UNIVERSITY OF TRENTO

School of International Studies

&

Interdepartmental Centre for Research and Training in Economics and Management (CIFREM)

ESSAYS ON SPATIAL INEQUALITIES IN INCOME AND EDUCATION

ECONOMETRIC EVIDENCE FROM PAKISTAN

PhD Dissertation

Sofia Ahmed

May 2011

SUPERVISOR

Prof. Christopher Leslie Gilbert
Department of Economics
Faculty of Economics
University of Trento, Italy

EXAMINERS

Prof. Maria Sassi
Department of Management Studies
Faculty of Economics
University of Pavia, Italy

Prof. Richard Palmer-Jones
School of International Development
University of East Anglia, United Kingdom

Prof. Janette F. Walde
Department of Statistics
Faculty of Economics and Statistics
University of Innsbruck, Austria

ACKNOWLEDGEMENTS

This indeed is a very blessed moment in my life and I would begin by thanking my supervisor Dr. Christopher Gilbert for his immense support, encouragement, and faith in my research endeavours. Despite his extremely busy schedule, our regular Thursday morning meetings and prompt e-mail replies over the past three years have been of great assistance in easing my research stress and erasing my fears. I have been very fortunate to work with and learn from him, and I will always remain indebted to him for all his efforts in facilitating the completion of this dissertation. Thank you, Chris.

I am also grateful to many people in the faculty at the University of Trento for their encouragement, concern, and input to my work. In particular I would like to thank Dr. Vincent Della Sala, Dr. Luigi Bonatti, Dr. Richard Pomfret, and Mark Beittel. Amongst the administrative staff, I am extremely thankful to Rosaria Astarita and Silvia Tomaselli at the School of International Studies, and Nicole Bertotti at CIFREM for all of their prompt assistance.

The completion of a doctoral program is a physically draining and a mentally straining process. In this regard, my PhD years were made bearable and actually fun because of my friends and fellow PhD colleagues at the University of Trento. I am extremely indebted to my friends Venera Demukaj, Sibilla di Guida, Eranda Gjomema, Andrea Ciambra, Dharmaraj N. Krishnan, Flavia Zanon, Mariya Shisheva, Alice Gadler, Viktoria Akchurina, with whom I have lived, studied, worked, travelled, laughed, and cried. I have learnt a lot from our intellectual discussions on life, our academic identities and current affairs. They are the biggest gift of my PhD years and an integral part of my life now. Thank you all for being so wonderful.

A penultimate acknowledgement goes to my family members without whom the achievements in my life would be meaningless. I am extremely grateful to my parents for raising me with the love for knowledge, creativity and ambition, for instilling the qualities of humility and hard work in me, and for their sacrifices and hard work in order to provide me with the best of everything. Their dreams and aspirations to see me as a successful individual have served as the main motivation behind all my quests to achieve excellence in academics and at work. Moreover, I would like to thank my elder brother and sister (both of whom are PhDs themselves) for being the best siblings in every possible manner. They have always been my role models and a constant source of love, guidance, and inspiration.

ABSTRACT

From the industrial revolution to the emergence of the so-called knowledge economy, history has shown that economic development has taken place unevenly across regions. National economies are a complex mix of varying types of geographical configurations which in turn involve varying economic systems, infrastructure, and human capital. In this context recent literature in regional growth and development has highlighted how crucial it is to analyse socio-economic phenomena through the lens of spatial concepts such as density, distance, neighbourhood, and agglomeration. This dissertation emphasizes this fact in three interrelated studies that have analyzed the spatial aspects of income and human capital inequalities in Pakistan.

The first study investigates the evolution and trend of earnings income distribution in Pakistan between 1993 and 2006—a period characterised by macroeconomic reforms and a decent average GDP growth rate. Specifically, it shows the extent to which changes in the returns to human capital have contributed towards changes in earnings inequality across Pakistan. The second study utilizes exploratory spatial data techniques to analyze the extent of spatial clustering of economic inequalities, growth and development across Pakistani districts between 1998 and 2006. The final study then investigates for the first time income convergence across Pakistani districts between 1998 and 2005 using spatial and non spatial econometric techniques.

The main empirical findings from the first study reveal that the returns to low levels of education have declined while the returns to higher education levels have increased. Moreover these returns are much higher for females as compared to males. It has also been shown that earnings inequality has consistently remained higher within provinces as compared to inequality between them. Finally this study demonstrates that females in general and rural females in particular are most sensitive to policies that can affect earnings inequality and labour market outcomes in Pakistan. The findings from the second study demonstrate how neighbouring districts share each other's development levels, confirming that economic geography does matter for regional inequalities, growth, and development across Pakistan. Finally, the convergence analysis carried out in the third study, demonstrates that conditional convergence may be observed across Pakistan once spatial effects have been taken into account.

In summary, this dissertation contributes to the literature and policy debate on economic inequalities and convergence in developing economies and in particular to district level research in Pakistan, by applying counter factual regression analysis and spatial econometric techniques for the first time to Pakistani data.

TABLE OF CONTENTS

Chapter 3. Does Economic Geography Matter for Pakistan? A Spatial Exploratory Analysis

Chapter 4. Are Districts across Pakistan catching up? A Spatial Econometric Analysis of Regional Convergence

Chapter 5. Conclusions

INTRODUCTION

"The bottom line is that Pakistan made little social progress for given rates of per capita income and growth relative to comparator groups, lowering the welfare of the population compared to that under more broadly based development. Pakistan is an interesting illustration that growth alone is not enough for broader development under circumstances of high social polarization. It may help us understand why economic growth is not always reliably associated with social and institutional progress".

The Political Economy of Growth without Development: A Case Study of Pakistan

(William Easterly, 2001)

1. Motivation

Since independence in 1947, Pakistan's economic growth trends have been paradoxical. One of the reasons for this is that the periods of high growth have not resulted in the improvement of its human development indicators. With a population of over 170 million people, Pakistan is the 6th most populous country in the world. The United Nations projects that by 2050, Pakistan's population is expected to double to about 350 million people making it the world's 3rd or 4th most populous country. Currently the country is also experiencing a demographic transition representing an increase in the population of the working age group 15 to 64 years only (also known as demographic dividend). It has been claimed that "up to 40 percent of the success of the East Asian Economic miracle can be attributed to the fact that the region took advantage of its demographic dividend" (Bloom; Canning, Sevilla, 2002:45). In the case of Pakistan, whether this opportunity of a demographic dividend can be translated into the country becoming a regional or a global power crucially depends on the characteristics of its population. By 2050, it has also been estimated that the size of the potential workforce will be approximately 221 million and Pakistan will lose its opportunity to take advantage of this demographic dividend if it is unable to invest in its human capital and provide employment to this emerging workforce (Nayab, 2007; Cohen, 2008).

At the same time, with the existing political insecurities, the country has become unattractive for its existing human capital supplies. It suffers from a constant brain drain and the large diaspora of highly educated Pakistanis throughout the world has reduced incentives to contribute towards the growth and development of the country. Moreover the current education system in Pakistan is in a truly deplorable state. Out of an estimated population of 180 million, 25 million children do not go to school. This implies that roughly 1 out of 10 children out of school in this world is now a Pakistani (Education Emergency Report, Ministry of Education, Pakistan (MOEP), 2011). Under these circumstances, there exists a grim chance that the government will be able to achieve the millennium development goals on education by 2015. It has also been estimated that the economic cost of not educating Pakistan is the equivalent to the economic cost of one flood every year, with the difference being that this is a ‘self-inflicted disaster’ (MOEP, 2011:16). Consequently today, human capital enhancement strategies matter more for Pakistan as compared to any other time. These strategies are required not only to realize its potential demographic dividend but also to increase educational enrolments, raise educational standards, reduce the disincentives for human capital migration, ethnic polarization and the extent of elite domination in the society. They are needed to achieve sustained increases in economic growth and development, improve quality of governance, and to raise standards of civil society awareness in order to combat extremist and violent elements currently influencing the population.

Furthermore, like most developing countries today, Pakistan faces the daunting challenge of overcoming its unequal regional growth. Administratively, Pakistan is divided into 4 provinces which comprise of 113 districts, 420 sub-districts and 48,344 villages. The spatial concentration of economic activities (particularly services sector activities) together with human capital abundance in Punjab, have resulted in it achieving the status of the most privileged province of Pakistan in terms of public investments in social sector developments. With increasing agglomeration of manufacturing and research and development activities in the ‘core’ areas of Punjab and some parts of Sindh, it can be expected that localized human capital investments could lead to further geographic stratification of the population and result in widening the existing regional human development disparities.

In the background of the above mentioned challenges, Pakistan’s potential labour force struggles with unequal opportunities for human capital development and employment across its regions. These inequalities become aggravated at a regional level as institutional sources of education and employment are unequally dispersed throughout the country. This calls for attention towards studying intra-country regional differences not only in order to identify the most neglected subgroups of the population in terms of literacy, innovation and employment, but also to assist in the development of policies that can alleviate these issues of income and human development inequalities.

Economists have pointed to a number of causes of regional disparities such as institutional ineffectiveness, governance issues, sluggish technological diffusion, endowment disadvantages, and the penalty of remoteness from core areas. This dissertation focuses on regional differences created by human capital and earnings inequalities in particular. By utilizing standard econometric tools along with exploratory and spatial econometric data analysis techniques, this project aims to investigate the link between human capital, economic inequalities, regional growth, spatial heterogeneity, and spatial dependence between Pakistani provinces and districts.

2. Economic Growth in Pakistan (1990-2010)

Most emerging market economies faced a stiff economic environment in the decade of the 1990s. Countries such as Mexico, Turkey and Argentina met chronic structural problems that poisoned their domestic currencies, eventually resulting in extremely high twin deficit levels and damaged socio-economic systems. In fact mid 1990s were characterised by failed episodes of structural adjustment and privatization programs, and poor stabilization regimes and in 1997, the East-Asian financial crisis took place. Under these international economic and financial developments, problems for Pakistan increased as it faced possible economic sanctions after its first successful nuclear testing in 1998. This resulted in a shattered confidence of international and domestic investors in the domestic market and the State Bank of Pakistan resorted to freezing of foreign exchange bank accounts as it feared capital flights. This situation completely reversed with the advent of the 9/11 incident as it led to an increased flow of remittances for most developing countries including Pakistan. The fiscal year 2002-03 witnessed a sharp recovery in economic growth as Pakistan experienced a severe decline in its GDP growth rate in 2001 due to catastrophic drought conditions at the domestic front coupled with the world economic down turn in the after math of the 9/11 incident. In South Asia, Pakistan was the only country that achieved more than 5 percent growth in 2002-03 which demonstrated its resilience to external shocks (Ahmed, 2005).

The second half of the 2000s was an eventful period for Pakistan in terms of economic and political developments. In 2005, Pakistan was amongst the fastest growing economies in the world (see Table 1). Workers' remittances increased by 323 percent from \$0.98 billion in the year 2000 to \$ 4.17 billion in 2005, and Pakistani rupee gained strength as the balance of payments situation improved. The year 2005 also marked beginning of an inflation era particularly, double digit high food inflation which put an extra-ordinary burden on the fixed income and other poorer segments of the population. In 2006, despite the massive destruction caused by a catastrophic earthquake and its related expenditures, in 2006 the economy of Pakistan continued to gain further traction. However, price hike for essential food items continued indicating 'a future domestic food crisis' coupled with a high negative growth rate of 8.4 percent of electricity and gas distribution marking the beginning of an "energy crisis" in Pakistan.

An unstable domestic law and order situation, soaring oil, food and other commodity prices, softening of external demand and an international financial crisis made the fiscal years 2007-08 and 2008-09 rather difficult for the Pakistani economy as it witnessed rising inflation, a growing fiscal deficit, an energy supply crisis, and widening trade and current account deficits. Intensification of security challenges internally and on the borders with Afghanistan in 2009 have exacted an extremely high cost on the economy not only in terms of direct costs of fighting the war against terrorism but also in terms of reduction in investment flows and market confidence since 2001 (Pakistan Economic Survey 2009). During 2009-10, a total of 1,906 terror attacks were recorded in the country according to the National Crisis Management Cell, Ministry of Interior and its cost amounted to about 6 percent of GDP in 2009-10. As a front line state in the global "War against Terror", official estimates suggest that there has been a loss of 43 billion US dollars to the economy between 2001 and 2010 (see Table 2).

Despite the various above mentioned economic and political challenges at the national and international level, Pakistan has maintained a fairly decent average GDP growth rate of 4.6 percent over the past two decades (see Figure 1). Although, the growth in real GNP persistently decelerated during the 1990s—as it declined from an average of 5.7 percent in the 1980s to an average of 4.2 in the first half of the 1990s and 3.5 percent in its second half, it again rose to an average of 5.1 percent in the 2000s. While many low income countries were stagnating, a considerable reduction of the number of people below the poverty line over the years is a plausible achievement. On the human capital front, Pakistan's domestic professional *elite* is at par with those in the industrialized world. It also has an educated and entrepreneurial Diaspora that run small businesses in various countries throughout the world, skilled workers in the Gulf States, and high ranking official in almost all international organizations (Easterly, 2001). Pakistan was also amongst the few countries that depicted positive growth rates even during the peak of the global credit crunch in 2008-09.

3. The Paradox of Growth

Yet Pakistan has performed quite poorly when one observes its social sector development indicators over the years. In 2010, Pakistan ranked 125th out of 169 countries in the Human Development Index (HDI) rankings, keeping its place in the medium human development category as defined by the annual UNDP Human Development Report. Despite the level of its economic growth, it has always stayed even below the average HDI value of South Asia (see Table 3). It must be noted that the Human Development Index only provides part of the picture of a country's socio-economic development since it does not take into account the status of political participation and gender inequalities in that country. With the current trends of political instability and gender biases, perhaps Pakistan's performance would be even worse highlighting the fact that an exclusive pursuit of economic growth is inadequate to ensure human development[1].

Although Pakistan's level of *total* public expenditure has been quite comparable to international standards, but within that the actual amount of expenditure on social sectors has been very low throughout its history (Joekes *et al*, 2000). Until recently for its level of income, Pakistan's indicators like infant mortality, and female primary and secondary enrolment rates were among the lowest in the world (Easterly, 2001). In 2010, Pakistan ranked the lowest in the categories of Adult Literacy Rate (for both sexes), and expected years of schooling of children in South Asia (World Bank, 2010). Pakistan is also more politically unstable, more corrupt, more violent, less democratic, and "less respectful of human rights" (as demonstrated by various indices) when compared with countries that have similar income levels. (Easterly, 2001:3).

Despite declining absolute poverty trends, inequalities in income, consumption, and human development conditions have also continued to rise. It has been widely demonstrated how inequality threatens a country's poverty reduction capacity and its sustainability of growth (World Bank, 2006). By excluding a proportion of a country's population (e.g. poor, elderly, women, illiterate, disabled etc), it constricts human potential through which "entrepreneurship, innovation, investment and productivity growth can occur to sustain GDP growth" (Hussain, 2008:4). Health and education in particular have been extremely neglected social sectors as demonstrated by low levels of public expenditures on them (Joekes *et al*,

[1] It should also be noted that the current rankings were estimated before the catastrophic flood came in July 2010. The flood directly affected 20 million people mostly by destruction of property, livestock, and infrastructure including hospital and school buildings.

2000). This gross neglect of human capital has been amongst the major causes that have prevented Pakistan from developing into a knowledge-based economy. As many as 11 National Education Commissions were held between 1947 and 1993, but education has remained “unabashedly elitist”. (Easterly 2001:19). Pakistan’s levels of literacy, high school enrolment rates, and schooling quality have been amongst the lowest in the South Asia indicating its significance in government policy (see Table 4). Although the literacy rate has increased from 46 percent in 1999 to 56 percent in 2010, the gender gap remains high and the total population still consists of more than 50 million illiterate adults (Economic Survey of Pakistan, 2010). Regional human capital disparities have worsened even more after Pakistan’s direct involvement in the war on terror as educational attainment for “virtually a whole generation of school-going age in the affected areas of NWFP and FATA has been jeopardized or severely undermined” (Economic Survey of Pakistan 2010:7).

In the next section, this chapter analyzes this paradox of Pakistan’s growth in a historical context to highlight how political developments (particularly political instability) over the years have jeopardized the social development process in the country.

3.1 *Political developments in Pakistan—A historical account*

When the British left the subcontinent after agreeing to divide it into two independent states in 1947, Pakistan inherited a heavy baggage of unresolved social, economic and political issues. Neither the British nor the Pakistani leaders were certain of the newly created state boundaries, the composition of a highly multi-ethnic population that it consisted of, and the type of development policies it would have to pursue (Naseem, 2008). As Pakistan lost its founder—Mohammad Ali Jinnah, and its first Prime Minister—Liaqat Ali Khan within the first 4 years of its creation, the country faced a leadership dilemma since the very beginning. Within its first decade, 7 Prime Ministers changed between 1947 and 1958.

Two major divisions emerged during the first decade. While the first was based on geography, between West Pakistan (Pakistan) and East Pakistan (Bangladesh), the second was between the Army and the Civil Service—the two public elite services of Pakistan (Naseem, 2008). Pakistan’s military—in particular the Army—has always remained an influential institution and its most important interest group. It has power over the country’s politics, foreign policy and over the years it has also begun to dictate its economic policies

(Zaidi, 2005). Pakistan had two military regimes lasting for 14 years in its first 25 years since independence. Ironically, military regimes in Pakistan have always been characterized with more political stability as compared to civilian regimes. Direct military intervention in Pakistan's democratic system started when the Army staged its first coup in 1958 and General Ayub Khan began his decade long military rule.

The Ayub regime (1958-69) is mainly characterized by a period of economic growth, inequality, and increased dependence on economic and military assistance from the United States. A highly regulated policy framework for import substitution led industrialization in the consumer goods sector emerged in 1960s. This policy was combined with tariff protections for manufacturers of consumer goods and import controls on competing imports enabling the emergence of an industrial elite class that made large profits without having to achieve high levels of efficiency. In addition to these protectionist measures, exporters were offered incentives such as tax rebates, tax exemptions, and accelerated depreciation allowances which enabled the rich industrial elite to make large amounts of profits without having to worry about international competitiveness, improving efficiency, and diversifying production into high value added goods (Hussain, 2003).

Moreover in order to secure a power base, the government transferred rents to the industrial elite via subsidies and tariff manipulations. From then onwards, the ruling elite became accustomed to seeking rents from the government which started to increase Pakistan's economic burden from a very early stage (Ibid). The 1960s policy of concentrating national income in the hands of the upper income groups was based on the assumption that since the rich tend to save a larger proportion of their income, higher national savings could be achieved through unequal income distribution. While the policy of distributing income in favour of the elite was successful it failed to raise domestic savings over time. As noted by Griffin (1972), 15 percent of the resources that were annually generated in the rural sector were transferred to the urban industrialists. Out of these transferred resources, 63 to 85 percent "went into increased urban consumption" (Hussain, 2003:38). The requirement of foreign aid shot up as the savings from elite incomes proved to be inadequate. The debt servicing burden rose as the interest rates on loans increased. In 1960-61 debt servicing as a percentage of foreign exchange earnings was 4.2 per cent and it increased to 34.5 per cent by 1971-72. Its magnitude continued to rise for the next three decades and by the year 2000, it stood at 40 per cent (Ibid).

Although land reforms and industrialization helped to achieve rapid economic growth, by the end of the 1960s, vast regional income and sectoral disparities had emerged. Small groups of elite families not only dominated industry, banking, and insurance sectors of Pakistan but also exercised considerable power over government agencies that sanctioned industrial projects. It has been noted that 46 per cent of the value added in the large scale-manufacturing sector originated in firms that were owned by only 43 families (Hussain, 2003). While a monopolistic class of industrialists emerged, the majority of Pakistan's population experienced an absolute decline in its living standards as real wages in the industry declined by 25 percent between1960-67 (Griffin & Khan, 1972). By 1971-72 rural poverty had deteriorated to the extent that it has been estimated that 82 percent of rural households were unable to provide 2100 calories per day to each household member (Naseem, 1977).

This period also marked the beginning of worsening regional disparities not only between East and West Pakistan then but also within (today's) Pakistan. Infrastructure development and facilities were markedly different between provinces as the relatively well endowed (with infrastructure facilities) provinces of Punjab and Sindh were able to attract a larger proportion of industrial investment. Another cause of rising intra-country regional income disparity came from the differential impact of developments in the agriculture process, the so-called 'Green Revolution'. Regions with larger areas of irrigation systems benefitted more from the mechanization of agriculture processes and the adoption of high yielding seed varieties. Again Punjab and Sindh disproportionately benefited more from it as compared to Khyber Pakhtoonkhwa (previously known as North West Frontier Province) and Balochistan. Studies have shown that due to the use of large farm machinery and expensive seeds, the Green Revolution favoured large, well established farmers over the small and medium sized ones. Moreover, land ownership also remained concentrated in the hands of few landlords. In 1972, 30 per cent of total farm area was owned by large landowners (owning 150 acres and above). These two features led to an economic polarization of the population and landlessness in rural Pakistan (Hussain, 1988). Wealth continued to be concentrated in the hands of big landlords while incomes of the poor peasantry –most of which with smaller areas of land and unable to buy more—declined. Moreover the poor peasant's dependence on the big landlord intensified as the latter became the former's source of agricultural inputs (seeds, fertilizers etc). Although the high agriculture growth rate from the Ayub Era could not be sustained in future, the inequality enhancing mechanisms that emerged continued to persist over the next 4 decades (see Table 5).

Next came the Bhutto era (1973-1977) which was characterized by widespread nationalization. While it was aimed at constraining the powers of the small group of industrial elite that had emerged in the 60s, it actually increased size of public sector and widened the ‘resource base of the regime for the practice of traditional form of power through state patronage’ (Hussain, 2003: 41). Although Bhutto initiated pro-poor labour intensive programs such as the National Development Volunteer and the People Work Program to generate employment for the educated unemployed and the rural poor, both the programs could not be sustained due to budgetary constraints. The financial constraints had been caused by large sums of budget being directed towards non development expenditures such as the military to win the confidence of the military and to run the national industries that were running on losses.

However the decreasing GDP growth rates and growing debt servicing burden from the Bhutto era were unable to have a crippling impact on the Zia regime (1977-89) because of a) increased financial support from the west particularly due to the Soviet invasion of Afghanistan and Pakistan’s role, and b) increased remittances from the Middle East due to its consumption and construction boom. These developments helped ease budgetary pressures, balance of payment pressures and also social pressures by benefitting about 10 million people that too from the lower middle and working class (Hussain, 2003). Despite the rise in GDP during the Zia regime, investments in social and economic infrastructure remained inadequate. As the Afghan War ended and financial assistance from the western allies stopped, the Pakistani economy plunged towards an economic recession in the 1990s.

As many as 8 Prime Ministers changed in a short span of 11 years between 1988 and 1999. During the 1990s political instability, corruption and a worsening law and order situation adversely affected the already declining GDP growth rate. It was a period of institutional degradation as political instability, corruption and the law and order situation worsened. To finance their rising unproductive expenditures, successive governments chose to reduce development expenditure which reduced from an average of 7.4 per cent of GDP in 1973-77 (Bhutto regime) to only 3.5 per cent of GDP in 1997-2000 (last Nawaz Sharif Regime). As the GDP growth rates declined—from 6.3 percent in 1980s to 4.2 percent in

1990s—employment growth rate also remained at low levels. A study in 1987 had already argued that in the decade of 1990s, Pakistan would be unable to sustain the high GDP growth rate of the preceding 3 decades due to structural constraints originating from infrastructure deterioration, slow export growth and low rate of savings (Hussain, 1988). Problems of employment aggravated, growth rate of labour productivity fell and real wages in both agriculture and industry declined in the 1990s. An NHDR/PIDE (2001) survey reveals that slower economic growth rates combined with declining employment elasticities and real wages during the 1990s had important consequences for poverty aggravation. Increasing number of families working in the agriculture and manufacturing sectors, were pushed into poverty as "second family members of households on the margin of poverty could not get adequate wage employment" (Hussain, 2003:46). Inequality levels also rose as Pakistan's Gini coefficient increased from 26.85 in 1992-93 to 30.19 in 1998-99 (Federal Bureau of Statistics, 2001).

In 1999, Pakistan again came under a military regime when General Pervez Musharraf staged the 4th military coup in Pakistan's history. It was a politically stable period mainly due to the fact that Musharraf was not only the head of the state but also the de facto head of the government. While this period has been lauded internationally for its extraordinary performance in the macroeconomic sphere, income disparities across different segments of the population and regions continued to widen. Domestic critics attribute the macro-economic success to debt rescheduling done in the last years of the Nawaz Sharif regime, increased remittances after 9/11, and the increased aid flows Pakistan received by being a front line partner in the international war against terror (Naseem, 2008).

This section attempted to shed light on the trends of increasing poverty, inequality and tendency for loan dependence in Pakistan in a historical context. Throughout its history, Pakistan's main motive of adopting adjustment programmes has been to obtain short term foreign liquidity by the IFIs and other donors and "not for the desire for longer term structural change" (Hussain, 1999:9). Moreover, Hussain (2003) concludes that economic policies initiated in the Ayub regime (1960s) induced vast social and regional disparities for the rest of Pakistan's history. Structural constraints and institutional degradation continued as successive governments engaged in resource allocation based on political patronage instead of economic efficiency. Rising budget deficits that rose from the Bhutto regime increased even more during the Zia regime as state funds were allocated for unproductive political activities instead of investing them in human capital or institutional upgrading. Hence irresponsible governance accentuated the already existing poverty, inequality and employment issues.

Easterly (2001) demonstrates that over the years Pakistan grew much more than other countries that shared its initial level of income, but achieved more or less the same amount of social progress or even less. The study also compares countries that grew at the same rate as Pakistan regardless of the initial income levels and concludes that other moderately growing countries achieved higher rates of social progress than Pakistan which had higher levels of growth than them. While Pakistan's case demonstrates how per capita income growth is possible even when it is not accompanied by human capital accumulation (as measured by enrolment and educational attainment statistics) on the other hand, it also highlights how economic growth cannot always be associated with social and institutional progress (Easterly, 1999 & 2001; Pritchett, 1999; Benhabib and Spiegel, 1994). Currently the country is experiencing a multi-faceted crisis with economic and financial dimensions coupled with deploring human conditions. This has been exacerbated by its status as a front line state in the so called war against terror and frequent natural disasters such as the floods of July 2010.

This section has demonstrated how the absence of far sighted economic management policies, lack of domestic ownership and capacity, political instability, ethno-linguistic fractionalization, regional inequalities, budget misallocations away from productive development projects and towards non productive projects have jointly contributed towards the failure of various well intentioned government and donor led development programs. Moreover, when regional disparities combine with political and ethnic issues they aggravate political and social instabilities even more (Burki *et al*, 2010). In this context, there is an immediate requirement of re-evaluating Pakistan's growth strategy that also takes into account spatial dimensions of development and focuses on the strengthening of institutions (Haque, 2006).

4. Growth, Distribution and Space

This section establishes the place of this dissertation in the current literature on regional economic inequalities (Chapters 2 and 3) and convergence (Chapter 4), and presents the issues which form its foundation.

4.1 How does income inequality affect economic growth?

On a theoretical level the relationship between inequality and growth is ambiguous hence the actual impact of inequality on growth has always remained an empirical question. The 'classical' economic approach asserts that inequality enhances incentives for increased efficiency, growth and capital accumulation in various ways. For example since different classes in a society have different propensities to save, inequality in income distribution may increase the rate of savings and capital accumulation leading to accelerated growth (Kuznets, 1955; Kaldor, 1956). Moreover, inequalities may signal opportunities to people to improve their position in the income distribution by taking risks and innovating in their entrepreneurial activities (Siebert, 1998; Bell and Freeman, 2001). Wider income inequalities can be helpful in the sense that they may signal strong incentives to work effort, innovation and skill development. Hence very low levels of income inequality may actually weaken the incentive to invest in human capital, and thereby adversely affect economic growth prospects (ILO, 2008).

However relatively recent research has demonstrated that inequalities are actually detrimental for economic growth through various channels (Leoni and Pollan, 2003; Canadas, 2008):

1. Unequal distribution of income leads to a pressure for redistribution through distortionary taxes which in turn reduce growth (Knowles, 2001).
2. Rising inequalities increase fertility levels which can negatively influence human capital investments and eventually growth (Knowles, 2001).
3. Usually the poor entirely depend on their initial wealth to undertake important investments, and due to imperfections in the credit and insurance markets they may be unable to invest in profitable projects (Galor and Zeira, 1993; Banerjee and Newman, 1993; Aghion *et al*, 1999; Levine, 2004). Imperfections in the capital markets also reduce investments in human capital thereby reducing growth (Knowles, 2001).
4. In polarized societies, the rich may influence economic policies in their favour neglecting the middle and lower income groups. This rent seeking often distorts resource allocation in an economy and not only threatens growth but also a society's social cohesion and political fabric. (Leon, 2007; Grossman, 2003; Easterly, 2001; Barro, 2000; Benabou, 1996; Alisina and Rodrick, 1994; Persson and Tabellini, 1994; Perotti, 1993; Bertola, 1993).

5. Inequality and corruption are highly interrelated as unequal distribution of wealth and income creates incentives for certain high-income groups to intervene in the political process and democratic governance (You and Khagram, 2005). High degrees of inequality provide richer groups with sufficient resources to offer bribes even to the highest ranking officials in the government. In this context it has been noted that the cost for richer individuals not bribing is much higher in societies with high levels of inequalities as compared to the more equal ones. This is because a competition merges within the rich class of who is able to bribe more to successfully evade taxes and avoid those actions that may be harmful to their personal businesses (Glasear, Scheinkmann and Shleifer, 2003). When political power arises from excessively large income and wealth inequalities, it enables the rich to protect their interests via anti-competitive measures (such as receiving licenses for monopolies) by exercising their political influence on economic policies. Such mismanagement not only leads to inflation inducing distortions and prevents the introduction of new technology; they also reduce consumer welfare and eventually economic growth (Scarpetta and Tressel, 2002).
6. Wealthy households have a combined adverse effect on the allocation of public resources. It has been noted that wealth decreases the opportunity cost of lobbying and increases the chance that like-minded people will get together and influence government activity (Zhang, 2008). For instance the wealthier elite lobbies often divert public spending away from education. When public spending on tertiary education is higher than on primary or secondary education, children from poorer households have lesser chances of obtaining secondary education which is a pre requisite to university education. It has also been noted that the richer segments of the society obtain public contracts for construction for their own companies or influence public expenditure by granting subsidies to specific industries or goods which they deal with or consume (ILO, 2008).
7. Differences in distribution of income and public goods, and opportunities related to education, distribution of land and assets, and political influence imply a narrow set of opportunities in society which decreases the productivity of resources (Bourguignon, Ferreira, and Menedez, 2003; Ferreira, 2001; Sen, 1992). Such inequalities can also lead to increased political/ social conflict, crime and other illegal activities that prevent investment and weaken property rights. If inequalities are very high, the poor may attempt to engage in riots and revolutions threatening the stability of political and

government institutions (Rodriguez, 2000; Banergee and Duflo, 2000; Alesina and Perotti, 1996; Benhabib and Rustichini, 1996). Resources are wasted as the poorer segments of a society engage in riots instead of devoting their energies to productive activities (Canadas, 2008).

8. Inequalities negatively influence social capital such as civil society networks that facilitate contracts and help in maintaining social stability (Caramuta, 2005; Nan, 2000; Knack and Keefer, 1997; Kawachi *et al*, 1997).

Moreover, income inequality can have different effects in the short and long run, and on urban and rural areas (Partridge, 2005; Forbes, 2000). For the case of the U.S. a positive relationship has been estimated between inequality and growth in the urban areas, while the contrary has been found for rural areas. Fallah and Partridge (2006) shows that this relationship is different for urban and rural areas because factors such as agglomeration economies and labour specialization intensify market rewards and attract more skilled labour and generate growth in cities. However in rural areas greater income inequality weakens social cohesion due to lack of anonymity and in turn economic growth (Canadas, 2008; Fallah and Partridge, 2006).

All in all it has been widely accepted that reducing income inequality can foster long term economic growth by enabling more individuals to participate in an economy to their maximum potential. More equal societies provide wider employment opportunities, increased access to borrowing, greater product consumption and an investor friendly environment (Gehring and Kulkarni, 2006).

4.2 How do human capital inequalities affect economic growth?

A separate branch in economic literature highlights the role of human capital in explaining the relationship between inequality and economic growth. Various studies have demonstrated that in most parts of the world the process of local growth is heterogeneous (see Paci and Usai, 1999, 2000; 2001; Castella and Domanech, 2002 for case studies on Italy). Although not seen as crucial, the role of human capital is important since the distribution of income is mainly driven by the distribution of human capital across or within countries. Golmm and Ravikuman (1992), Saint-Paul and Verdier (1993) or Galor and Tsiddon (1997) are among those few who present models in which the sources of inequality are driven by inequalities in human capital distribution. There exists a dyadic relation between human capital accumulation and inequality with both affecting each other. It has also shown that human capital inequality negatively affects economic growth rates not only through inefficient resource allocation but also through reduced investment rates (Bird-Sal and Londono, 1997; Lopez *et al*. 1998).

