in order to capture firm age specific differences and to avoid multi-collinearity, I include only the absolute value of age in the model specification[53].

A second control variable is foreign ownership. It is a dummy variable that takes a value of 1 if foreign share in the firm is greater than 0 and 0 otherwise. In developing countries foreign firms—because of having advantages over local firms in terms of better technological advances, management strategies, skilled labor force, training and development programs, etc.,—the former are more productive. The expected sign on the coefficient of foreign ownership is hence positive (Commander and Svejnar 2011). In a similar fashion, the third control variable is export; a dummy that takes a value 1 if the firm's export is greater than 0 and 0 otherwise. Literature also shows that exporting firms are more productive than non-exporting ones in an industry because the former need better technology and management and a skilled labor force to compete in foreign markets.

Another firm characteristic variable is firm ownership (i.e. government vs. private). Firm ownership in developing countries, whether it owns by government or by private individuals, matters the most in determining firm productivity; state-owned enterprises are less likely efficient than the private ones. The latter need to cope with dynamic markets and competitive environment by continuously investing in technologies, modern machineries, market research, etc., to survive in the industry while on the other hand, the

[52] Data are taken from the WBES. See next section for more details.
[53] I also use log of age in the model, but that does not change the estimates.

former, which are hugely subsidized by the government, do not take care of the efficiency loss.

In addition to firm's conventional input and characteristic variables, this study also includes two dummies, D_j for industry and D_c for country, in order to control for other unobservable—economic, political, and geographical—shocks at the industry and country levels. Since firms in different industries use different technologies and skilled labors in their production process, and by including the industry dummies, it may capture the heterogeneous industrial effects. In order to control for other economic, political, and geographical differences across countries in South Asia Region (SAR), I not only include the country dummies but also correct the standard errors by clustering them across regions within a country.

4.5.4 *Measurement of Variable of Interest*

In order to estimate the relationship between electricity infrastructure and firm productivity in SAR, I take a number of proxies, both subjective-qualitative and subjective-quantitative in nature, for the quantity and quality of electricity supply; including electricity shortages, losses due to electricity shortages, electricity perceived as an obstacle for business operations, etc. The World Bank enterprise surveys' dataset is the only source of information for such type of cross-country analysis. According to Allcott, Collard-Wexler, and O'Connell (2014), governments of developing countries are often reluctant to publish and provide data related to public infrastructure, and even if they do provide data, the reliability of such type of data is often very low.

The first and main measure of the electricity infrastructure, the subjective-qualitative measure, is whether electricity is perceived as an obstacle. This is measured as the level of intensity (on a 5-point Liket-type scale) of electricity as an obstacle for doing business operations as perceived by managers or owner of a firm. This measure has some advantages over others. It represents electricity supply in terms of its actual availability and quality better than other measures because if the satisfaction level of a manager or owner towards electricity supply is high then it is less likely that they complain about it. Furthermore, no matter what government invests on huge power projects; at the end of the day, it is the utilities of these projects that business community uses. And the evaluation of the electricity networks in the managers' point of view reflects the reliability of the power infrastructure. On the other hand, such types of huge investment on power projects are highly predisposition to corruption in developing countries and the monetary values that would be documented in paper, are less reliable figures for the actual investment on infrastructure or infrastructural public capital stocks.

Electricity generation capacity, which is one of the widely used proxies for power network in macro (cross-country) analysis, could not be used for micro-level analysis in developing countries for a number of reasons. First, there are huge transmission and distribution networks' losses in developing countries and by not taking these losses and inefficiencies into account in the estimation, the electricity generation capacity, itself, is a very poor proxy for electricity supply. Second, since the power generation capacity in developing countries varies very little from year to year, electricity generation capacity as an independent variable is rather meaningless. Third, at more disaggregated level (for instance at district level within a country), it is very hard to trace the electricity networks from the point of power production to the final users because distribution of electricity is not necessarily prioritized based on the condition of proximity to the production plants.

Having said that, using perception data on electricity infrastructure based on a Likert-type scale (on which 0, 1, 2, 3, and 4 represent no, minor, moderate, major, and very-severe obstacle respectively) is not free from limitations. The main limitation of this measure is its categorical nature. Though such type of information has widely been used as a continuous variable, this requires a strong assumption of the equality of distance between each two consecutive points on the scale. In other words, the distance between points 1 and 2 is equal to the distance between points 3 and 4 and if this assumption were not satisfied, then the point average estimates based on the assumption would be misleading. In order to avoid the issue, the subjective-qualitative measure should be used as an indicator variable by taking the no-obstacle category as a reference group, and then compare the estimates of dummies for each category with respect to the reference group. The expected sign of the coefficient on each category would be negative and the magnitudes of their coefficients, compared to reference group, increase as the level of the severity of electricity obstacle increases. For further extension and robustness checks, I also use another perception-based measure for the electricity infrastructure in which managers or owners of a firm were asked to rank the electricity as an obstacle among a set of other obstacles[54] in the business environment. This is an indicator variable which holds a value of 1 if the firm identified electricity is the biggest obstacle among other obstacles and 0 otherwise.

[54]Other elements in the business environment are: access to finance; access to land; business licensing and permits; corruption; courts; crime, theft and disorder; customs and trade regulations; inadequately educated workforce; labor regulations; political instability; practices of competitors in the informal sector; tax administration; tax rates; and transport.

There are two further issues related to subjective-qualitative measures identified by Stefan (2015): First, manger responses may be correlated with firm's characteristics. For instance, managers of a poorly performing firm might blame external factors (e.g. power failure) for the firm's inefficiency rather than internal ones (e.g. management failure). In such case, the estimates based on subjective-qualitative measures as a proxy for electricity infrastructure would be biased downwardly. Second, it is quite challenging to test whether these subjective-qualitative measures, being used as a proxy for the electricity supply for a business firm, truly represent the actual electricity supply by the public grids, otherwise the relationship would be spurious.

In order to deal with the first problem, a new variable—firm performance—enters into the model by interacting with the variable of interest; the former is an indicator variable that holds a value of 1 for high performing firm and 0 otherwise. The indicator variable is calculated by the following way. First a ratio of firm's current sales to the sales three years ago is calculated and then based on the median ratio, the sample is divided into two halves. The first half contains high performing firms (their ratios are greater than or equal to the middle ratio) and other half contains low performing firms (their ratios are less than the middle ratio). For handling the second problem, I expand the analysis to subjective-quantitative (literature refer it as objective) measures, in which, I also use a set of subjective-quantitative measures in the model. This set includes the number of power outage in a month, the duration of power outage (in hours), and the percentage of sales lost due to power outage. For more detail, see Table 4.4 on page 142.

4.6 Results and Discussion

4.6.1 *Main Results*

The results from the augmented production function (4.14) are given in Table 4.2. I begin my empirical analysis by running the first baseline regression that contains only the common determinants of total factor productivity (labor, capital, and materials), a set of firm characteristics (age, foreign ownership, government ownership, large firm, and exporting firm), and finally a set of dummy variables (country, industry, and the size of the firm's located-city) to capture unobserved heterogeneity across countries, industries and cities. The coefficients of all private inputs; labor, capital, and materials with their magnitudes; 0.712, 0.033, and 0.448 respectively, are in their anticipated range and statistically significant at the 1% level.

Firm characteristics also show expected signs and they are statistically significant except government ownership. Unlike Commander and Svejnar (2011) findings, the negative sign of the government ownership indicator in this study, however insignificant, hints us somehow toward governments' mismanagement and inefficiencies in business operation in the South Asian countries. With reference to the coefficient of firm age[55], it is positive and statistically highly significant, indicating that a more experienced firm performs better in terms of TFP than a less experienced one and vice versa. In other words, a firm that has been in business operation for 100 years, ceteris paribus, seemingly generates an additional of about 0.30% of revenue than a new entrant in the industry. With respect to the coefficients of firms' others characteristics, whether a firm has foreign stakes, whether it is a large one, and whether it is an exporting firm, they are all positive and statistically significant at the 5% level. There are, of course, many reasons for their well performance in terms of TFP, but here I mention some of the most obvious ones. Foreign firms might have more sophisticated production technology, larger firms might have strengths of having better and more production capabilities and can utilize economies of scale, and finally exporting firms might have better production, marketing and cost reduction strategies in coping with international market competition and by learning from abroad. These findings are consistent with the study by Dinh, et al. (2010). The reported standard errors in column 1 of Table 4.2 are the OLS robust standard errors which are corrected for heteroscedasticity, but when run the regression by the OLS with default standard errors, there is no significant difference between the p-values produced by the both strategies; however I will use a more conservative approach for the error structure when the variable of interest, electricity infrastructure, enters into the full model.

[55]In a separate regression, I included the squared term of firm age but it was statistically insignificant, so to avoid multicollinearity issue, the squared term is not included in the specification.

Table 4.2: Regression Analysis of Electricity Infrastructure on Firm Sales

Dependent variable: Log (Sales)	Baseline	Overall Sample	Overall Sample	Own/Share Generator	No Generator
	(1)	(2)	(3)	(4)	(5)
Log (Labor)	0.712***	0.712***	0.712***	0.667***	0.700***
	(0.017)	(0.017)	(0.040)	(0.033)	(0.065)
Log (Capital)	0.033**	0.032***	0.032**	0.045***	0.026
	(0.006)	(0.006)	(0.015)	(0.013)	(0.020)
Log (Material)	0.448***	0.449***	0.449***	0.490***	0.414***
	(0.012)	(0.012)	(0.039)	(0.036)	(0.050)
Foreign	0.301***	0.296***	0.296***	0.384***	0.012
	(0.098)	(0.098)	(0.082)	(0.122)	(0.161)
Government	-0.027	-0.029	-0.029	0.041	-0.050
	(0.168)	(0.168)	(0.173)	(0.215)	(0.263)
Large	-0.107**	-0.106***	-0.106	-0.130**	-0.008
	(0.039)	(0.039)	(0.064)	(0.049)	(0.126)
Firm Age	0.003***	0.002***	0.002**	0.001	0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Export	0.126***	0.127***	0.127**	0.053	0.269***
	(0.027)	(0.027)	(0.051)	(0.051)	(0.085)
Reference group (No Obstacle)					
Minor Obstacle		-0.105***	-0.105**	-0.212***	-0.078
		(0.027)	(0.041)	(0.062)	(0.059)
Moderate Obstacle		-0.101***	-0.101*	-0.230***	-0.114
		(0.026)	(0.054)	(0.065)	(0.073)
Major Obstacle		-0.077***	-0.077	-0.245***	-0.059
		(0.028)	(0.050)	(0.062)	(0.060)
Very Severe Obstacle		- 0.109***	- 0.109*	-0.286***	-0.058
		(0.030)	(0.063)	(0.088)	(0.062)
Constant	4.265***	4.343***	4.343***	4.162***	4.717***
	(0.109)	(0.110)	(0.283)	(0.317)	(0.336)
Country-Year Dummy	Yes	Yes	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes	Yes	Yes
City-Size Dummy	Yes	Yes	Yes	Yes	Yes
Observations	17,331	17,331	17,331	9,287	7,750
Clusters	-	-	78	78	77
R^2	0.8814	0.8815	0.8815	0.8741	0.9026

Notes: Model estimated by OLS. Robust standard errors in (1) and (2) and clustered at the region level in (3) to (5) are in parentheses. Region is the second tier of administrative units (e.g. states in India and provinces in Pakistan). ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: log (sales). Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Two dummies for missing values of capital and material are included. Data Source: WBES (various issues).

In a second step, I extend the model by including electricity infrastructure, which is the variable of interest, to estimate the conditional correlation of electricity infrastructure with firm productivity. I begin to test this hypothesis by including a subjective-qualitative measure for electricity infrastructure in the main regression, assuming that it measures to both the quantity and the quality of available electricity networks for the city/town in which the firm is operating. This subjective-qualitative measure enters into the model as a set of four dummies for the Likert-scale values ‘very severe obstacle’, ‘major obstacle’, ‘moderate obstacle’, and ‘minor obstacle’. Firms that reported that electricity is not an obstacle for doing business represent the reference group. The coefficient of each group is interpreted with respect to the reference group. The OLS estimates with robust standard errors are reported in column 2 of Table 4.2. Compared to the base category (in which electricity is not an obstacle), the OLS estimates of all four categories are negative and significant at the 1% level. Interestingly, the magnitudes on the coefficients of all categories are very similar indicating that the electricity infrastructure based on managers’ perception is closely related to firms' output. On average the percentage of sales volume of firms which complain that electricity is an obstacle is around 8-12%[56] lower than for the base category. In addition, the F-statistic = 0.56 with p-value = 0.6428 of the null hypothesis—whether the coefficients of all four non-reference categories (minor, moderate, major and very severe) are equal—indicates that the coefficients are statistically not different across all four categories. From a policy recommendation perspective, it is vital to note that the quality of electricity supply that satisfy the managers/owners of a firm matters. The absence of electricity supply quality negatively associated with firm sales independent from the level of severity of supply shortage the firm has identify as an obstacle.

The OLS estimation based on default standard errors assumes that the errors are independent and identically distributed (iid), and when errors are corrected for heteroscedasticity, the second half of the assumption, i.e. identically distributed, is relaxed but the errors which are serially correlated within a region in this case, could not be relaxed. In my sample the assumption of intra-cluster zero correlation is inappropriate because errors might be correlated within region of each country and the interpretation of the findings based on only OLS default errors or robust standard errors are misleading. I tackle this problem by using a more conservative method suggested by Cameron and

[56] $100*(e^{0.077}-1) = 8$ and $100*(e^{0.109}-1) = 11.5$

Miller (2015). According to them, errors might be correlated within a region because of some unobservable shocks at regional level that affect all firms in the region in a similar way, and therefore both the usual default and heteroskedastic-robust standard errors do not take into account the within region errors. In this case the additional piece of data in the region do not provide any additional information which cause high t-statistics and compressed confidence intervals of point estimates, leading the null hypothesis to be failed to reject. Making my findings more meaningful, I assume that errors are independent over individual firms but correlated across regions. I run regressions by clustering standard errors by regions throughout the rest of the study, except mentioned otherwise. Column 3 of Table 4.2 above shows the regression results with regionally clustered standard errors. With exception of private capital and firm size, the significance levels of the coefficients of the control variables remain almost the same. However, this exercise does affect the significance levels of the variables of interest. The interpretation of the point estimates of all categories, except major-category, compared with no-obstacle category still remains unchanged at the 10% significance level.

The relationship of electricity infrastructure with firm productivity would be different for a firm if it had an internal capacity to substitute the publicly provided electrical power by its own power generation at the plant level. Such type of firms which could produce power at their own disposal may or may not be better off than the rest of the firms because despite the power interruption they could continue their production and diminish their fixed production cost by employing idle capital and labor. However, the cost of fuels for power generation at the plant level may be higher than the cost for purchasing electricity in the short run. And in the long run, investment in power generation at plant level may crowd out other private and potentially more productive investment. To test the question whether or not a firm would be better off if it could substitute the electricity which it gets from public grids with their own power generation at the plant level to fulfill the power demand, I split the sample into two based on whether a firm owns/shares a generator. Columns (4) and (5) show the results. Doing this exercise, some of the control variables show slightly different values though, the overall conditional correlation between the main variable of interest and firm productivity remains unchanged. For instance, a noticeable change of the sign on the coefficient of foreign ownership in column (5) is observed, but it is statistically insignificant. One obvious reason for this counterintuitive sign would seem to be that there are very few foreign firms in the sample which do not own/share generators, and therefore with a slight variation in this dummy variable, the regression couldn't capture its true relationship with firm productivity.

Now turning the attention to the variable of interest, once again negative coefficients for all four groups in the both specifications, (4) and (5), are observed. However, the coefficients in regression (4) are statistically significant and increasing in the ascending order of their ratings on the scale. Compared with the base category, these findings indicate that both types of firms, with and without own/shared generator facilities, seemingly suffer by power outages from public grid. But the level of this suffrage is much higher for firms that own/share generators compared to those which do not. More precisely, in the presence of a generator at plant level, the percentage of sales of the firms, which identify electricity is a severe obstacle, is ostensibly 33.11%[57] lower than those which report electricity is not an obstacle.

There may be several reasons why also firms with own or share generator facilities also report obstacles in electricity supply which significantly affects their productivity. First, electricity generation and distribution at government disposal in developing countries costs enterprises much less than their own power production. Second, it is somehow related to the first one that the firms that have own power generation capacity—try to continue their business operations in anticipation of reducing their fixed cost by employing their idle capital and labor—suffer the most due to high cost of fuels which they use for their own power generation. Third, there might be some reverse causation. For instance, nonproductive-firms, which own/share generators mention electricity is the biggest obstacle for their business operations, simply blame on government for their inefficiencies and mismanagement. To test this selection bias and reverse causality, I test my finding based on different checks in the robustness section.

4.6.2 *Robustness Checks and Extensions*

In this empirical analysis of the conditional correlation between electricity infrastructure and firm productivity in the South Asian economies, I so far used a subjective-qualitative measure as a proxy for electricity infrastructure which may be criticized for its potential measurement error and of not representing the actual public capital stocks of electricity infrastructure in the economies. In this subjective-qualitative measure, firm managers/owners were asked to rank their perception of the role of electricity supply for their business operations on a 5-point Likert scale. If this does not fully represent the actual

[57] $100*(e^{0.286}-1) = 33.11$

electricity infrastructure (capital stock) for business firms, then the measurement errors cause endogeneity problem which tends to produce biased estimates. To lessen the severity of the problem, I test my findings based on different alternative measures and methodologies.

First, I use a different indicator variable that holds a value of 1 if a firm mentioned that electricity was the biggest obstacle for doing business compared to a set of elements in the business environment, which were identified by the World Bank in the enterprise surveys, and 0 otherwise. Column 1 of Table 4.3 reports the results. The coefficient on the variable of interest is still negative but statistically insignificant at the conventional levels. This negative coefficient on the electricity infrastructure indicates that productivity of firms that marked electricity was the biggest obstacle for their business operation is lower than the productivity of the rest of firms in the South Asian countries. This exercise further strengthens the evidence that the business suffrage is conditionally correlated with the electric supply failure, as the managers/owners of the firms in developing countries perceived.

For the sake of meaningful findings and interpretations, and rather relying on the subjective-qualitative measures only, I run two regressions by including alternative subjective-quantitative measures as proxies for the electricity infrastructure. The first quantitative measure is the percentage of sales losses due to power outages that enters into the production function as a proxy for the electricity infrastructure. Literature suggest that the quantitative measures from the World Bank enterprise surveys data sets measure the electricity infrastructure in terms of quantity (supply) and quality in a more meaningful way than the qualitative measures. The results from regression (2) in Table 4.3 indicate that, on average, the percentage of total annual sales losses due to power outages and the annual sales of a firm are negatively correlated and the conditional correlation, however economically small (i.e. 0.5%), is statistically significant at the 5% level.

For the second regression, I include two subjective-quantitative measures in the model, logs of the average number of power outages in a typical month and the average duration of power outages in hours. In specification (3) of Table 4.3, I report the findings which show a negative correlation between the variables of interest and firm output, however the coefficients of both variables are statistically insignificant. This estimates is in the line of

(Moyo 2013) study. Furthermore, running two separate regressions[58] for these subjective-quantitative measures makes no significant difference in the findings.

In addition to the above exercises that provide some evidence against endogeneity issues and measurement errors, another criticism may stress reverse causation between the variables of interest and firms' sales. To put it differently, the direction of this correlation may flow from the left-hand side variable to the variables of interest, but not the other way round. This might sometimes happen not because of the existence of true relationship between the two variables, but rather because higher performing firms demand more electricity units to fulfill their production targets under a tight electricity supply constraint from public grids, and therefore the managers/owners of such firms identify overly the severity-level of the obstacle for electricity as a results of their frustrations. If there exists a selection bias in this way, then there must be some systematic differences between high and low performing firms in selecting the options on the scale. To test this issue, a dummy variable for high performing firms (HPF) is created based on the median ratio of firms' current sales to theirs' sales three years ago. The dummy takes a value of 1 if the ratio for a firm is greater than or equal to the median ratio of the sample and 0 otherwise. It enters into the regression model by interacting with the categorical variable, the variable of interest. The base category consists of those low performing firms (LPF) that identified electricity is not an obstacle for their business operations. The estimated coefficients of the regression are given in column 1 of Table 4.10 on page 146. It shows that, on average, the high-performing-firms' productivity is greater than the low-performing-firms' productivity at the 10% confidence level, once I control for all other background variables. Furthermore, the interaction terms also indicate that there is no significance difference between high and low performing firms in selection of the options on the rating scale of electricity as an obstacle. Interestingly, the negative signs of all coefficients on the categories remain almost same. Compared with the reference group, the coefficients on moderate and very-severe groups are statistically significant and robust, indicating that there is no strong evidence for selection bias in term of firm performance.

[58]Results are not included in the Tables.