Since policy makers have realized the importance of distribution of human capital and its externalities, the creation and distribution of these benefits calls for increased assistance at the regional level. Lucas (1988), Nijkamo and Poot (1998), Martin and Sunley (1998) have emphasized the importance of human capital in fostering regional growth. Human capital investment (expenditures) in health, education and training have all been acknowledged as generating externalities that increase the quantity and quality of regional labour force. However, this empirical work on public investments in human capital often overlooks spatial interaction among regions. Most regions are highly interconnected through family networks, transport, trade, migration and regional labour markets. They often exhibit a high degree of interaction due to which unemployment and other human capital issues may be highly spatially correlated (Molho, 1995; Acemoglu, 2001)[2].

It is also being argued that human capital levels are diverging (Berry and Glaeser, 2005) and its concentration is likely to continue to occur in certain regions only (Florida, 2002a, Berry and Glaeser, 2005). Proximity and distance play an important role in inequality literature because it has been noted that countries with a lower market access may also have lower education levels. As a result some authors have shown that the already peripheral countries in the world may continue to become remote over time (see Redding and Schott, 2003).These results indicate that policies aimed at increasing growth should focus not only on education but also on region wise distribution of human capital enhancing infrastructure and institutions that can cater to a larger section of the population.

4.3 Economic growth, income and human capital inequalities in the light of New Economic Geography literature

Although the new endogenous growth theory emphasizes the role of knowledge spillovers in macroeconomic growth, it tends to neglect the regional dimension. Recent evidence suggests that knowledge spillovers tend to be spatially bounded or localized. It has been shown that regions are not isolated but are a part of a core-periphery system. Disregarding of interactions

[2] See Zhang et. al. (2006) and Karlson and Haynes (2002) for a review of spatial perspectives on regional labour markets.

between individual regions in the system can lead to an adverse impact of regional policies (Baldwin *et al*, 2003 and Midelfart *et al*, 2004)[3]. Therefore, economists have turned their attention towards the so-called new economic geography and spatial patterns of inequality and the dynamics of geographic income disparities have come back under the lime light (Krugman, 1991; Fujita *et al*, 1999; Krugman and Venebles, 1995; Ottaviano and Puga, 1998; Puga and Venebles, 1997, 1999; Darlauf and Quah 1999; The World Bank—World Development Report, 2009). The acceptance of new economic geography and agglomeration is represented by the application of mechanisms such as clusters, innovation networks, knowledge externalities and other similar spillovers (Maskell and Malmberg, 1999; Moulaert and Sekia, 2003; Henry and Pinch, 2006). This according to some authors has stimulated a new wave of so-called 'third wave' regional industrial policies (Benneworth and Hospers 2006; Larosse, 2004; Bradshaw and Blakely, 1999) which aim to promote knowledge based growth through regional technology exchange and innovation policies (Hospers, 2006; Charles, 2003; Lagendijk and Hassink, 2001).

With this increased interest on regional development issues and enhancement of spatial data, empirical analysis the use of spatial econometric techniques has also gained popularity (Arbia, 2006). Studies such as Quah (1996) and Moreno and Trehan (1997) have demonstrated how geographical spillovers and physical location are just as important as other macroeconomic factors in growth studies. Spatial externalities model social interaction that introduces dependence among the agents of a social system (Anselin, 2003a; Arbia, 2006). Among the various kinds of spatial externalities, international and regional level studies have confirmed the existence of a positive spatial correlation between knowledge spillovers of human capital, Research and Development, and other types of training activities and economic growth (Feldman, 1994; Anselin *et al*, 1997). Further more from the findings of Nelson and Phelps (1966) to Benhabib and Spiegal (1994), studies have shown that economies located closer to a technology leader benefit more and grow faster[4]. Recent results have also revealed that human capital has a positive and a significant effect on total factor productivity growth when interacted with the distance to the technology leader usually measured in terms of income per capita (Pede, Florax and L.F de Groot, 2006). This is because technology is also spatially bounded, and is not completely exogenous and freely

[3] For surveys on the effectiveness of regional policy see: Rodriguez-Pose and Fratesi (2004) and Ederveen et al (2002)

[4] Other studies that examine the links between human capital, development and growth include Bils and Klenow (2000) and Mankiw et. al (1992).

available in a region (Lawson and Lorenz, 1999; Jaffe *et al*, 1993). The application of spatial econometric methods to study regional convergence has also demonstrated how previous studies were incomplete without taking spatial interdependence between regions into account (Fingleton and Lopez-Bazo, 2006; Rey and Montouri, 1999).

5. Objectives and Research questions

This dissertation was motivated by the need for an analysis on the spatial aspects of income and human capital inequalities at lower geographic aggregation levels in Pakistan. As it will be demonstrated in the literature reviews of the upcoming chapters, with the exception of a few, most existing studies on earnings/income inequality and regional growth in Pakistan have not explicitly considered spatial analysis. In the light of the recent development in the fields of spatial econometrics and new economic geography, the overall objective of this dissertation is to re-examine previous findings on regional convergence and economic inequalities particularly those associated with the returns to human capital across Pakistan by employing standard econometric and spatial econometric techniques.

Specifically, this dissertation attempts to address the following main research questions which also constitute the basis of the research carried out for it:

1. To what extent does education explain changes in earnings inequality?
2. How have the dynamics of earnings inequality changed in the last two decades?
3. Can spatial clustering of income and average education levels explain inequality in income and education attainment across Pakistani districts?
4. Has regional income converged across Pakistani districts?

6. Outline of the study

This dissertation consists of three main substantive chapters broadly based on the themes of returns to human capital and earnings inequality, spatial analysis of inequalities in income and human development conditions, and regional convergence of per capita income across Pakistani districts.

Specifically, the overall aim of Chapter 2 is to analyze how changes in earnings inequality in Pakistan can be attributed to changes in education levels and the returns to it. In order to do this it first studies the dynamics of earnings inequality between 1993 and 2006 by calculating, comparing and decomposing some General Entropy Indices and Gini coefficient of earnings inequality within and across the provinces. It extends this analysis by complementing it with regression based approach to analyzing inequality decomposition. Moreover, since differences in gender play a vital role in labour force participation, labour supply, and wages, it takes into account the role of gender differences among returns to individual characteristics by investigating earnings inequalities due to gender differences. Finally, it identifies the contribution of changes in the structure of earnings and socio-demographic characteristics to earnings inequality in Pakistan using counterfactual analysis, and highlights the effect of unequal human capital characteristics of individuals on their earnings over time. One of the empirical findings of this chapter is that of rising income inequality within provinces. This finding led towards a deeper investigation of convergence, income, earnings and human capital inequalities at a geographic aggregation level lower than the commonly used provinces i.e. *districts*. District level analysis is beneficial because it allows for a better exploitation of the geographical characteristics of socio-economic data, and a deeper analysis of the spatial effects such as regional spillovers and spatial regimes. With this finding the dissertation shifts its focus towards carrying out a spatial analysis of the evolution of income inequalities and varying human development levels across 98 Pakistani districts.

Chapter 3 therefore investigates whether spatial clustering of income and average education levels can explain the distribution of income and education across Pakistani districts. By employing exploratory spatial data analysis techniques it describes and illustrates spatial distributions and identifies spatial outliers of human development levels in Pakistani districts in general and district wise differences in income and average education attainment levels in particular. One of the main findings of this chapter is that the distribution of district wise income inequality and education levels does indeed exhibit a significant tendency for similar levels of inequality in income and education levels to cluster. The detection of significant spatial effects implies that districts should not be viewed as independent observations as commonly done in regression analysis, and calls for the utilization of spatial econometric techniques that explicitly model this fact.

In the light of the above mentioned findings from Chapter 2 and 3, Chapter 4 investigates for the first time regional convergence (absolute and conditional) of per capita income across Pakistani districts between 1998 and 2005 using spatial and non-spatial techniques. Chapter 5 concludes this dissertation. It summarizes its main findings, discusses its limitations and possible future extensions to it, and presents policy recommendations that emerge from this dissertation in a political economy context.

In summary, this dissertation contributes to the literature and policy debate on economic inequalities, convergence and growth in emerging market economies. As it will be demonstrated in the literature review sections of Chapter 2 and 3, there is dearth of literature on inequality and growth in Pakistan that takes spatial phenomena such as agglomerations and regional spillovers into account. Therefore the goal of this dissertation is to reconsider existing research on inequalities and growth in Pakistan by taking spatial effects explicitly into account. Its empirical findings contribute to policy debate on economic inequalities, convergence and growth in South Asia, and in particular to district level research in Pakistan by applying counter factual regression analysis and spatial econometric techniques for the first time to Pakistani data.

APPENDIX

Table 1. Comparative Real GDP Growth Rates (%)									
Region/Country	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09
World GDP	4.7	2.3	3	4	5.3	4.9	5.4	4.9	-1.3
Euro Area	3.5	1.4	0.8	0.5	2	1.4	2.6	2.6	-4.2
United States	3.8	0.3	2.4	3	3.9	3.2	3.3	2.2	-2.8
China	8	7.3	8	9.3	9.5	9.9	10.7	11.4	10.7
South Asia									
India	5.4	4.2	4.4	7.5	7.3	9.2	9.2	9.2	4.5
Bangladesh	5.9	5.3	4.4	5.4	5.4	6.3	6.7	5.6	5
Sri Lanka	6	-1.4	3.7	5.9	5.2	6	7.5	6.3	2.2
Pakistan	2.2	3.4	5.1	6.4	9	5.8	6.8	4.1	2.4
Source: World Economic Outlook & World Development Report (Various Issues)									

Table 2. Estimated Loss To Economy from Terrorism (2005-2009)							
Rs billion	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	Total (2005-10)
Direct Costs	67	78	83	109	114	262	712
Indirect Costs	192	223	278	376	564	707	2,304
Total	259	301	361	484	678	969	3,052
In US $ bn	4.4	5.0	6.0	7.7	8.6	11.5	43.0
Source: Finance Division, Government of Pakistan (2009)							*: July-April

Table 3. Human Development Index (1980-2010)			
Year	Pakistan	South Asia	World
1980	0.311	0.315	0.455
1990	0.359	0.387	0.526
2000	0.440	0.440	0.57
2005	0.460	0.481	0.598
2008	0.484	0.504	0.619
2010	0.490	0.516	0.624
Source: World Bank, World Development Reports (2002, 2010)			

Table 4. Comparative Education Statistics								
Country	Adult Literacy Rate (%)				Expected Years of Schooling			
	1980	1990	2000	2010	1980	1990	2000	2010
China	65.5	77.8	90.9	94.2	8.3	8.9	9.7	11.4
India	40.8	48.2	61	68.3	6.3	7.8	8.4	10.3
Sri Lanka	86.8	86.8	90.7	90.8	9.9	11.2	--	12
Bangladesh	29.2	35.3	47.5	56.5	4.4	5	7	8.1
Nepal	20.6	33	48.6	60.3	4	7.3	9	9.6
Pakistan	25.7	25.7	42.7	54.2	3.2	4	5.3	6.8

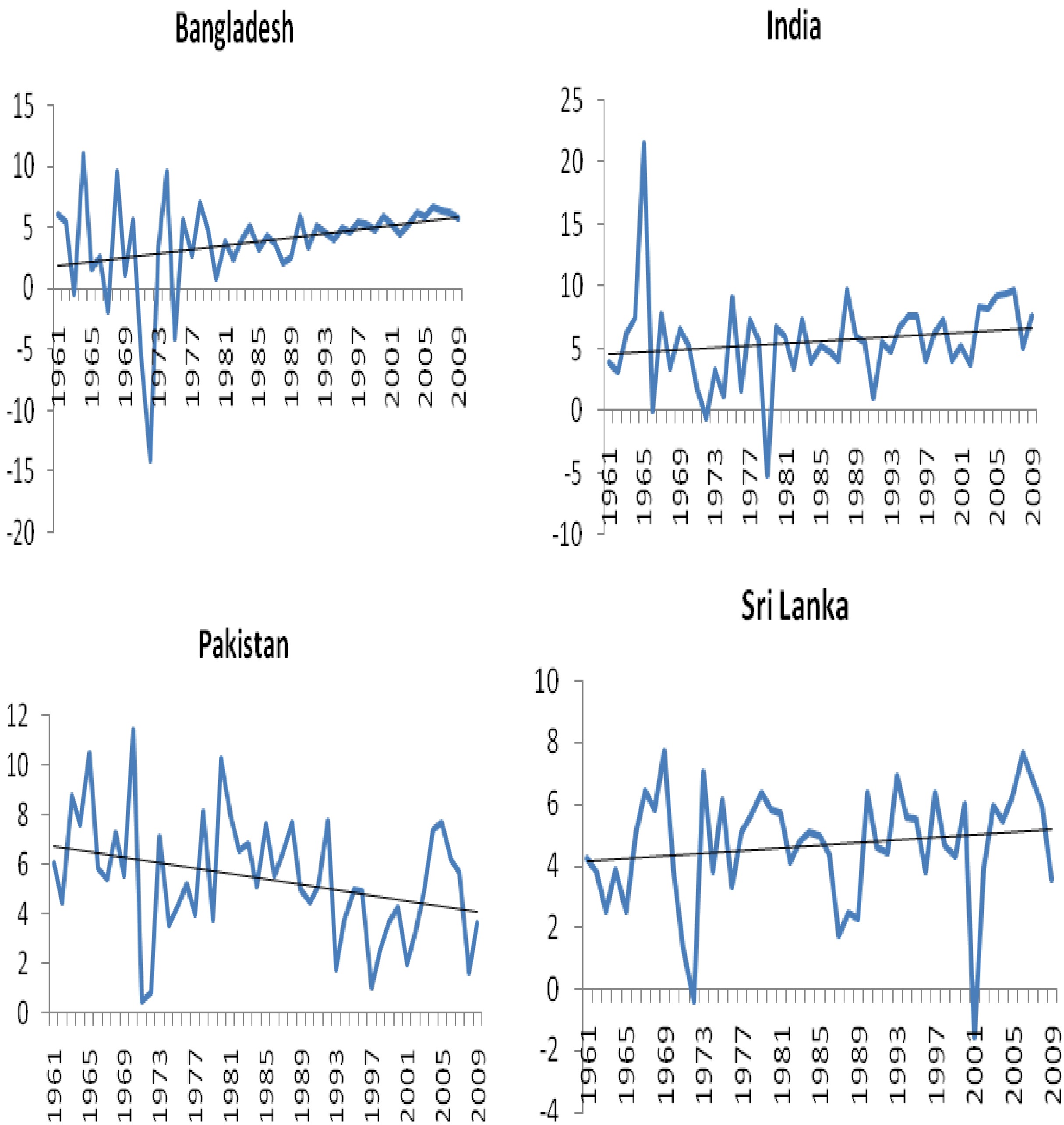
Note: Adult Literacy Rate (both sexes) is the percentage of people 15 years and above who can read and write simple statements with understanding.

Source: World Development Report (2002, 2010)

Table 5. Growth and Income Distribution in Pakistan		
Period	Growth (%)	Income Distribution
1960s	6.8	Improved
1970s	4.8	Worsened
1980s	6.5	Slightly Improved
1990s	4.6	Worsened
2000s	5.6	Worsened

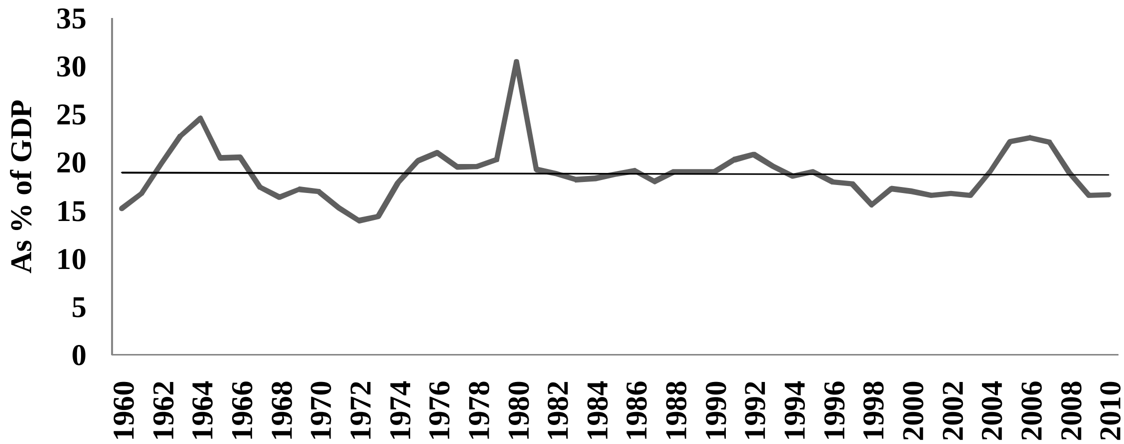
Source: Zakir and Idrees (2009)

Figure 1. Economic Growth in South Asian Countries—A Comparison



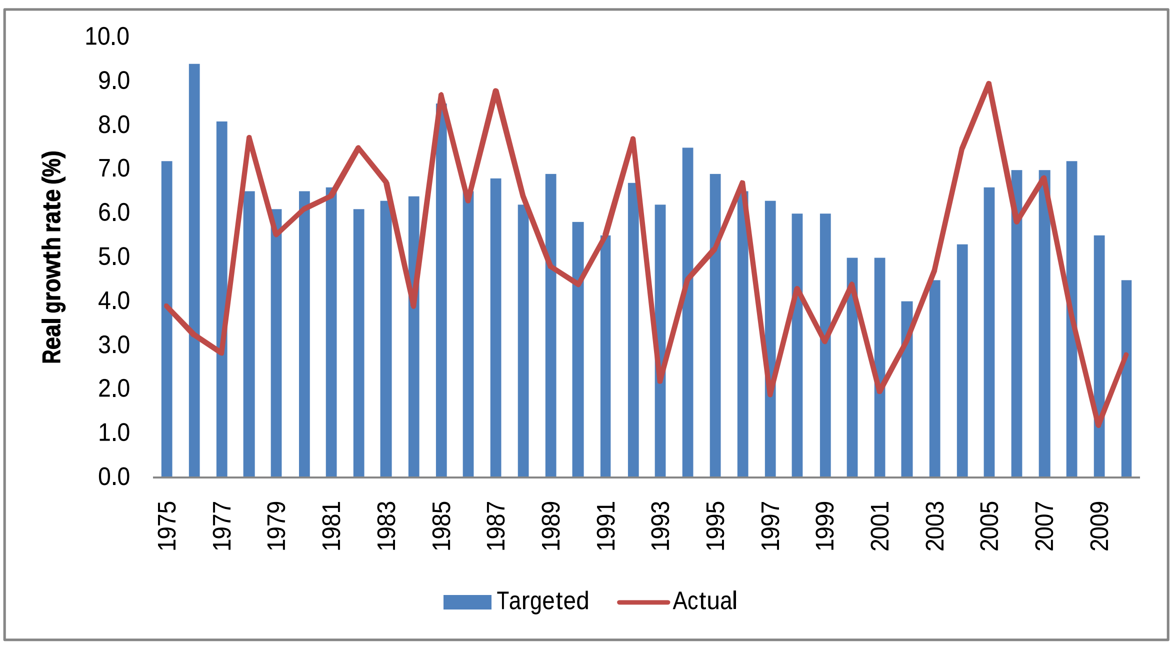
Source: World Development Report (various issues)

Figure 2. Investment to GDP Ratio (1960-2010)



Source: Pakistan Economic Survey (various issues)

Figure 3. Targeted versus Actual Economic Growth



Source: Economic Survey of Pakistan (various Issues)

Table 6. Overall Infrastructure Quality

Country	Rank
Indonesia	96
India	89
China	66
Pakistan	87
Thailand	41
Korea	20
Taiwan	19
Malaysia	27
Hong Kong	3

Source: Global Competitiveness Report (2010)

Table 7. Pakistan: State of Governance

Indicators	World Ranking out of 102 Countries
Judicial Independence	77/102
Property Rights	71/102
Favouritism in decisions of government officials	52/102
Irregular payments in tax collection	77/102
Corruption	77/102

Source: Global Competitiveness Report (2010)

Table 8. Transport Infrastructure Quality (Rank out of 125 countries)

	Railways	Road	Air
Pakistan	51	65	76
India	20	89	65
Indonesia	60	94	68
Thailand	52	35	26
Malaysia	19	24	27
China	27	50	80

Source: Global Competitiveness Report (2010)

Table 9. Consumption Inequalities across South Asia (Gini Coefficient)

Year	Bangladesh	India	Nepal	Pakistan	Sri Lanka	South Asia
1980-84	25.9	31.4	30	32	27.6	30.84
1985-89	28.8	31.5	-	22.2	35.8	31.46
1990-94	28.2	31.7	-	31.2	30.1	31.27
1995-99	33.5	37.8	36.66	32.99	34.4	36.86
2000-04	31.7	36	47.2	30.6	33.2	35.16

Source: Wagle (2008)

Table 10. Property Rights, markets and Technology (out of 139 countries)

	Property Rights	Corruption	Education & Training	Labour Market Efficiency	Technology Readiness
Malaysia	37	44	49	35	40
Thailand	87	71	59	24	68
China	43	47	60	38	78
Indonesia	74	65	66	84	91
India	61	80	85	92	86
Pakistan	99	108	123	131	109

Source: Global Competitiveness Report (2011)

Chapter 2

Can Education Explain Changes in Earnings Inequality?

Decomposition of Earnings Inequality in Pakistan (1993-2006)

Abstract

This study investigates the evolution of earnings income distribution in Pakistan between 1993 and 2006. In particular it identifies the contributions of changes in the structure of socio-economic and demographic characteristics on earnings inequality in Pakistan. In order to do so it employs a number of statistical and econometric inequality decomposition techniques. First, it analyzes the dynamics of earnings inequality by calculating, comparing and decomposing some General Entropy Indices and Gini coefficient of earnings inequality within and across provinces, and over time. This descriptive analysis is complemented with a regression based approach to inequality decomposition in order to identify the effect of changes in human capital characteristics on earnings inequality. The decomposition demonstrates the effect of unequal human capital characteristics of individuals on their earnings over time in Pakistan[5].

The main empirical findings from this study reveal that the returns to education are convex in Pakistan, and are higher for females as compared to males. Moreover, it has been shown that earnings inequality has consistently remained higher within provinces as compared to inequality between them. Finally, this study demonstrates that females in general and rural females in particular are most sensitive to policies that can affect earnings inequality and labour market outcomes in Pakistan.

Keywords: Inequality, education, decomposition, counter-factual analysis

[5] **Acknowledgments**: I'm thankful to Dr. Richard Pomfret (University of Adelaide), and Dr. Vaqar Ahmed (Ministry of Planning and Development, Pakistan) for their invaluable comments and suggestions. I am also indebted to Pakistan Institute of Development Economics (PIDE) and Rabia Awan at the Federal Bureau of Statistics (Government of Pakistan, Islamabad) for their data support. I also thank the participants in the seminar series session held at the Planning Commission, Islamabad (May 2011) for their comments on the working paper version of this chapter.

1. Introduction

The process of economic development in a country is characterized by changes in the nature of its economic activities, individual economic behaviour, and its distribution of income. While these structural changes experienced by economies are similar in nature, the "combination, sequence, and the timing of (these) changes...are always unique and unprecedented" (Bourguignon *et al*, 2005: 1). The empirical results from a recent study on the microeconomics of income distribution for six emerging economies in Latin America and East Asia reveals that all the economies experienced similar socio-economic and demographic changes during their development process but the evolution of inequality in them followed very different patterns (Bourguignon *et al*, 2005)[6]. The motivation for this chapter comes from this finding, in order to investigate the evolution of income distribution of wage earners in Pakistan between 1993 and 2006—a period characterised by macroeconomic reforms and a decent average GDP growth rate. Specifically it addresses the question of how have changing human capital characteristics and the returns to them, contributed towards the changing dynamics of income distribution.

Wage earners instead of the total working population have been selected in particular because during the period which this chapter focuses on, earnings inequality rose by 34 percent as the proportion of wage earners increased from 40 to 62 percent and for the first time began to dominate the overall inequality pattern amongst the total employed population in Pakistan[7]. Moreover while in 1993, 51 percent of the Gini coefficient of earnings inequality could be explained by differences in the levels of education amongst wage earners, in 2006 it raised to almost 62 percent of the Gini coefficient. Altogether differences in education levels explain nearly a quarter (22 percent) of the total increase in earnings inequality between 1993 and 2006. In this chapter, I argue that a large portion of this observed change in earnings inequality can be attributed to changes in education and the returns to it. In this way, policies seeking to reduce the overall income inequality in Pakistan will have to focus more on achieving greater equality in the earnings of paid employees.

Recent literature on income inequality has shown that three fundamental forces are responsible for changing income distribution dynamics:

[6] They all experienced an increase in "the average years of schooling, the share of urban population, and the participation of women in the labour force...while the average household size fell" (Bourguignon *et al*, 2005:5).
[7] Total employed population refers to self employed and wage employed in the formal sector. Earnings inequality in this chapter refers to inequality amongst wage earners only.

i) Labour-choice effect: the effect of changes in the probabilities of participation and occupational choice decisions.

ii) Price effect: the changes in the returns to population characteristics[8].

iii) Population effect: the effect of changing population characteristics (Bourguignon *et al* 2005)[9].

Keeping this in view, this chapter utilizes a regression based decomposition analysis similar to that used in Bourguignon *et al* (2005) to identify the contributions of changes in the structure of socio-economic and demographic characteristics on earnings inequality in Pakistan by simulating the effect on the distribution of earnings of the observed changes in structure of and returns to human capital characteristics. This allows us to investigate the effect of unequal human capital characteristics of individuals on their earnings over time across Pakistan.

Considering the fact that 1 in 10 of the world's out of school children is now a Pakistani (Pakistan Ministry of Education, 2011), an influx of young workers into the labour force implies a bourgeoning unskilled labour force of Pakistan. In this context, the results on earnings inequality from this chapter have important implications for Pakistan's labour and education policies, both of which require government attention on an emergency basis. The chapter begins with a literature review which forms the basis for this chapter and that highlights why is there a need for studies that combine literatures on returns to human capital characteristics and income inequality in Pakistan in Section 2. It is followed by a political economy analysis of education in Pakistan in Section 3 to brief the reader about the inequalities and quality of education in Pakistan since it influences the interpretation of the empirical results. Section 4 discusses the data and methodology that have been utilized in this chapter. It elaborates the statistical and econometric techniques—inequality measures and Mincerian Earnings Function (MEF)—that have been utilized as the basis for the micro-econometric decomposition exercise using counterfactuals. Section 5 presents and examines the empirical results obtained from the analysis of static and partial decompositions of earnings inequality, and the regression estimates obtained from MEFs. It discusses the evolution of earnings inequality, and demographic and labour force characteristics in Pakistan between 1993 and 2006. It also provides the most recent estimates for returns to different

[8] Population characteristics in this chapter are considered as human capital characteristics such as education and experience.

[9] It measures the changes in the distribution of age, gender, ethnicity, and location.

education levels (by gender and region) and gender wage gap in Pakistan. Section 6 concludes the chapter and discusses the policy implications of the results obtained in Section 5.

2. Literature Review

Although intra-country income inequality and the human capital model have been two popular research fields for Pakistan, very few studies have actually examined the impact of changing returns to human capital characteristics on the increasing income inequality over the years. This section first highlights the two separate literatures on income inequality and returns to human capital in Pakistan, and then builds a case for a need to study how the two phenomena affect each other[10].

A wide range of literature has investigated the private economic returns to education for different countries worldwide (e.g. see Pscacharopoulos, 1994; Psacharopoulos and Patrinos, 2004; Li, 2003 for China; Kingdon and Unni, 2001 for India) by applying the human capital model developed by Becker (1964) and enhanced by Mincer (1974)[11]. Most estimates for the Mincerian Earnings Function (MEF) for Pakistan however are outdated (Shabbir, 1994; Ashraf and Ashraf, 1993a, 1993b; Nasir, 1998; Siddiqui and Siddiqui, 1998; Nasir, 1999; Nasir and Nazli, 2000). The more recent studies have complemented their Mincerian estimates with a gender disaggregated analysis of earnings or occupational choice decisions of wage earners in Pakistan (Jamal *et al*, 2003; Nasir, 2002; Hyder, 2007; Abass and Peck, 2007). Another category of studies has examined the returns to education using not just the standard OLS analysis, but also by applying the Heckman procedure to deal with sample selectivity issues (Ahmed, 2005; Hyder, 2007; Aslam, 2007; Aslam, 2009; Abass and Peck, 2007), 2SLS estimates as instrumental variables to deal with measurement errors and endogeneity (Abass and Peck 2007; Aslam 2007), and household fixed effects estimation to "control for unobserved family-specific heterogeneity" (Aslam 2007). Moreover, with the exception of Jamal *et al* (2003) and Khan and Toor (2003) most MEF estimates on Pakistan are static estimates of the rate of return to education in one particular year and hence neglect

[10] Human capital is a multifaceted concept which comprises various types of investments in people. Although the choice of human capital proxy selection is a controversial topic, usually education related measures have been considered as closest proxies if not direct measures of human capital. This chapter refers to education & experience related measures when it uses the term human capital.

[11] The semi-logarithmic approach (or its modified version) is used in which the natural logarithm of monthly earnings is a linear function of experience, its square and completed years of schooling (or levels of schooling in the modified version).

the dynamics of increasing earnings income inequality (for example, Hyder, 2007; Aslam, 2007; Abbas and Peck, 2007). Several findings have consistently emerged from the above mentioned studies on the MEF for Pakistan. First, returns to education in Pakistan are low as compared to other developing countries. Second, education emerges as a productivity enhancing mechanism rather than a screening mechanism. Third, the returns to education increase with the level of education indicating a convex earnings function for Pakistan[12]. Moreover, although females have a significantly higher economic incentive to invest in education since they experience higher returns to education than males, the total earnings for males continue to be much higher than that of females suggesting the existence of possible gender discrimination. It should be noted here that these conclusions mainly emerge from studies that analyse the returns to education for workers in the non-agricultural wage employment since their incomes are most sensitive to changes in their educational qualifications (Aslam, 2007).

With regards to research on income inequality in Pakistan, most studies have utilized General Entropy measures and the Gini coefficient to estimate income inequality, while taking the *household* as a unit of measurement (Ahmad and Ludlow, 1989; Jafri and Khattak, 1995; World Bank, 2003; Anwar, 2003). Moreover, with the exception of a few studies like Idrees (2006) and Idress and Zakir (2009), the relatively recent studies on economic inequalities in Pakistan have analysed consumption expenditures instead of income as a measure for economic inequality (Haq, 1998; Jamal, 2003; Anwar, 2003). The overall conclusion from studies on the incidence of income inequality is that income inequality per capita household has: exhibited fluctuating trends between 1964 and 1987, gradually been rising since 1987, consistently remained higher in urban areas, and has been more severe than consumption inequalities in Pakistan (Idrees, 2006; Idress and Zakir, 2009).

However since 1993, the employment scene has considerably changed from being dominated by the non wage sector towards an expansion of the wage employment sector making the waged work force the largest part of the total employed work force today. This can be attributed to the changing structure of the economy, the increasing size of the services sector in Pakistan, occupational move from the agriculture labour force to non-farm wage employment, and the increased demand for educated labour, and the rural-urban migration (in particular the drift of rural youth towards urban areas) which have further increased pressure

[12] Convex returns to education means that higher levels of education have higher returns to education as compared to the lower levels of education.

on the expansion of urban social services and labour market. Hence, as earnings from employed individuals increasingly constitute a major part of the total household income in Pakistan (particularly in the urban regions), the overall pattern of household income inequality is increasingly being dominated by earnings income inequality amongst wage earners. This calls for a deeper investigation on the effects of changing socio-economic characteristics of Pakistani labour force on its income inequality. In order to do so this chapter applies the regression based inequality decomposition technique for the first time to Pakistani micro data to explicitly examine the highly connected phenomena of changing returns to human capital characteristics together with the increasing earnings inequality.

Different methods have been developed to decompose income inequality by subgroups, income sources, causal factors, and by socio-demographic characteristics (see Pyatt 1976; Sharrocks, 1980, 1982, and 1983; Feilds, 2000, 2003; Mourduch and Sicular 2002; Bourguignon *et al* 2005). In the context of Pakistan, Adams and Alderman (1992) has performed an inequality decomposition analysis but it is limited to a few districts in rural Pakistan, Nasir and Mahmood (1998) have examined the earnings inequality in Pakistan caused by the changes in the returns to education using the HIES data from 1993-94, while Idrees (2006) has decomposed earnings inequality using the General Entropy measures and the Gini coefficient by levels of education, region, provinces, gender and occupation. As demonstrated by this discussion of literature, a limited amount of economic research on Pakistan has investigated the changes in earnings inequality caused by microeconomic factors. This study attempts to fill this void by carrying out an analysis of earnings income inequality in Pakistan using both descriptive and regression based approaches to inequality decomposition and investigates the effect of the changing returns to human capital on earnings inequality in particular.