Table 4.3: Robustness Checks—Alternative Qualitative and Quantitative Measures

Dependent Variable: Log (Sales)	Alternative Measure -I	Alternative Measure -II	Alternative Measure -III
	(1)	(2)	(3)
Log (Labor)	0.726 ***	0.703***	0.740***
	(0.041)	(0.050)	(0.042)
Log (Capital)	0.031**	0.038*	0.037**
	(0.015)	(0.020)	(0.018)
Log (Material)	0.440***	0.473***	0.428***
	(0.039)	(0.060)	(0.048)
Foreign	0.297***	0.205*	0.240*
	(0.087)	(0.119)	(0.124)
Government	0.010	-0.068	0.084
	(0.184)	(0.260)	(0.259)
Large	-0.105	-0.196***	-0.184***
	(0.067)	(0.063)	(0.064)
Firm Age	0.003**	0.003***	0.002*
	(0.001)	(0.001)	(0.001)
Export	0.123	0.083	0.117**
	(0.056)	(0.058)	(0.057)
Electricity: the biggest obstacle	-0.064		
	(0.044)		
Electricity: % of sales lost due to power outage		-0.005**	
		(0.002)	
Log (average number of power outage			-0.011
per month (in numbers))			(0.022)
Log (average duration of power outage (in hours))			-0.035
			(0.044)
Constant	4.483***	3.969***	4.532***
	(0.287)	(0.393)	(0.339)
Country-Year Dummy	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes
City-Size Dummy	Yes	Yes	Yes
Observation	16,388	10,267	11,489
Clusters	74	73	74
R^2	0.8835	0.9044	0.8870

Notes: Model estimated by OLS. Robust standard errors in parentheses, clustered at the region level. Region is the second tier of administrative units (e.g. states in India and provinces in Pakistan). ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: log (sales). Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Two dummies for missing values of capital and material are included. Data Source: WBES (various issues).

Likewise, one can also argue that firm performance is not the only driving force behind the selection bias, but it may be the level of energy consumption. High electricity-intensive firms (HEIF) could be affected more by power outages than low electricity-intensive firms (LEIF), and therefore managers/owners' perception on Likert-type scale (electricity as an obstacle) toward the electricity supply would be different for these two types of firms. To tease out the effect, I create another dummy variable to capture the effects of these two types of firms. Due to data limitation in the level of electricity consumption at the firm level, I mimic the previous procedure to create a dummy, HEIF, for this variable as well. First, I estimate the ratio of total electricity costs to total sales for each firm and then set values of the dummy to 1 for firms with a ratio are greater than or equal to the median ratio of the sample (calculated in the same way) and 0 otherwise. In regression 2 of Table 4.10 on page 146, this dummy variable enters along with the categorical variable of interest and their interaction terms. My findings from this exercise reveal that the productivity of low electric-intensive firms is significantly higher (with 65%, exp(0.499) = 1.647) than their counterparts. Also, there exists no significance difference between high and low electricity-intensive firms in selection of the severity-level of the obstacle for electricity infrastructure. With reference of my variable of interest, the coefficients on all four categories are still negative and the conventional significant levels, this time, hold for the minor, moderate, and major categories with the p-values of 0.013, 0.021, and 0.096 respectively. To this point, there is sufficient evidence to believe that electricity infrastructure, observed by the final user of the infrastructure (the firms), is conditionally correlated with firm productivity in the six South Asian nations.

In extension of the above exercises, I extend the analysis by replacing the dependent variable, log of sales, to sales growth and employment growth. Both growth rates are calculated by the formula[59] which is taken from the study by Dinh, et al. (2010). The conventional inputs are excluded from the model. Table 4.13 on page 149 represents the estimated results, which is a replication exercise of the column (3) to (5) in Table 4.2. Columns (1) to (3) and (4) to (6) show the estimates of the models when the dependent variable is sales growth and employment growth, respectively. These findings also show that there exists no significance association between firm performance (employment growth and sales growth) and mangers/owners' perception toward the electricity as an

[59]Sales growth = $(ln\_sales_{i,t} - ln\_sales_{i,t-3})/3$ and Employment growth = $(ln\_worker_{i,t} - ln\_worker_{i,t-3})/3$.

obstacle which further supports the evidence that there is no systematic bias in perception data.

In the previous main analysis, I used a pooled regression model with country and industry fixed effects to test the hypothesis based on the assumption that the model contains, on average, the same coefficients for all explanatory variables for individual country or industry. But if the true effect of electricity supply on firm productivity with reference to country or industry context varies, then the estimated coefficients using pooling data would be biased (Moyo 2013). To test whether the shortage of electricity supply or losses due to electricity shortages affect firm productivity at country or industry level differently, I run separate regressions for each country and industry. The results are shown in Table 4.11 and Table 4.12 respectively. Regression estimates for Indian firms based on subjective-qualitative measure show that the electricity supply shortage is negatively associated with firm productivity while for other countries, the coefficients of interest, though negative in most of the cases, are insignificant at the conventional significance levels. However, the coefficients on the subjective-quantitative measure, i.e. the % of sales loss due to power outage, are negative in most of the case, but significant only for three south Asian countries; Pakistan, India, and Sri Lanka.

Now I run nine separate regressions for different industrial groups[60]. Results are shown in Table 4.12 on page 148. In industries wise estimation, both qualitative and quantitative measures show negative relationship with firm sales in many industries but the negative conditional correlation significantly holds only for the following group of industries: a) manufacturing of wood and wood products including furniture (33) and manufacture of paper and paper products, printing and publishing (34); b) manufacture of non-metallic mineral products (36) and basic metal industries (37); c) manufacture of fabricated metal products, machinery and equipment (38); and d) services including; retail, wholesale, services of motor vehicles, hotel and restaurants (50). On the other hand, the negative relationship between the percentage of sales loss due to power outage (subjective-quantitative measure) and firm sales holds only for the two aforementioned group of industries; a) and b).

Finally, from the above analysis, we observe that private capital in some cases is very sensitive to the inclusion of other control variables. This might be because of the existence of a high correlation between variables; private capital and cost of materials in the sample.

[60]Groups, by using ISIC Rev. 2 codes, are made in such a way that each group contains a sufficient number of observations for the analysis. For industries and their codes refer Table 4.7, in Appendix D.

In order to test whether the coefficients on the variable of interest are sensitive to private capital, I run a regression by excluding private capital. The results are shown in the last column of Table 4.10 on page 146[61]. Once again, the magnitudes with negative sign and the significance level of the coefficients on all categories, i.e. variable of interest, remain almost unchanged. Actually this time the significance level of the moderate category is even improved (from 10% to the 5% level).

Based on these robustness checks, my findings reveal sufficient evidence that the complaints about the electricity infrastructure posed by managers/owners of firms represent the actual electricity infrastructure stocks in term of quantity (supply) or quality. Firms mentioning that the electricity infrastructure is not an obstacle in doing business operations are performing significantly better than those firms that complain about the electricity infrastructure severely.

4.7 Conclusion

The importance of manufacturing and service sectors in South Asia has been widely recognized for the purpose of generating economic growth and development, creating employment opportunities, and reducing poverty. For these sectors, access to quality public infrastructure is one of the prerequisite for doing business in the region. Among other public infrastructure networks and facilities, electricity is the most crucial input identified by business enterprises in South Asia Region (SAR). Current literature to date has tried to study the relationship between electricity infrastructure and firm productivity by implementing different methodologies in order to avoid econometric issues such as endogeneity, measurement errors, and omitted variable biases. These studies used data at the levels, cross-country aggregated (macro) level and firm (micro) level.

In this study, I have investigated how electricity infrastructure and total factor productivity and output of business enterprise are correlated, conditional on various controls, in the six South Asian countries. The study has utilized both types of measures, subjective-qualitative and subjective-quantitative, which are taken from the World Bank enterprise surveys. Unlike other papers, this study has used the subjective-qualitative measure as a categorical variable rather than a continuous one, the latter requiring a strong assumption of equal distance among all intervals on the perception-based Likert-type scale.

[61]Coefficients on all conventional are excluded from the table for space purpose.

The findings, based on different alternative measures, indicate that there exists a negative relationship between the quality of electricity infrastructure and firm productivity in the South Asian countries. Subjective-qualitative measures seem to capture this conditional correlation better than the subjective-quantitative ones. For instance, the average percentage of sales for those firms that mentioned electricity as an obstacle (qualitative measure)—irrespective of the severity levels—was higher than those which mention electricity is not an obstacle by 11.5 percent. On the other hand, the percentage of total annual sales lost due to power outages (quantitative measure) and firms' annual sales are negatively correlated and statistically significant at the 5% level.

These findings, based on different exercises for robustness checks and further model extensions, reveal that the results are not prone to a likely systematic bias by mangers'/owners' perception towards the electricity infrastructure, because the correlation between sales/employment growth and the qualitative measure (the severity level of obstacle for the electricity infrastructure, as managers/owners of firms identified on a 5-point Likert-type scale) is insignificant.

My overall findings reveal that even though there exists evidence for a negative relationship between electricity as an obstacle and firm productivity, the relationship, economically and/or statistically, is not too strong. I conclude that the World Bank enterprise surveys—the only source of data for cross country analysis in South Asia, provide the basic insights about how business enterprises face infrastructure obstacle in their daily business—could not be used for generalizing the results for policy recommendations. In this context, further research is needed by looking into the causal relationship of electricity infrastructure on firm productivity by employing external data sources on the quantity and quality of electricity infrastructure which could better measure the actual infrastructure capital stocks in the economy.

4.8 Appendix to Chapter 4

4.8.1 *Tables*

Table 4.4: Variables Definition

Main variable	**Description**
Log (Sales)	Log of annual sales adjusted with the real effective exchange rate (Index based on 2005 price levels) from local currency.
Log (Labor)	Log of total number of permanent full time employees (both production and non-production) at the end of fiscal year.
Log (Capital)	Log of net book value of capital (machinery, vehicles, and equipment) adjusted with the real effective exchange rate (Index based on 2005 price levels) from local currency.
Log (Materials)	Log of cost of raw materials adjusted with the real effective exchange rate (Index based on 2005 price levels) from local currency.
Firm Age	The difference of the year when firm started its operation and the given year of survey was conducted.
Foreign	A dummy variable holds a value 1 if a firm does have at least some foreign share.
Government	A dummy variable holds a value 1 if a firm does have at least some government share.
Export	A dummy variable, holds a value 1 if more than one percentage of a firm's sales come directly or indirectly (selling other domestic exporters) from exports.
Country region	An indicator variable represents the second level of administrative political boundaries in each country.
Number of Power Outages	Average number of power outage per month (in numbers).
Duration of Power Outages	Average duration of power outage in hour.
Power Loss (% of Sales)	Losses due to power outage (as percentage of sales).
Electricity Obstacle	An indicator variable which firms recorded their level of electricity as an obstacle while doing business in the region based on a ranking scale of 0 to 4; where 0 means electricity is not an obstacle while 4 means it is a very severer obstacle.
Own Generator	A dummy variable holds a value 1 if firm owns or shares a generator, 0 otherwise.
Firm performance	Ratio of current sales to the sales of three years ago. A High performing firm is one that has a value of equal to or greater than the median value of the ratio.
Electric-Intensive firm	Absolute Electricity Intensity (AEI) is a ratio which is: Electricity cost/total sales. Both variables are adjusted by the real effective exchange rate (Index based on 2005 price levels). An Electric-intensive firm is one that has a value of equal to or greater than the median value of the ratio (absolute electricity intensity).

Source: Word Bank Enterprise Surveys' questionnaire.

Table 4.5: Comparative Statistics for Electricity Infrastructure: South Asia and World

Country/Region	Number of Power Outages per Month (numbers)	Average Duration of Power Outage (hours)	Annual Sales Losses due to Power Outages (%)	Firms own/share a generator (%)	Electricity from own Generator (%)	Days to Obtain an Electrical Connection (days)	Firms identified Electricity as a Major Constraint (%)
World	**6.4**	**4.5**	**4.7**	**32.4**	**19.7**	**32.0**	**32.0**
East Asia & Pacific	4.4	6.4	1.8	36.4	28.6	47.7	17.2
Eastern Europe & Central Asia	2.0	3.5	2.9	21.3	11.9	26.6	17.9
High income non-OECD	1.3	2.8	0.8	22.4	9.5	30.7	31.1
High income: OECD	0.4	2.9	0.9	13.1	4.7	42.7	21.8
Latin America & Caribbean	2.8	2.6	3.4	27.6	20.8	22.0	37.5
Sub-Saharan Africa	8.3	5.8	8.6	49.8	25.3	29.3	37.9
Middle East & North Africa	17.6	9.7	6.9	41.0	34.8	41.1	40.5
South Asia	**25.4**	**5.3**	**10.9**	**45.4**	**24.4**	**55.1**	**46.1**
Afghanistan (2014)	11.5	3.8	9.6	48.0	38.3	111.3	65.8
Bangladesh (2013)	64.5	1.2	5.5	62.8	26.1	84.7	52.0
Bhutan (2015)	0.4	8.1	3.7	9.5	10.2	21.3	14.1
India (2014)	13.8	2.0	3.7	46.5	8.8	21.9	21.3
Nepal (2013)	8.7	3.6	17.0	50.5	41.3	21.3	68.8
Pakistan (2013)	75.2	16.9	33.8	65.4	41.4	82.8	75.3
Sri Lanka (2011)	4.1	1.5	3.0	35.1	4.8	42.4	25.6

Source: Word Bank Enterprise Surveys (various issues).

Table 4.6: World Bank Enterprise Surveys' Countries and Years

Country	Year
Bangladesh	2007, 2013
Bhutan	2009
India	2005, 2013
Nepal	2009, 2013
Sri Lanka	2011
Pakistan	2002, 2007, 2013

Source: WBES (various issues).

Table 4.7: Categories of Manufacturing and Service Industries

Manufacturing	Code
Manufacture of Food, Beverages and Tobacco	31
Textile, Wearing Apparel and Leather Industries	32
Manufacture of Wood and Wood Products, Including Furniture + Manufacture of Paper and Paper Products, Printing and Publishing (34)	33
Manufacture of Chemicals and Chemical, Petroleum, Coal, Rubber and Plastic Products	35
Manufacture of Non-Metallic Mineral Products, except Products of Petroleum and Coal + Basic Metal Industries (37)	36
Manufacture of Fabricated Metal Products, Machinery and Equipment	38
Other Manufacturing Industries	39
Services	
Retail + Wholesale + Services of motor vehicles + Hotel and restaurants	50
Construction + Transport + IT + Other services	300

Notes: Code 34 is combined with 35 and code 37 is combined with 38. Retail, Wholesale, Services of motor vehicles, and Hotel and restaurants are combined (50). Construction, Transport, IT, Other services are combined (300). Source: WBES (various issues).

Table 4.8: Distribution of Firms According to Manager Response toward Electricity as an Obstacle

Electricity: Obstacle Category	**Overall Sample**		**Own/Share Generator Sample**		**No Generator Sample**	
	Number of Firms	Percent	Number of Firms	Percent	Number of Firms	Percent
No obstacle	3,510	20.25	2,347	30.28	1,140	12.28
Minor obstacle	3,474	20.05	1,664	21.47	1,782	19.19
Moderate obstacle	3,763	21.71	1,299	16.76	2,429	26.15
Major obstacle	3,997	23.06	1,381	17.82	2,515	27.08
Very severe obstacle	2,587	14.93	1,059	13.66	1,421	15.30
Total	17,331	100	7,750	100	9,287	100

Source: Author's calculation based on WBES data.

Table 4.9: Distribution of Firms Identifying the Severity Level of Electricity Obstacle in South Asia

	No obstacle	**Minor obstacle**	**Moderate obstacle**	**Major obstacle**	**Very severe obstacle**
South Asia	20.25	20.05	21.71	23.06	14.93
Pakistan	13.59	14.13	16.71	26.69	28.88
India	26.31	23.97	22.19	18.02	09.51
Bangladesh	03.40	8.26	27.12	39.03	22.19
Sri Lanka	17.74	36.06	18.91	20.47	06.82
Bhutan	43.78	38.15	10.44	06.02	01.61
Nepal	10.82	09.13	14.30	31.01	34.74

Source: Author's calculation based on WBES data.

Table 4.10: Additional Robustness Checks

Dependent Variable: Log (Sales)	High Performing Firm (HPF)	High Electricity Intensive Firm (HEIF)	Regression Without Private Capital
	(1)	(2)	(3)
HPF in (1) or	0.112*	-0.499***	
HEIF in (2)	(0.064)	(0.052)	
Electricity: Reference group (No Obstacle)			
Minor	-0.065	-0.143**	-0.103**
	(0.062)	(0.056)	(0.040)
Moderate	-0.132*	-0.132**	-0.099**
	(0.072)	(0.056)	(0.054)
Major	-0.099	-0.098*	-0.081
	(0.066)	(0.054)	(0.050)
Very Severe	-0.157**	-0.123	-0.111*
	(0.074)	(0.086)	(0.063)
Interaction: Reference group (LPF # No Obstacle in (1)) or (LEIF # No Obstacle in (2))			
HPF # Minor in (1) or	-0.061	0.059	
HEIF # Minor in (2)	(0.083)	(0.074)	
HPF # Moderate in (1) or	0.053	0.049	
HEIF # Moderate in (2)	(0.083)	(0.066)	
HPF # Major in (1) or	0.025	0.056	
HEIF # Major in (2)	(0.082)	(0.068)	
HPF # Very Severe in (1) or	0.097	0.049	
HEIF # Very Severe in (2)	(0.082)	(0.102)	
Constant	4.403***	4.887***	4.537***
	(0.256)	(0.282)	(0.324)
Other Controls	Yes	Yes	Yes
Country-Year Dummy	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes
City-Size Dummy	Yes	Yes	Yes
Observation	16,016	17,331	17,331
Clusters	78	78	78
R^2	0.8841	0.8860	0.8811

Notes: Model estimated by OLS. Robust standard errors in parentheses, clustered at the region level. Region is the second tier of administrative units (e.g. states in India and provinces in Pakistan). ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: log (sales). Other controls include: log (labor), log (capital), log (material), government, foreign, large, firm age, and export are included in (1) to (3) except in (3) log (capital) is not included. LPF and LEIF stands for low performing firm and low electricity intensive firm respectively. Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Two dummies for missing values of capital and material are included. Data Source: WBES (various issues).

Table 4.11: Country-Specific Regression Analysis of Electricity Infrastructure on Firm Sales

Dependent Variable: Log (Sales)	Pakistan	India	Bangladesh	Sri Lanka	Bhutan	Nepal
	(1)	(2)	(3)	(4)	(5)	(6)
Subjective-Qualitative Measure						
Electricity: Reference group (No Obstacle)						
Minor	0.052	-0.130***	-0.176	-0.151	0.267	-0.084
	(0.107)	(0.029)	(0.138)	(0.158)	(0.177)	(0.172)
Moderate	0.170	-0.111***	-0.160	-0.147	-0.021	-0.098
	(0.101)	(0.029)	(0.127)	(0.186)	(0.225)	(0.138)
Major	0.064	-0.095***	-0.153	-0.058	-0.157	-0.099
	(0.106)	(0.030)	(0.123)	(0.175)	(0.223)	(0.126)
Very Severe	0.058	-0.170***	-0.151	-0.172	0.795	-0.191
	(0.099)	(0.039)	(0.125)	(0.270)	(1.205)	(0.121)
Observation	2,053	10,804	2,880	513	249	832
Subjective-Quantitative Measure						
Electricity: % of sales lost due to power outage	-0.005*	-0.004***	0.001	-0.038*	0.020	-0.003
	(0.004)	(0.002)	(0.003)	(0.020)	(0.023)	(0.003)
Observation	1,709	6,337	1,273	322	47	579

Notes: Model estimated by OLS. Robust standard errors are in parentheses. ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: log (sales). Other controls include: log (labor), log (capital), log (material), government, foreign, large, firm age, export, country-year, industry, and city size fixed effects are included in (1) to (6). Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Two dummies for missing values of capital and material are included. Data Source: WBES (various issues).

Table 4.12: Industry-Specific Regression Analysis of Electricity Infrastructure on Firm Sales

Industrial Codes[a]	31	32	33	35	36	38	39	50	300
Dependent Variable: Log (Sales)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Subjective-Qualitative Measure									
Electricity: Reference group (No Obstacle)									
Minor	-0.060	0.041	-0.158	-0.021	-0.078	-0.033	0.017	-0.247**	-0.149
	(0.115)	(0.082)	(0.124)	(0.065)	(0.070)	(0.044)	(0.149)	(0.117)	(0.197)
Moderate	-0.043	0.200**	-0.244***	-0.069	-0.133*	-0.087	0.066	-0.291*	-0.094
	(0.103)	(0.076)	(0.087)	(0.074)	(0.074)	(0.054)	(0.153)	(0.147)	(0.230)
Major	-0.098	0.172**	-0.327***	-0.008	-0.040	-0.105**	0.043	-0.205	-0.307
	(0.091)	(0.071)	(0.112)	(0.062)	(0.086)	(0.041)	(0.149)	(0.191)	(0.300)
Very Severe	0.012	0.148*	-0.345***	-0.077	-0.327**	-0.117**	0.301	-0.588**	-0.295
	(0.102)	(0.078)	(0.093)	(0.074)	(0.137)	(0.054)	(0.219)	(0.260)	(0.395)
Clusters	74	71	59	73	61	70	19	56	49
Observation	1,701	3,415	917	2,096	1,688	3,584	575	2,222	1,133
Subjective-Quantitative Measure									
Electricity: % of sales lost due to power outage	0.002	-0.004	-0.011*	-0.000	-0.011***	-0.002	-0.004	-0.003	-0.013
	(0.002)	(0.003)	(0.006)	(0.004)	(0.003)	(0.004)	(0.014)	(0.004)	(0.011)
Clusters	67	60	54	64	54	61	17	49	41
Observation	1,071	2,369	384	1,252	912	2,202	399	1,116	562

Notes: (a) see Table 4.7 for industrial codes. Model estimated by OLS. Robust standard errors in parentheses, clustered at the region level. Region is the second tier of administrative units (e.g. states in India and provinces in Pakistan). ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: log (sales). Other controls include: log (labor), log (capital), log (material), government, foreign, large, firm age, export, country-year, industry, and city size fixed effects are included in (1) to (9). Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Two dummies for missing values of capital and material are included. Data Source: WBES (various issues).