3. The Political Economy of Education in Pakistan

Education has remained an extremely neglected sector in Pakistan. Despite acknowledging it as one of the most important poverty alleviating measures, the government has not been successful in improving the consistently low net enrolment rates. This can be mainly attributed to the low level of public spending on education which has remained between 1.7 to 2.5 percent of the Gross Domestic Product (GDP) over the past two decades. This level of investment in education is much less than 4 percent of the GDP, as prescribed by UNESCO for developing countries. While in 1993, 2.4 percent of the total GDP was spent on education; in 2006 it declined to 2.1 percent of the GDP. In the fiscal year 2009-10, only 2 percent of the GDP was spent on education (Economic Survey of Pakistan, 2009-2010). This is quite contradictory to the aims of the government expressed in the current Pakistan Education Policy (released in September 2009), which aims at steadily increasing the education spending to 7 percent of the GDP[13].

The education system in Pakistan is complex and is characterized by a mixture of public and private schools with highly diverse standards between and within themselves. Amongst the public schools, there exist the 'standard' public schools/ collages and 'model' public schools/ colleges, the latter having more sophisticated campuses, faculty, and curriculum. There is a similar situation amongst the private educational institutions which have shown a mushroom growth. While some cater to the elite population in the urban areas, the others target the middle incomed class. With the limited resources directed towards education by the government, the private sector is playing an increasingly important role in the delivery of educational services in the country. The sector has expanded from about 3,300 institutions in 1983, to about 90,000 in 2009 and now caters to around 40 per cent of total enrolment. More than half of urban children go attend private educational institutions and in tertiary education alone, private sector universities account for 25 % of total enrolment (Planning Commission, 2011; Pakistan Ministry of Education, 2011). There also exists a separate category of schools called 'madressahs' which are actually religious institutions. Since only a few established 'madressahs' in the main cities teach English and Science, they are not considered to be a part of the formal education system[14] (Khan and Toor, 2003; Jamal *et al*, 2003).

Despite the above mentioned issues and, financial and institutional constraints, the overall education profile in terms of education attainment levels amongst individuals has improved since 1993. The total literacy rate has increased from 35 percent in 1993 to 53 percent in 2006 and is currently estimated to be about 57 percent in 2010, but remains to be one of the lowest in the South Asia region. This improvement may be attributed to the expanded role of the private sector and media, infrastructure development, expansion of the Information Technology (IT) and mobile technologies sector, non-governmental educational projects regarding educational awareness, and rural-urban migration leading to greater

[13]Higher Education budgets and facilities are an exception, since they have received immense importance under the Higher Education reforms, inequalities still continue to prevail.
[14] The latest education census of Pakistan has documented all the major madressahs of Pakistan.

exposure to non-agricultural work. It should also be noted that this seemingly overall improvement in the education profile has been accompanied by increasing disparities between enrolment levels of girls and boys, and between urban and rural areas. Differences in provincial budgetary allocations, ethnic factors, cultural and social norms, and the recent internal displacement of the people (due to the war on terror) have further aggravated the existing educational disparities (Sarmad *et al*, 1988; Khan and Toor, 2003).

Finally, although Pakistan aims to achieve the target of 100 percent primary school enrolment of girls and boys by 2015 (Millennium Development Goal, MDG no.2), its net primary enrolment rate in 2008 stood at only 66 percent for Pakistan[15]. At present the government is spending less than 1.5 percent of its total budget on schooling. It has also been noted that about 30,000 school buildings are in such a dangerous condition that they threaten the wellbeing of children, while almost 21,000 schools currently have no building what so ever. Analysts predict that unlike Bangladesh and India, Pakistan will not be able to achieve the MDG no. 2, as almost 1 in 10 of the World's out of school children is now a Pakistani (Ministry of Education-Pakistan, 2011).

4. Data and Methodology

4.1 Data

This study uses micro data from the Household Income and Expenditure Survey for 1992-93 and the Pakistan Social and Living Standards Survey for 2005-06, both of which are collected by the Federal Bureau of Statistics (FBS) of Pakistan. The year 2006 was chosen, since the 2005-06 was the latest micro data set when this chapter was being written. To analyse the temporal changes, I decided to go backwards at least by a decade to take into account the macro-economic reforms that characterise the 1990s , and found the micro data from the HIES (1992-93) to be a large data set with the least amount of missing observations. Other reasons that makes this an interesting time period for analysis include the facts that Pakistan saw nine Prime Ministers change between 1992 and 2006, and that this period is also characterized by expansion of the private sector and information communication technologies (ICTs).

[15] Latest UNESCO estimates for Pakistan obtained from: http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=121&IF_Language=eng&BR_Country=5860

The HIES (1992-93) data set has a sample size of 14,594 households. It is a nationally representative survey with a sample frame covering all urban and rural areas of the four provinces of Pakistan defined by the 1981 Population Census excluding the Federally Administered Tribal Areas (FATA), military restricted areas and the districts of Kohistan, Chitral and Malakand in Khyber Pakhtoonkhwa (KP). The population of the excluded areas constitutes about 4 percent of the total population.

The PSLM survey for 2005-06 provides district level welfare indicators for a sample size of about 15,453 households. It is annually produced by the FBS since 2004. It is the only socio-economic micro data that is representative at the provincial and at the district level. Moreover, the sample size of the district level data is also substantially larger than the provincial level data contained in micro data surveys such as Household Income and Expenditure Survey (HIES) of Pakistan and the Labour Force Survey (LFS) of Pakistan. This enables researchers to draw socioeconomic information which is representative at lower administrative levels as well. The data is statistically comparable with the Pakistan Census Data (1998). It provides data on districts in all four provinces of Pakistan namely; Punjab, Sindh, Khyber Pakhtoonkhwa (KP), and Balochistan. The federally administered tribal areas (FATA region) along the Afghan border in the North West and Azad Kashmir are not included in the data. The PSLM is divided into two parts. The first part contains data on socio-economic characteristics such as education, health, population welfare, immunization, pre/post natal care, family planning, water supply, and sanitation, and the second part contains household income and expenditure data. Finally, this study has also utilized macro data provided in the Economic Survey of Pakistan and the Labour Force Survey of Pakistan for various years between 1990 and 2009.

The analysis carried out in this chapter is confined to wage workers between the age of 15 and 65 with a positive income. For each year the earnings data was collected for the 'past one month' from all those in the sample who could report it on a monthly basis. With this a sample of 9,070 employed waged workers between the ages of 15 and 65, has been analysed for the year 1992-93, and a sample of 14,463 employed waged workers between the ages of 15 and 65 has been analysed for the year 2005-06. All earnings data from 2006 was deflated with the Pakistani consumer price index (CPI) of 1993. Further details on variable specification along with their descriptive statistics have been provided in Table 2.

4.2 Methodology

This chapter first analyzes the dynamics of earnings inequality between 1993 and 2006 by calculating, comparing, and decomposing some General Entropy Indices and the Gini coefficient of earnings inequality. The analysis is further disaggregated by population groups: region (urban and rural), gender, and levels of education to quantify how individual characteristics have affected inequality. However, such scalar indices limit the decomposition results to pure descriptive results because of their lack of control for endogeneity since there can be the case that the 'variables used to explain existing inequality may themselves be partly determined by income patterns' (Mourduch and Sicular, 2002: 94)[16].

As a remedy, the descriptive analysis has been complemented with the regression based approach to inequality decomposition. A labour market model is estimated at two points in time (1993 and 2006) using the Mincerian Earnings Function (MEF). The MEF is first estimated using years of education as a continuous variable. However, since the return to education is not the same for each successive year of education, I have incorporated splines of education in the MEF to examine the additional earnings associated with each successive level of education. The estimations utilize the Standard Ordinary Least Squares (OLS) procedure. Moreover in order to take into account the role of gender differences among returns to individual characteristics and gender discrimination in the labour market, this chapter also investigates earnings inequalities due to gender differences by employing the Oaxaca (1973) and Blinder (1973) decomposition techniques. Finally, an empirical framework similar to that proposed in Bourguignon *et al* (2005) has been utilized to identify the contributions of changes in the structure of earnings and socio-demographic characteristics in Pakistan by simulating the effect on the distribution of earnings from the observed changes in: returns to education and the demographic structure.

4.2.1 Inequality measures

Inequality decomposition techniques can be broadly categorized into macro and micro economic approaches. The microeconomic techniques identify the contribution of changes in the characteristics of the groups under study (males and females, regions, levels of education,

[16] For a further analysis of income inequality decomposition literature see Bourguignon *et al* (2005) and Heshmati (2004).

location etc) to the changes in inequality. A scalar measure is employed and decomposed into[17]:

1) change in inequality due to changes in the relative mean income of various groups under study
2) change in inequality due to changes in their shares in the total population
3) change in inequality due to changes within those groups

This study uses Gini coefficient, Pyatt's decomposition of Gini coeffcient (1976), and the General Entropy measures to calculate the extent of earnings income inequality 'between' and 'within' the groups under study.

The Gini coefficient is the most popular version of measuring inequality and ranges from 0 (perfect equality) to 1 (perfect inequality). It is derived from the Lorenz curve which sorts the population under study from the poorest to richest and displays the cumulative proportion of the population on the horizontal axis and the cumulative proportion of income on the vertical axis. Formally, assume x_i to be a point on the X-axis, and y_i to be a point on the Y-axis. Then:

$$Gini = 1 - \sum_{i=1}^{N}(x_i - x_{i-1})(y_i + y_{i-1}) \tag{1}$$

On the other hand, General Entropy Measures (GE) are inequality indices that differ in their sensitivity (based upon a weight α) to income differences in different parts of the income distribution (Sharrocks, 1980). Formally the value of a GE measure is:

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{N}\sum_{i=1}^{N}\left[\frac{y_i}{\bar{y}}\right]^{\alpha} - 1\right] \tag{2}$$

where $\bar{y}$ is the mean income. The values of the GE measures range between zero and infinity, with zero implying an equal distribution while higher values indicate inequality. The parameter α, represents the weight given to distances between different income levels at different parts of the income distribution. The more positive α is, the more sensitive GE(α) is to income differences at the top of the distribution; the more negative α is, the more sensitive it is to differences at the bottom of the distribution. GE(0) is the mean logarithmic deviation, GE(1) is the Theil index, and GE(2) is half the square of the coefficient of variation. Therefore while the Theil index is more sensitive than Mean Logarithm Deviation to income

[17] See Sharrocks (1980) for desirable properties of an inequality measure.

differences at the top of the distribution, the Gini coefficient is most sensitive to income differences at the mode of the distribution.

4.2.2 Decomposing earnings inequality by gender

Studies on gender discrimination in Pakistan's labour market confirm that men continue to earn higher wages even in cases where men and women have similar individual characteristics (Siddique *et al* 2006; Nasir and Nazli, 2000; Sabir and Aftab, 2007).

For estimating the extent of gender pay gap during the time under study, the Oaxaca (1973) and Blinder (1973) decomposition technique is employed[18]. Consider the following example as a simple illustration of how Oaxaca-Blinder decomposition operates. Assume that wages only depend upon years of education (Ed) and the years of experience (Exp) of a worker, and that the relationship between these characteristics is linear. The male (m) and female (f) wage equations would then be:

$$W_m = \alpha_m + \beta_m Ed_m + \gamma_m Exp_m \tag{3}$$

$$W_f = \alpha_f + \beta_f Ed_f + \gamma_f Exp_f \tag{4}$$

where β and γ are the amount by which an extra year of education and experience will raise male and female wages. The estimated regression uses the mean of the sample means. Hence if W, Ed and Exp are sample averages, the above equations exactly hold. The gap in average wage can then be written as:

$$W_m - W_f = \alpha_m + \beta_m Ed_m + \gamma_m Exp_m - (\alpha_f + \beta_f Ed_f + \gamma_f Exp_f) \tag{5}$$

If we add and subtract $\beta_m Ed_f + \gamma_m Exp_f$, we obtain:

$$W_m - W_f = \beta_m (Ed_m - Ed_f) + \gamma_m (Exp_m - Exp_f) + (\alpha_m - \alpha_f) + Ed_f (\beta_m - \beta_f) + Exp_f (\gamma_m - \gamma_f) \tag{6}$$

---

[18] I have also utilized the dummy variable approach by performing a regression analysis using the MEF in which gender is incorporated as a dummy variable to capture the effect of discrimination.
Previously Oaxaca-Blinder approach has only been utilized by Siddique and Siddique (1998) and Ashraf and Ashraf (1993), the results both have become out dated by now. Recent studies which have carried out a gender disaggregated analysis of income inequality have employed the dummy variable approach via the Mincerian Earning Function (e.g. Nasir and Nazli 2000; Aslam 2007) or the quantile regression approach (Sabir and Zehra 2007).

The total wage gap between men and women consists of an explained and an unexplained part. The first three terms in the equation represent the part of the wage gap that occurs due to differences in the 'average' characteristics between men and women. It is referred to as the 'explained' wage gap. It implies that if men and women have the same mean characteristics, these three terms would be zero. The last two terms represent the 'discriminatory' part of the wage gap which is referred to as the 'unexplained' wage gap or 'discrimination'. If the coefficients of the explanatory variables are the same for men and women, the gap occurs entirely due to differences in characteristics and there would be no discrimination. Since they are rarely the same, this part explains the differences due to unobserved variables that may influence labour productivity and how these differences are remunerated during the wage determination process in the labour market.

4.2.3 Mincerian earnings function

The rates of return to education in Pakistan have been estimated using a modified Mincer's wage equation also known as the 'human capital model'. Consider the earnings equation specified below:

$$\ln Y_i = \alpha + \beta_1 Exper_i + \beta_2 (Exper_i)^2 + \beta_3 Yrsed_i + \beta_4 (Yrsed_i)^2 + \beta_5 Male_i + \beta_6 Urban_i + \beta_7 Punjab_i + \beta_8 Sindh_i + \beta_9 KP_i + \beta_{10} Manuf_i + \beta_{11} Serv_i + u_i \tag{7}$$

where ln Y is the natural logarithm of monthly earnings of an individual. It is regressed on the individual's total years of work experience, the square of experience (to capture non linearity), total years of schooling, and the square of schooling (to capture non linearity). Moreover, β_5 to β_{11} are the estimated coefficients of discrete dummy variables: Urban, indicating whether the individual belongs to an urban area; Male, indicating whether the individual is male (to capture the gender effect); Punjab, Sindh and KP, indicating whether the individual belongs to any one of these provinces (to capture the provincial/spatial effect[19]); and Manuf and Serv, indicating whether the individual belongs to the manufacturing or services industries (to capture the industry effect[20]). The residual term

[19] Balochistan was used as the reference category.
[20] Agriculture was used as the reference category.

captures any other determinants of earnings such as unobserved ability of individuals. The definitions of the variables used in the earnings equation are elaborated in Table (1) along with their means and standard deviations in Table (2).

Since it would be misleading to assume that the rate of return to all education levels is uniform, a modified MEF is used to include different educational splines. The overall structure of the education system in Pakistan consists of primary, secondary (middle), lower secondary (matric), upper secondary (intermediate), bachelors (BA/BSc), and post graduation (MA/MSc or above) education levels. The primary education consists of five years of schooling; middle requires three more years up to grade 8; after completing two more years after middle, students obtain the Secondary School Certificate (Matriculation /Matric); another two years of non technical formal education i.e. until grade 12 makes students eligible for the Higher Secondary School Certificate (Intermediate). Further two years of education can make students obtain a Bachelor's degree (BA) or an Honours degree (BSc) after four years of education. This can be followed by a Master's program after which most students have completed sixteen years of education. This can be followed by an M. Phil which requires further two years of study or a doctorate degree requiring three years of study after a Master's degree. This information is incorporated into the MEF using the following specification:

$$\ln Y_i = \alpha + \beta_1 Exper_i + \beta_2 (Exper_i)^2 + \beta_3 Prim_i + \beta_4 Middle_i + \beta_5 Matric_i + \beta_6 Inter_i + \beta_7 BA_i + \beta_8 Postgrad_i + \beta_9 Male_i + \beta_{10} Urban_i + \beta_{11} Punjab_i + \beta_{12} Sindh_i + \beta_{13} NWFP_i + \beta_{14} Manuf_i + \beta_{15} Serv_i + u_i \quad (8)$$

Again the dependent variable is the natural logarithm of monthly earnings. β_3 to β_8 are the estimated coefficients from primary to post graduate levels of education.

The estimated coefficients are utilized to calculate the private rate of return to each level of education. The rate of return to the j^{th} level of education (r_j) is estimated by subtracting the coefficient of a previous year of education from a successive year of education and dividing it by the number of years of schooling it takes to complete the j^{th} level of education ($r_j = (\beta_j - \beta_{j-1})/n_j$[21].

[21] Note that the coefficients have to be first adjusted by ($e^{coefficient} - 1$) .This is because the "value of the coefficient of a dummy variable in the semilogarithmic regression equation is not a good estimate of the effect of that variable on the variable being explained for large values of the coefficient" (Siphambe, 2000: 292).

4.2.4 Econometric issues

Sample selectivity bias

Since this study only focuses on waged workers (between the age of 15- 65), its estimates of returns to education for wage workers are on a potentially non random sample from the population and hence possibly biased (lnY is only observed for wage workers). As a remedy Heckman's two-step procedure for sample selectivity correction can be utilized (Heckman 1979)[22]. However, while estimating earnings equations using this procedure, I was unable to find good instruments[23]. It has also been noted that the "assumptions required to validate OLS estimation of standard wage equations are not more demanding than those required to validate the results of the selection bias procedure" (Bourguignon *et al*, 2005:96) . This is because the Heckman procedure makes strong assumptions about the orthogonality between the error terms of the MEF and the wage participation model. Therefore following Ferreira and Paes de Barros (2005), I have also chosen not to present the selection-bias corrected results, and assume that errors are independently distributed. Even without the application of this procedure, it will be demonstrated (see Section 5.3) that the earnings estimates from this study are in line with the results from other recent studies on Pakistani labour market including those that have tried to use the selection-bias correction procedure (e.g Aslam, 2007; 2008).

Finally, omitted variable bias issues are also endemic to applied econometric analysis. As previously mentioned data constraints made the extraction of appropriate instruments difficult and I have not utilized any instrumental variables.

4.3 Micro-econometric decomposition of earnings inequality via Bourguignon et al (2005)

This part of the methodology–based on the work of Legovini *et al* in Bourguignon *et al* (2005)—may be seen as an extension of the Oaxaca-Blinder methodology which decomposes the effects of discrimination among two different groups of people into: differences in means caused by differences in the characteristics of the individuals in the two groups (the population effect), and differences in the remuneration of these groups (price effect).

[22] This involves finding the probability that an individual chooses wage work via a probit regression equation. A selection variable lambda (the Inverse Mills Ratio) is obtained from the probit equation and is included in the regressors of the main earnings equation. The new equation then provides results for the wage workers only and its estimates are free of selection bias.
[23] Micro data on developing countries often lacks the required information to find appropriate instruments (Jamal *et al*, 2003)

However, now the decomposition is made on the full distribution rather than on means and for every individual in the sample. It allows us to measure the relative importance of the different sources (such as changing demographics, changes in the returns to human capital characteristics) on the change in income distribution of earners across Pakistan.

As the first step, a standard labour market model is estimated for individual earners in 1993 and 2006. The model includes earnings functions for men and women in wage employment located in urban and rural locations in the four provinces of Pakistan. The simulation is performed by re-estimating the vector of incomes, changing one microeconomic factor at a time. For example the following vector of earnings is estimated for 1993:

$$\widehat{y_{93}} = \widehat{a_{93}} + \widehat{b_{93}}X_{93} \tag{9}$$

In order to determine the effect of the changes in the price of X on the distribution of y, the estimated parameter for 1993 is replaced with that for 2006 to obtain a new vector of y:

$$y_{93}^{\widehat{b_{06}}} = \widehat{a_{93}} + \widehat{b_{06}}\, X_{93} \tag{10}$$

The difference between vectors (9) and (10) describes the changes in income which can be attributed to changes in b across the entire distribution. This is carried out for each parameter and explanatory variable in the earnings equations for both 1993 and 2006. This way it is possible to obtain a detailed description of the effect of each of the microeconomic determinants on the distribution of individual incomes (Legovini *et al*, 2005).

Conventional measures of income inequality have also been estimated for each simulated vector of income to obtain estimates of the proportional contribution of each factor. The Gini coefficient, log mean deviation (E_0), and Theil Index (E_1) have been used to summarize our results. Since these measures give different weights to different parts of the income distribution, the differences in their results highlights the part of the distribution that is responsible for the change.

4.4 Empirical specification of the labour-market model utilized in counterfactual based decomposition analysis

4.4.1 Earnings equation

For this part a much simpler specification of the earnings function has been utilized in order to make the 'simulation stage of the decomposition feasible' (see Bourguignon *et al*

2005:95). Earnings are a function of skills (proxied by education and experience), and are controlled for regional variation. The earnings equation has been separately estimated for 7 labour categories on the basis of gender, urban or rural location, and four provincial dummies.

As a first step the following OLS equation is estimated:

$$\log(y) = \alpha + Edu\beta_1 + Edu^2\beta_2 + Exp\beta_3 + Exp^2\beta_4 + R\delta + \varepsilon \tag{11}$$

where: y is the real individual monthly earning; Edu is the years of schooling[24]; Exp is the years of work experience, and R consists of four regional dummies for Punjab, Khyber Pakhtoonkhwa (KP), Sindh and Balochistan.

Next to investigate what determines differences across distributions; counterfactual distributions have to be simulated. The counterfactual is obtained by simulating the fact that what would happen to the earnings of each individual in the sample if the returns to each observed characteristic had been those which were observed at time *t-1* rather than the actual returns observed at time *t* (and vice versa).

Therefore following the methodology of Legovini *et al* (2005), to simulate changes in years of education and experience, the observations have to be clustered according to gender and location. Then the distribution of these factors (mean and standard deviation) in each cluster is estimated, and finally the distribution of the 1993 cluster is replicated into the corresponding cluster in 2006 and vice versa.

Hence for each *x* in cluster *c*, the following transformation is applied:

$$x_{c93}^{dist06} = (x_{c93} - \mu_{c93})\frac{\sigma_{c06}}{\sigma_{c93}} + \mu_{c06} \tag{12}$$

where the μs and the σs are the means and the standard deviations in each cluster[25].

Finally to treat the residual, the distribution of residuals from one year is imported to another year through the *rank-preserving transformation* procedure. If we assume that the residuals are normally distributed with a zero mean, then this transformation is equivalent to

[24] The rate of return across education levels is assumed to be constant for decompositions.
[25] "To simulate regional distribution, the weights from the household surveys are used and the observations of one survey are reweighed for example 1993 with the weighs of 2006. This technique ensures that the resulting regional distribution of the population across all regions matches the observed one in 2006 and vice versa" (Legovini *et al*, 2005: 285).

"multiplying the residual observed at time t by the ratio of standard deviations observed at time *t'* and *t*" (Juhn, Murphy, and Pierce 1993; Bourguignon *et al*, 2005:35).

Having set up the basic foundation of this technique, it is now possible to carry out decomposition exercise. Assume D (y) to be a measure of income distribution, where y is earnings. Let β consist of estimated parameters in the earnings equation; let X contain the explanatory variables of education, experience, and regional location, and let ε be the error terms in the earnings equations. With this D (y) can be expressed as D (β, X, ε). The following decomposition exercise allows us to estimate the effects in the joint distribution of income by changing one or more elements contained in D (.).The price effect has been estimated by modifying β_t (estimated returns to education, experience, and location in the earnings equations). The population effect is estimated by modifying the structure of X_t (e.g. the distribution of the years of schooling and experience). The effect of unobservable factors is estimated by simulating the distribution of residuals, as described in the previous paragraph.

Let y be income and in the initial time period (1993) and y' be income in the final time period (2006). To explain the changes in income distribution between the initial year and the final year we use the following specification:

$$\Delta D = D\ (y') - D\ (y) = \ D\ (\beta', X', \varepsilon') - \ D\ (\ \beta, X, \varepsilon) \tag{13}$$

This change in distribution of income can be decomposed into three effects. The effect of changing prices, changing unobservable factors (after changing prices), and finally changing the Xs (after having changed prices and unobservable factors).

5. Results

5.1 Changing demographics and labour force characteristics

This section provides a description of the changes in the socio-demographic structure of Pakistan between 1993 and 2006 (see Table 3)[26]. In 1993 the total population was estimated to be at 120.84 million, with 82.77 million persons living in rural areas and 38.07 million living in urban areas. The total labour force stood at a 33.80 million persons and the labour

[26] The information provided here is obtained from the Economic Survey of Pakistan (1992-1993, 2005-06 and 2006-07) and the Labour Force Survey of Pakistan (1990-91 and 2005-06).

force participation rate was 27.97 percent (see Figure 1). By 2006, the total population was estimated to be 151.55 million, out of which 55.05 million persons constituted its labour force. About 18.9 percent of this labour force consisted of females. Although their participation rate has increased immensely as compared to 1993 but as compared to the South Asian region, it has continued to remain low. For example, in 2006, the female participation rate stood at 12.6 percent in Balochistan, 13 percent in KP, and 9.1 percent in Sindh. It is believed that female participation is usually under reported due to cultural factors, under reporting in non wage sectors, and unavailability of suitable jobs especially in rural areas (Labour Force Survey, 1991)[27].

As for the Pakistani labour force, the proportion of currently active population (as measured by the Crude Activity Rate) has tremendously increased[28]. Out of this the fraction of young people in the labour force (15-35 age group) has increased by 21 percent as family size increased by 6 percent during this period. The higher participation of younger age groups into the labour force generates a proportional increase in wage employment as they are less likely to opt for self employment. Hence, wage work replaced self employment as the preferred employment category between 1993 and 2006 as the proportion of paid employees has risen by 13 percent and their real wages grew by 19 percent. The skill premium has increased by 13 percent for average skilled workers (measured by those having completed 10 years of schooling).

Another major feature that has characterises structural changes in Pakistan is how the education structure of the employed population has changed as the total proportion of population having passed grade 10 (known as the matriculate level in Pakistan) increased from 9 percent in 1993 to 22 percent in 2006. Moreover, the total proportion of persons having passed secondary education increased by 57 percent, while those with an education level below primary have reduced by 5 percent between 1993 and 2006. Education attainment increased for males in rural and urban areas as the percentage of males having completed primary education grew by 25 percent, secondary education grew by 96 percent and basic

---

[27] An increasing amount of women are entering the non agricultural labour force in Pakistan. However, since most of them (in rural areas and towns) own or run micro enterprises, they do not enter the formal economy. There are very few estimates of the informal economy in Pakistan. According to a recent study, about 30 percent of the GDP remained unmeasured as a percentage of the total GDP (Arby, Malki and Hanif, 2010).The existence of a large informal sector can bias the results based on data on the formal economy. However this issues is remains out of this chapter's scope.

[28] The Labour Force Survey of Pakistan defines the 'crude activity rate' as the currently active population expressed as a percentage of the total population in Pakistan.

tertiary education (Matric and above) grew by 171 percent in urban areas. In rural areas the percentage of males having completed primary increased by 32 percent, secondary education increased by 101 percent, and those with education levels of Matric and above rose by 244 percent. Although female enrolment rates have considerably improved, their education attainment at all levels dropped as compared to 1993 (See Figure 2 and Table 6).

5.2 Earnings inequality in Pakistan

Earnings inequality increased by 34 percent between 1993 and 2006[29]. This paper has decomposed earnings inequality into the following subgroups: location (urban/rural), gender (males/females), level of education (primary, middle, matric, inter, bachelors, post-graduation), to observe how it has changed 'between' and 'within' them. Although the discussion of results below only highlights inequality as measured by the changes in Gini coefficient, Table 4 reports the results using the Theil index, mean log deviation, and the modified coefficient of variation as well[30].

i. *Earnings inequality by location (urban/rural)*: In 1993, earnings inequality 'within' rural and urban areas accounted for 50 percent of the total inequality; while the earnings inequality 'between' urban and rural locations was about 25 percent of the total inequality. By 2006, 'within' location inequality accounted for 48 percent and the between location inequality was 28 percent of the total inequality. Inter-temporal changes in the Gini coefficient show that between the 13 years under consideration, within location earnings inequality increased by 28 percent; while earnings inequality between rural and urban locations increased by 48 percent.

ii. *Earnings inequality by gender*: In 1993, within group earnings inequality amongst males and amongst females accounted for 87 percent of total earnings inequality; while only 4 percent of earnings inequality was attributed to earnings inequality between males and females. In 2006, within earnings inequality amongst males and females reduced to about 73 percent of the total inequality, while earnings inequality between males and females as a proportion of the total inequality increased to 16 percent. Hence, between 1993 and 2006, earnings inequality within males and females

[29] As measured by percentage change in the Gini coefficient
[30] I used Pyatt's (1976) methodology to decompose the Gini coefficient.

increased by 10 percent, while earnings inequality between males and females increased even more.

iii. *Earnings inequality by the level of education*: In 1993 earnings inequality ‘within’ each of the 7 education levels, accounted for 17 percent of total earnings inequality, while earnings inequality ‘between’ different education levels constituted 51 percent of the total earnings inequality. In 2006, ‘within’ education levels earnings inequality reduced to 12 percent of the total Gini coeffcient, while earnings inequality between various education levels increased to 61 percent of the total Gini coefficient. Altogether between 1993 and 2006, earnings inequality within educational levels decreased by 10 percent, while earnings inequality amongst individuals with different education levels increased by 57 percent.

The dynamics of earnings inequality between 1993 and 2006 demonstrate that within group inequalities continue to constitute a greater part of the total earnings inequality over the years. Within group inequality is even more than between group inequality when measured by General entropy measures since they are sensitive to different levels of the earnings income distribution. This is in accordance with international evidence on income inequality which suggests that in most developing countries within group earnings inequality usually constitute about 75 percent of total income inequality (World Bank Handbook on Inequality, 2005). Moreover, although within group inequality has increased (at a decreasing rate) for location and gender subgroups, it has decreased for inequality within different levels of education. Finally, earnings inequality between groups has continued to rise at an increasing rate suggesting widening disparities between them.

5.3 OLS Results

5.3.1 The MEF estimates for Pakistan

The estimated results of equations 7 and 8 have been reported in Table 7 for waged workers between 15 to 65 years of age. The coefficients are highly significant and bear signs in accordance with the human capital model for both 1993 and 2006. The overall estimated results also indicate a good fit of the model. Between 1993 and 2006, the marginal returns to an additional year of schooling increased by 38 percent. The positive and significant coefficient of ‘male’ in both the years demonstrates how males continue to earn more than females over time. With respect to location, earnings continued to remain significantly higher in urban areas than rural areas as the coefficient to ‘urban’ increased by 24 percent between 1993 and 2006. These results are similar to the findings of Jamal *et al* (2003) and Nasir and Nazli (2000).

5.3.2 MEF estimates by province

Since significant inter provincial differences are observed in the returns to individual characteristics (equations 7 and 8), four separate regression equations have been estimated for each of the four provinces of Pakistan to examine the spatial effect in differences to returns over time (See Tables 6.4, 6.5, 6.6, 8, 9, 10, 11 for details) .

It is observed that between 1993 and 2006, the maximum change in the returns to an education level was associated with the lower secondary level ‘middle’ in Punjab, Sindh, and Balochistan, as returns to middle increased from 2 to 6 percent. However, in KP, the returns to primary changed the most, as they rose from 2 percent in 1993 to 7 percent in 2006. The returns to post graduation were highest in Balochistan in 1993 (22 percent) and in 2006 (73 percent). This finding is in accordance with the findings of Psacharopoulos and Patrinos, (2002) which suggests that low levels of development (as present in Balochistan) are associated with higher returns to education. In Balochistan high levels to post graduation levels may be due to a greater demand for skilled labour because of increasing development activities or due to the fact that the supply of skilled labour is still low as compared to other provinces.

5.3.3 MEF estimates by level of education

Between 1993 and 2006—except for primary—the returns to all the rest of education levels rose substantially. In 1993, the private returns to education were 3, 2, 5, 7, 9, 14, and 21 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. However in 2006, the private returns to education were 2, 3, 6, 9, 14, 20, and 26 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. Hence tertiary education has been most profitable over the years as the increase in returns were the largest for the three highest education levels. Rates of return to primary and middle education levels have been low probably because the earnings differential between workers who have completed primary and middle, and workers with no education has been small. This suggests that the monetary incentives of having primary or lower secondary education are very low. Workers with these educational levels have increasingly been pushed into very low paying jobs or into the activities in the informal sector and have to now compete with illiterate working population. This phenomenon has been noted as a dominant feature of developing countries experiencing an influx of workers in their labour markets (Siphambe, 2000).