Table 4.13: Regression Analysis of Electricity Infrastructure on Sales and Employment Growth

	Sales Growth			Employment Growth		
	Overall sample	Own/Share Generator	No Generator	Overall sample	Own/Share Generator	No Generator
	(1)	(2)	(3)	(4)	(5)	(6)
Reference group (No Obstacle)						
Minor Obstacle	0.010	0.002	0.014	0.003	0.004	-0.000
	(0.007)	(0.007)	(0.009)	(0.004)	(0.006)	(0.005)
Moderate Obstacle	0.007	0.001	0.003	0.006	0.005	0.004
	(0.006)	(0.008)	(0.007)	(0.006)	(0.006)	(0.007)
Major Obstacle	0.010	-0.004	0.023**	0.000	-0.001	-0.000
	(0.007)	(0.008)	(0.010)	(0.005)	(0.005)	(0.006)
Very Severe Obstacle	-0.000	-0.010	0.006	0.001	-0.003	0.005
	(0.007)	(0.010)	(0.009)	(0.004)	(0.005)	(0.006)
Country-Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes	Yes	Yes	Yes
City Size Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observation	16,033	8,547	7,260	17,076	9,368	7,439
Clusters	78	77	77	78	78	78
R^2	0.017	0.014	0.027	0.030	0.027	0.042

Notes: Model estimated by OLS. Robust standard errors in parentheses, clustered at the region level. Region is the second tier of administrative units (e.g. states in India and provinces in Pakistan). ***p<0.01; **p<0.05; *p<0.1. Dependent Variable: sales growth for (1) to (3) and employment growth (4) to (5). Log values of vector X is obtained by adding 1 to each value of the variables, in order to avoid zeros, before transformation, i.e. log(x+1). Data Source: WBES (various issues).

Table 4.14: Development Indicators

	Proportion of Population Living below $1.25 Purchasing Power Parity Per Day	Poverty Gap Ratio (%)	Annual growth rate of GDP per person employed (%)	Prevalence of Underweight Under 5 Years Children (%)	Proportion of Population Below Minimum Level of Dietary Energy Consumption (%)	Net Enrolment Ratio in Primary Education (%)	Literacy Rate (% of Population Aged 15-24, Who can Read and Write)	Child Mortality Rate (Under-5 per 1,000 Live Births)	Maternal Mortality Ratio (Number of Maternal Deaths per 100,000 Live Births)
Country/Region	2011	2011	2014	2015	2014-2016	2015	2015	2013	2013
World			1.8	14	10.9	91.5	91.3	45	210
Northern Africa	1.5	0.3	1.2	4	<5	99.5	91.3	25	69
Sub-Saharan Africa	46.8	19.2	1.7	20	23.2	79.7	74.1	92	510
Latin America and the Caribbean	4.6	2.2	-0.1	2	5.5	93.6	98.3	18	85
Eastern Asia	6.3	1.3	6.4	2	<5	96.8	99.7	13	33
Southern Asia	23.4	4.9	3.3	28	15.7	94.8	87.4	55	190
South-Eastern Asia	12.1	2.2	2.8	16	9.6	94.3	98	29	140

Data Source: United Nations Statistics Division (2015).

4.8.2 *Figures*

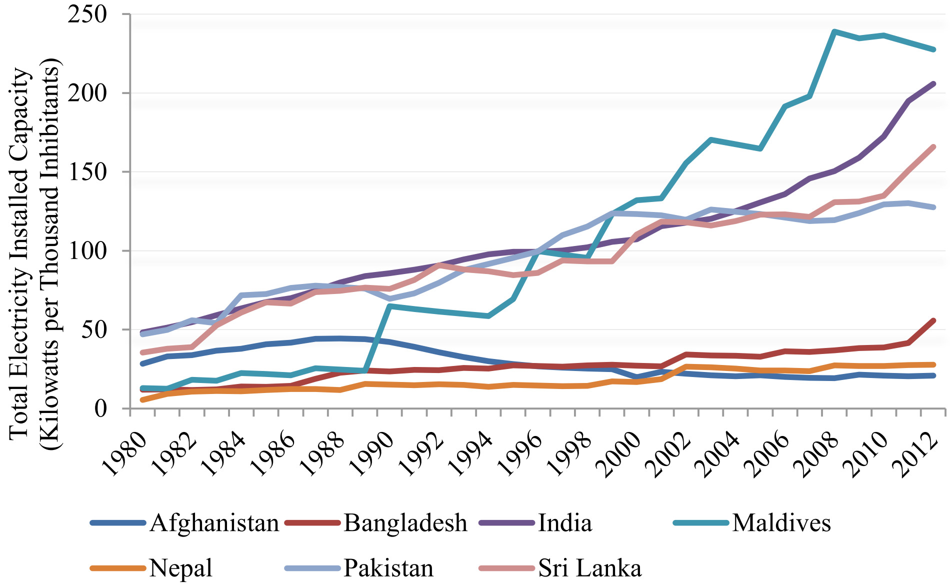


Figure 4.1: Total Installed Electricity Generation Capacities in South Asia, 1980-2010
Data Source: Electricity data from International Energy Agency and population data from World Development Indicators.

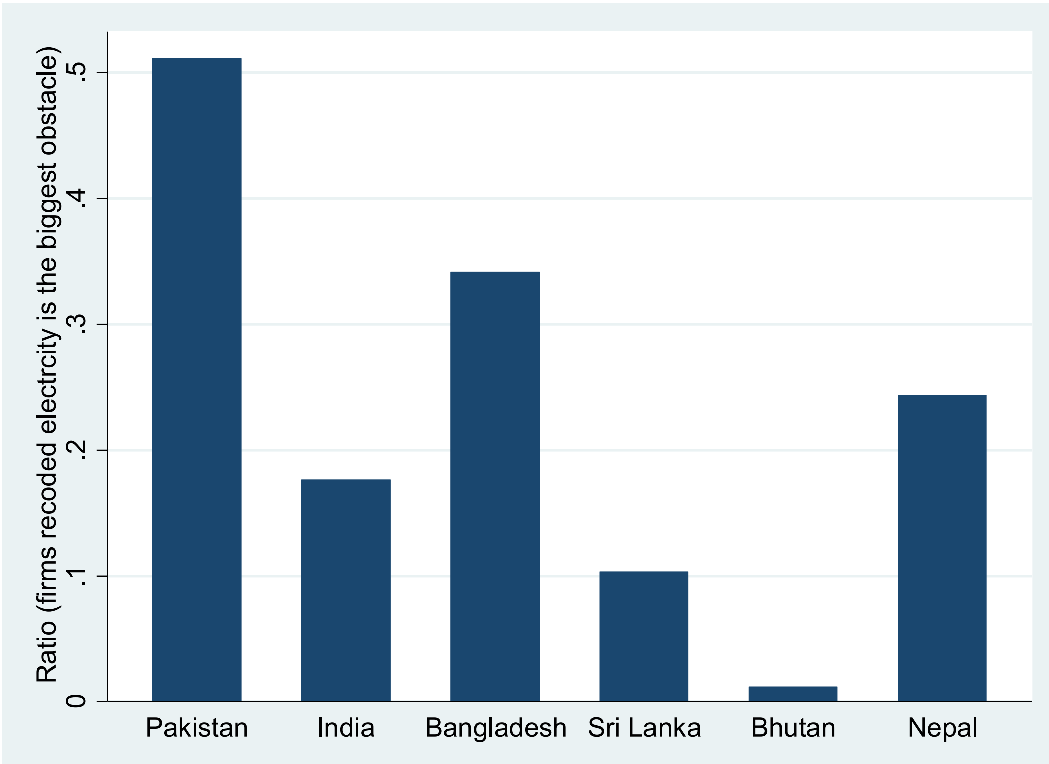


Figure 4.2: Electricity is the Biggest Obstacle Compared to Other Set of Obstacles.
Notes: Other elements in the business environment are: access to finance; access to land; business licensing and permits; corruption; courts; crime, theft and disorder; customs and trade regulations; inadequately educated workforce; labor regulations; political instability; practices of competitors in the informal sector; tax administration; tax rates; and transport. Data Source: WBES data (various issues).

Bibliography

Adger, W. Neil, Juan M. Pulhin, Jon Barnett, Geoffrey D. Dabelko, Grete K. Hovelsrud, Marc Levy, Úrsula Oswald Spring, and Coleen H. Vogel, eds., *Human Security. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge, United Kingdom and New York, USA: Cambridge University Press, 2014).

Afzal, Muhammad, Muhammad Gulfam Arshed, and Kafeel Sarwar, "Education, Health, Food Inflation and Economic Growth in Pakistan," Pakistan Economic and Social Review, 51 (2013), 109-138.

Afzal, Muhammad, Muhammad Shahid Farooq, Hafiz Khalil Ahmad, Ishrat Begum, and M. Abdu Quddus, "Relationship Between School Education and Economic Growth in Pakistan: Ardls Bounds Testing Approach to Cointegration," Pakistan Economic and Social Review, 48 (2010), 39-60.

Ahmed, Manzoor, "Fiscal Decenteralisation and Politcal Econmy of Poverty Reduction: Theory and Evidence from Pakistan," in *Economics, Finance and Business, School of*, Thesis (Doctoral), (Durham University, 2013).

Ahrend, Rudiger, Emily Farchy, Ioannis Kaplanis, and Alexander C Lembcke, "What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries," OECD Publishing, 2014.

Ali, Liaqat, and Muhammad Ismaeel Ramay, "Effects of Human Capital and Trade Orientation on Output and Total Factor Productivity in Pakistan," Research Journal of Applied Sciences, Engineering and Technology, 8 (2014), 1594-1606.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell, "How Do Electricity Shortages Affect Productivity? Evidence from India," National Bureau of Economic Research Working Paper No. 19977, 2014.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell, "How Do Electricity Shortages Affect Industry? Evidence from India," American Economic Review, 106 (2016), 587-624.

Amjad, Rashid, Ejaz Ghani, Musleh ud Din, and Tariq Mahmood, "Export Barriers in Pakistan: Results of a Firm Level Survey," The Lahore Journal of Economics, 17 (2012), 103–134.

Anderson, RM, "Positive Sexuality and Its Impact on Overall Well-Being," Bundesgesundheitsbl, 56 (2013), 208-214.

Andrés, Luis, Dan Biller, and Matías Herrera Dappe, "Reducing Poverty by Closing South Asia's Infrastructure Gap," (Washington, DC: The World Bank, 2013).