5.3.4 MEF estimates by gender

In 1993, the private returns to education for male waged workers were 2, 2, 5, 7, 10, and 15 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. By 2006, the returns were estimated to be 2, 2, 3, 5, 8, 12, and 17 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. Therefore, males primarily experienced a rise in the returns to higher secondary, bachelors, and post graduation education levels. Similarly in 1993, the private rate of return to education levels for females was estimated to be 9, 3, 12, 17, 15, and 28 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. By 2006, they were 5, 5, 19, 20, 33 and 51 percent for primary, middle, matric, inter, bachelors, and post graduation respectively. As evident in both the years, the highest returns were associated with post graduation education level (i.e. having completed masters or above) for both the genders. However, the coefficients at all education levels were significantly higher for females than for males.

While the returns to experience for males, decreased by 4 percent; they increased by 17 percent for females between 1993 and 2006. As expected, males and females in urban areas earned higher than their counter parts in rural areas. This disparity rose by 48 percent for men, and 79 percent for women, between 1993 and 2006. The returns to ‘urban’ region alone remained higher for men than for women. It should be noted that despite the higher returns to education for females, men continue to earn more in both rural and urban areas. Wage working females are predominantly associated with the services sector, and in particular with low paying jobs in social and community service areas. In 1993, 62 percent of wage working females belonged to the services industry as compared to 54 percent females in 2006. While these jobs provide flexibility in timings and require limited travelling, they are also associated with low wages and short term contracts. Higher female participation in these jobs may in part be due their own job preference but the discrimination in the labour market could also be playing a role in forcing females to opt for them as demonstrated by Aslam (2008) and Riboud *et al* (2006).

5.3.5 Gender wage gap

The results from Oaxaca-Blinder decomposition carried out for male and female wage differentials are reported in Tables (5.1) and (5.2). The estimates from the MEF (which utilized the dummy variable approach to estimating wage differentials) suggest that females (and within females, urban females) have higher returns to an additional year of schooling than males. Provincial dummies in 1993 and 2006 suggest that urban females in Balochistan experience maximum gains from an additional year of schooling.

In 1993, the net gender wage difference (between the OLS estimates of logarithm of wages) was 0.399. Out of 0.399, only 0.0027 was 'explained' by better male characteristics (when standardized by male means), while 0.396 of the gap remained 'unexplained'. This means that 99 percent of the gender wage was unexplained in 1993 indicating a large role of gender discrimination in wages. In 2006, the gender wage difference increased to 1.167, out of which 0. 278 was 'explained' and 1.069 was 'unexplained'. Hence by 2006, 64 percent of the wage difference was unexplained. These results indicate that the 'unexplained' component of gender wage difference has declined by 35 percent between 1993 and 2006. While gender wage discrimination has declined, it still constitutes a large part of the gender was difference. However, in order to be more confident about our results we have to consider the number of days worked by both males and females in the two years. Since our data suggests that only 2620 females as compared to 19009 males worked for more than 300 days in 1993 and only 1668 females as compared to 14441 males worked for more than 300 days in 2006, the large discrimination may be a result of factors that have influenced low female participation in the labour force.

5.4 Characterization of changes in earnings inequality and decomposition results

As the first step of the simulation exercise, a simple Mincerian function (equation 11) has been re-estimated again, but this time only for four categories: urban males, urban females, rural males, and rural females in order to capture gender and regional effects. Detailed results have been presented in Table 5. As discussed in Section 5.3.3, the returns to education function have become more convex during 1993 and 2006 implying that marginal returns to education are lowest for individuals with lower educational qualifications. The wage gap related to skills (measured by the returns to various years of education) has widened (see

Tables 4 and 5) as the returns to low and medium levels of education declined, while returns to higher education levels increased for both urban and rural males and females[31].

This widening wage gap in returns to various education levels can be attributed to different demand and supply factors. As the supply of primary qualified wage earners has increased over the year, the wage rewards to primary education have fallen. On the demand side, it is conjectured that *skill biased technological change* could be a major reason behind greater variance in wages (Bound and Jhonson, 1992; Juhn *et al*, 1993; Autor *et al*, 1999; Fräßdorf *et al*, 2008; Colclough *et al*, 2009). This refers to a shift in production technology that increases the relative productivity and hence the demand of skilled labourers as compared to the unskilled[32]. This factor bias has put "technological change at the centerstage of the income distribution debate" (Violante, 2008)[33]. Traditionally technical change is considered to be factor neutral, but at least in its adoption phase many developed and developing countries have experienced increased returns to skills after trade liberalization (Harrison, 1997; Beyer *et al*, 1999; Robertson, 2000 Berman *et al*, 1998; 2003). Studies on the impact of trade liberalization in Pakistan have also found evidence of shifts in demand towards skills (Berman and Machin, 2000; Sabir, 2004).

Changes in the structure of returns to education coupled with improvement in the distribution of overall education can generate large increases in income inequality. A comparison of education attainment across our sample, suggests that percentage of workers with secondary education or above was much larger in 2006 than in 1993. This was accompanied by much larger marginal rates of return to education at higher levels of education as compared to lower levels in 2006 than in 1993. Therefore convexification of returns to education in Pakistan has also been unequalizing over time. Altogether changes in returns to various educational levels account for almost a 23 percent increase in earnings income inequality between 1993 and 2006. This has led to an increasing number of workers with low education levels at the bottom of the income distribution. Bourguignon *et al* (2005) explains that poorer workers experience greater constraints to achieving higher levels of

[31] Only in 1993, the coefficient for returns to education for rural females was not statistically significant from zero. However estimates for each level of education for rural females in 1993 show that only the coefficient for primary education was statistically different from zero.

[32] In-flowing capital embodies in-flowing technology, which is assumed to be skill-biased because the new technology is mainly designed in the industrialized world, which is skill intensive, and because there is evidence that new technology is skill-biased within the industrialized world (Berman *et al*, 1998).

[33] "A large number of economic models in the literature provides a foundation for Skill Biased Technological Change (for surveys, see Acemoglu, 2002; Aghion, 2002; Hornstein *et al*., 2005). The central tenet of all these theories is technology-skill complementarity" (Violente, 2008).

education due to demand side factors (such as inability to afford schooling) and supply side factors (such as unavailability of schools and educational infrastructure). With rapidly changing labour market situations, such groups suffer the most as they are unable to catch up with the changing demand situation in the labour market.

An important implication of our findings for Pakistan is that over the years, the positive yet reduced returns to primary education have reduced the poverty reducing element of attaining primary education since the wage increment associated with it is much lower than before. An important policy concern here is to think about whether individuals who have only attained primary education (because it was free) or who can only afford to obtain primary education have any incentive at all to acquire education. However it still remains important to take into account the non-monetary benefits of acquiring basic education (increased awareness, reduced crime, improved civic sense, lower fertility rates etc) and its positive socio-economic externalities. This implication also puts pressure to expand the education system at the secondary level and above in order to increase the average duration of schooling[34].

Furthermore, experience was more valued in 2006 than in 1993 perhaps as a result of a shift in the demand for goods and services requiring more experience due to the skill biased technological growth. The effect of experience on earnings remained positive and concave. The returns to experience have risen over the 13 years, but not as much as for males (urban and rural) as they have for females (rural and urban). While males experienced a 1 percent increase in earnings for an additional year of work, females experienced a 4 percent increase. The returns to experience increased by 8 percent for urban males, 55 percent for urban females, 19 percent for rural males, and 43 percent for rural females between 1993 and 2006. Therefore although 'experience' as a characteristic has also had an unequalizing effect on earnings income distribution (as the Gini coefficient for years of experience increased from 0.34 to 0.35), but this has not been as large as the effect of returns to education.

The regional effect (rural/urban and the effect of belonging to the provinces) has also been unequalizing. The returns to provinces show that in 1993, regional effects (measured by the effect of belonging to one of the four provinces of Pakistan, with Balochistan as the

[34] This chapter does not discuss the effect of schooling quality on the demand for schooling. Colclough (2009) argues that if the returns to primary education are declining (partly) due to declining schooling quality, then the demand for primary education "as a terminal stage" will be reduced.

reference category) were negative and highly significant. In 2006 however, although regional effects still remained negative but some of them became statistically insignificant. This change confirms uneven growth across the provinces (and urban & rural regions) and that their conditions have extensively diverged from each other. It can also been taken as an evidence of the fact that the influence of regional differences on wage determination have reduced over time across Pakistan.

Finally to analyse the possible effects of unobservable factors, we observe that the variance of the residuals of the earnings function for males and females (both in rural and urban areas) increased during 1993-2006. An intuitive explanation for this is that the unobservable talents (cognitive skills and non cognitive factors) of males and females such as innate ability increased due to the expansion of media, information technology sector, contributing to a raise in the general awareness levels.

5.5 Decomposing changes in earnings inequality

Four counterfactuals have been simulated for four population sub groups: urban males, urban females, rural males, and rural females. Since decompositions are path dependent, this study considers both 1993 and 2006 alternatively as base years. While Table 2 reports the simulated change in earnings inequality by simulating the 2006 model on the 1993 distribution of earnings, Table 3 reports the results from the simulation of the 1993 model on the distribution of earnings in 2006.

The first simulated change demonstrates what individual earnings would have been in 1993 (or 2006) if the returns to each of the observed individual characteristics had been those observed in year 2006 (or 1993). However, the returns to unobserved factors and the distribution of the individual characteristics were kept unchanged. This counterfactual shows the change in the earnings distribution due to the ‘price effect’. The second simulated change incorporates the effect of changing unobservable factors (after having changed prices) by performing a counterfactual on the distribution of the random term. It imports the distribution of the residuals from 2006 (or 1993) to 1993 (or 2006) through the rank preservation process. The third simulated change highlights the effect of changing the distribution of education and experience (after having changed prices and residuals) on the distribution of earnings. It simulates the distribution of individual endowments of education and experience in 2006 (1993) on the distribution of earnings in 1993 (2006). Finally, the fourth simulated change shows the effect of changing the returns to education only (keeping all the other elements of the equation the same) on earnings inequality.

To understand the ex-ante changes in the distribution and returns to individual characteristics, 1993 is kept as the base year and 2006 coefficients are replicated into the 1993 distribution[35]. After performing simulation 1, urban females experience the least increase in their mean wages (15 percent), but the largest increase in their wage inequality (73 percent). Simulation 2 and simulation 3 affects rural males the most as their wage inequality increases by 103 percent and 65 percent respectively (see Table 15 and Figure 4). To examine the effect of only modifying education related parameters, simulation 4 shows that except for rural females, it has a highly unequalizing impact for urban males, urban females, and rural males whose wage inequality increases by 23 percent, 63 percent, and 53 percent respectively. Rural females however experience a slight reduction of 0.8 percent under this simulation.

To observe ex-post changes, we replicate 1993 coefficients and distribution into that of 2006. By simply importing the 2006 returns to individual characteristics into 1993 (simulation 1), urban women experience the maximum reduction in their wage inequality (23 percent), while the Gini coefficient for rural males increases the most (46 percent). After combining simulation 1 with a rank preservation process (simulation 2) rural males experience the maximum increase in their wage inequality (19 percent). Under this simulation, rural females experience the maximum increase their mean incomes (94 percent) which is accompanied by a 19 percent reduction in their wage inequality. After combining the two previous simulations with a procedure of replicating the distribution of education and experience (the Xs) of 2006 into 1993, we perform simulation 3 which turns out to be the most equalizing for all four population groups under consideration. It results in reducing wage inequality by 3 percent for urban males, 44 percent for urban females, 19 percent for rural males, and 40 percent for rural females. Finally, only replicating the returns to education of 2006 into the distribution of 1993 (simulation 4), proves to be equalizing only for urban females and rural females, whose wage inequality reduces by 23 percent and 16 percent

[35] Here the results have been discussed in terms of changes in the Gini coefficient here. Tables 18 and 17 (in the appendix) report the changes in earnings distribution through Gini coefficient, Theil Index, and log mean deviation.

respectively. On the other hand it proves to be unequalizing for males as it increases wage inequality for urban males by 40 percent and for rural males by 45 percent.

The impact of the simulations on changes in mean real monthly earnings has also been examined. Table 14 reports the changes in mean monthly wages because of the price effect after having replaced the vector coefficients of 1993 by the vector of coefficients of 2006 while keeping the residual constant in part(a) and the vice versa in part(b). It shows the evolution of monthly wages of urban males and females, and rural males and females, while keeping their individual characteristics constant. In part (a) it is observed that when 1993 is the base year, rural males and females benefit the most by having the 2006 returns to individual characteristics. Part (b) however, demonstrates that rural females and urban females would have experienced a substantial increase of 746 percent and 892 percent in their monthly wages if the returns to their characteristics were the same as that in 1993. This is explained by the improvement in the distribution of their characteristics. For example by 2006, female educational enrolment has increased intensively (100% increase or more) at all education levels (Figure 2). Moreover, rural females benefit from the maximum increase in their mean earnings under all four simulations when 2006 is considered as the base year. However, when 1993 is considered as the base year, the simulation measuring the price effect increases the mean earnings the most only. Under the other three conditions (simulations 2, 3, and 4), urban women experience the maximum increase in their mean earnings (See Table 15).

Results from the decomposition exercise demonstrate that the rise in earnings inequality of individuals in Pakistan between 1993 and 2006 is a consequence of strongly unequalizing effects of education and regions, and weakly unequalizing effects of experience and unobservable factors. First, as shown by the price effect (simulation 1), changes in the returns to education have particularly affected earnings inequality for urban females and rural males the most over time (Table 14). Two other factors that have contributed towards increased overall earnings inequality include: the rise in the variance of unobserved individual characteristics, and changes in the structure of schooling and experience (socio-demographic changes depicted in simulation 3). The empirical results also demonstrate that females in general and rural females in particular are most susceptible to changes in policies that can affect earnings inequality (See Table16). Finally, as shown by simulation 4, changes in education related parameters can highly affect wage inequality for all four segments of the population, but affect rural females the most.

It should be noted here that this analysis remains preliminary in three ways. First, it has only analysed the distribution of earnings income. Second, it has not taken into account the effect of changing labour force participation decisions over the years on income inequality. Third, this analysis has decomposed changes in earnings inequality at an individual level only. In most developing countries, income is a household concern and hence it is equally important to analyse the effect of structural changes in an economy on household poverty and inequality. These caveats are being addressed in a working paper version of this chapter.

6. Conclusions and Policy Recommendations

Between 1993 and 2006 earnings inequality in Pakistan rose by approximately 34 percent as the proportion of wage earners in its working population increased from 40 to 62 percent[36]. This steady rise in earnings income inequality now dominates the overall inequality pattern amongst its total employed population. Since nearly a quarter (22 percent) of the total increase in earnings inequality between 1993 and 2006 is due to differences in education levels, this chapter investigated the extent to which can this observed change in earnings inequality be attributed to changes in returns to human capital across Pakistan between 1993 and 2006? In this way it makes a timely contribution towards the literature on earnings inequality in Pakistan and has presented the latest estimates for returns to education across Pakistan, taking into account gender and spatial differences.

The Pakistani labour force increased by over 60 percent between 1993 and 2006 during which the proportion of wage earners increased by 13 percent. The chapter first examined how the dynamics of their earnings inequality over the past decade by location, gender and education levels. Between 1993 and 2006, earnings inequality within rural and urban regions has increased more than between rural and urban regions. Similarly it remained higher amongst women as compared to men in both the years under study. However earnings inequality between different levels of education has increased much more than within different education levels reflecting the inequality generating effect of education levels. These findings of greater within group inequalities are typical of emerging market economies

[36] When this chapter was written, the latest available household data set (PSLM) was used to carry out the analysis for the fiscal year 2005-06. The next PSLM survey was carried out in 2007-08 but was not made available officially until late 2009.

during their development phase and are similar to the results from previous studies (see WB, 2005, for global inequalities; Idrees, 2006, for the case of Pakistan).

Next the chapter investigated how have the returns to human capital changed over the period under investigation with the aim of updating existing evidence on returns to skills and gender discrimination in the Pakistani labour market. The following main results emerge from this empirical analysis:

1) The marginal returns to an additional year of schooling increased by 38 percent, and remained significantly higher in urban areas, and in manufacturing and services sectors between 1993 and 2006. Moreover except for primary education, the returns to each successive level of education have substantially increased between 1993 and 2006 with female returns being higher as compared to males throughout.
2) To examine the spatial differences in the returns to education over time, four separate Mincerian earnings functions (MEFs) were estimated for each of the provinces. Balochistan has had the highest returns to all education levels in 1993 and in 2006. This could be explained by the overall enrolment rates in Balochistan which have always remained the lowest in Pakistan.
3) When MEF estimates were obtained for each of the seven educational levels in Pakistan, it has been observed that the returns to all education levels rose substantially between 1993 and 2006. While the returns to primary education have not substantially risen for Punjab and Sindh (the two provinces with the largest amount of primary graduates), tertiary education become even more profitable in 2006 as compared to 1993 in all four provinces. The falling marginal returns to primary education are a dominant feature of emerging market economies that are experiencing an influx of workers. With an increasing proportion of population between the ages 15 and 65 (i.e. working population) and due to skill biased technology change, this chapter demonstrated that Pakistan too is now experiencing this phenomenon. However despite increased convexity of returns to skills, the returns to primary are still higher in Pakistan as compared to other developing countries (Aslam, 2009). This perhaps relates to a possible un-met demand for low-skilled labour in some industries. Low schooling quality is another reason which may explain falling returns to primary education in Pakistan. It has been observed that having completed primary educated for many individuals may not guarantee the acquisition of basic literacy and numeracy skills (Colclough *et al*, 2009).

4) The findings from this chapter clearly demonstrate that the returns to skills of females are much higher than the returns to skills for men in the Pakistani labour market (Section 5.3.4). However the Oaxaca-Blinder decomposition results demonstrate a large element of gender discrimination (Section 5.4). These results are consistent with the findings of other similar studies on Pakistan (Jamal *et al*, 3003, Abbas, 2004; Aslam, 2007; 2008; 2009). Although higher returns to skill for females should imply that parents would favour education amongst girls as compared to boys, it is not so in reality. Various cultural factors in Pakistan (such as parents living with sons when old) have contributed towards a gender bias against female education. Moreover in a male dominated society, females find it particularly hard to enter the labour market after having suffered from under investment in their education.
5) Unequal returns to various levels of education have not only contributed towards increasing earnings inequality between provinces but also earnings inequality within them. The uneven spread of gains associated with increased distribution of education can also threaten social solidarity, as ethnicity may be perceived as a determinant of wage earnings. Clearly this calls for a detailed spatial analysis of educational and income inequalities (see Chapter 3 for details).

All of the above results are validated by the evidence from recent studies on Pakistan such as Jamal *et al* (2003) and Aslam (2007; 2008; 2009), and studies on emerging market economies similar to Pakistan (in terms of earnings inequality patterns in the past 2 decades) such as Kingdon *et al* (2008), Vasudeva-Datta (2004) for India, Li (2003) for the case China, and Bouillon *et al* (1999; 2005) for Mexico. Moreover, micro-econometric decomposition technique has been utilized to separate the effects of changing labour market and human capital characteristics on the evolution of earnings income distribution over the past decade. Micro-simulations were carried out for the following four counter factual questions respectively, for urban males, urban females, rural males, and rural females in Pakistan:

1) How would have individual earnings changed in 1993 (or 2006) if the returns to each of the individual characteristics had been those observed in the year 2006 (or 1993)? This was carried out to extract the price effect.
2) After having changed the returns, what would have happened to individual earnings in 1993 (or 2006) if the distribution of the unobservable factors had been those in the year 2006 (or 1993)? This was carried out to examine the role of unobservable individual characteristics.

3) How would the earnings income distribution have changed if the distribution of individual characteristics (education and earnings) in 1993 (or 2006) were replaced by the distribution of individual characteristics in 2006 (or 1993)? This was carried out to analyse the effect of individual endowments on earnings.
4) How would have the distribution of earnings be changed if only the returns to education in 1993 (or 2006) were exchanged with the returns to education in 2006 (or 1993)? This was carried out to separate the effect of education from all other factors on earnings income.

The following results emerged from the ex-ante exercise i.e. results for 1993 using 2006 coefficients and distribution:

1) From the first simulation, urban females would have been affected the most. They would have experienced the maximum increase in their earnings inequality while their mean wages would experience the least amount of rise as compared to urban males, rural males, and rural females.
2) Simulations 2 and 3, resulted in increasing the wage inequality for rural men by the largest amount.
3) Finally replacing the returns to education for 1993 with those in 2006 proved to be highly unequalizing for rural males, urban males, and urban females.

The following results emerged from the ex-post exercise i.e. results for 2006 using 1993 coefficients and distribution:

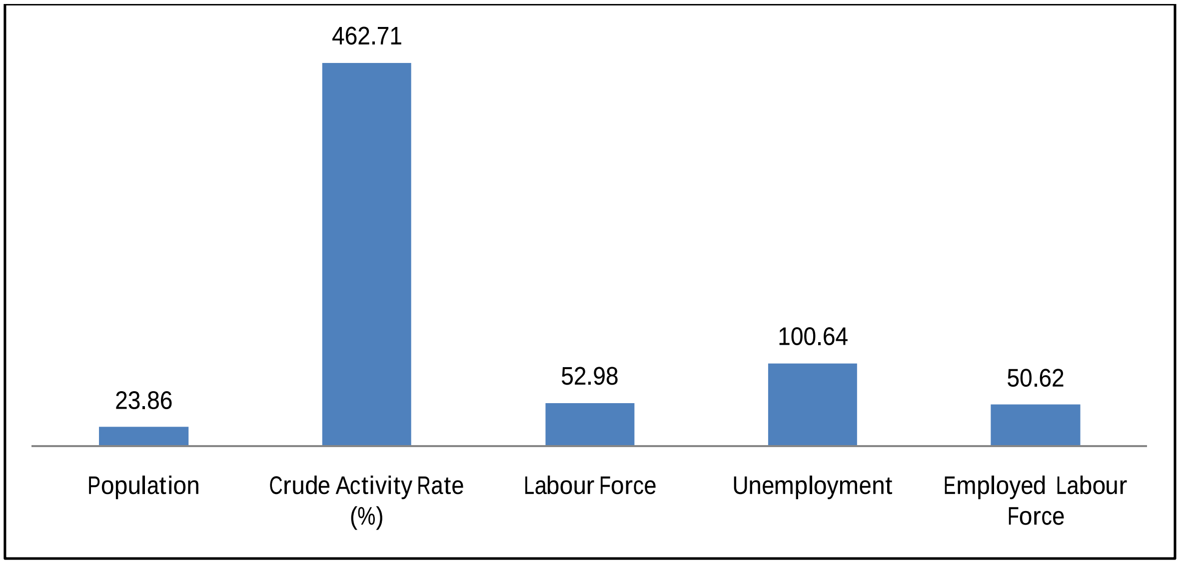
1) Simulation 1 attributed the maximum benefits to urban females in terms of reducing their wage inequality but also resulted in the largest increase in the wage inequality amongst rural males.
2) Simulation 2 resulted in reducing wage inequality for both urban females and rural females, but again resulted in generating higher wage inequality for rural males.
3) Simulation 3 appeared to have the most equalizing effect for all population groups under study, as it reduced wage inequality for all of them.
4) Simulation 4 resulted in reducing wage inequalities for rural and urban females, while increasing wage inequality for rural and urban males.

The counterfactual analysis has highlighted the importance of education, in influencing the labour market returns and the distribution of earnings income in Pakistan. Most importantly, the micro simulations have demonstrated how females in general and rural females in particular (due to the high returns to their skills) are most sensitive to changes in policies that can affect labour market outcomes. This inequality reducing role of education supports the case for enhanced public investments in female education in Pakistan (Kingdon and Soderbom, 2008).

At the same time the unequalizing effects of returns to education also demonstrate that educational progress does not necessarily imply an improvement in the distribution of earnings income. Although this analysis has just been carried for the distribution of earnings income, several studies have observed similar results for the distribution of total income (see Bourguignon *et al*, 2005 for examples on Indonesia and Mexico). Hence if earnings-education profile continues to disadvantage workers with lower levels of education (convex), increased distribution of basic education might not even contribute much towards poverty reduction as presumed by the Millennium Development Goals. The policy implication for this finding lies in targeted government interventions that can facilitate in distributing the gains of education across different groups in the Pakistani labour market (see Chapter 5 for details on policy prescriptions). However no matter what the consequences of education are for income distribution, it generates various other positive externalities such as crime reduction and intergenerational benefits (O' Donoghue, 1999). Moreover, since it has been noted that returns to lower education levels are still high in Pakistan as compared to other similar emerging market economies, educating the children from households belonging to the lower ends of the income distribution, may outweigh the convexification of returns to education (Legovini *et al*, 2005).

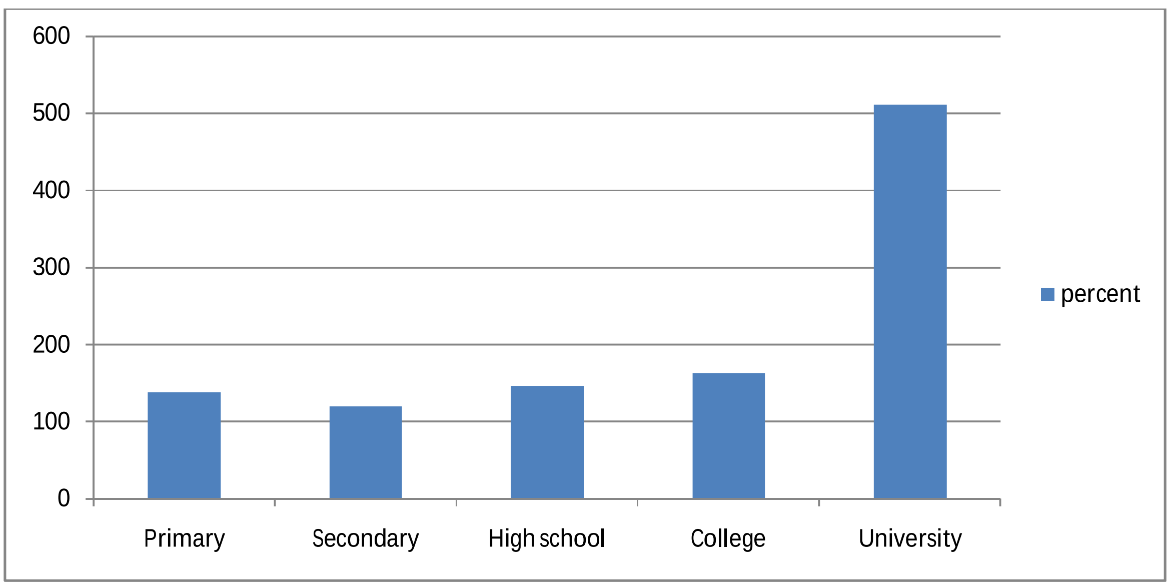
APPENDIX

Figure 1: Evolution of Population and Labour Force in Pakistan, 1993-2006 (% change)



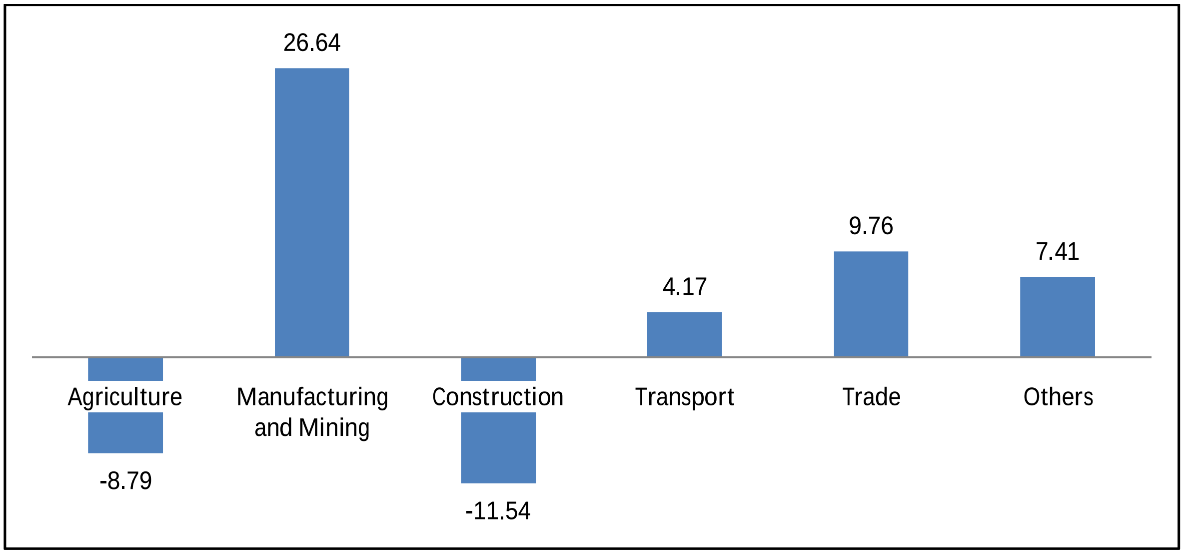
Source: Author's Calculations, Economic Survey of Pakistan 2005-06 and 2006-07

Figure 2: Change in Female Enrolment in Educational Institutions, 1993-2006 (%)



Source: Author's Calculations, Economic Survey of Pakistan 2005-06 and 2006-07

Figure 3: Change in the Distribution of Employed Persons by Industry, 1993-2006 (%)



Source: Author’s Calculations, Economic Survey of Pakistan 2005-06 and 2006-07

Table 1. Definition of variables used in Mincerian Earnings Function

Variable	Definition
Lwage	Log of real monthly wage
Exper	Total years of labour market experience (Age-yrsed-5), 5 years was replaced by the age of starting school if provided in the data.
Exper2	Square of years of labour market experience
Yrsed	Number of completed years of education
Yrsed2	Square of completed years of education
No_ed	Equals 1 if an individual has had no schooling or completed less than 5 years of it, 0 otherwise
Primary	Equals 1 if an individual has completed 5 years of schooling , 0 otherwise
Middle	Equals 1 if an individual has completed 6,7 or 8 years of education, 0 otherwise
Matric	Equals 1 if an individual has completed 9 or 10 years of education, 0 otherwise
Inter	Equals 1 if an individual has completed 11 or 12 years of education, 0 otherwise
Bachelors	Equals 1 if an individual has completed 13 or 14 years of education, 0 otherwise
Pgrad	Equals 1 if an individual has completed 16 years of education or more (Masters and above), 0 otherwise
Urban	Equals 1 if an individual lives in an urban location, 0 otherwise
Male	Equals 1 if an individual is a male, 0 otherwise
Punjab	Equals 1 if an individual is located in Punjab, 0 otherwise
Sindh	Equals 1 if an individual is located in Sindh, 0 otherwise
KP	Equals 1 if an individual is located in KP, 0 otherwise
Manuf	Equals 1 if an individual is employed in the Manufacturing sector, 0 otherwise
Serv	Equals 1 if an individual is employed in the Tertiary/Services sector, 0 otherwise

Note: 1) Descriptive statistics are computed for wage workers between 15-65 years of age.
2) 'No_ed' is the reference category for educational splines, 'Balochistan 'for province, and 'Agriculture' for industry of occupation.