——, "Infrastructure Gap in South Asia: Infrastructure Needs, Prioritization, and Financing," World Bank Policy Research Working Paper No. 7032, 2014.

Anttila-Hughes, Jesse Keith, and Solomon M Hsiang, "Destruction, Disinvestment, and Death: Economic and Human Losses Following Environmental Disaster," SSRN Working Paper, 2013.

Aschauer, David Alan, "Does Public Capital Crowd out Private Capital?," Journal of Monetary Economics, 24 (1989a), 171-188.

——, "Is Public Expenditure Productive?," Journal of Monetary Economics, 23 (1989b), 177-200.

——, "Public Investment and Productivity Growth in the Group of Seven," Economic Perspectives, 13 (1989c), 17-25.

Asgary, Ali, Muhammad Imtiaz Anjum, and Nooreddin Azimi, "Disaster Recovery and Business Continuity After the 2010 Flood in Pakistan: Case of Small Businesses," International Journal of Disaster Risk Reduction, 2 (2012), 46-56.

Aterido, Reyes, and Mary Hallward-Driemeier, "The Impact of the Investment Climate on Employment Growth: Does Sub-Saharan Africa Mirror Other Low-Income Regions?," World Bank Policy Research Working Paper No. 5218, 2010.

Austin, Peter C., and Elizabeth A. Stuart, "Moving towards Best Practice When Using Inverse Probability of Treatment Weighting (IPTW) Using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies," Statistics in medicine, 34 (2015), 3661-3679.

Ayyagari, Meghana, Asli Demirgüç-Kunt, and Vojislav Maksimovic, "How Important Are Financing Constraints? The Role of Finance in the Business Environment," The World Bank Economic Review, 22 (2008), 483-516.

Aziz, Babar, Tasneem Khan, and Shumaila Aziz, "Impact of Higher Education on Economic Growth of Pakistan," MPRA Working Paper No. 22912, 2010.

Baldacci, Emanuele, Benedict Clements, Sanjeev Gupta, and Qiang Cui, "Social Spending, Human Capital, and Growth in Developing Countries," World Development, 36 (2008), 1317-1341.

Bardasi, Elena, and Mark Taylor, "Marriage and Wages: A Test of the Specialization Hypothesis," Economica, 75 (2008), 569-591.

Barro, Robert J, "Government Spending in a Simple Model of Endogenous Growth," Journal of Political Economy, 98 (1991), S103-S125.

Becker, Sascha O, and Andrea Ichino, "Estimation of Average Treatment Effects Based on Propensity Scores," The Stata Journal, 2 (2002), 358-377.

Bellak, Christian, "How Performance Gaps Between Domestic Firms and Foreign Affiliates Matter for Economic Policy," Transnational Corporations, 13 (2004), 29-56.

Benhabib, Jess, and Mark M Spiegel, "The Role of Human Capital in Economic Development Evidence from Aggregate Cross-Country Data," Journal of Monetary Economics, 34 (1994), 143-173.

Xinhuanet,"Under-Age Marriage Agonizes Girls in Post-Flood Pakistan," (http://news.xinhuanet.com/english2010/indepth/2011-07/31/c_131020928.htm, Bhatti, Jamil).

Bhutta, Roomana Naz, Imtiaz Ahmad Warich, ArslanIjaz Bhutta, Nouman Ijaz Bhutta, and Muhammad Ali, "Dynamics of Watta Satta Marriages in Rural Areas of Southern Punjab Pakistan," Open Journal of Social Sciences, 3 (2015), 166.

Bils, Mark, and Peter J Klenow, "Does Schooling Cause Growth?," American Economic Review, 90 (2000), 1160-1183.

Blanquicett, Sami Gabriel Coavas, and Laura Gómez Duarte, "For Love or for Money? A Study of the Marriage Wage Premium in Colombia," Ecos de Economía: A Latin American Journal of Applied Economics, 20 (2016), 70-89.

Bloom, David E, and David Canning, "Schooling, Health, and Economic Growth: Reconciling the Micro and Macro Evidence," Unpublished, 2005.

Bloom, David E, David Canning, and Jaypee Sevilla, "The Effect of Health on Economic Growth: A Production Function Approach," World Development, 32 (2004), 1-13.

Bowlby, John, *Attachment and Loss: Attachment* (New York: Basic Books, 1969).

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," Journal of Development Economics, 97 (2012), 339-351.

Burki, Abid A, and Mushtaq A Khan, "Agglomeration Economies and Their Effects on Technical Inefficiency of Manufacturing Firms: Evidence from Pakistan," International Growth Centre Working Paper No. March, 2013.

Calderón, César, and Luis Servén, "The effects of infrastructure development on growth and income distribution," World Bank Working Paper No. WPS3400, 2004.

Cameron, A Colin, and Douglas L Miller, "A Practitioner's Guide to Cluster-Robust Inference," Journal of Human Resources, 50 (2015), 317-372.

Canning, David, and Peter Pedroni, "The Effect of Infrastructure on Long Run Economic Growth," The World Bank, 2004.

Carlin, Wendy, and Mark Schaffer, "Understanding the Business Environment in South Asia: Evidence from Firm-Level Surveys," World Bank Policy Research Working Paper No. 6160, 2012.

Carlin, Wendy, Mark E Schaffer, and Paul Seabright, "Where are the Real Bottlenecks? A Lagrangian Approach to Identifying Constraints on Growth from Subjective Survey Data," Centre for Economic Policy Research Discussion Paper 5719, 2006.

Cockburn, John, Yazid Dissou, Jean-Yves Duclos, and Luca Tiberti, eds., *Public Infrastructure and Economic Growth in Pakistan: A Dynamic CGE-Microsimulation Analysis* (Cham, Heidelberg, New York, Dordrecht, London: Springer, 2013).

Cohan, Catherine L., and Steve W. Cole, "Life Course Transitions and Natural Disaster: Marriage, Birth, and Divorce Following Hurricane Hugo," Journal of Family Psychology, 16 (2002), 14-25.

Cohan, Catherine L., Steve W. Cole, and Robert Schoen, "Divorce Following the September 11 Terrorist Attacks," Journal of Social and Personal Relationships, 26 (2009), 512-530.

Cole, Stephen R, and Miguel A Hernán, "Constructing Inverse Probability Weights for Marginal Structural Models," American Journal of Epidemiology, 168 (2008), 656-664.

Commander, Simon, and Jan Svejnar, "Business Environment, Exports, Ownership, and Firm Performance," The Review of Economics and Statistics, 93 (2011), 309-337.

Cornaglia, Francesca, and Naomi E Feldman, "Productivity, Wages, and Marriage: The Case of Major League Baseball," IZA DP No. 5695, 2011.

Dastidar, Sayantan Ghosh, Sushil Mohan, and Monojit Chatterji, "The Relationship between Public Education Expenditure and Economic Growth: The Case of India," Scottish Institute for Research In Economics Discussion Paper, 2013.

Datta, Saugato, "The Impact of Improved Highways on Indian Firms," Journal of Development Economics, 99 (2012), 46-57.

de Hoon, Sean, Renske Keizer, and Pearl Dykstra, "The Male Marriage Wage Premium in Cross-National Perspective," LIS Working Paper Series, No. 642, 2015.

Deen, Samar, "Pakistan 2010 Floods. Policy Gaps in Disaster Preparedness and Response," International Journal of Disaster Risk Reduction, 12 (2015), 341-349.

Devarajan, Shantayanan, Vinaya Swaroop, and Heng-Fu Zou, "The Composition of Public Expenditure and Economic Performance," Journal of Monetary Economics, 37 (1996), 313-344.

Dinh, Hinh T, and George RG Clarke, eds., *Performance of Manufacturing Firms in Africa: An Empirical Analysis* (World Bank, Washington, DC: 2012).

Dinh, Hinh T, Dimitris Mavridis, and Hoa Nguyen, "The Binding Constraint on Firms' Growth in Developing Countries," World Bank Policy Research Working Paper No. 5485, 2010.

Dollar, David, Mary Hallward-Driemeier, and Taye Mengistae, "Investment Climate and Firm Performance in Developing Economies," Economic Development and Cultural Change, 54 (2005), 1-31.

Dube, Wisdom Garikai, Andrew Phiri, and Mohsen Bahmani-Oskooee, "Nutrition and Economic Growth in South Africa: A Threshold Co-Integration Approach," Journal of Economic Studies, 42 (2015), 138-156.

Duflo, Esther, "Women Empowerment and Economic Development," Journal of Economic Literature, 50 (2012), 1051-1079.

Dutz, Mark A, and Stephen D O'Connell, "Productivity, Innovation and Growth in Sri Lanka: An Empirical Investigation," World Bank Policy Research Working Paper No. 6354, 2013.

Easterly, William, and Sergio Rebelo, "Fiscal Policy and Economic Growth," Journal of Monetary Economics, 32 (1993), 417-458.

Ellison, Glenn , and Edward L. Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," Journal of Political Economy, 105 (1997), 889-927.

Esfahani, Hadi Salehi, and María Teresa Ramírez, "Institutions, Infrastructure, and Economic Growth," Journal of Development Economics, 70 (2003), 443-477.

Estache, Antonio, "Infrastructure and development: A survey of recent and upcoming issues," in *Annual World Bank Conference on Development Economics. World Bank, Washington DC*, (2007).

Fan, Shenggen, Peter Hazell, and Sukhadeo Thorat, "Government Spending, Growth and Poverty in Rural India," American Journal of Agricultural Economics, 82 (2000), 1038-1051.

Farzanegan, Mohammad Reza, "Education Spending and Productivity in Iran: Where have All the Education Expenditures Gone?," in *Politica and Economic Development: ERF 17th Annaul Conference*, (Renaissance Antalya Beach Resort & Spa, Turkey: Economic Research Forum, 2011).

Feng-ying, GUO, "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 (2013), 019.

Field, Erica, and Attila Ambrus, "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh," Journal of Political Economy, 116 (2008), 881-930.

Finlay, Jocelyn, "Fertility Response to Natural Disasters: The Case of Three High Mortality Earthquakes.," World Bank Policy Research Working Paper No. 4883, 2009.

Fisher-Vanden, Karen, Erin T Mansur, and Qiong Juliana Wang, "Electricity Shortages and Firm Productivity: Evidence from China's Industrial Firms," Journal of Development Economics, 114 (2015), 172-188.

Fuller, Brandon, and Paul Romer, "Urbanization as Opportunity," World Bank Policy Research Working Paper No. 6874, 2014.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr, "The Golden Quadrilateral Highway Project and Urban/Rural Manufacturing in India," World Bank Policy Research Working Paper No. 6620, 2013.

Government of Pakistan, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2004-05," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2005).

——, "Pakistan Mouza Census 2008," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2008a).

——, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2006-07," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2008b).

——, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2008-09," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2010).