Table 2. Descriptive Statistics of Variables used in Earnings Equations

Variable	1993		2006	
	Mean	SD	Mean	SD
Lwage	7.373	0.656	7.107	0.915
Exper	12.677	16.652	15.005	18.604
Exper2	438.005	869.290	571.256	1071.258
Yrsed	1.928	3.758	3.728	4.633
Yrsed2	32.876	38.032	35.361	62.195
No_ed*	0.760	0.427	0.533	0.499
Primary*	0.100	0.300	0.214	0.410
Middle*	0.102	0.439	0.089	0.285
Matric*	0.050	0.218	0.085	0.279
Inter*	0.020	0.138	0.037	0.188
Bachelors*	0.013	0.115	0.025	0.155
Pgrad*	0.006	0.078	0.018	0.132
Urban*	0.393	0.488	0.393	0.488
Male*	0.517	0.500	0.503	0.500
Punjab*	0.444	0.497	0.390	0.488
Sindh*	0.249	0.432	0.248	0.432
KP*	0.192	0.394	0.215	0.411
Manuf*	0.010	0.097	0.056	0.230
Serv*	0.620	1.214	0.117	0.322
	N = 9,070		N = 14,463	

N= Total number of employed waged workers (males and females)
The variables with superscript (*) are binary 0/1 variables and their means represent the proportions of ones in the sample

Table 3. Characteristics of the Labour Force in Pakistan, 1993 and 2006

Characteristic	1993			2006			% Change
	Pakistan	Urban	Rural	Pakistan	Urban	Rural	
Average no. of persons per household (%)	6.40	6.66	6.30	6.83	6.65	6.93	6.72
Average no. of earners per household (%)	1.70	1.64	1.72	2.07	1.91	2.16	21.76
Average monthly income per household (Rs 1993)	3590	4976	3070	3643	4249	3321	1.47
Earners by employment status (%)							
Employer	1.67	3.09	1.16	1.06	1.78	0.72	-36.53
Self-employed	31.67	21.64	35.26	27.39	21.87	29.97	-13.51
Paid employee	40.07	62.15	32.17	45.43	62.22	37.58	13.38
Unpaid family helper	23.80	9.96	28.74	23.74	10.28	29.58	-1.39
Economically inactive	2.75	3.10	2.62	2.65	3.85	2.09	-3.64
Distribution of monthly household Income by source (%)							
Wages & salaries	33.28	46.34	25.33	35.33	48.81	25.57	6.16
Self employment	41.41	29.47	48.68	39.71	23.42	51.53	-4.11
Age structure (%)							
15-35 age-group	28.92	32.87	27.33	35.03	37.68	33.29	21.13
35-55 age group	16.37	16.35	16.38	15.66	16.91	14.84	-4.34
55-65 age-group	4.56	4.34	4.64	4.45	4.47	4.44	-2.41
Education structure (%)							
Primary & below	20.18	24.09	18.60	19.05	18.92	19.11	-5.60
Middle	6.08	9.91	4.52	9.59	12.65	8.18	57.73
Matric & above	8.91	19.83	4.44	22.20	39.50	14.26	149.20
Earnings of wage employees (Rs 1993)							
Average monthly real earnings	3385			4045			19.50
Earnings premium (ratio)							
Average-to-low skilled earnings	1.27			1.44			13.39
High-to-low skilled earnings	3.36			2.98			-11.31

Source: Author's calculations based on HIES 1992-93 and HIES 2005-06

Table 4. Earnings Inequality in Pakistan 1993-2006

	1993				2006				Percentage Change			
Indicator	Gini	GE(0)	GE(1)	GE(2)	Gini	GE(0)	GE(1)	GE(2)	Gini	GE(0)	GE(1)	GE(2)
Earnings Inequality	0.35	0.22	0.25	0.52	0.46	0.41	0.46	1.21	34.29	86.36	84	132.69
Static Decomposition of Between and Within-group Inequality in Earnings												
By Location (Urban/Rural)												
Within group	0.176	0.201	0.235	0.501	0.225	0.375	0.422	1.181	27.841	86.567	79.574	135.73
Between group	0.09	0.017	0.017	0.016	0.133	0.036	0.035	0.035	47.777	111.77	105.88	118.75
By Gender (Male/Female)												
Within group	0.304	0.217	0.250	0.516	0.335	0.383	0.433	1.20	10.197	76.497	73.7	132.55
Between group	0.014	0.002	0.002	0.002	0.074	0.029	0.024	0.02	428.57	1350	1100	900
By Education (levels)												
Within group	0.066	0.157	0.176	0.416	0.059	0.277	0.292	0.99	-10.61	76.433	65.909	137.98
Between group	0.18	0.062	0.076	0.101	0.282	0.135	0.164	0.226	56.66	117.74	115.79	123.76

Source: Author's Calculations.
Note: GE(0) is the mean log deviation, GE(1) is the Theil Index, GE(2) is the modified coefficient of variation

Table 5.1

Blinder-Oaxaca decomposition (1993)	Number of obs	=	9070
	Model	=	linear
Group 1: males	N of obs 1	=	8477
Group 2: females	N of obs 2	=	593

lwage	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
overall						
group_1	7.398877	.0068436	1081.14	0.000	7.385463	7.41229
group_2	6.999727	.035879	195.09	0.000	6.929405	7.070048
difference	.3991498	.0365259	10.93	0.000	.3275605	.4707392
endowments	.0027897	.0281785	0.10	0.921	-.0524392	.0580185
coefficients	.396779	.0301393	13.16	0.000	.3377071	.4558509
interaction	-.0004188	.0189378	-0.02	0.982	-.0375362	.0366986

Source: Author’s estimates from HIES 1992-93

Table 5.2

Blinder-Oaxaca decomposition (2006)	Number of obs	=	14463
	Model	=	linear
Group 1: males	N of obs 1	=	12170
Group 2: females	N of obs 2	=	2293

lwage	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
overall						
group_1	7.29192	.0066346	1099.07	0.000	7.278916	7.304924
group_2	6.124576	.0237419	257.96	0.000	6.078043	6.171109
difference	1.167344	.0246515	47.35	0.000	1.119028	1.21566
endowments	.2789452	.0249169	11.20	0.000	.2301091	.3277814
coefficients	1.069887	.0206987	51.69	0.000	1.029319	1.110456
interaction	-.1814884	.0204632	-8.87	0.000	-.2215954	-.1413813

Source: Author's estimates from HIES 2005-06

Table 6.1 Marginal returns to educational levels in Pakistan (%)

Level	1993	2006
Primary	2.63	1.61
Middle	1.77	2.45
Matric	5.20	4.89
Inter	6.86	7.85
Graduation	9.38	10.41
Post-graduation	13.64	17.69

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 6.2 Marginal returns to educational levels for males in Pakistan (%)

Level	1993	2006
Primary	2.63	3.62
Middle	1.77	4.21
Matric	5.20	6.06
Inter	6.86	9.48
Graduation	9.38	12.98
Post-graduation	13.64	20.45

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 6.3 Marginal returns to educational levels for females in Pakistan (%)

Level	1993	2006
Primary	7.02	10.13
Middle	2.86	10.70
Matric	8.04	26.37
Inter	11.88	29.27
Graduation	13.49	44.96
Post-graduation	21.60	73.35

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 6.4 Marginal returns to educational levels in Punjab (%)

Level	1993	2006
Primary	3.22	5.83
Middle	2.04	5.90
Matric	11.34	8.78
Inter	8.56	14.35
Graduation	12.17	21.30
Post-graduation	19.10	34.95

Table 6.5 Marginal returns to educational levels in Sindh (%)

Level	1993	2006
Primary	3.51	3.56
Middle	2.04	13.75
Matric	5.62	6.52
Inter	6.70	8.18
Graduation	9.09	13.56
Post-graduation	11.41	24.04

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 6.6 Marginal returns to educational levels in Khyber Pakhtoonkhwa –KP (%)

Level	1993	2006
Primary	0.84	6.65
Middle	1.02	5.07
Matric	5.40	9.10
Inter	6.52	13.98
Graduation	8.15	17.24
Post-graduation	12.86	24.73

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 6.7 Marginal returns to educational levels in Balochistan (%)

Level	1993	2006
Primary	7.02	10.13
Middle	2.86	10.70
Matric	8.04	26.37
Inter	11.88	29.26
Graduation	11.56	44.96
Post-graduation	21.60	73.36

Source: Author's estimates from HIES 1992-93 and HIES 2005-06

Table 7. Earnings functions for wage earners in Pakistan

	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	P>|Z|
Exper	0.057	0.000	0.057	0.000	0.064	0.000	0.055	0.000
	0.002		*0.002*		*0.001*		*0.002*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	0.019	0.000	--		0.026	0.000	--	
	0.002				*0.003*			
Yrsed2	0.003	0.000	--		0.004	0.000	--	
	0.000				*0.000*			
Primary	--		0.122	0.000	--		0.104	0.000
			0.016				*0.017*	
Middle	--		0.137	0.000	--		0.221	0.000
			0.010				*0.020*	
Matric	--		0.429	0.000	--		0.453	0.000
			0.017				*0.018*	
Inter	--		0.616	0.000	--		0.742	0.000
			0.025				*0.024*	
Bachelor	--		0.855	0.000	--		1.063	0.000
			0.025				*0.025*	
Pgrad	--		1.180	0.000	--		1.438	0.000
			0.029				*0.029*	
Male	0.446	0.000	0.444	0.000	1.030	0.000	1.088	0.000
	0.022		*0.022*		*0.015*		*0.016*	
Urban	0.187	0.000	0.188	0.000	0.178	0.000	0.232	0.000
	0.011		*0.011*		*0.011*		*0.012*	
Manuf	-0.111	0.000	-0.110	0.000	0.126	0.000	0.155	0.000
	0.020		*0.020*		*0.018*		*0.019*	
Serv	-0.048	0.000	-0.048	0.000	0.169	0.000	0.226	0.000
	0.004		*0.004*		*0.018*		*0.019*	
Punjab	-0.213	0.000	-0.214	0.000	-0.131	0.000	-0.115	0.000
	0.018		*0.018*		*0.016*		*0.017*	
Sindh	-0.133	0.000	-0.135	0.000	-0.083	0.000	-0.081	0.000
	0.019		*0.019*		*0.017*		*0.018*	
KP	-0.250	0.000	-0.251	0.000	-0.160	0.000	-0.139	0.000
	0.020		*0.020*		*0.019*		*0.020*	
Constant	6.124		6.213		4.837		4.996	
R square	0.418		0.417		0.533		0.475	
N	9070		9070		14463		14463	

Note: Standard errors are in the second rows, *Source*: Author's estimates.

Table 8. Earnings functions for wage earners in Punjab

PUNJAB								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.060	0.000	0.060	0.000	0.066	0.000	0.065	0.000
	0.003		*0.003*		*0.003*		-0.003	
Exper2	-0.001	0.000	-0.001	0.000	*-0.001*	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	*0.018*	0.000	--	--	*0.017*	0.000	--	--
	0.004				*-0.005*			--
Yrsed2	*0.004*	0.000	--	--	*0.005*	0.000	--	
	0.000				*0.000*			
Primary	--	--	0.147	0.000	--	--	0.270	0.000
			0.027				*-0.028*	
Middle	--	--	*0.155*	0.000	--	--	*0.405*	0.000
			0.015				*-0.032*	
Matric	--	--	*0.457*	0.000	--	--	*0.656*	0.000
			0.028				*-0.030*	
Inter	--	--	*0.717*	0.000	--	--	*1.036*	0.000
			0.042				*-0.044*	
Bachelor	--	--	1.005	0.000	--	--	1.418	0.000
			0.048				*-0.046*	
Pgrad	--	--	*1.415*	0.000	--	--	*1.930*	0.000
			0.055				*-0.046*	
Male	0.516	0.000	*0.508*	0.000	1.175	0.000	*1.186*	0.000
	0.033		*0.033*		*0.022*		*-0.022*	
Urban	0.187	0.000	*0.186*	0.000	0.155	0.000	*0.146*	0.000
	0.018		*0.018*		*0.020*		*0.020*	
Manuf	-0.086	0.013	*-0.076*	0.032	*0.184*	0.000	*0.192*	0.000
	0.035		*0.035*		*0.031*		*-0.031*	
Serv	-0.034	0.000	-0.035	0.000	*0.229*	0.000	0.227	0.000
	0.007		0.007		*0.032*		0.032	
Constant	5.742		5.856		*4.530*		4.536	
R square	0.404		0.405		*0.588*		0.594	
N	3812		3812		5426		5426	

Note: Standard errors are in the second rows, *Source*: Author's estimates.

Table 9. Earnings functions for wage earners in Sindh

SINDH								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.055	0.000	0.055	0.000	0.058	0.000	0.056	0.000
	0.003		*0.003*		*0.002*		*0.002*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		0.000	
Yrsed	0.029	0.000	--	--	0.006	0.192	--	--
	0.004		( )		*-0.005*		( )	
Yrsed2	0.002	0.000	--	--	0.005	0.000	--	--
	0.000		( )		0.000		( )	
Primary	--	--	0.166	0.000	--	--	0.191	0.000
			0.027				*0.025*	
Middle	--	--	0.158	0.000	--	--	0.389	0.000
			0.017				*0.034*	
Matric	--	--	0.462	0.000	--	--	0.553	0.000
			0.030				*0.029*	
Inter	--	--	0.609	0.000	--	--	0.757	0.000
			0.040				*0.035*	
Bachelor	--	--	0.842	0.000	--	--	1.140	0.000
			0.038				*0.037*	
Pgrad	--	--	1.076	0.000	--	--	1.664	0.000
			0.044				*0.040*	
Male	0.459	0.000	0.456	0.000	0.992	0.000	0.991	0.000
	-0.038		*0.038*		*-0.019*		*0.024*	
Urban	0.267	0.000	0.268	0.000	0.265	0.000	0.266	0.000
	0.020		*0.020*		*-0.019*		*0.019*	
Manuf	-0.134	0.000	-0.140	0.000	0.014	0.582	0.022	0.393
	0.035		*0.035*		*-0.026*		*0.026*	
Serv	-0.047	0.000	-0.047	0.000	0.075	0.004	0.071	0.000
	0.007		*0.008*		*0.026*		*0.026*	
Constant								
R square								
N	2630		2630		4679		4679	

Note: Standard errors are in the second rows, *Source*: Author’s estimates.

Table 10. Earnings functions for wage earners in Khyber Pakhtoonkhwa (KP)

KHYBER PAKHTOONKHWA								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.053	0.000	0.054	0.000	0.080	0.000	0.077	0.000
	0.004		*0.004*		*0.004*		*0.004*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	0.013	0.016	--	--	0.052	0.000	--	--
	0.005		*( )*		*0.007*		*( )*	
Yrsed2	0.004	0.000	--	--	0.002	0.000	--	--
	0.000		*( )*		*0.000*		*( )*	
Primary	--	--	0.035	0.363	--	--	0.311	0.000
			0.038				*0.046*	
Middle	--	--	0.081	0.000	--	--	0.372	0.000
			0.023				*0.048*	
Matric	--	--	0.431	0.000	--	--	0.692	0.000
			0.039				*0.043*	
Inter	--	--	0.583	0.000	--	--	1.053	0.000
			0.058				*0.059*	
Bachelor	--	--	0.767	0.000	--	--	1.295	0.000
			0.059				*0.058*	
Pgrad	--	--	1.123	0.000	--	--	1.677	0.000
			0.073				*0.054*	
Male	0.380	0.000	0.361	0.000	0.493	0.000	0.551	0.000
	0.049		*0.051*		*0.049*		*0.048*	
Urban	0.157	0.000	0.139	0.000	0.105	0.000	0.088	0.001
	0.025		*0.025*		*0.028*		*0.027*	
Manuf	-0.073	0.109	-0.068	0.132	0.300	0.000	0.326	0.000
	0.046		*0.045*		*0.079*		*0.077*	
Serv	-0.054	0.000	-0.044	0.000	0.241	0.002	0.246	0.001
	0.010		*0.010*		*0.078*		*0.076*	
Constant	6.025		6.114		4.852		4.848	
R square	0.380		0.380		0.420		0.450	
N	1599		1599		2246		2246	

Note: Standard errors are in the second rows, *Source*: Author's estimates.

Table 11. Earnings functions for wage earners in Balochistan

BALOCHISTAN								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.052	0.000	0.051	0.000	0.057	0.000	0.055	0.000
	0.004		*0.004*		*0.004*		*0.003*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	0.004	0.510	--	--	0.043	0.000	--	--
	0.006				*0.007*			
Yrsed2	0.003	0.000	--	--	0.002	0.000	--	--
	0.000				*0.000*			
Primary	--	--	0.070	0.110	--	--	0.222	0.000
			0.044				*0.040*	
Middle	--	--	0.101	0.001	--	--	0.339	0.000
			0.027				*0.045*	
Matric	--	--	0.275	0.000	--	--	0.563	0.000
			0.043				*0.039*	
Inter	--	--	0.360	0.000	--	--	0.875	0.000
			0.065				*0.053*	
Bachelor	--	--	0.590	0.000	--	--	1.052	0.000
			0.065				*0.055*	
Pgrad	--	--	1.051	0.000	--	--	1.264	0.000
			0.073				*0.057*	
Male	0.040	0.598	0.034	0.661	0.596	0.000	0.638	0.000
	0.077		*0.077*		*0.067*		*0.067*	
Urban	0.056	0.037	0.052	0.052	0.104	0.000	0.109	0.000
	0.027		*0.027*		*0.025*		*0.025*	
Manuf	-0.181	0.000	-0.158	0.001	0.030	0.487	0.042	0.316
	0.045		*0.046*		*0.043*		*0.042*	
Serv	-0.090	0.000	-0.092	0.000	0.101	0.015	0.102	0.012
	0.010		*0.010*		*0.041*		*0.041*	
Constant	6.744		6.851		5.447		5.447	
R square	0.390		0.390		0.350		0.370	
N	1029		1029		2112		2112	

Note: Standard errors are in the second rows, *Source*: Author's estimates.

Table 12. Earnings functions for male waged workers

MALES								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.060	0.000	0.060	0.000	0.066	0.000	0.057	0.000
	0.002		*-0.002*		*-0.001*		*0.002*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	0.016	0.000	--		0.021	0.000		
	0.002				*-0.003*			
Yrsed2	0.004	0.000	--		0.004	0.000		
	0.000				*0.000*			
Primary	--		*0.109*	0.000	--		*0.081*	0.000
			0.016				*0.016*	
Middle	--		*0.136*	0.000	--		*0.207*	0.000
			0.010				*0.019*	
Matric	--		0.428	0.000	--		0.402	0.000
			0.017				*0.017*	
Inter	--		0.622	0.000	--		0.694	0.000
			0.024				*0.023*	
Bachelor	--		0.893	0.000	--		0.977	0.000
			0.025				*0.024*	
Pgrad	--		*1.235*	0.000	--		1.328	0.000
			0.028				*0.028*	
Urban	*0.176*	0.000	*0.176*	0.000	*0.190*	0.000	0.250	0.000
	0.011		*0.011*		*0.010*		*0.011*	
Punjab	-0.207	0.000	-0.208	0.000	-0.081	0.000	-0.068	0.000
	0.018		*0.018*		*0.015*		*0.016*	
Sindh	-0.127	0.000	-0.130	0.000	-0.090	0.000	-0.093	0.000
	0.018		*0.018*		*0.015*		*0.016*	
KP	*-0.253*	0.000	*-0.254*	0.000	-0.171	0.000	*-0.148*	0.000
	0.020		*0.020*		*0.017*		*0.019*	
Constant	6.423		6.520		6.010		6.233	
R square	0.404		0.403		0.426		0.341	
N	8477		8477		12170		12170	

Note: Standard errors are in 2nd rows, *Source*: Author’s estimates.

Table 13. Earnings functions for female waged workers

FEMALES								
	1993				2006			
Variable	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|	Coeff (a)	p>|Z|	Coeff (b)	p>|Z|
Exper	0.055	0.000	0.054	0.000	0.083	0.000	0.063	0.000
	-0.007		0.007		*0.005*		*0.005*	
Exper2	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
	0.000		*0.000*		*0.000*		*0.000*	
Yrsed	0.065	0.000			0.069	0.000		
	0.013				*0.011*			
Yrsed2	0.002	0.037			0.005	0.000		
	0.001				*0.001*			
Primary	--		*0.365*	0.000	--		*0.237*	0.000
			0.135				*0.070*	
Middle	--		*0.243*	0.000	--		*0.340*	0.001
			0.068				*0.100*	
Matric	--		0.822	0.000	--		1.094	0.001
							0.071	
Inter	--				--		1.251	0.000
							0.086	
Bachelor	--				--		1.734	0.000
							0.085	
Pgrad	--		*1.708*	0.000	--		*2.206*	0.000
			0.128				*0.099*	
Urban	*0.193*	0.000	*0.199*	0.000	0.209	0.000	0.310	0.000
	0.063		*0.063*		*0.038*		*0.040*	
Punjab	*-0.630*	0.000	-0.647	0.000	-0.423	0.000	-0.454	0.000
	0.126		*0.127*		*0.103*		*0.110*	
Sindh	-0.538	0.000	-0.544	0.000	-0.177	0.095	-0.221	0.050
	0.131		*0.131*		*0.106*		*0.113*	
KP	-0.565	0.000	-0.576	0.000	-0.122	0.294	-0.129	0.297
	0.136		*0.136*		*0.116*		*0.124*	
Constant	6.361		6.433		4.644		5.051	
R square	0.444		0.449		0.441		0.372	
N	593		593		2293		2293	

Note: Standard errors in the 2nd rows, *Source*: Author's estimates from HIES 1992-93 & HIES 2005-06

Table 14. Mincerian Estimates for Pakistan (wage workers between 15-65 years)

	Urban Males		Urban Females		Rural Males		Rural Females	
Variable	**1993**	**2006**	**1993**	**2006**	**1993**	**2006**	**1993**	**2006**
Exper	0.064***	0.069***	0.052***	0.079***	0.053***	0.063***	0.058***	0.083***
	0.002	*0.002*	*0.008*	*0.007*	*0.003*	*0.002*	*0.015*	*0.008*
Exper2	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	0.000	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
Yrsed	0.015***	0.010***	0.052***	0.082***	0.021***	0.040***	0.064*	0.048**
	0.003	*0.004*	*0.015*	*0.016*	*0.004*	*0.004*	*0.034*	*0.019*
Yrsed2	0.004***	0.005***	0.003**	0.005***	0.003***	0.002***	0.005	0.007***
	0.000	*0.000*	*0.001*	*0.001*	*0.000*	*0.000*	*0.004*	*0.001*
Punjab	-0.119***	-0.034	-0.581***	-0.296**	-0.308***	-0.130***	-0.790***	-0.657***
	0.023	*0.024*	*0.142*	*0.130*	*0.027*	*0.019*	*0.254*	*0.174*
Sindh	-0.024	0.005	-0.492***	-0.132	-0.265***	-0.184***	-0.702**	-0.335***
	0.024	*0.024*	*0.144*	*0.133*	*0.029*	*0.019*	*0.288*	*0.177*
KP	-0.209***	-0.185***	-0.514***	-0.061	-0.309***	-0.170***	-0.730**	-0.322***
	0.027	*0.028*	*0.153*	*0.149*	*0.029*	*0.022*	*0.280*	*0.191*
Constant	6.441	6.108	6.538	4.707	6.645	6.080	6.409	4.913
R squared	0.45	0.456	0.41	0.464	0.237	0.332	0.423	0.297
N	4731	6195	413	1100	3746	6184	180	1193

Table 15. Simulated change in typical mean earnings: Price Effect

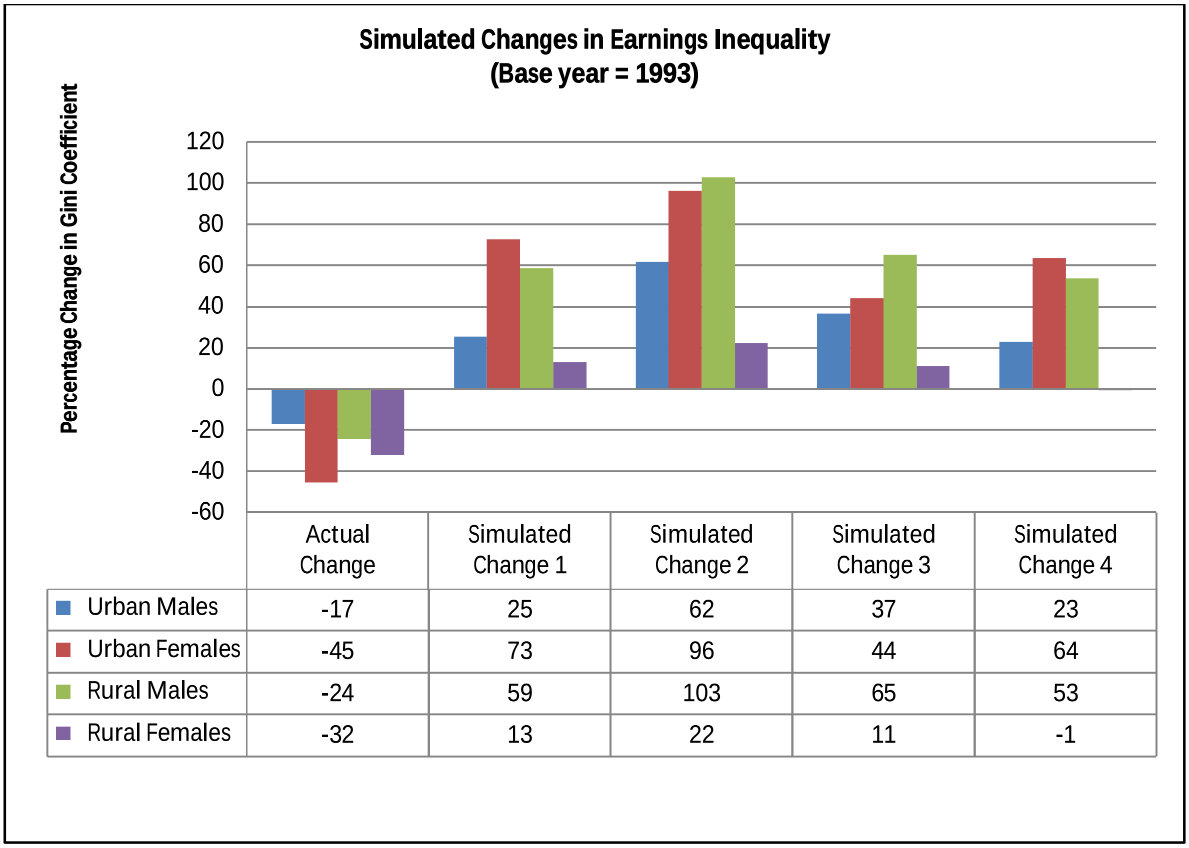
a) Percentage increase in mean income from 1993 to 2006 because of price effect	
Urban males	41
Urban females	15
Rural males	48
Rural females	58
b) Percentage increase in mean income from 2006 to 1993 because of price effect	
Urban males	53
Urban females	746
Rural males	83
Rural females	892
Source: Author’s calculations	

Table 16. Changes in earnings inequality

Base Year = 1993				
	Urban Males	Urban Females	Rural Males	Rural Females
Simulation 1	↑	* ↑	↑	↑
Simulation 2	↑	↑	* ↑	↑
Simulation 3	↑	↑	* ↑	↑
Simulation 4	↑	* ↑	↑	↑
Base Year = 2006				
Simulation 1	↑	** ↓	* ↑	↓
Simulation 2	↑	** ↓	* ↑	↓
Simulation 3	↓	** ↓	↓	↓
Simulation 4	↑	** ↓	* ↑	↓

Note: * stands for largest increase in earnings inequality after performing a simulation, ** stands for largest decrease in earnings inequality after performing a simulation.

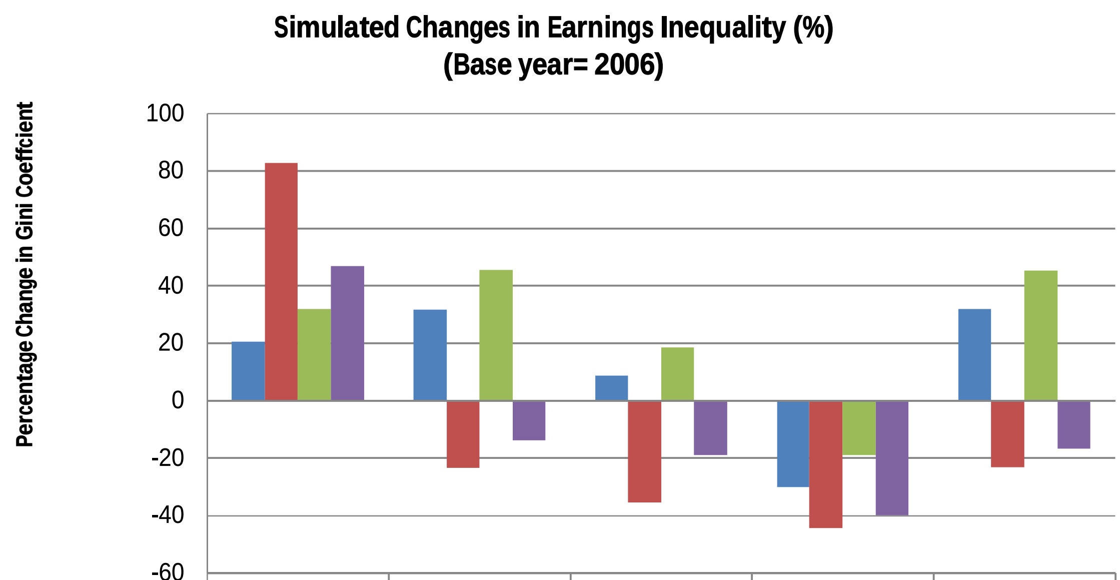
Figure 4: Simulated Earnings Inequality in 1993, using 2006 Coefficients



	Actual Change	Simulated Change 1	Simulated Change 2	Simulated Change 3	Simulated Change 4
Urban Males	-17	25	62	37	23
Urban Females	-45	73	96	44	64
Rural Males	-24	59	103	65	53
Rural Females	-32	13	22	11	-1

Source: Author's calculations using HIES 1992-93 and HIES 2005-06. This table reports GINI coefficient results.

Figure 5: Simulated Earnings Inequality for 2006, using 1993 Coefficients



	Actual Change	Simulated Change 1	Simulated Change 2	Simulated Change 3	Simulated Change 4
Urban Males	21	32	9	-30	32
Urban Females	83	-23	-35	-44	-23
Rural Males	32	46	19	-19	45
Rural Females	47	-14	-19	-40	-16

Source: Author's calculations using HIES 1992-93 and HIES 2005-06. This table reports GINI coefficient results.

Table 17. Distribution of Earnings Income, Substituting 1993 Values into 2006 Distribution

Indicator	Inequality Measure		
	Gini	**E(0)**	**E(1)**
Actual value 2006			
Urban Males	0.352	0.203	0.229
Urban Females	0.527	0.478	0.626
Rural Males	0.315	0.160	0.171
Rural Females	0.559	0.544	0.817
Actual value 1993			
Urban Males	0.292	0.136	0.143
Urban Females	0.288	0.134	0.150
Rural Males	0.239	0.091	0.096
Rural Females	0.380	0.245	0.297
Actual change (as a % of total change)			
Urban Males	20.821	49.339	60.129
Urban Females	82.931	255.199	317.322
Rural Males	32.154	76.007	79.253
Rural Females	46.860	121.848	175.295
Simulated change 1 (absolute value)			
Urban Males	0.465	0.416	0.474
Urban Females	0.404	0.327	0.401
Rural Males	0.459	0.402	0.462
Rural Females	0.483	0.455	0.666
Simulated change 1 (as a % of total change)			
Urban Males	31.835	104.520	107.487
Urban Females	-23.280	-31.598	-35.979
Rural Males	45.603	151.761	169.591
Rural Females	-13.553	-16.414	-18.534
Simulated change 2 (absolute value)			
Urban Males	0.383	0.263	0.285
Urban Females	0.341	0.216	0.235
Rural Males	0.374	0.251	0.272
Rural Females	0.453	0.385	0.498
Simulated change 2 (as a % of total change)			
Urban Males	8.737	29.399	24.624
Urban Females	-35.303	-54.778	-62.452
Rural Males	18.699	56.947	58.507

Rural Females	-18.860	-29.184	-39.047
Simulated change 3 (absolute value)			
Urban Males	0.247	0.122	0.112
Urban Females	0.293	0.157	0.149
Rural Males	0.256	0.130	0.121
Rural Females	0.336	0.212	0.201
Simulated change 3 (as a % of total change)			
Urban Males	-29.959	-39.982	-50.840
Urban Females	-44.338	-67.155	-76.204
Rural Males	-18.753	-18.968	-29.429
Rural Females	-39.772	-61.075	-75.424
Simulated change 4 (absolute value)			
Urban Males	0.465	0.416	0.473
Urban Females	0.405	0.326	0.397
Rural Males	0.458	0.399	0.459
Rural Females	0.467	0.422	0.624
Simulated change 4 (as a % of total change)			
Urban Males	31.977	104.461	106.927
Urban Females	-23.100	-31.803	-36.595
Rural Males	45.428	149.703	167.974
Rural Females	-16.488	-22.384	-23.681

Note: E(0) is log mean deviation, E(1) is Theil Index. *Source*: Author’s calculations based on HIES 1993 and 2006.

Table 18. Distribution of Earnings Income, Substituting 2006 Values into 1993 Distribution

Indicator	Inequality Measure		
	Gini	E(0)	E(1)
Actual value 1993			
Urban Males	0.292	0.136	0.143
Urban Females	0.288	0.134	0.150
Rural Males	0.239	0.091	0.096
Rural Females	0.380	0.245	0.297
Actual value 2006			
Urban Males	0.352	0.203	0.229
Urban Females	0.527	0.478	0.626
Rural Males	0.315	0.160	0.171
Rural Females	0.559	0.544	0.817
Actual change (as a % of total change)			
Urban Males	-17.233	-33.038	-37.550
Urban Females	-45.334	-71.847	-76.038
Rural Males	-24.331	-43.184	-44.213
Rural Females	-31.908	-54.924	-63.675
Simulated change 1 (absolute value)			
Urban Males	0.365	0.239	0.278
Urban Females	0.495	0.434	0.519
Rural Males	0.378	0.255	0.291
Rural Females	0.429	0.326	0.375
Simulated change 1 (as a % of total change)			
Urban Males	25.336	75.837	95.004
Urban Females	72.544	222.847	246.141
Rural Males	58.654	180.841	204.195
Rural Females	12.829	33.139	26.386
Simulated change 2 (absolute value)			
Urban Males	0.471	0.418	0.591
Urban Females	0.563	0.589	0.787
Rural Males	0.483	0.436	0.598
Rural Females	0.465	0.392	0.475
Simulated change 2 (as a % of total change)			
Urban Males	61.540	206.683	314.351
Urban Females	96.056	338.249	424.540
Rural Males	102.667	380.324	525.748

Rural Females	22.301	59.706	60.026
Simulated change 3 (absolute value)			
Urban Males	0.398	0.308	0.463
Urban Females	0.414	0.321	0.481
Rural Males	0.394	0.299	0.446
Rural Females	0.423	0.322	0.401
Simulated change 3 (as a % of total change)			
Urban Males	36.533	126.520	224.189
Urban Females	44.080	138.420	220.581
Rural Males	65.099	229.333	367.033
Rural Females	11.194	31.413	35.095
Simulated change 4 (absolute value)			
Urban Males	0.35801	0.23131	0.26991
Urban Females	0.47168	0.38821	0.4734
Rural Males	0.36594	0.24043	0.27515
Rural Females	0.37709	0.25449	0.30558
Simulated change 4 (as a % of total change)			
Urban Males	22.74067	69.88102	89.13181
Urban Females	63.77778	188.7178	215.5159
Rural Males	53.42753	164.7324	187.874
Rural Females	-0.86753	3.809912	2.909679

Note: E(0) is log mean deviation, E(1) is Theil Index. *Source*: Author's calculations based on HIES 1993 and 2006.