——, "Pakistan Floods 2010: Early Recovery Plan for the Health Sector," (National Disaster Management Authority, 2011a).

——, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2010-11," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2011b).

——, "Pakistan Millennium Development Goals Report 2013," (Islamabad, Pakistan: Ministry of Planning, Development and Reform, 2013).

——, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2012-13," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2014).

——, "Pakistan Social and Living Standards Measurement (PSLM) Survey 2014-15," (Pakistan Bureau of Statistics, Islamabad, Pakistan, 2016).

Gramlich, Edward M, "Infrastructure Investment: A Review Essay," Journal of Economic Literature, 32 (1994), 1176-1196.

Greenberg, Jeff, Tom Pyszczynski, and Sheldon Solomon, "The Causes and Consequences of a Need for Self-Esteem: A Terror Management Theory.," in *Public Self and Private Self*, (New York: Springer, 1986).

Hallward-Driemeier, Mary, and Reyes Aterido, "Comparing Apples with…Apples: How to Make (More) Sense of Subjective Rankings of Constraints to Business," World Bank Policy Research Working Paper No. 5054, 2009.

Hallward-Driemeier, Mary, Scott Wallsten, and Lixin Colin Xu, "Ownership, Investment Climate and Firm Performance," Economics of Transition, 14 (2006), 629-647.

Hamamatsu, Yuri, Yosuke Inoue, Chiho Watanabe, and Masahiro Umezaki, "Impact of the 2011 Earthquake on Marriages, Births and the Secondary Sex Ratio in Japan," Journal of Biosocial Science, 46 (2014), 830-812.

Hamamatsu, Yuri, Chiho Watanabe, and Masahiro Umezaki, "Secondary Sex Ratio and Marriage After Disaster: A Reply to O'donnell and Behie," Journal of Biosocial Science, 47 (2015), 417-420.

Hansel, Tonya Cross, Paul A. Nakonezny, and Joseph Lee Rodgers, "Did Divorces Decline After the Attacks on the World Trade Center?," Journal of Applied Social Psychology, 41 (2011), 1680-1700.

Hatkoff, Craig, and Isabella Hatkoff, *Owen and Mzee: The True Story of a Remarkable Friendship* (New York: Scholastic Inc., 2016).

Hazan, Cindy, and Phillip R Shaver, "Attachment as an Organizational Framework for Research on Close Relationships," Psychological Inquiry, 5 (1994), 1-22.

Holl, Adelheid, "Highways and Productivity in Urban and Rural Locations," Universitat de Barcelona-Institut d'Economia de Barcelona Working Paper, 2013.

Hong, Houqi, and Sadiq Ahmed, "Government Spending on Public Goods: Evidence on Growth and Poverty," Economic and Political Weekly, 44 (2009), 102-108.

Hulten, Charles R, Esra Bennathan, and Sylaja Srinivasan, "Infrastructure, Externalities, and Economic Development: A Study of the Indian Manufacturing Industry," The World Bank Economic Review, 20 (2006), 291-308.

Jahan, Selim, E Jespersen, S Mukherjee, M Kovacevic, A Bonini, C Calderon, C Cazabat, YC Hsu, C Lengfelder, and S Lucic, "Human Development Report 2015: Work for Human Development," (New York, USA: United Nations Development Programme, 2015).

Jakobsson, Niklas, and Andreas Kotsadam, "Does Marriage Affect Men's Labor Market Outcomes? A European Perspective," Review of Economics of the Household, 14 (2016), 373-389.

Jamal, Haroon, and Amir Jahan Khan, "Trends in Regional Human Development Indices," Social Policy and Development Centre Research Report No. 73, 2007.

Joffe, Marshall M., Thomas R. Ten Have, Harold I. Feldman, and Stephen E. Kimmel, "Model Selection, Confounder Control, and Marginal Structural Models," The American Statistician, 58 (2004), 272-279.

Jones, Stephen, "Infrastructure Challenges in East and South Asia," IDS bulletin, 37 (2006), 28-44.

Khaleek, Shahid Qayyum, and Abid Hussain, "Surge of Divorce in Lahore, its Exploring Factors and Possible Solutions, Pakistan," Advances in Social Sciences Research Journal, 2 (2015), 143-152.

Khanna, Tina, Ravi Verma, and Ellen Weiss, "Child Marriage in South Asia: Realities, Responses and the Way Forward," International Center for Research on Women Unpublished, 2013.

Khilji, Shaista E., "Editor's Perspective: Does South Asia Matter? Rethinking South Asia as Relevant in International Business Research," South Asian Journal of Global Business Research, 1 (2012), 8-21.

Khilji, Shaista E., and Chris Rowley, eds., *Globalization, Change and Learning in South Asia* (Oxford, UK: Chandos Publishing, 2013).

Kneller, Richard, and Florian Misch, "The Effects of Public Spending Composition on Firm Productivity," Economic Inquiry, 52 (2014), 1525-1542.

Kousky, Carolyn, "Informing Climate Adaptation: A Review of The Economic Costs of Natural Disasters," Energy Economics, 46 (2014), 576-592.

Krueger, Alan B, and Mikael Lindahl, "Education for Growth: Why and for Whom?," National Bureau of Economic Research Working Paper No. 7591, 2000.

Kumar, Meeta, and Mihir Pandey, eds., *Economic Challenges to Make South Asia Free from Poverty and Deprivation* (New Dehli: Academic Foundation, 2010).

Lacovone, Leonardo, Vijaya Ramachandran, and Martin Schmidt, "Stunted Growth: Why don't African Firms Create More Jobs?," Center for Global Development Working Paper No. 353, 2014.

Lin, Hui-Lin, Hsiao-Yun Li, and Chih-Hai Yang, "Agglomeration and Productivity: Firm-Level Evidence from China's Textile Industry," China Economic Review, 22 (2011), 313-329.

Lincoln, Anne E., "Gender, Productivity, and the Marital Wage Premium," Journal of Marriage and Family, 70 (2008), 806-814.

Lorentzen, Peter, John McMillan, and Romain Wacziarg, "Death and Development," Journal of Economic Growth, 13 (2008), 81-124.

Lowe, Sarah R., Jean E. Rhodes, and Arielle A. J. Scoglio, "Changes in Marital and Partner Relationships in the Aftermath of Hurricane Katrina: An Analysis With Low-Income Women," Psychology of Women Quarterly, 36 (2012), 286-300.

Lunt, Mark, "Propensity Analyses in Stata: Revision 1.1," 2014. Available at: http://personalpages.manchester.ac.uk/staff/mark.lunt/. (Accessed on 5 January, 2017).

Malhotra, Anju, Ann Warner, Allison McGonagle, and Susan Lee-Rife, "Solutions to End Child Marriage: What the Evidence Shows," International Center for Research on Women, Washington, DC, 2011.

Mamatzakis, Emmanuel C, "An Analysis of the Impact of Public Infrastructure on Productivity Performance of Mexican Industry," Centre for Economic Studies and Information and Forschung Working Paper No. 2099, 2007.

MapAction, "Pakistan - Flood Extent (02 Sept 2010) and Flood Losses (06 Sept 2010)," 2010. Available at: https://mapaction.org/. (Accessed on 16 February 2017).

McDonald, Scott, and Jennifer Roberts, "Growth and Multiple Forms of Human Capital in an Augmented Solow Model: A Panel Data Investigation," Economics Letters, 74 (2002), 271-276.

Mekdad, Yousra, Aziz Dahmani, and Monir Louaj, "Public Spending on Education and Economic Growth in Algeria: Causality Test," International Journal of Business and Management, 2 (2014), 55-70.

Mikulincer, Mario, Victor Florian, and Gilad Hirschberger, "The Existential Function of Close Relationships: Introducing Death into the Science of Love," Personality and Social Psychology Review, 7 (2003), 20-40.

Moyo, Busani, "Power Infrastructure Quality and Manufacturing Productivity in Africa: A Firm Level Analysis," Energy Policy, 61 (2013), 1063-1070.

Munnell, Alicia H, "Why has Productivity Growth Declined? Productivity and Public Investment," New England Economic Review, (1990), 3-22.

——, "Policy Watch: Infrastructure Investment and Economic Growth," The Journal of Economic Perspectives, 6 (1992), 189-198.

Nabi, Ijaz, Abdul Malik, Rabin Hattari, Turab Husain, Adeel Shafqat, Sana Anwaar, and Ammar Rashid, "Economic Growth And Structural Change In South Asia: Miracle

Or Mirage?," London School of Economics and Political Science Working Paper No. 10/0859, 2010.

Nakonezny, Paul A., Rebecca Reddick, and Joseph Lee Rodgers, "Did Divorces Decline after the Oklahoma City Bombing?," Journal of Marriage and Family, 66 (2004), 90-100.

Nasrullah, Muazzam, Rubeena Zakar, and Muhammad Zakria Zakar, "Child Marriage and its Associations with Controlling Behaviors and Spousal Violence against Adolescent and Young Women in Pakistan," Journal of Adolescent Health, 55 (2014), 804-809.

National Institute of Population Studies - NIPS/Pakistan, and ICF International, "Pakistan Demographic and Health Survey 2012-13," (Islamabad, Pakistan: NIPS/Pakistan and ICF International, 2013).

National Institute of Population Studies - NIPS/Pakistan, and Macro International, "Pakistan Demographic and Health Survey 2006-07," (Islamabad, Pakistan: NIPS/Pakistan and Macro International, 2008).

National Legislative Bodies/National Authorities, "Pakistan: VIII of 1961, Muslim Family Laws Ordinance, 1961," (1961).

Nour, Nawal M., "Child Marriage: A Silent Health and Human Rights Issue," Reviews in Obstetrics & Gynecology, 2 (2009), 51-56.

O'Donnell, M. H., and A. M. Behie, "Adapting to Disaster? A Commentary On Hamamatsu et al. (2014)," Journal of Biosocial Science, 47 (2014), 413-416.

——, "Adapting to Disaster? A Commentary on Hamamatsu et al. (2014)--a Brief Reply," Journal of Biosocial Science, 47 (2015), 421-422.

Olsson, Lennart, Maggie Opondo, Petra Tschakert, Arun Agrawal, Siri H. Eriksen, Shiming Ma, Leisa N. Perch, and Sumaya A. Zakieldeen, "Human Security," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, (Cambridge, United Kingdom and New York, USA: Cambridge University Press, 2014).

Oppenheimer, Valerie Kincade, "A Theory of Marriage Timing," American Journal of Sociology, 94 (1988), 563-591.

Palit, Amitendu, and Gloria Spittel, eds., *South Asia in the New Decade : Challenges and Prospects* (Singapore: World Scientific Publishing Co. Pte. Ltd, 2013).

Parsons, Jennifer, Jeffrey Edmeades, Aslihan Kes, Suzanne Petroni, Maggie Sexton, and Quentin Wodon, "Economic Impacts of Child Marriage: A Review of the Literature," The Review of Faith & International Affairs, 13 (2015), 12-22.