Chapter 3

Does Economic Geography Matter for Pakistan? A Spatial Exploratory Analysis

Abstract

Generally, econometric studies of income inequality consider regions as independent entities, ignoring the likely possibility of spatial interaction between them. This interaction may cause spatial dependency or clustering, which is referred to as spatial autocorrelation. This chapter analyzes the spatial clustering of income, income inequality, education, and growth by employing spatial exploratory data analysis (ESDA) techniques for the first time to data on Pakistani districts. By detecting outliers and clusters, ESDA allows policy makers to focus on the geography of socio-economic regional characteristics. Global and local measures of spatial autocorrelation were computed using the Moran's *I* and the Geary's *C* index to obtain estimates of the spatial autocorrelation of spatial disparities across 98 districts. The overall finding is that the distribution of district wise income inequality, income, education attainment, growth, and development levels, exhibit a significant tendency for inequality and similar levels of education to cluster in Pakistan (i.e. the presence of spatial autocorrelation is confirmed)[37].

Key words: Pakistan, spatial effects, spatial exploratory analysis, spatial disparities

[37] **Acknowledgements:** I'm thankful to Dr. Harry Kelejian (University of Maryland), Dr. Giuseppe Arbia (University G. D'Annunzio of Chieti, Spatial Econometric Association), Dr. Richard Pomfret (University of Adelaide), and the participants at the Advanced Summer School in Spatial Econometrics, Rome (May-June, 2009), for their technical support and valuable comments on this chapter.
I would also like to thank Khydija Wakeel and Muhammad Qadeer at the Planners Resource Centre Pakistan, for their generous support in providing me the shape files that have been used for producing cluster maps in this chapter. Finally, I would like to acknowledge, the data management staff at the Pakistan Institute of Development Economics (PIDE) Islamabad, Federal Bureau of Statistics, Islamabad, and Dr. Amir Jahan Khan and Dr. Haroon Jamal (Sustainable Policy Development Centre, SPDC Karachi) for their extremely generous data support.

1. Introduction

From the industrial revolution to the emergence of the so-called knowledge economy, history has shown that economic development has taken place unevenly across regions. A region's economy is a complex mix of varying types of geographical locations comprising different kinds of economic structures, infrastructure, and human capital. In this context recent literature in regional sciences has highlighted how crucial it is to analyse socio-economic phenomena in the light of spatial concepts such as geography, neighbourhood, density, and distance (Krugman, 1991; Krugman and Venebles, 1995; Quah, 1996; Baldwin et al, 2003; van Oort, 2004; Kanbur and Venebles, 2005; World Development Report, 2009). Keeping these recent developments in view, this chapter identifies, measures, and models the temporal relationship between space, economic inequalities and growth for the case of Pakistan[38]. Specifically, by using data at district level from 1998 to 2007, it utilizes spatial exploratory techniques to determine the effect of distance and contiguity among 98 of Pakistan's administrative districts on their human capital characteristics and inequalities[39]. This way it provides some of the first spatially explicit results for clustering of socio-economic characteristics across Pakistani districts[40].

Most of the existing research on Pakistan's economy is based on a provincial level, and it neglects the role of social interactions within the provinces[41]. This chapter in particular investigates whether spatial clustering of income and average education levels can explain their distribution across Pakistani districts. District level research has become even more important as Pakistan has taken a major step towards fiscal decentralization with the enaction of the 18th Constitutional Amendment[42]. Moreover the 7th National Finance Commission Award has allowed the transfer of more funds from the federation to the provinces which will

[38] Economic inequalities refer to education, earnings, and income inequalities in particular.

[39] Examples of studies similar to this chapter include: Rey and Montouri (1999) on convergence across USA, Balisacan and Fuwa (2004) for income inequality in Philipines, Dall'erba (2004) analyses productivity convergence across Spanish regions over time, Dominicis, Arbia and de Groot (2005) analyses spatial distribution of economic activities in Italy, Pose and Tselios (2007) investigates education and income inequalities in the European Union, and Celebioglu and Dall'erba (2009) analyses spatial disparities in growth and development in Turkey.

[40] The only other exception includes Burki *et al* (2010) that has explicitly considered spatial dependencies in its analysis. However it only studies 56 districts.

[41] Exceptions include Jamal and Khan (2003a, 2003b), Jamal and Khan (2008a, 2008b), Naqvi (2007),Arif *et al* (2010), Siddique (2008) and a few others. Except for Jamal and Khan (2003a, 2003b), Jamal and Khan (2007a, 2007b), most of them only study selected districts/villages from the same province e.g. Naqvi (2007) only analyses the districts/villages of Punjab.

[42] The 18th Amendment to the Constitution of Pakistan (8th April 2010) has removed the power of the President of Pakistan to unilaterally dissolve the Parliament, removed the limit on a Prime Minister serving more than two terms, and has increased provincial autonomy to a very large extent.

now have more authority over the provision of health, educational and physical infrastructure facilities. This fundamental shift towards the division of power between the centre and the provinces bears significant implications for the country's long term policy planning, management and implementation. As education and other public and social services become the sole domain of the provinces, there is a need for increased research on how will the provinces build their capacities and generate resources for efficient public service delivery in social sectors across their districts.

Furthermore, economic inequalities in particular are being considered in this chapter because like most emerging market economies today, Pakistan is also characterised with increasing inequalities in the spatial distribution of its key socio-economic characteristics such as education, health, physical infrastructure, etc (Burki *et al*, 2010). While some districts have the state of the art physical and human capital infrastructure, others have made little or no progress at all. This phenomenon is in line with the findings of the World Bank's World Development Report (2009) that has demonstrated how and why the clustering or concentration of people and production usually takes place in certain favourable areas (coasts, cities, etc) during the growth process in any country. For the case of Pakistan, the most developed districts are located in Northern and Central Punjab. Moreover it has been noted that districts with population density of more than 600 persons per square km are characterized by industrial clusters, superior education and health infrastructure and better sanitation facilities that serve as attractive pull factors, e.g., Karachi, Lahore, Peshawar, Charsadda, Gujranwala, Faisalabad, Sialkot, Mardan, Islamabad, Multan, Swabi, Gujrat and Rawalpindi (Khan, 2003). On the other hand, districts with lower population density (or below 30 persons per square km) are characterized by prevalence of various push factors such as: absence of job opportunities due to lower education and health facilities, poor agricultural endowments, barren or mountainous topography, and limited presence of industrial units (Khan, 2003). Moreover, since a region's (a town, city or country) prosperity and development is ultimately also shared by its neighbours, recent evidence has also demonstrated that neighbourhoods do matter in Pakistan. For example, the highly (and medium) concentrated districts (except for Swat and Muzzaffargarh) are mostly clustered around metropolitan cities of Karachi and Lahore (Burki *et al*, 2010).

In the light of the above mentioned issues, this study investigates the spatial clustering of economic inequalities, growth and development across Pakistani districts by utilizing ESDA techniques. The chapter is organized as follows. Section 2 describes the data; Section

3 gives an overview of the methodology; Section 3 explains the global and the local spatial autocorrelation detection techniques; Section 4 provides an analysis of the results after having applied the ESDA techniques on district income and education data; finally Section 5 summarizes and evaluates.

2. Data

This study uses micro data from the Pakistan Social and Living Standards Measurement survey (PSLM) which is annually produced by the Federal Bureau of Statistics (FBS) of Pakistan since 2004. It is the only socio-economic micro data that is representative at the provincial and at the district level. Moreover, the sample size of the district level data is also substantially larger than the provincial level data contained in micro data surveys such as Household Income and Expenditure Survey (HIES) of Pakistan and the Labour Force Survey (LFS) of Pakistan. This has enabled researchers to draw socioeconomic information which is representative at lower administrative levels as well.

The survey for 2004-05 provides district level welfare indicators for a sample size of about 76,500 households. It provides data on districts in all four provinces of Pakistan namely; Punjab, Sindh, Khyber Pakhtoonkhwa (KP), and Balochistan. The federally administered tribal areas (FATA region) along the Afghan border in the north-west and Azad Kashmir are not included in the data. The PSLM is divided into two parts. The first part contains data on socio-economic characteristics such as education, health, population welfare, immunization, pre/post natal care, family planning, water supply, and sanitation while the second part contains household income and expenditure data.

To analyse the spatial differences in growth rates and development levels over time, this chapter has also utilized the district level data from the 1998 Population Census of Pakistan. This is the latest available census and provides detailed macro-data on all the socioeconomic and demographic characteristics of Pakistani districts. Since the data from PSLM (2004-05) is statistically comparable with the Pakistan Census Data (1998), with some margin of sampling error, the two data sets together provide a decent gap of 7 years to analyse the temporal changes in income and development characteristics across Pakistan.

3. Methodology

Due to the abundance in data collected at a provincial or a rural/urban disaggregation, most socio-economic studies on Pakistan, are a province based analysis. Pakistani provinces however have extreme 'within' diversity in terms of their economic structures, development levels, cultures, language, natural resources and geography. Hence regional policy making requires analyzing socio–economic issues at an even smaller geographical disaggregation. For this reason, the spatial unit of analysis in this study will be the 'districts' of Pakistan. In terms of geographical disaggregation Pakistan (excluding the Federally Administered Tribal Area (FATA) region and Azad Kashmir) has 4 levels consisting of 4 provinces (Punjab, Sindh, Khyber Pakhtoonkhwa (KP), and Balochistan), 107 districts, 377 sub-districts, and 45653 villages (see Figure 1 and 2). A lower level unit of analysis is not being used because data on regional scales below the district level in Pakistan suffers from reliability issues. Finally, due to data constraints, this chapter analyzes 98 out of 107 districts in Pakistan.

3.1 Spatial economic analysis

A fundamental concept in geography is that proximate locations often share more similarities than locations far apart. This idea is commonly referred to as the 'Tobler's first law of geography' (Tobler, 1970). Classical statistical inference such as conventional regressions are inadequate for an in-depth spatial analysis since they fail to take into account spatial effects and problems of spatial data analysis such as spatial autocorrelation, identification of spatial clusters and outliers, edge effects, modifiable areal unit problem, and lack of spatial independence (Arbia, Benedetti, and Espa, 1996; Beck, Gleditsch, and Beardsley, 2006; Franzese and Hays, 2007)[43]. Moreover, as an uneven distribution of socio-economic economic characteristics is shaping the economic geography of most countries, spatial analysis also has increasing policy relevance (World Development Report—WDR, 2009). These reasons together necessitate the use of spatial exploratory and explanatory methods that can explicitly take spatial effects into account.

[43] Modifiable Areal Unit Problem: When attributes of a spatially homogenous phenomenon (e.g people) are aggregated into districts, the resulting values (e.g totals, rates and ratios) are influenced by the choice of the district boundaries just as much as by the underlying spatial patterns of the phenomenon.

3.2 Spatial effects

Spatial effects can be divided into two main kinds: spatial dependence and spatial heterogeneity. Spatial heterogeneity refers to the display of instability in the behaviour of the relationships under study. This implies that parameters and functional relationships vary across space and are not homogenous throughout data sets. Spatial dependence on the other hand, refers to the lack of independence between observations often present in cross sectional data sets. It can be considered as a functional relationship between what happens at one point in space and what happens in another. If the Euclidean sense of space is extended to include general space (consisting of policy space, inter-personal distance, social networks etc) it shows how spatial dependence is a phenomenon with a wide range of application in social sciences. Two factors can lead to it. First, measurement errors may exist for observations in contiguous spatial units. The second reason can be the use of inappropriate functional frameworks in the presence of different spatial processes (such as diffusion, exchange and transfer, interaction and dispersal) as a result of which what happens at one location is partly determined by what happens elsewhere in the system under analysis.

Assuming non-stationarity or structural stability over space is a highly unrealistic assumption when the variable under study belongs to different locations across space. Along the lines of temporal autocorrelation often found in time series data, spatial autocorrelation also violates the standard assumption of independence among observations. Hence standard regression analysis that does not compensate for spatial dependency can yield possibly biased estimators and unreliable significance tests. As a remedy spatial autocorrelation statistics have been devised in order to detect, measure, and analyze the degree of dependency among observations.

3.3 Quantifying spatial effects

Spatial dependence puts forward the need to determine which spatial units in a system are related, how spatial dependence occurs between them, and what kind of influence do they exercise on each other. Formally these questions are answered by using the concepts of neighbourhood expressed in terms of distance or contiguity.

Boundaries of spatial units can be used to determine contiguity or adjacency which can be of several orders (e.g. first order contiguity or more). Contiguity can be defined as linear contiguity (i.e. when counties which share a border with the county of interest are immediately on its left or right), rook contiguity (i.e. counties that share a common side with the county of interest), bishop contiguity (i.e. counties share a vertex with the county of interest), double rook contiguity (i.e. two counties to the north, south, east, west of the county of interest), and queen contiguity (i.e. when counties share a common side or a vertex with the county of interest) (LeSage, 1999). Other common conceptualizations of spatial relationships include inverse distance, travel time, fixed distance bands, and k-nearest neighbours.

The most popular way of representing a type of contiguity or adjacency is the use of the binary contiguity (Cliff and Ord, 1973; 1981) expressed in a spatial weight matrix (**W)**. In spatial econometrics **W** provides the composition of the spatial relationships among different points in space. The spatial weight matrix enables us to relate a variable at one point in space to the observations for that variable in other spatial units of the system. It is used as a variable while modelling spatial effects contained in the data. Generally it is based on using either distance or contiguity between spatial units. Consider below a spatial weight matrix for three units:

$$\mathbf{W} = \begin{bmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{bmatrix}$$

where w_{ij} may be the inverse distance between two units i and j or it may be 0 and 1 if they share a border or a vertex. The **W** matrix displays the properties of a spatial system and can be used to gauge the prominence of a spatial unit within the system. The usual expectation is that values at adjacent locations will be similar.

3.4 The spatial weight matrix for Pakistan

The choice of the W matrix representation and its conceptualization has to be carefully based on theoretical reasoning and the historical factors underlying the concept or phenomenon under study. For example for cluster detection and influence analysis inverse distance is the most appropriate measure due to the distance decay effect, but when we are assessing the geographic distribution of a region’s commuters, travel time or cost would be a better choice.

This paper has employed two W matrices for Pakistan[44]. It is the first time a W matrix has been utilized for a spatial analysis of Pakistan. The first matrix is a simple binary contiguity W matrix (referred to as *BC matrix* from now onwards) based on the concept of Queen Contiguity i.e. if a district *i* shares a border *or* a vertex with another district *j*, they are considered as neighbours, and $w_{i,j}$ takes the value 1 and 0 otherwise. This matrix is also zero along its diagonal implying that a district cannot be a neighbour to itself. Hence it is a symmetric binary matrix with a dimension of 98x98 (98 being the total number of the districts being analyzed). This matrix precisely tells us the influence of geographically adjacent neighbours on each other. A simple binary contiguity matrix is a standard starting point and its influence is often compared with other types of W matrices.

The second W matrix developed for Pakistan is one based on inverse average road distance from district *i* to the closest district *j* which has a 'large city' in it (referred to as *ID matrix* from now onwards). Out of the 98 districts being studied there are only 14 that come under the category of a district with a 'large size' city as per the classification of the coding scheme for the PSLM survey. These include Islamabad as the federal capital city; Lahore, Faisalabad, Rawalpindi, Multan, Gujranwala, Sargodha, Sialkot, and Bahawalpur as districts with a 'large size' city in Punjab; Karachi, Hyderabad and Sukkur in Sindh; Peshawar in the North West Frontier Province and Quetta in Balochistan. This matrix is a symmetric non-binary matrix, again with a dimension of 98x98.

The reason for selecting road distance instead of train distance as is normally done in most studies on urban area analysis is that in Pakistan, the road network is much better developed than the railway network . As a result, Pakistan's transport system is primarily dependent on road transport which makes up 90 percent of national passenger traffic and 96 percent of freight movement every year (The Economic Survey of Pakistan, 2007-08).

Inverse distance matrices have more explanatory power as partitions of geographic space especially when the phenomenon under study involves the exchange or transfer of information and knowledge (in our case income and education). It establishes a decay function that weighs the effect of events in geographically proximate units more heavily than those in geographically distant units. Since a country is not a plain piece of land, Euclidean distance calculations or distance as 'the crow flies' make little economic sense when we are

[44] Usually two or more weights matrices are utilized in spatial exploratory and econometric studies as a robustness measure. It is way of demonstrating whether the strength of spatial effects is robust to changing definitions of neighbourhood.

trying to investigate the effect of distance from districts with a large city on regional wages. The effect of the density of country's infrastructure network is an important influence. For this reason this chapter has utilized the Google Maps service of distance calculation. It not only provides the Euclidean or the straight line distance between districts using their longitude and latitude information but also the maximum and minimum road distance to reach from one district to another carefully taking into consideration the existing road network of Pakistan. The distance used in this paper is the inverse of the average of the maximum and the minimum road distance between a district and its closest large city district.

Finally both the matrices are row-standardized i.e. each weight is divided by its row sum. Row standardization is recommended whenever the distribution of the variables under consideration is potentially biased due to errors in sampling design or due to an imposed aggregation scheme.

4. Exploratory spatial data analysis

Spatial effects are incorporated in spatial modelling which typically aims to look for "associations instead of trying to develop explanations" (Haining 2003: 358)**.** This chapter applies exploratory spatial data analysis (ESDA) techniques to district wise data on income, education, growth and development levels in order to detect the presence of spatial dependence. ESDA describes and visualizes spatial distributions, "identifies spatial outliers, detects agglomerations and local spatial autocorrelations, and highlights the types of spatial heterogeneities" (Oort 2004, 107; Haining 1990; Bailey and Gatrell 1995; Anselin 1988; Le Gallo and Ertur 2003).The particular techniques employed in this study include the calculation of Moran's *I* statistic and Geary's *C* statsitic. The global Moran's I demonstrates the spatial association of data collected from points in space and measures similarities and dissimilarities in observations across space in the whole system (Anselin, 1995). However in the presence of uneven spatial clustering, the Local Indicators of Spatial Association have been utilized. They measure the contribution of individual spatial units to the global Moran's I statistic (Anselin, 1995). The study will also generate Moran scatter plots to demonstrate the spatial distribution of district wage and education levels across Pakistan.

4.1 Measures of spatial autocorrelation:

i) Global spatial autocorrelation

Spatial autocorrelation occurs when the spatial distribution of the variable of interest exhibits a systematic pattern (Cliff and Ord, 1981). Positive (negative) spatial autocorrelation occurs when a geographical area tends to be surrounded by neighbours with similar (dissimilar) values of the variable of interest. As previously mentioned, this paper utilizes two measures Moran's *I* and Geary's *C* statistics to detect the global spatial autocorrelation present in the data[45]. The Moran's *I* is the most widely used measure for detecting and explaining spatial clustering not only because of its interpretative simplicity but also because it can be decomposed into a local statistic along with providing graphical evidence of the presence of absence of spatial clustering.

It is defined as:

$$I = \frac{n}{S_0} \cdot \frac{\sum_i^n \sum_j^n w_{i,j} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i^n (y_i - \bar{y})^2} \qquad (1)$$

where y_i is the observation of variable in location i , $\bar{y}$ is the mean of the observations across all locations, n is the total number of geographical units or locations, $w_{i,j}$ is one of the elements of the weights matrix and it indicates the spatial relationship between location *i* and location j.

S_0 is a scaling factor which is equal to the sum of all the elements of the W matrix :

$$S_0 = \sum_i^n \sum_j^n w_{i,j} \qquad (2)$$

S_0 is equal to n for row standardized weights matrices (which is the preferred way to implement the Moran's *I* statistic), since each row then adds up to 1. The first term in equation (1) then becomes equal to 1 and the Moran's *I* simplifies to a ratio of spatial cross products to variance.

[45] Another well known measure of spatial autocorrelation is Getis and Ord's *G* statistic, see Anselin (1995a, p.22-23).

Under the null hypothesis of no spatial autocorrelation, the theoretical mean of Moran's *I* is given by

$$E(I) = -1/(n-1) \tag{3}$$

The expected value is thus negative and will tend to zero as the sample size increases as it is only a function of n (the sample size). Moran's *I* ranges from -1 (perfect spatial dispersion) to +1 (perfect spatial correlation) while a 0 value indicates a random spatial pattern. If the Moran's *I* is larger than its expected value, then the distribution of *y* will display positive spatial autocorrelation i.e. the value of *y* at each location *i* tends to be similar to values of *y* at spatially contiguous locations. However, if *I* is smaller than its expected value, then the distribution of *y* will be characterized by negative spatial autocorrelation, implying that the value of *y* at each location *i* tends to be different from the value of y at spatially contiguous locations. Inference is based on z-values computed as

$$z_I = \frac{I-E(I)}{sd(I)} \tag{4}$$

i.e. the expected value of *I* is subtracted from *I* and divided by its standard deviation. The theoretical variance of Moran's *I* depends on the assumptions made about the data and the nature of spatial autocorrelation. This paper will present the results under the randomization assumption i.e. each value observed could have equally occurred at all locations[46]. Under this assumption z_I asymptotically follows a normal distribution, so that its significance can be evaluated using a standard normal table (Anselin 1992a). A positive (negative) and significant z- value for Moran's *I* accompanied by a low (high) p-value indicates positive (negative) spatial autocorrelation[47].

The second measure of spatial autocorrelation that has been utilized is the Geary's C which is defined as:

$$C = \frac{(N-1)\sum_i \sum_j w_{i,j}(X_i - X_j)^2}{2W \sum_i (X_i - \bar{X})^2} \tag{5}$$

[46] The other two assumptions include the assumption of normal distribution of the variables in question (normality assumption) or a randomization approach using a reference distribution for I that is generated empirically (permutation assumption). For details and formulas of the randomization assumption, see Sokal *et al*. 1998).

[47] Negative spatial autocorrelation reflects lack of clustering, more than even the case of a random pattern. The checkerboard pattern is an example of perfect negative spatial autocorrelation.

where N is the number of spatial units (districts in our case); X is the variable of interest; $w_{i,j}$ represents the spatial weights matrix, where W is the sum of all $w_{i,j}$. The value of Geary's C lies between 0 and 2. Under the null hypothesis of no global spatial autocorrelation, the expected value of C is equal to 1. If C is larger (smaller) than 1, it indicates positive (negative) spatial autocorrelation. Geary's C is more sensitive to local spatial autocorrelation than Moran's I. Inference is based on z-values, computed by subtracting 1 from C and dividing the result by the standard deviation of C:

$$z_c = \frac{c-1}{sd(c)} \qquad (6)$$

The standard deviation of C is computed under the assumption of total randomness, implying that z_c is asymptotically distributed as a standard normal variate (Anselin, 1992a; Pissati, 2001).

Finally, the results of the Moran's *I* and Geary's *C* are dependent on the specification of the weights matrix. Although interpretations change depending on whether the matrix was based on the use of physical distance or economic distance, a "pattern of decreasing spatial autocorrelation with increasing orders of contiguity (distance decay) is commonly witnessed in most spatial autoregressive processes regardless of the matrix specification" (van Oort, 2004: 109).

ii) Local spatial autocorrelation

Since the Moran's *I* and Geary's *C* are global statistics based on simultaneous measurements from many locations, they only provide broad spatial association measurements, ignore the location specific details, and cannot identify which local spatial clusters (or *hot spots*) contribute the most to the global statistic. As a remedy, local statistics commonly referred to as 'Local Indicators of Spatial Association (LISA)'are used along with graphic visualization techniques of the spatial clustering such as a Moran's Scatterplot (Fotheringham et al, 2000; Haining, 2003).

The Moran scatterplot is derived from the global Moran *I* statistic. Recall that the Moran's I formula when we use a row standardized matrix can be written as

$$I=\frac{\sum_i^n (y_i-\bar{y})\,(\sum_i^n w_{i,j}\,(y_j-\overline{y}))}{\sum_i^n (y_i-\bar{y})^2} \qquad (7)$$

This is similar to the formula for a coefficient of the linear regression b, with the exception of $(\sum_i^n w_{i,j} \; (y_j - \overline{y}))$, which is the so-called spatial lag of the location i.

Therefore *I* is formally equivalent to the regression coefficient in a regression of a location's spatial lag (Wz) on the location itself. This interpretation is used by the Moran's scatterplot, enabling us to visualize the Moran's I in a scatterplot of Wz versus z, where $z = y_i - \bar{y})/(y_i)$.Moran's *I* is then the slope of the regression line contained in the scatterplot. A lack of fit in this scatterplot indicates local spatial associations (local pockets/non-stationarity). This scatterplot is centered on 0 and is divided in four quadrants that represent different types of spatial associations.

However since graphical evidence alone does not give the significance levels of the spatial clustering for which we resort to complementing the Moran scatterplot with a local statistic. Local statistics or indicators can reveal the locations that display significant deviation from spatial randomness in the presence of global spatial autocorrelation (hot spots) and the significant outliers in a diagnostic analysis for local stability. Anselin (1995b) defines a LISA as a statistic that satisfies the following two requirements:

1) The LISA for each observation gives an indication of the spatial clustering of similar values around that observation;
2) The sum of all LISA's for all observations is proportional to a global indicator of spatial association

We use the local Moran's *I* statistic which satisfies the above requirements for our analysis. Each local Moran *I* for a particular location indicates the extent of spatial clustering around it and the sum of all local Moran's *I*'s is equal to the global Moran's *I*. The Local Moran's *I* can be defined as:

$$I_i = (y_i - \bar{y}) \sum_i^n w_{i,j} \, (y_j - \bar{y}) \qquad (8)$$

The null hypothesis tested in this case is that there is no association between the value observed at a location i and values observed in its neighbours i.e. values of I_i's are zero. Positive (negative) local spatial autocorrelation exists when we obtain positive (negative) values for I_i and z-scores which indicate the clustering of similar (dissimilar) values of *y* around location *i*.

5. Empirical Results

5.1 *Spatial autocorrelation of income inequality*

Our first empirical estimation involves calculating measures of spatial dependence for the Gini coefficient of district per capita income in the year 2005. Table 1 shows the results of Moran's test and Geary's *C* statistic for district income inequality levels using the two weights matrices. In both of the cases, the null hypothesis of no spatial dependence of income inequality between districts is rejected at the significance level of 1% as the measures demonstrate a weakly positive spatial autocorrelation amongst district inequality levels (0.21 under BC matrix specification and 0.25 under ID matrix specification). The results for Geary's C statistic have been reported in Table 2 in Appendix.

Table 1: Global autocorrelation results for income inequality—Moran's I (2005)

Weight Matrix	**I**	**II**
$i \neq j$	$w_{i,j} = 0 \text{ or } 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
$i = j$	$w_{i,i} = 0$	
Moran's *I*	0.211	0.257
E(I)	-0.010	-0.010
Sd(I)	0.074	0.103
Z	2.985	2.601
p-value	0.003	0.009

5.2 Local spatial autocorrelation of income inequality

The Moran scatterplot (in Figures 1 and 2) provides a more disaggregated view of the nature of the global autocorrelation. It not only provides us information on the presence of clusters in the data but also the outliers contained in it. This scatterplot is divided into four quadrants, each of which represents a different type of spatial association:

- The upper right quadrant represents spatial clustering of a district with a high income level around neighbours that also have high incomes. This quadrant is also called the High-High zone (HH) since Z-score and Wz (the spatial lag) both have high values. In general these are locations that have a positive value for the local Moran's I.

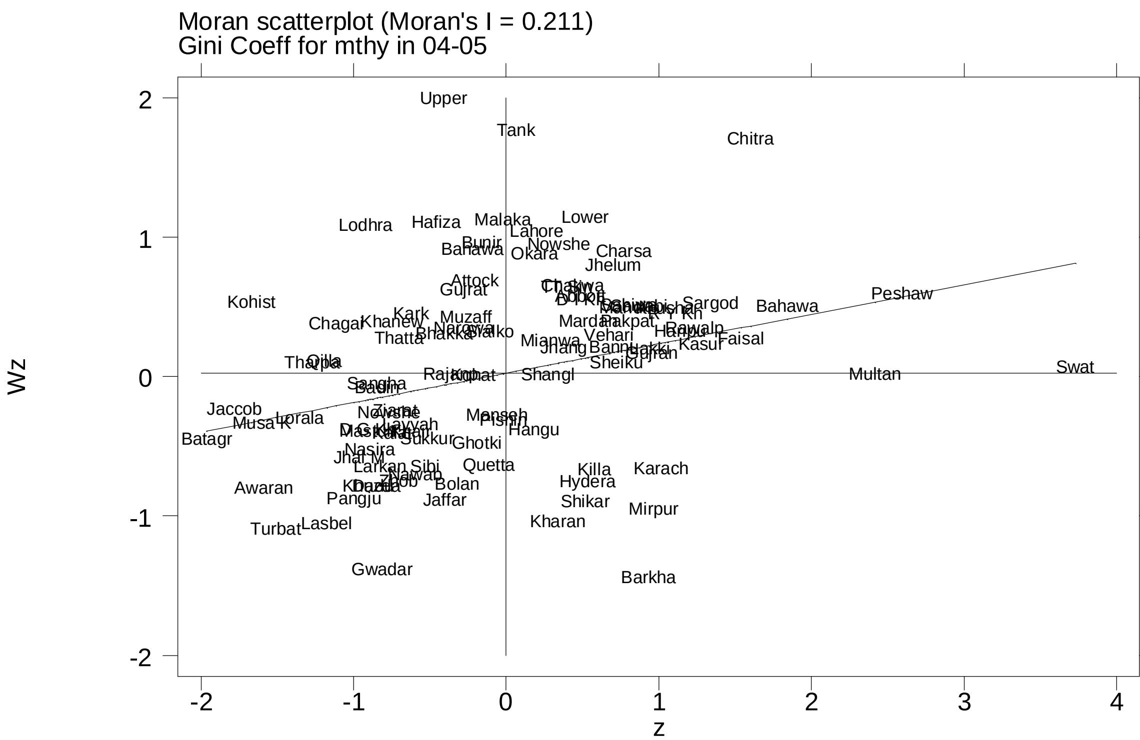
- The upper left quadrant represents spatial clustering of a district with a low income level (or any other economic variable under study) which is neighbours to districts with high income levels. This quadrant is also called the Low-High zone (LH) since Z-score is low while Wz has high values indicating a low outlier among neighbours with high values. In general these are locations that have a negative value for the local Moran's I.

- The lower left quadrant represents spatial clustering of a district with a low income level around neighbours that also have low incomes. This quadrant is also called the Low-Low zone (LL) since Z-score and Wz both have low values. In general these are locations that have a negative value for the local Moran's I.

- The lower right quadrant represents spatial clustering of a high income district with neighbours that have low income levels. This quadrant is also called the High-Low zone (HL) since the Z-score is high while Wz has low values indicating a high outlier among neighbours with high values. In general these are locations that have a negative value for the local Moran's I.

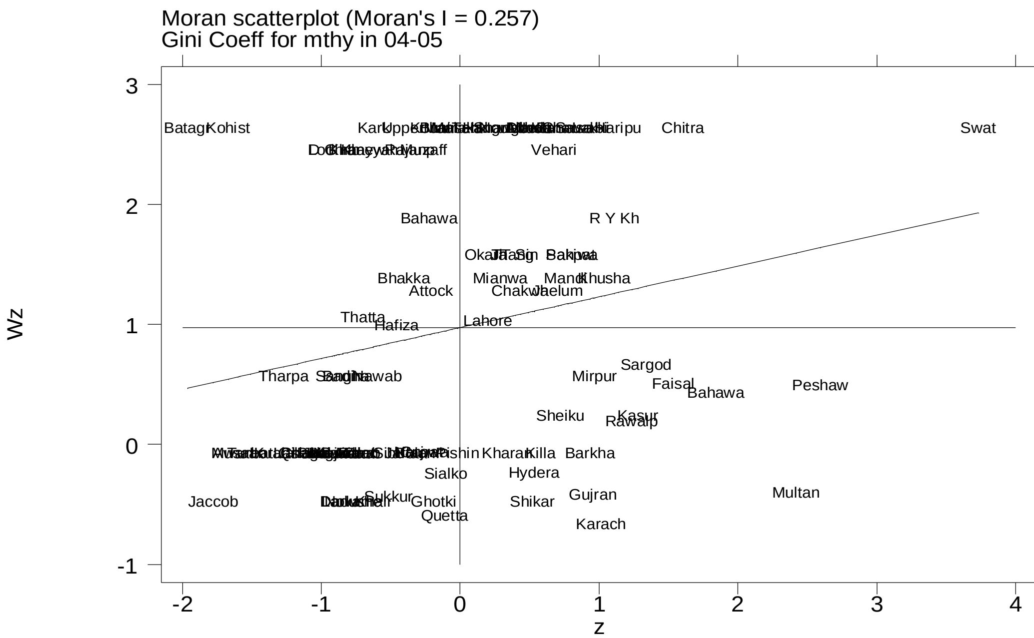
Figure 1 illustrates the Moran scatterplot for Gini coefficient of district per capita incomes using the binary contiguity weights matrix. It shows a positive global Moran's I (z-score = 2.98), which is represented by the slope of the black line. Due to the weakly positive spatial autocorrelation, we are unable to detect any substantial clusters of high (or low) inequality districts in particular for the year 2005.