Pasha, Hafiz A, Aisha Ghaus, and Salman Malik, "The Economic Cost of Power Outages in the Industrial Sector of Pakistan," Energy Economics, 11 (1989), 301-318.

Pasha, Hafiz A., Khalida Ghaus, Haroon Jamal, Muhammad Asif Iqbal, Muhammad Sabir, Khattak Umar, Nadeem Ahmed, Zia-ul-Haq Khan, Farooque A. Siddiqui, Naveed Aamir, Tabinda Areeb, Manzoor Hussain Memon, Rafea Anis, Rabia Sidat, Sumrin Kalia, Meher M. Noshirwani, and Muhammad Rizwanullah Khan, "Social Development in Pakistan: Annual Review 2014-15," (Karachi, Pakistan: Social Policy and Development Centre, 2016).

Pastor, Manuel, and Rachel Morello-Frosch, "Integrating Public Health and Community Development to Tackle Neighborhood Distress and Promote Well-Being," Health Affairs, 33 (2014), 1890-1896.

Pesaran, M Hashem, Yongcheol Shin, and Richard J Smith, "Bounds Testing Approaches to The Analysis of Level Relationships," Journal of Applied Econometrics, 16 (2001), 289-326.

Ping, HUA, "How Does Education at all Levels Influence Productivity Growth? Evidence from the Chinese Provinces," CERDI-IDREC, CNRS-Université d'Auvergne Documents, Ec 2005.15, 2005.

Pissarides, Francesca, Miroslav Singer, and Jan Svejnar, "Objectives and Constraints of Entrepreneurs: Evidence from Small and Medium Size Enterprises in Russia and Bulgaria," Journal of Comparative Economics, 31 (2003), 503-531.

Polastro, Riccardo, Aatika Nagrah, Nicolai Steen, and Farwa Zafar, "Inter-Agency Real Time Evaluation of the Humanitarian Response to Pakistan's 2010 Flood Crisis," in *Madrid: DARA*, (UNICEF, 2011).

Pollmann-Schult, Matthias, "Marriage and Earnings: Why Do Married Men Earn More than Single Men?," European Sociological Review, 27 (2011), 147-163.

Popa, Ana-Maria, "The Impact of Social Factors on Economic Growth: Empirical Evidence for Romania and European Union Countries," Romanian Journal of Fiscal Policy, 3 (2012), 1-16.

Prati, Gabriele, and Luca Pietrantoni, "Marriage Following the 1997 Umbria-Marche (Italy) Earthquake," Disaster Prevention and Management, 23 (2014), 12-18.

Pritchett, Lant, "Mind your P's and Q's," World Bank Policy Research Working Paper No. 1660, 1996.

Pyszczynski, Tom, Jeff Greenberg, and Sheldon Solomon, "A Dual-Process Model of Defense against Conscious and Unconscious Death-Related Thoughts," Psychological Review, 106 (1999), 835-845.

Raj, Anita, "When the Mother is a Child: The Impact of Child Marriage on the Health and Human Rights of Girls," Archives of Disease in Childhood, 95 (2010), 931-935.

Reinikka, Ritva, and Jakob Svensson, "Coping with Poor Public Capital," Journal of Development Economics, 69 (2002), 51-69.

Research and Development Solutions, "Childhood Immunization in Pakistan," in *Policy Briefs Series No. 3*, (USAIDS, 2012).

Röller, Lars-Hendrik, and Leonard Waverman, "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach," American Economic Review, 91 (2001), 909-923.

Rothenberg, Alexander David, "Essays on Spatial Development," in *Economics*, UC Berkeley Electronic Theses and Dissertations, (University of California, Berkeley, 2012).

Rubonis, Anthony V, and Leonard Bickman, "Psychological Impairment in the Wake of Disaster: The Disaster–Psychopathology Relationship," Psychological Bulletin, 109 (1991), 384-399.

Ruther, Matt, "The Fertility Response to September 11th: Evidence from the Five Boroughs," Population Studies Centre, University of Pennsylvania, Philadelphia, 2010.

Sahoo, Pravakar, and Ranjan Kumar Dash, "Economic growth in South Asia: Role of Infrastructure," Journal of International Trade & Economic Development, 21 (2012), 217-252.

Sanchez-Triana, Ernesto, Dan Biller, Ijaz Nabi, Leonard Ortolano, Ghazal Dezfuli, Javaid Afzal, and Santiago Enriquez, *Revitalizing Industrial Growth in Pakistan: Trade, Infrastructure, and Environmental Performance* (Washington, DC: World Bank, 2014).

Sanchez-Robles, Blanca, "Infrastructure Investment and Growth: Some Empirical Evidence," Contemporary Economic Policy, 16 (1998), 98-108.

Satake, Kenji, and Brian F Atwater, "Long-Term Perspectives on Giant Earthquakes and Tsunamis at Subduction Zones," Annual Review of Earth and Planetary Sciences, 35 (2007), 349-374.

Schwab, Klaus, and Xavier Sala-i-Martín, "Global Competitiveness Report 2014–2015," (Geneva: World Economic Forum, 2014).

Sharma, Chandan, and Sanjay Sehgal, "Impact of Infrastructure on Output, Productivity and Efficiency: Evidence from the Indian Manufacturing Industry," Indian Growth and Development Review, 3 (2010), 100-121.

Shtudiner, Ze'ev, "The Marriage Premium and Productivity: The Case of NBA Players," International Journal of Economic Sciences, 4 (2015), 53-65.

Singh, Anoop, Tooraj Jamasb, Rabindra Nepal, and Michael Toman, "Cross-Border Electricity Cooperation in South Asia," World Bank Policy Research Working Paper No. 7328, 2015.

Smith, Kirk R., Alistair Woodward, Diarmid Campbell-Lendrum, Dave D. Chadee, Yasushi Honda, Qiyong Liu, Jane M. Olwoch, Boris Revich, and Rainer Sauerborn, "Human Health: Impacts, Adaptation, and Co-Benefits," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, (Cambridge, United Kingdom and New York, USA: Cambridge University Press, 2014).

Solomon, Sheldon, Jeff Greenberg, and Tom Pyszczynski, "A Terror Management Theory of Social Behavior: The Psychological Functions of Self-Esteem and Cultural Worldviews," Advances in Experimental Social Psychology, 24 (1991), 93-159.

Stefan, Reppen, "Finding Obstacles to Growth: Is Infrastructure a Binding Constraint on African Light-manufacturing Firms?," International Journal of Economics and Financial Issues, 5 (2015), 820-832.

Straub, Stephane, "Infrastructure and Development: A Critical Appraisal of The Macro Level Literature," World Bank Policy Research Working Paper No. 4590, 2008.

Straub, Stéphane, and Akiko Terada-Hagiwara, "Infrastructure and Growth in Developing Asia," ADB Economics Working Paper No. 231, 2010.

Sumarto, Sudarno, and Indunil De Silva, "Being Healthy, Wealthy, and Wise: Dynamics of Indonesian Subnational Growth and Poverty," Munich Personal RePEc Archive Working Paper No. 57824, 2014.

Taymaz, Erol, and Aykut Lenger, "Multinational Corporations as a Vehicle for Productivity Spillovers in Turkey," Danish Research Unit for Industrial Dynamics Working Paper No. 04-09, 2004.

Timilsina, Govinda R, Michael Toman, Jorge Karacsonyi, and Luca de Tena Diego, "How Much could South Asia Benefit from Regional Electricity Cooperation and Trade?," World Bank Policy Research Working Paper No. 7341, 2015.

Transparency International, "Corruption Perceptions Index 2016," Transparency International, 2016. Available at: http://www.transparency.org/news/feature/corruption_perceptions_index_2016. (Accessed on 17-03-2017).

UNESCO, "EFA Global Monitoring Report 2012:," in *Youth and Skills - Putting Education to Work*, (Paris: UNESCO Publishing, 2012).

——, "EFA Global Monitoring Report 2015:," in *Education for All 2000-2015 - Achievements and Challenges.*, (Paris: UNESCO Publishing, 2015).

United Nations, "Transforming Our World: the 2030 Agenda for Sustainable Development. Resolution 70/1 Adopted by the General Assembly on 25 September 2015," (New York, 2015).

United Nations Statistics Division, "Progress Reports: MDG Report Statistical Annex 2015," (United Nations, 2015).

Vogelstein, Rachel B, *Ending Child Marriage: How Elevating the Status of Girls Advances US Foreign Policy Objectives* (United States: Council on Foreign Relations, 2013).

Waite, Linda J., and Evelyn L. Lehrer, "The Benefits from Marriage and Religion in the United States: A Comparative Analysis," Population and Development Review, 29 (2003), 255-275.

Wang, Hung-Jen, and Peter Schmidt, "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels," Journal of Productivity Analysis, 18 (2002), 129-144.

Weil, David N, "Accounting for the Effect of Health on Economic Growth," The Quarterly Journal of Economics, 122 (2007), 1265-1306.

Weiss, William M, Thomas D Kirsch, Shannon Doocy, and Paul Perrin, "A Comparison of the Medium-Term Impact and Recovery of the Pakistan Floods and the Haiti Earthquake: Objective and Subjective Measures," Prehospital and Disaster Medicine, 29 (2014), 237-244.

White, Lynn, and Stacy J Rogers, "Economic Circumstances and Family Outcomes: A Review of the 1990s," Journal of Marriage and Family, 62 (2000), 1035-1051.

Wijayatunga, Priyantha, D Chattopadhyay, and PN Fernando, "Cross-Border Power Trading in South Asia: A Techno Economic Rationale," ADB South Asia Working Paper Series, Working Paper No. 38, 2015.

Wilson, Chris M, and Andrew J Oswald, "How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence," IZA Discussion Paper No. 1619, 2005.

Wodon, Quentin, Minh Cong Nguyen, and Clarence Tsimpo, "Child Marriage, Education, and Agency in Uganda," Feminist Economics, 22 (2016), 54-79.

World Bank, "Why is South Asia Vulnerable to Climate Change?," 2009. Available at: http://go.worldbank.org/OJ4FWPUB10. (Accessed on 6 April, 2017).

——, "Overview: South Asia," 2016. Available at: http://www.worldbank.org/en/region/sar/overview#1. (Accessed on 17 March, 2017).

World Bank Group, "Global Economic Prospects, January 2017 Weak Investment in Uncertain Times," (Washington, DC: World Bank, 2017).

World Health Organization, "Developing Sexual Health Programmes: A Framework for Action," (Department of Reproductive Health and Research, WHO, 2010).

Xu, Xian, and Jingbing Feng, "Earthquake Disasters, Marriage, and Divorce: Evidence from China 2000-2011," Disaster Prevention and Management, 25 (2016), 59-74.