Figure 1. Spatial autocorrelation of district income inequality using the BC matrix



Similarly, Figure 2 (see below) also shows a Moran scatterplot for Gini coefficient of district per capita incomes, however it has utilized an inverse distance weights matrix instead (also see Table 2). It has a slightly higher value for the higher value for the global Moran's I (z-score = 2.65) since the clusters here are not based on geographic contiguity but on geographic proximity to the provincial capital. The overall spatial autocorrelation is although statistically significant, it still remains weak. Significant LISA statistics for district income inequality have been reported in Tables 2a and 2b in the Appendix.

Figure 2. Spatial autocorrelation of district incomes using the inverse distance matrix



5.3 Spatial autocorrelation of district education levels

The role of human capital in generating growth is important since the distribution of income is mainly driven by the distribution of human capital within a country (Golmm and Ravikuman, 1992; Saint-Paul and Verdier, 1993; Galor and Tsiddon, 1997). Hence the operation of human capital externalities and knowledge spillovers plays an important role in generating regional dependencies and disparities. It has been demonstrated that regions located in an economic periphery experience lower returns to skill attainment and hence have reduced incentives for human capital investments and agglomerations. However spatial externalities do not spread without limits (Darlauf and Quah, 1999) as a result of which closely related economies or regions tend to have similar kinds of human capital externalities and technology levels as compared to the more distant ones (see Quah, 1996; Mion, 2004). This section investigates the spatial dimensions of education in Pakistan, the extent to which neighbouring districts share similar levels of education, and examines whether human capital, income and development level inequalities are spatially associated.

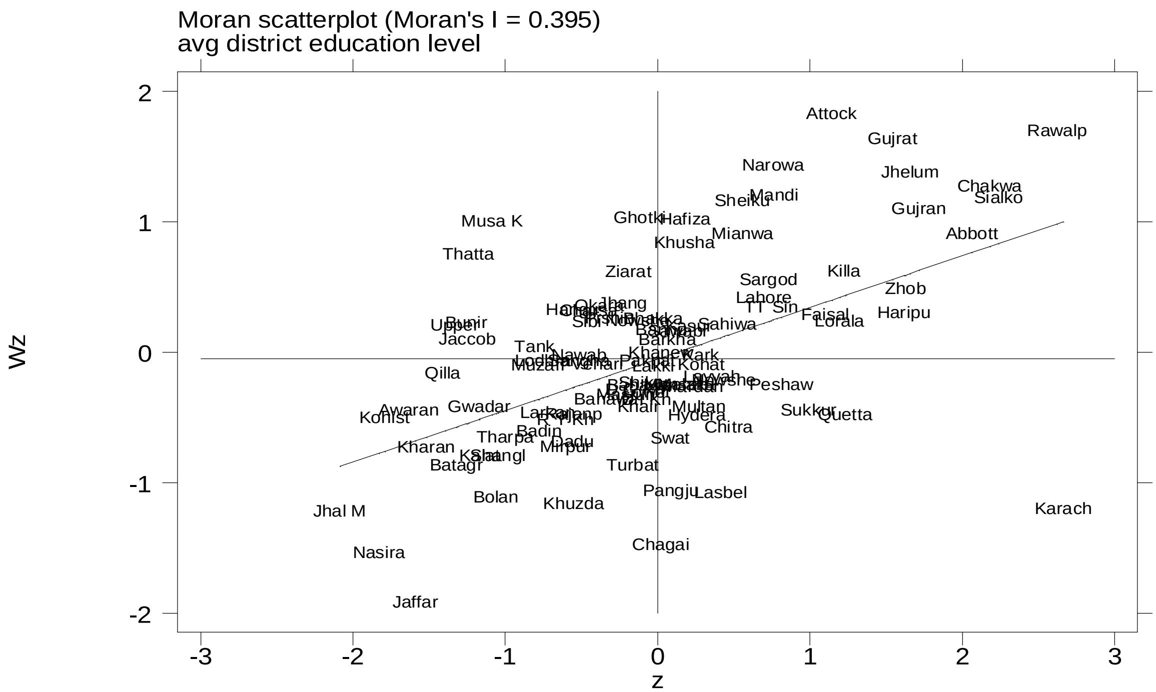
In order to do so, this chapter uses the average district wise education attainment level (which is measured as the average number of schooling years completed in a district) as a proxy for human capital. It is expected that neighbours of districts with high education attainment should also have high educational awareness and hence similar if not equal attainment levels. Again the chapter utilizes the Moran's *I* global and local indices along with a Moran scatterplot using the two weights matrices.

Results show that there exists a greater possibility of knowledge spillovers between districts that share a border, as compared to when they do not (see Table 2 and Tables 4a and 4b in Appendix). The global Moran's I for average district education level (measured as the average education attainment of a district's citizens) is positive and statistically significant when neighbourhood is defined in terms of contiguity, however it is negative and statistically insignificant when neighbourhood is defined in terms of proximity. These results imply that for a Pakistani district, sharing a border with a district whose individuals have a high (low) education level, 'may' result in increasing (decreasing) its own education levels.

Table 2: Global autocorrelation results for education attainment—Moran's I (2005)

Weight Matrix	I	II
$i \neq j$	$\boldsymbol{w_{i,j} = 0 \text{ or } 1}$	$\boldsymbol{w_{i,j} = \frac{1}{d_{i,j}}}$
$i = j$	$\boldsymbol{w_{i,i} = 0}$	
Moran's *I*	0.395	-0.003
E(I)	-0.010	-0.01
Sd(I)	0.075	0.103
Z	5.440	0.072
p-value	0.000	0.943

Figure 3. Spatial autocorrelation of district education levels using the BC matrix



The positive pattern for spatial autocorrelation for average district education levels demonstrated by the BC matrix shows more clusters with low education levels (in the case of Balochistan) and high education levels (in the case of Punjab) as compared to outliers. Districts in northern Punjab emerge in the High-High quadrant and confirm our assumption about high human capital districts being close to each other. These results have also been put forward in a recent study on agglomeration patterns of industries across Pakistani districts in a study by Burki and Khan (2010) and have been illustrated using a cluster map in Figure 9 in the Appendix.

The neighbouring districts of Karachi and Thatta emerge as the most significant outliers when we analyze the local Moran's I values using the BC and the ID matrices. While Karachi falls into the High-Low zone, Thatta falls in the Low-High zone. However, the fact that being a neighbour with Karachi (a district with one of the highest average education levels in Pakistan) does not translate in Thatta having improved human capital characteristics is not very surprising. Regional science and regional economics literature has demonstrated that the economic influence and knowledge spillover effects of coastal cities (such as Karachi) are quite different from the pattern of spillovers generated by landlocked regions (Glaeser *et al*, 1992; Henderson, 2003). The overall spatial pattern of autocorrelation is quite diffused when we use the ID matrix for analysis. However under both the neighbourhood structures Rawalpindi, Abbottabad, Chakwal and Jhelum emerge as a statistically significant cluster of districts with high average education attainment levels. The global spatial autocorrelation while using the ID matrix is negative but close to 0 and statistically insignificant (see Table 4b in Appendix). Again these findings indicate that we can reject the null hypothesis of no spatial association, or that a random pattern exists between districts for average education rates[48].

5.4 The dynamics of spatial autocorrelation of district inequality and education levels

The availability of district wise macro-data for the year 1998, allowed me to measure spatial association between district wise primary, secondary, and bachelors education levels for 1998. It has also been utilized in order to analyse the temporal change in the spatial distribution of district wise real per capita GDP growth rate, district wise per capita incomes, and district human development levels between 1998 and 2005.

Figures 3a, 3b, 4a, and 4b in the Appendix each demonstrates a Moran scatterplot which provides a disaggregated picture of the nature of spatial autocorrelation for district per capita income in 1998 and 2005, using the BC and ID matrix respectively. The spatial lag (Wz) in this situation is a weighted average of the incomes of a district’s neighbouring districts. The scatterplots in both the years using both the matrices demonstrate that the overall pattern of spatial dependence between districts has remained positive and statistically significant. However, the overall value of the global Moran’s I statistic has reduced from being 0.81 to 0.38 between 1998 and 2005 when the results are reported using the BC matrix. Similarly, the value of global Moran’s I statistic has reduced from being 0.91 to 0.51 between 1998 and 2005 under the results produced using the ID matrix (see Figures 4a and 4b in the Appendix). The LISA cluster maps further enhance our understanding of the temporal change in spatial association between district incomes (Figure 5 in Appendix). The clusters of High-High income levels have considerably reduced between 1998 and 2005 (as shown by dark grey regions), and the cluster of High-High income regions in 2005 shows how most of the districts of Punjab have the highest income levels in Pakistan.

[48] The Moran’s scatterplot using the ID matrix for average district education attainment level is provided in Figure 6 in the Appendix.

Table 3. Spatial Autocorrelation of per capita GDP growth rate between 1998—2005

GDP Growth Rate (1998-2005)		
	BC matrix	*ID matrix*
Moran's I	0.430	0.140
E(I)	-0.010	-0.010
Sd(I)	0.071	0.099
Z	6.204	1.524
P-value	0.000	0.128

Source: Author’s calculations

Furthermore a spatial analysis of the growth rate between 1998 and 2005, also indicates a positive and a statistically significant spatial autocorrelation pattern when neighbourhood is defined in terms of contiguity but a statistically insignificant pattern when neighbourhood is defined in terms of proximity as measured by the ID matrix (see Table 5). This implies that districts with a high (low) real GDP growth rate have been spatially associated with other districts which also have high (low) real GDP growth rates. This result motivates the spatial econometric analysis of beta-convergence across Pakistani districts by explicitly incorporating these spatial effects, in Chapter 4.

Moreover, since our macro-data set from 1998 provides district wise statistics on individual education attainment levels (measured as the percentage of individuals having completed an education level), it has allowed us to analyse how the distance from large neighbouring cities (or provincial capitals) affects the incentives to obtain education in a district. Table 4 demonstrates that whether neighbourhood is measured in terms of geographic proximity (using ID matrix) or in terms of geographic contiguity (using BC matrix), there exists a positive and highly significant spatial autocorrelation for levels of education below high-school (i.e primary, matric i.e. grade 10, and inter i.e. grade 12). However, for higher levels (Bachelors and above), geographic contiguity to a district with a high percentage of graduates, is more influential than the distance from the provincial capital or the nearest large city.

Table 4. Spatial autocorrelation for education levels, 1998

Primary Education			***Matric***			***Higher Education—Bachelors***		
	BC Matrix	ID Matrix		BC Matrix	ID Matrix		BC Matrix	ID Matrix
Moran's I	0.494	0.559	Moran's I	0.391	0.247	Moran's I	0.327	-0.014
E(I)	-0.010	-0.010	E(I)	-0.010	-0.010	E(I)	-0.010	-0.010
Sd(I)	0.075	0.103	Sd(I)	0.074	0.102	Sd(I)	0.074	0.102
Z	6.745	5.501	Z	5.443	2.523	Z	4.582	-0.038
P-value	0.000	0.000	P-value	0.000	0.012	P-value	0.000	0.969
Geary's C	0.497	0.983	Geary's C	0.610	0.703	Geary's C	0.610	1.643
E(c)	1.000	1.000	E(c)	1.000	1.000	E(c)	1.000	1.000
Sd(c)	0.079	0.244	Sd(c)	0.085	0.379	Sd(c)	0.086	0.392
Z	-6.401	-0.069	Z	-4.573	-0.783	Z	-4.538	4.193
P-value	0.000	0.945	P-value	0.000	0.434	P-value	0.000	0.000

Note: BC stands for Binary Continuity, ID for Inverse Distance. Author's own calculations.

Finally, spatial association between district development levels (as measured by the Human Development Index (HDI) calculated by the UNDP in NHDR, 2003) has reduced between 1998 and 2005 from 0.40 to 0.311, but is still positive and significant (see Table 5). These results again confirm the findings of the new economic geography literature that a region's development levels, depend on the development levels prevailing in its neighbouring regions.

Table 5. HDI Spatial Autocorrelation using the Binary Contiguity Matrix

District Human Development Index (HDI)		
	1998	*2005*
Moran's I	0.405	0.311
Standard deviation (I)	0.075	0.074
Z-value	5.573	4.341
P-value	0.000	0.000

Source: Author's calculations using data from NHDR (2003).

Finally, it should be noted that the opposite results obtained after using two different weight matrices poses a dilemma towards deriving policy implications from our empirical results. After having analyzed the matrices and results in detail, it seems that the logic upon which the inverse distance matrix is based on in this chapter could be making this matrix measure connectedness instead of proximity between districts. Future research work will focus on investigating this issue along with the formulation of weight matrices based on other definitions of proximity using the GIS software.

6. Conclusions

This chapter has performed an exploratory analysis of socio-economic disparities across Pakistan for the first time and has provided useful insights for the conduct of economic regional policy in Pakistan. It has investigated the spatial distribution of income inequality, income, education, growth and development levels for 98 districts between 1998 and 2005. The overall finding that emerges from this chapter is that the distribution of district wise income inequality, income, education attainment, growth, and development levels, exhibits a significant tendency to cluster in space (i.e. the presence of spatial autocorrelation is confirmed), and hence economic geography does matter for Pakistan.

Specifically the following main findings emerge using the contiguity matrix from this chapter:

1) The province of Punjab contains the largest cluster of high incomes districts in both 1998 and 2005.
2) Districts with a high (low) real GDP growth rate have been spatially associated with neighbouring districts (with which they share a border) which also have high (low) real GDP growth rates between 1998 and 2005.
3) District education levels reveal high spatial association.
4) Except for Lahore, none of the other 3 provincial capitals of Pakistan (Karachi, Peshawar, Quetta) have high knowledge spillovers. While this finding is not surprising for Karachi, since coastal cities have different spillover mechanisms as compared to landlocked cities, it indicates that infrastructure and cluster development can facilitate increased knowledge spillovers at least from the centers of economic activity in Pakistan if not from all large city districts.
5) There exists positive spatial autocorrelation dependence for education levels below bachelors (i.e. primary, matric i.e. grade 10, and inter i.e. grade 12).
6) Finally, an analysis of spatial association of district wise Human Development Indicators confirms that a district’s development levels are weakly associated with the development levels prevailing in its neighbouring districts.

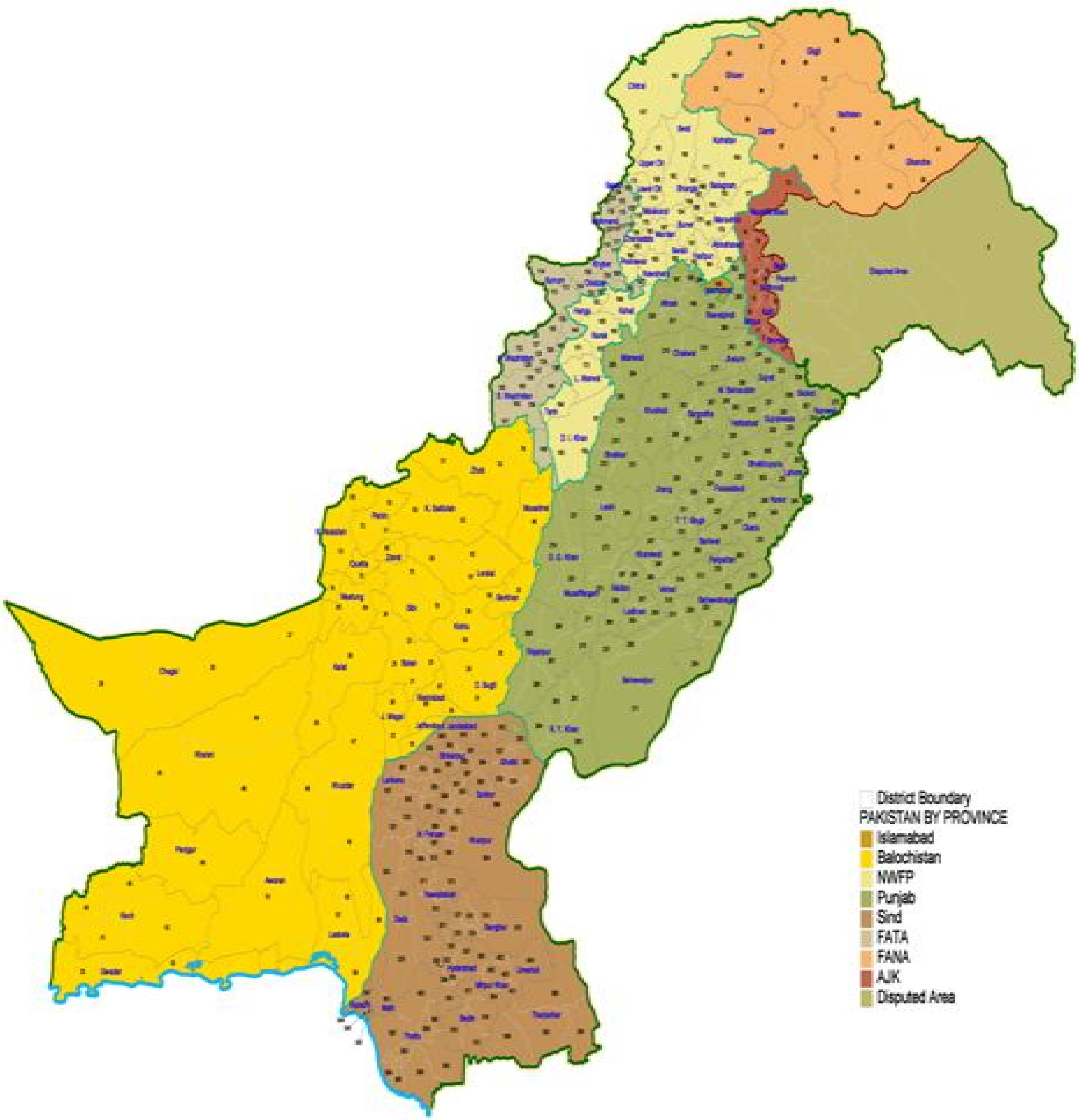
The methodological implication of the above mentioned results is that studies which utilize Ordinary Least Squares to investigate intra- Pakistan socio-economic issues could possibly be producing inaccurate statistical inferences. By assuming spatial-independence, they may produce estimates that are biased and over estimated, since our results show that observations for socio-economic district characteristics do tend to cluster in Pakistan. The main policy implication that emerges from our results is that growth and development policies need to focus on infrastructure and cluster development that can cater to large segments of the population. This is because the spatial pattern of income inequality, district incomes, education levels, and development levels shows how development in Pakistan is concentrated in Punjab (in particular Northern Punjab especially in terms of human development indicators).

Increasing public unrest in Balochistan due to insufficient public sector development activities in it has demonstrated that development in Punjab has taken place at the expense of other provinces especially Balochistan and Khyber Pakhtoonkhwa. Since geography of development matters, it is crucial to reduce these spatial inequalities via expansion of infrastructure and creation of ‘spatially blind’ institutions i.e. institutions that do not discriminate between regions when policies are implemented (World Bank—World Development Report, 2009; Celebioglu and Dall’erba, 2010). The presence of possible spatial spillovers also implies that cluster development can play an extremely important role in generating knowledge externalities, domestic commerce, and employment creation by bringing work and knowledge to people instead of them travelling to it. Pakistan already has many pseudo-clusters that have developed over time. Examples include the IT cluster ‘Karachi’, textile and leather cluster ‘Faisalabad’, automotive manufacturing cluster ‘Port Qasim’, furniture cluster ‘Gujranwala’, light engineering cluster ‘Gujrat’, sports and surgical cluster ‘Sialkot’, heavy industries cluster ‘Wah’ and even light weapons manufacturing cluster ‘Landikotal’. An emphasis on regional and industrial regeneration policies can play a crucial role in enhancing the regional advantages of these districts. Pakistan can learn from available models such as “National Advantage Model (capitalising on natural strengths of the country for instance cheap labour, youth bulge, natural resources, location), Networking Model (banking on and monetising the country’s huge network of SMEs both documented and un- documented), Regional cluster development model (clusters as described above) and Research-Industry relationship model (making of clusters in or around academic institutions)" for the development of new clusters based a region's comparative advantage (Planning Commission, 2011).

Finally, this chapter has highlighted the importance of additional research on Pakistan that takes into account spatial effects. Since this chapter has only considered spatial changes in socio-economic phenomena in 8 years between 1998 and 2005, as a next step I first plan to extend the data set to observe whether there exists a spatio-temporal pattern in the way inequality in income, education and development levels has spread across Pakistani districts. Moreover, an analysis of how growth and inequality have evolved over time and across space would also enrich this analysis. In particular, I plan to carry out a spatial econometric analysis of the effect of a district's inequality on its growth. While the presence of spatial clustering of income inequality in Pakistan could support the use of a spatial lag model to capture the spillover of inequality between districts, missing data on district incomes or omitted variables could also necessitate the use of a spatial error model (which reflects spatial autocorrelation in measurement errors) in analyzing the effect of inequality on district income levels.

APPENDIX

Figure 1. District Administrative Map of Pakistan[49]



[49] Punjab has a population density of 358.5. persons/square kilometer, Sindh of 216 persons/ sq, Khyber Pakhtoonkhwa of 238.1 persons/sq, while Balochistan has a population density of 4.9 persons/ sq. The more densely populated is a region, the more is the spatial interaction in it and hence a greater probability of regional spillovers.

Table 1. List of Districts

	PUNJAB		SINDH	67	Chitral
				68	Malakand Agency
1	Rawalpindi	35	Hyderabad	69	Shangla
2	Jhelum	36	Dadu	70	Bannu
3	Chakwal	37	Badin	71	Lakki Marwat
4	Attock	38	Thatta	72	D I Khan
5	Gujranwala	39	Mirpur Khas	73	Tank
6	Mandi Bahauddin	40	Sanghar	74	Bunir
7	Hafizabad	41	Tharparkar		
8	Gujrat	42	Sukkur		BALOCHISTAN
9	Sialkot	43	Ghotki	75	Quetta
10	Narowal	44	Khair pur	76	Sibi
11	Lahore	45	Nawab shah	77	Nasirabad
12	Kasur	46	Larkana	78	Kalat
13	SheikuhuPura	47	Jaccobabad	79	Pishin
14	Okara	48	Shikarpur	80	Qilla Abd
15	Faisalabad	49	Nowshero Feroz	81	Bolan
16	Jhang	50	Karachi	82	Pangjur
17	TT Singh			83	Barkhan
18	Sargodha		KP	84	Chagai
19	Khushab	51	Peshawar	85	Jaffarabad
20	Mianwali	52	Charsadda	86	Jhal Magsi
21	Bhakkar	53	Nowshera	87	Mastung
22	Multan	54	Kohat	88	Awaran
23	Khanewal	55	Kark	89	Gwadar
24	Lodhran	56	Hangu	90	Turbat
25	Vehari	57	Mardan	91	Kharan
26	Sahiwal	58	Sawabi	92	Ziarat
27	Pakpattan	59	Abbottabad	93	Khuzdar
28	Bahawalpur	60	Haripur	94	Killa Saif
29	Bahawalnagar	61	Mansehara	95	Lasbella
30	R Y Khan	62	Batagram	96	Loralai
31	D G Khan	63	Kohistan	97	Musa Khel
32	Muzaffar grah	64	Swat	98	Zhob
33	Layyah	65	Lower Dir		
34	Rajanpur	66	Upper Dir		

Figure 1a. Provincial Administrative Map of Pakistan

Figure 2. Administrative Set Up Of Pakistan

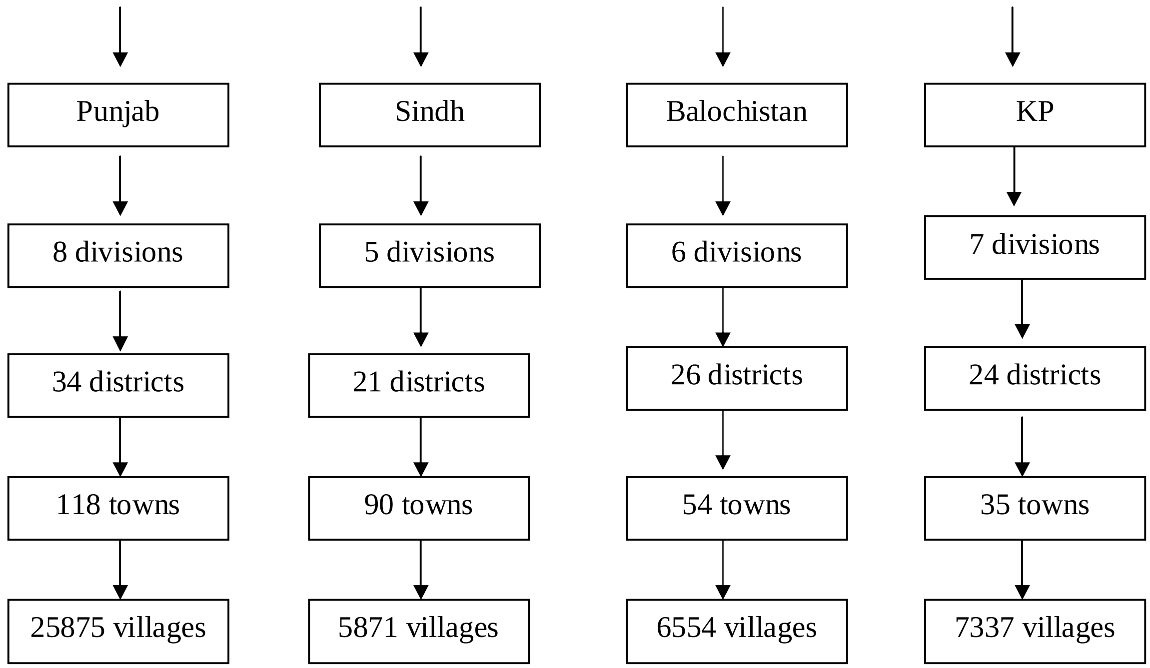


Table 1a. Sample Binary Contiguity Weight Matrix*

	Rawalpindi	Jhelum	Chakwal	Attock	Gujranwal	Mandi Bah	Hafizabad	Gujrat	Sialkot
Rawalpindi	0	1	1	1	0	0	0	0	0
Jhelum	1	0	1	0	0	1	0	1	0
Chakwal	1	1	0	1	0	0	0	0	0
Attock	1	0	1	0	0	0	0	0	0
Gujranwal	0	0	0	0	0	1	1	1	1
Mandi Bah	0	1	0	0	1	0	1	1	0
Hafizabad	0	0	0	0	1	1	0	0	0
Gujrat	0	1	0	0	1	1	0	0	1
Sialkot	0	0	0	0	1	0	0	1	0

Table 1b. Sample Inverse Distance Weight Matrix*

	Rawalpindi	Jhelum	Chakwal	Attock	Gujranwal	Mandi Bah	Hafizabad	Gujrat	Sialkot
Rawalpindi	0	0.003663	0.007813	0.011601	0	0	0	0	0
Jhelum	0.003663	0	0	0	0	0	0	0	0
Chakwal	0.007813	0	0	0	0	0	0	0	0
Attock	0.011601	0	0	0	0	0	0	0	0
Gujranwal	0	0	0	0	0	0	0.020121	0	0
Mandi Bah	0	0	0	0	0	0	0	0	0
Hafizabad	0	0	0	0	0.020121	0	0	0	0
Gujrat	0	0	0	0	0	0	0	0	0.016313
Sialkot	0	0	0	0	0	0	0	0.016313	0

*Full matrix is available in *Stata* format from the author upon request.

Table 2: Global autocorrelation results for income inequality—Geary's C (2005)

Weight Matrix	I	II
i ≠ j	$w_{i,j} = 0 \; or \; 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
$i = j$	$w_{i,i} = 0$	
Geary's C	0.824	1.458
E(C)	1.000	1.000
Sd(C)	0.082	0.324
Z	-2.138	1.413
p-value	0.033	0.158

Source: Author's Calculations

Table 2b. Global autocorrelation results for district per capita income— BC Matrix

Weight Matrix	1998	2005
i ≠ j	$w_{i,j} = 0 \; or \; 1$	$w_{i,j} = 0 \; or \; 1$
$i = j$	$w_{i,i} = 0$	
Moran's *I*	0.818	0.380
E(I)	-0.010	-0.010
Sd(I)	0.103	0.101
Z	8.048	3.856
p-value	0.000	0.000

Source: Author's Calculations

Table 3a. Local spatial autocorrelation for income inequality in 2005 (BC matrix)[50]

Moran's Ii (Gini Coeff for monthly in 2004-05)

dist	Ii	E(Ii)	sd(Ii)	z	p-value*
Barkhan	-1.393	-0.010	0.692	-1.997	0.046
Lodhran	-0.974	-0.010	0.485	-1.990	0.047
Mirpur Khas	-0.967	-0.010	0.562	-1.701	0.089
Kohistan	-0.832	-0.010	0.485	-1.696	0.090
Pangjur	0.930	-0.010	0.562	1.672	0.095
Khuzdar	0.736	-0.010	0.431	1.731	0.084
Batagram	0.967	-0.010	0.562	1.738	0.082
Gwadar	1.172	-0.010	0.562	2.101	0.036
Peshawar	1.460	-0.010	0.692	2.124	0.034
Bahawalpur	0.864	-0.010	0.391	2.234	0.025
Lasbella	1.298	-0.010	0.562	2.327	0.020
Turbat	1.732	-0.010	0.562	3.097	0.002
Awaran	1.349	-0.010	0.391	3.471	0.001
Chitral	2.701	-0.010	0.692	3.916	0.000

*2-tail test

Source: Author's Calculations

[50] Only the statistically significant LISA statistics have been reported here.

Table 3b. Local spatial autocorrelation for income inequality in 2005 (ID matrix)

Moran's Ii (Gini Coeff for mthy in 04-05)

dist	Ii	E(Ii)	sd(Ii)	z	p-value*
Batagram	-5.147	-0.010	0.984	-5.220	0.000
Kohistan	-4.373	-0.010	0.984	-4.433	0.000
Multan	-1.081	-0.010	0.430	-2.491	0.013
Lodhran	-2.249	-0.010	0.984	-2.275	0.023
D G Khan	-2.216	-0.010	0.984	-2.242	0.025
Khanewal	-1.822	-0.010	0.984	-1.841	0.066
Kark	-1.631	-0.010	0.984	-1.646	0.100
Faisalabad	0.720	-0.010	0.448	1.628	0.104
Sargodha	0.844	-0.010	0.519	1.646	0.100
Vehari	1.647	-0.010	0.984	1.684	0.092
Bannu	1.845	-0.010	0.984	1.885	0.059
Charsadda	2.020	-0.010	0.984	2.063	0.039
R Y Khan	2.059	-0.010	0.984	2.103	0.035
Sawabi	2.280	-0.010	0.984	2.327	0.020
Lakki Marwat	2.510	-0.010	0.984	2.561	0.010
Haripur	2.971	-0.010	0.984	3.030	0.002
Chitral	4.195	-0.010	0.984	4.274	0.000
Peshawar	1.192	-0.010	0.240	5.005	0.000
Swat	9.763	-0.010	0.984	9.932	0.000

*2-tail test

Source: Author's calculations

Figure 3a. Moran Scatterplot district real per capita income, 1998 (BC matrix)

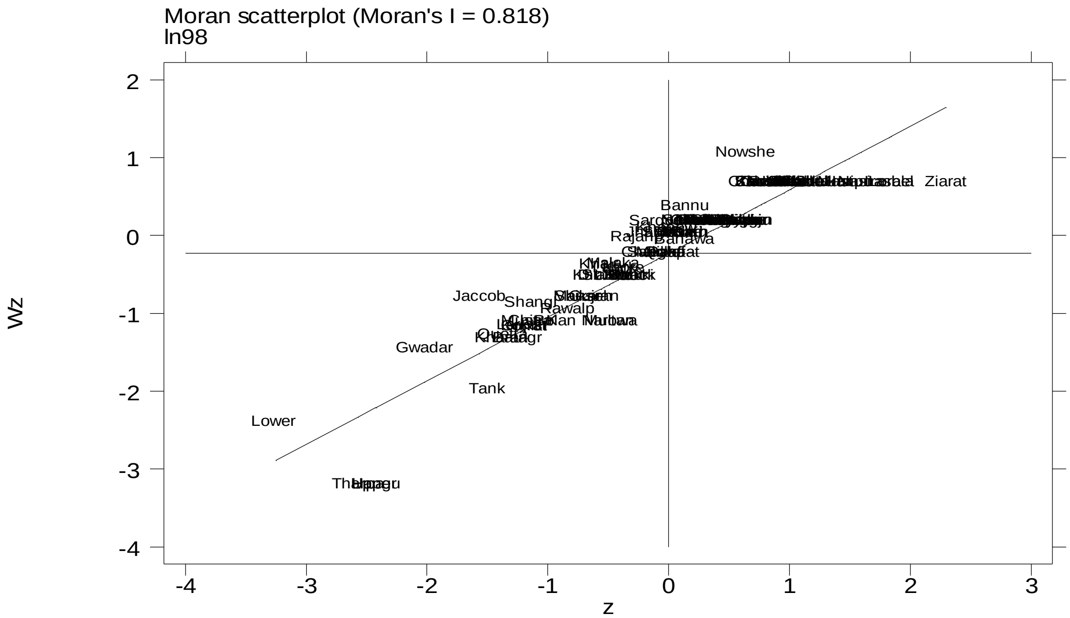


Figure 3b. Moran scatterplot district real per capita income, 2005 (BC matrix)

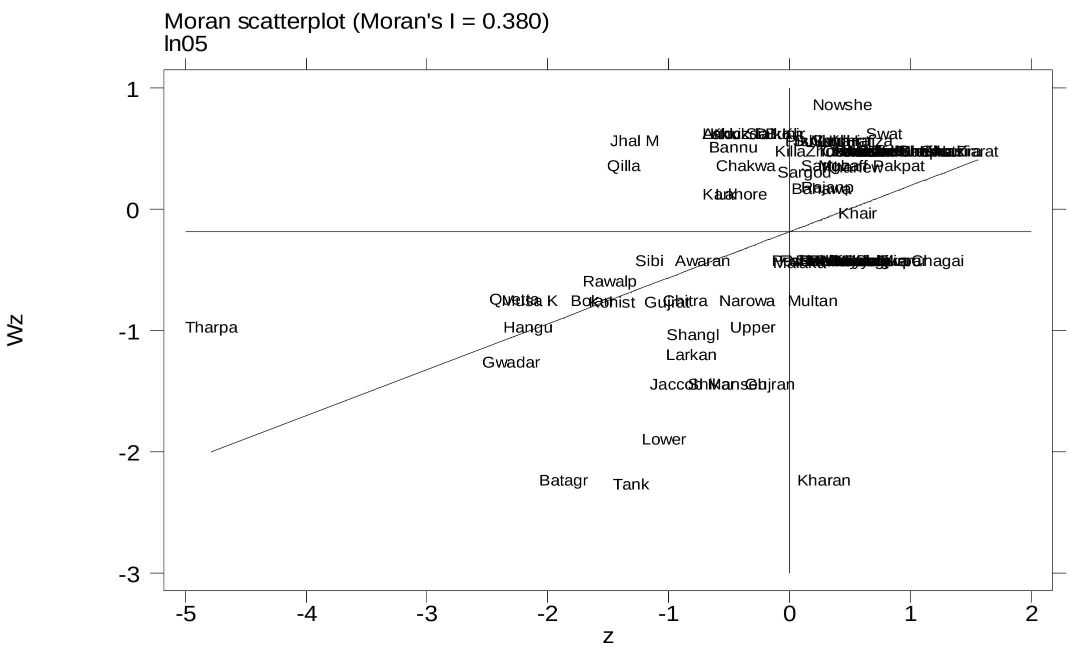


Figure 4a. Moran scatterplot district real per capita income, 2005 (ID matrix)

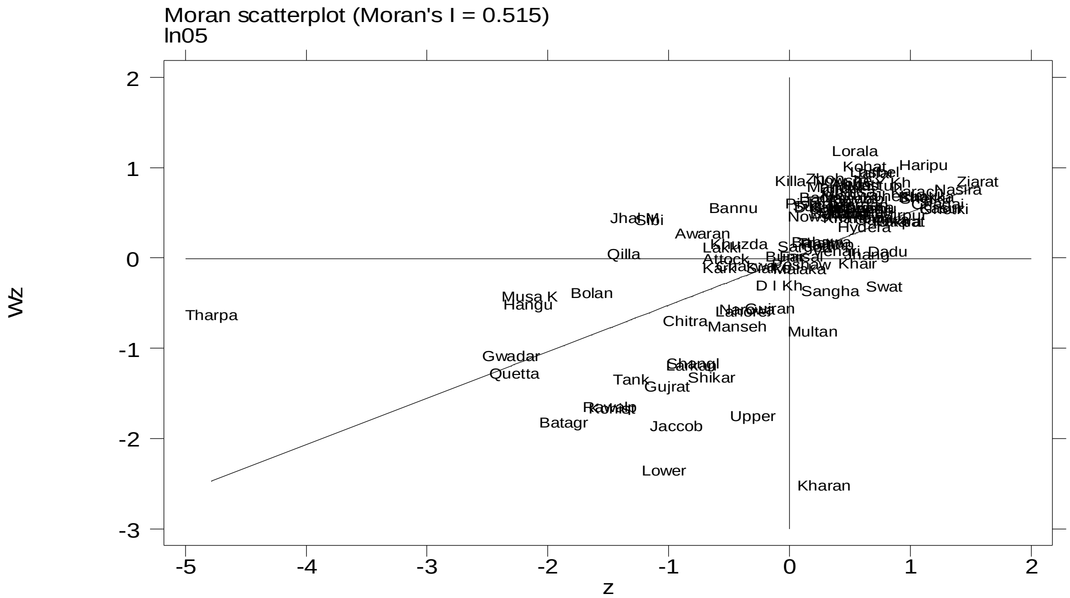


Figure 4b. Moran scatterplot district real per capita income, 1993 (ID matrix)

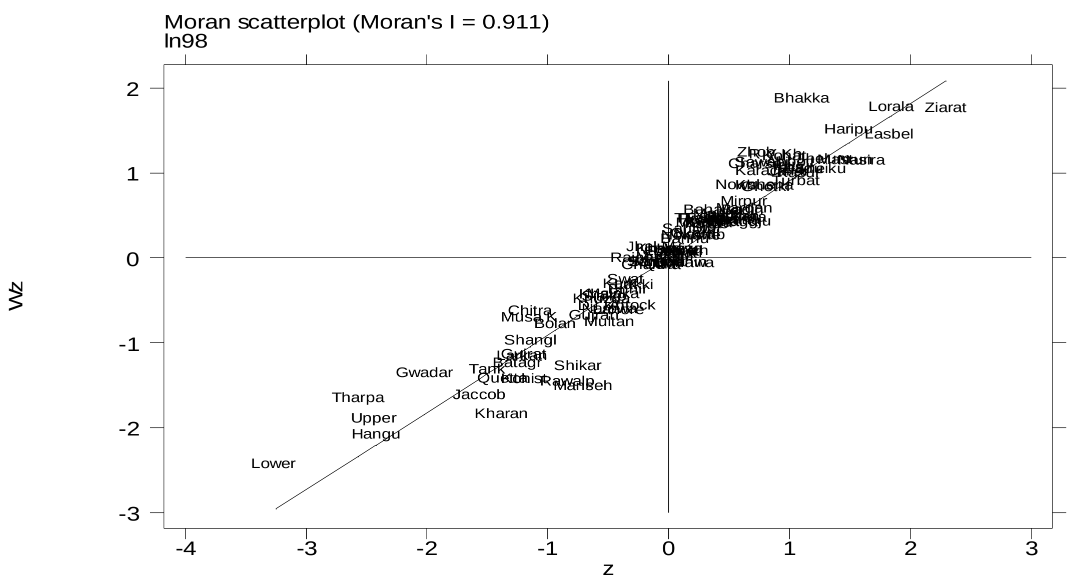
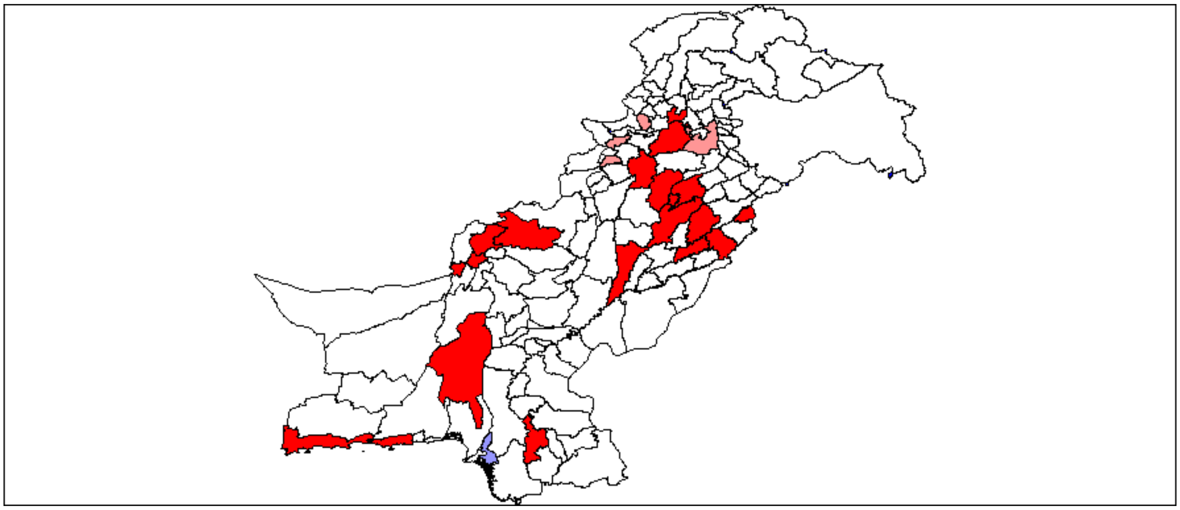
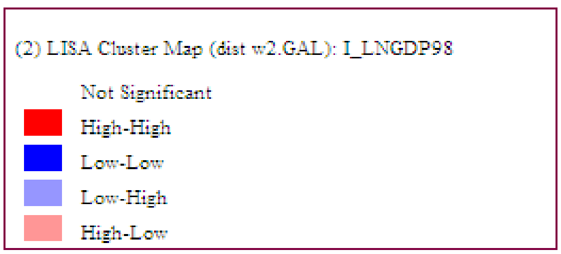


Figure 5. Local Moran statistics for per capita income

1998





2005

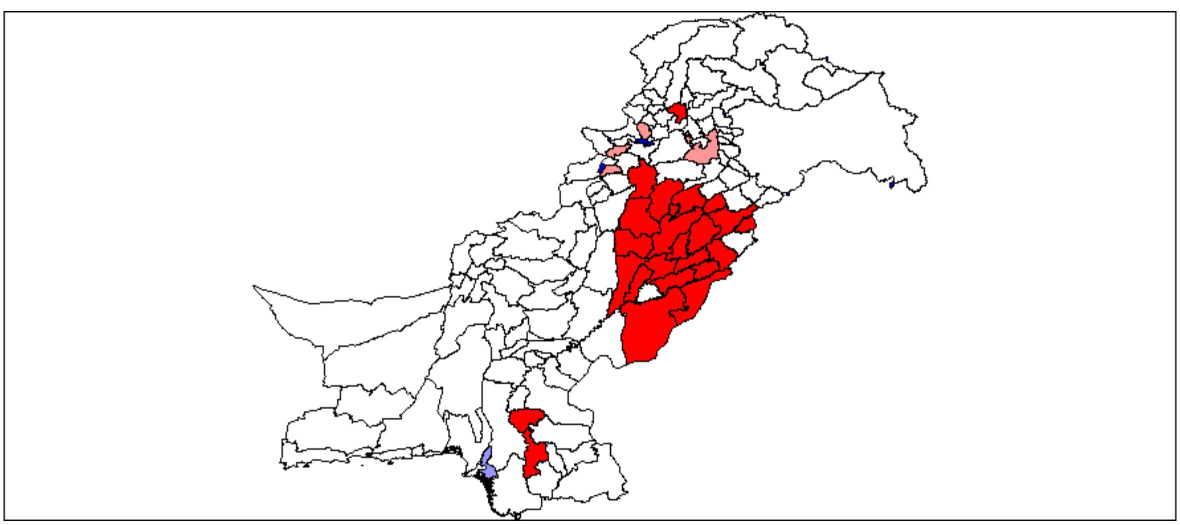


Figure 6. Moran scatterplot for average level of district education using the ID matrix

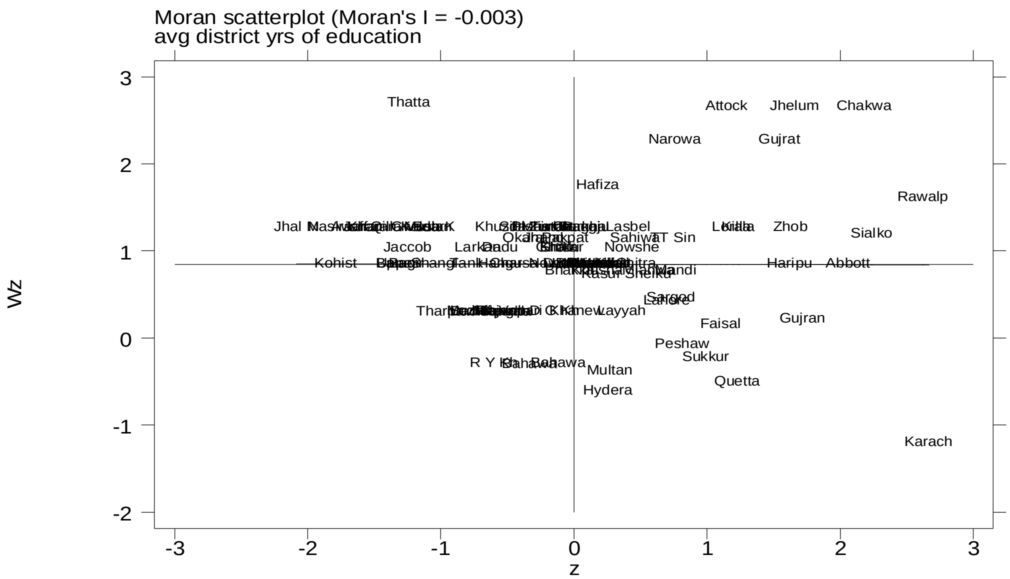


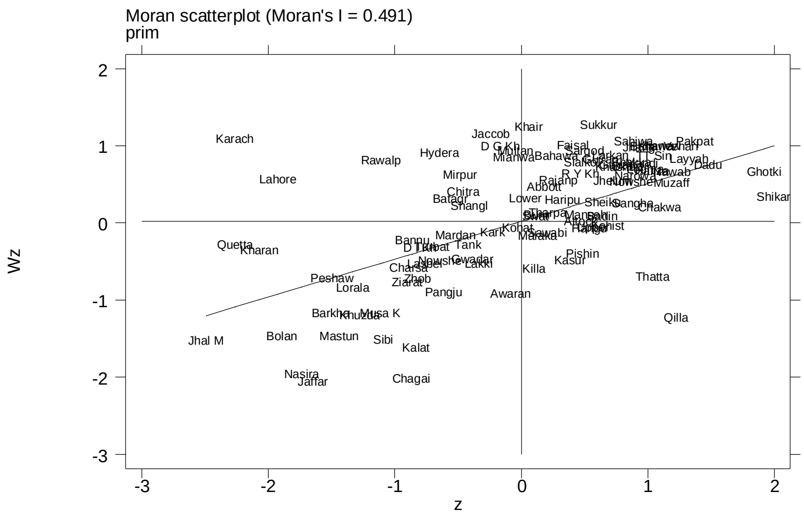
Table 4. Global autocorrelation results for education attainment—Geary's C (2005)

Weight Matrix	I	II
i ≠ j	$w_{i,j} = 0 \; or \; 1$	$w_{i,j} = \frac{1}{d_{i,j}}$
$i = j$	$w_{i,i} = 0$	
Moran's *I*	0.584	1.092
E(I)	1.000	1.000
Sd(I)	0.080	0.275
Z	-5.230	0.336
p-value	0.000	0.737

Source: Author's Calculations

Figure 7. Moran's scatterplot for primary education

i) Binary Contiguity Weights Matrix



ii) Inverse Distance Weights Matrix

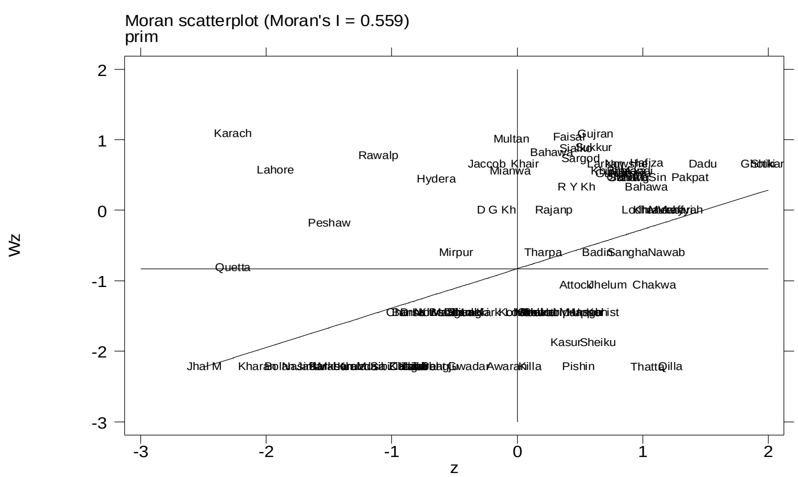
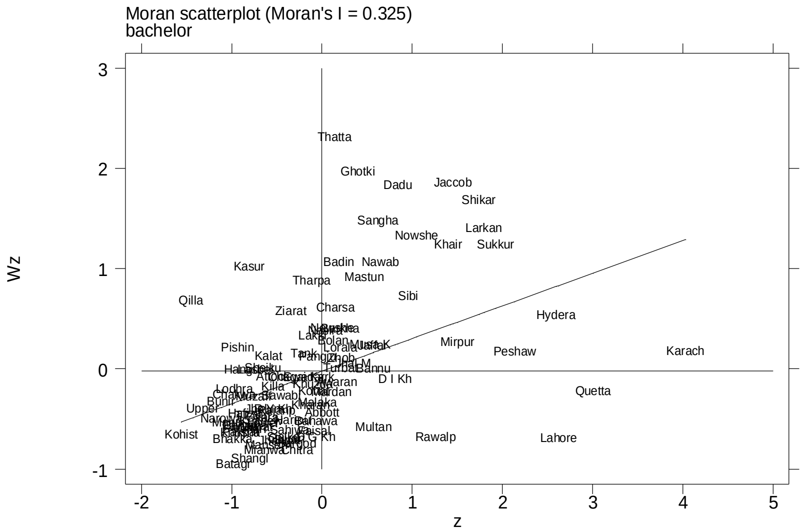


Figure 8. Moran's scatterplot for higher education (Bachelors) in 1998

i) ***Binary contiguity weights matrix***



ii) ***Inverse distance weights matrix***

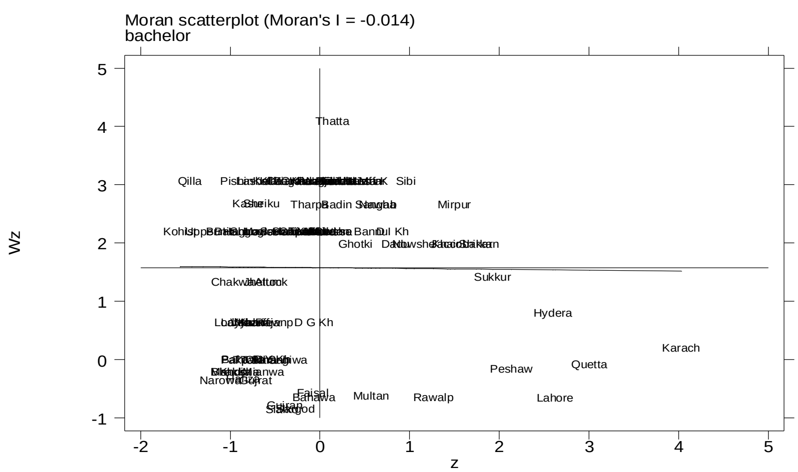
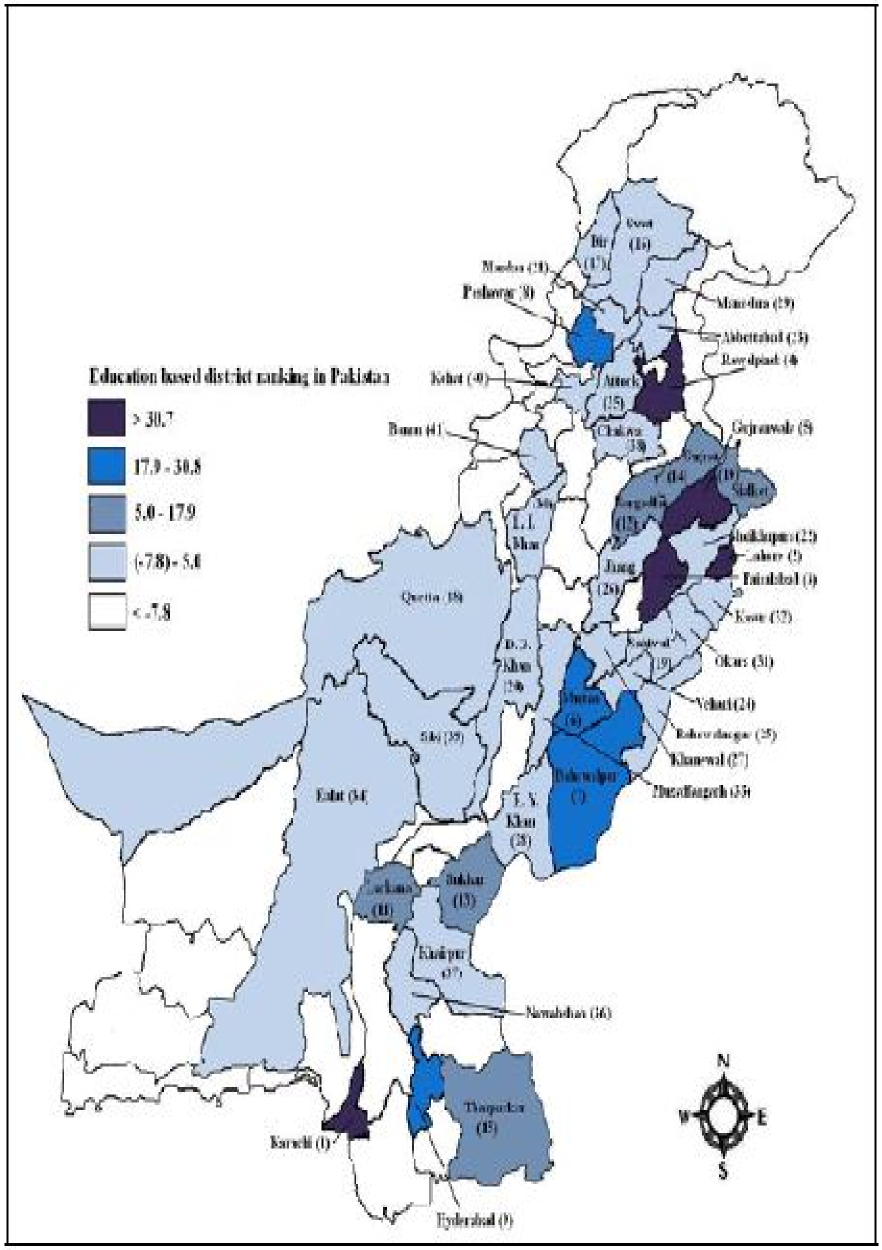


Figure 9. Education Based District Ranking in Pakistan[51]



Source: Burki and Khan (2010)

[51] A cluster map of education based district ranking. The darkest polygons are the districts with the highest rankings. Except for Karachi, only districts in Punjab have high education rankings.

Chapter 4

Are Districts across Pakistan catching up? A Spatial Econometric Analysis of Regional Convergence

Abstract

This study investigates for the first time convergence across 98 Pakistani districts between 1998 and 2005 using spatial and non-spatial econometric techniques. While standard econometric techniques utilized in this chapter demonstrate a lack of convergence, the inclusion of spatial effects leads to an evidence for possible convergence as highlighted by the results of the spatial error model. The results from this chapter show how ignoring spatial dependence in data may lead to model misspecification and eventually inaccurate inferences for the case of Pakistani districts. Several robustness checks have also been carried out to confirm our findings[52].

Keywords: Pakistan, spatial econometric analysis, growth, convergence

[52]**Acknowledgements:** I am indebted to Dr. Alejandro Canadas (Mount St Mary's University, Washington DC) for guiding me in the initial stages on how to compile and analyze spatial data. I'm also thankful to Dr. Harry Kelejian (University of Maryland), Dr. Giuseppe Arbia (University G. D'Annunzio of Chieti, Spatial Econometric Association), Dr. Roberto Basile (University of Naples), Dr. Vaqar Ahmed (Ministry of Planning and Development, Islamabad), the participants at the Advanced Summer School in Spatial Econometrics, Rome (May-June, 2009), and the participants at the School of Spatial Econometrics Trento (March 2011) for their invaluable suggestions and comments on this chapter.

I would also like to acknowledge the data management staff at the Pakistan Institute of Development Economics (PIDE) Islamabad, Federal Bureau of Statistics, Islamabad, and Dr. Amir Jahan Khan and Dr. Haroon Jamal (Sustainable Policy Development Centre, SPDC, Karachi) for their extremely generous data support.

1. Introduction

Do initially poorer economies grow faster in terms of per capita income and catch up to those that were initially richer? Economic theory refers to this issue as regional 'convergence', and it has consistently remained an attractive field of research for economists (Solow, 1956; Barro & Sala-i-Martin, 1995 for a review). Although Pakistan has been a part of the sample in studies on convergence across countries, a limited amount of research has been carried out on the convergence process across regions within the country. Moreover, like most developing countries today, Pakistan faces the daunting challenge of overcoming its unequal regional growth. While the growth hubs of the country continue to be situated in Karachi, some urban centres of Punjab and a few other districts along the river Indus, very few lesser developed districts have been successful in making a shift from periphery to core regions of the country. It has also been noted that economic geography matters for Pakistan as economic inequalities and development tend to cluster across Pakistan (see empirical findings in Chapter 3; Khan, 2003; Burki *et al*, 2010). In this context, analysing convergence and regional disparities in income distribution across Pakistan is a valuable exercise from a policy making perspective.

Despite an increasing agreement on the fact that a region's geographical location influences its growth performance via spatial interactions caused by technological spill-overs and factor mobility between regions (Vaya *et al*, 1998; Moretti, 2004; Rosenthal and Strange, 2005; Fingleton and Lopez-Bazo, 2006), empirical evidence on the spatial effects influencing convergence remains limited (Annekatrin, 2001). Most standard econometric studies neglect spatial effects across borders and assume that all the regions under study belong to the same growth cluster. Such studies overlook the fact that most of the spatial correlation is based in country specific effects (in cross-country analysis) or region specific effects (intra-country analysis). They either nationally weight their data or include a country dummy into the main regression specification to deal with spatial effects. In this way they manage to completely ignore the possible spatial correlation from the disturbance term, while failing to shed light on the process of convergence itself.

In the light of the above mentioned issues, this chapter aims to address the weaknesses of studies on inter/intra regional convergence that neglect spatial effects and contribute towards the literature on the spatial analysis of regional convergence by exploring the process of convergence across 98 Pakistani districts between 1998 and 2005 using spatial and non spatial econometric techniques[53]. Its main goal is to examine absolute and conditional convergence of real per capita income in Pakistan based on a district level aggregation. The advantage of working at a district level is that it allows for better exploitation of the geographical characteristics of socio-economic data and a deeper analysis of the spatial effects (regional spill-overs and spatial regimes) as compared to analysis carried out on a provincial level in Pakistan. It is the first study of its kind carried which analyses Pakistani districts and that too from a spatial econometric perspective.

The layout of the paper is the following: In Section 2, a review of convergence and growth concepts along with a description of the standard econometric techniques that have been used to study them is presented. Section 3 analyses convergence models from a spatial perspective and sheds light on the spatial econometric techniques that have been utilized in this chapter. Section 4 discusses the empirical results for convergence obtained from spatial and non-spatial techniques. Section 5, provides an alternative statistical methodology based on non-parametric analysis techniques to analyze convergence. Section 6 concludes.

2. Convergence and Growth concepts

Two forms of convergence hypothesis have been analyzed over the years in the growth literature (Sala-i-Martin, 1990; Barro and Sala-i-Martin, 1992). The first form is known as *Sigma* (σ) convergence which demonstrates the evolution of the dispersion of real per capita incomes across a group of regions or countries over time. A reduction in the dispersion of regional per capita incomes indicates σ-convergence between the regions. Measures such as standard deviation, coefficient of variation, and the log of per capita income have been utilized in literature to examine this form of convergence (Carlino and Mills, 1996a; Bernard and Jones, 1996a; Kuznets, 1955; Easterlin, 1960a, 1960b; Amos, 1988, 1989; Fan and Casetti, 1994). However, Rey and Montoury (1988) argues that σ-convergence analysis may "mask non-trivial geographical patterns that may also fluctuate over time" (Arbia, 2005:7-8). Therefore it is equally important to analyse the "geographical dimensions of income

[53] The analysis is carried out for the time period between 1998 and 2005. It should be noted here that the choice of the years for this chapter is data driven. District level analysis on Pakistan so far is rare since district level data is not annually produced.

distribution in addition to the dynamic behaviour of income dispersion" (Arbia, Basile and Salavtore, 2003:5).

The second form of convergence studied mainly by macroeconomists is referred to as the 'Absolute *Beta* (β)-convergence', which tests whether poorer regions grow faster than richer regions, and occurs when the former catch up to the per capita income levels of the latter. This is a neoclassical approach to β-convergence based on the Solow-Swan growth model which predicts that states (or regions within a country) would have similar levels of real per capita income in the long run under provided that they all share similar preferences and technology, and technological barriers that can prevent the flow of labour and capital do not exist. It is based on the assumption of decreasing marginal productivity of capital, implying that due to their high endowment of capital, richer regions would grow at a slower rate than poorer regions (Barro and Sala-i-Martin 1991, 1997, 2003). This chapter utilizes the following cross-sectional econometric specification to test absolute β-convergence:

$$growth_{i,T-t} = \alpha + \beta \log y_{i,t} + u_{i,t} \qquad (1)$$

where $y_{i,t}$ denotes the initial per capita income or GDP in region/country *i* at the beginning i.e. time *t* ; T denotes the last time period under study or *t+x* years; α is the intercept term and β the parameters that are estimated (where $\beta < 0$); and $u_{i,t}$ is the stochastic error term distributed iid $(0,\sigma^2)$. The dependent variable is measured as:

$$growth_i = \log\left(\frac{GDP2005}{GDP1998}\right) \qquad (2)$$

This standard model in the growth/convergence literature rests on a highly restrictive assumption that all the regions under study have the same steady-state income path which is a function of the initial conditions (Darlauf and Quah, 1999; Magrini, 2004). A negative value of β suggests beta convergence (Sala-i-Martin, 1996) implying that per capita income growth rates over the time period under study were negatively correlated with the initial income levels. The rate of convergence is calculated using the following formula:

$$C = -[\ln(1 - \beta P)]/P \qquad (3)$$

where P is the number of years (time period) being considered.

The absolute or the unconditional convergence hypothesis is confirmed usually only when applied to relatively similar economic units such as the provinces of the same country or countries belonging to the same region or political arrangement e.g. the EU (Arbia, 2005). For other cases, the absolute convergence model has been modified by relaxing the assumption of similar steady states for the regions under study. This is done by incorporating additional explanatory variables on the right hand side of equation (1), and the modified specification is referred to as the *conditional convergence* or an extended growth model (Barro and Sala-i-Martin, 1995; Mankiw et al, 1992):

$$growth_{i,T-t} = \alpha + \beta \log y_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \quad (4)$$

where $X_{i,t}$ is a vector of explanatory variables at the beginning of the time period *t*. The conditional β-convergence, analyses the income per capita of regions (or countries) with similar socioeconomic and political characteristics which converge to their own steady state income levels in the long run (Arbia, 2005)[54].Most studies first estimate σ–convergence and then β-convergence because while β-convergence is a necessary condition, it is not a sufficient condition for σ-convergence to take place (Barro and Sala-i-Martin, 1991; Sala-i-Martin, 1996; Young, Higgins and Levy, 2008).

However Quah (1993) argues that these traditional cross-sectional specifications do not reveal the actual growth dynamics because these models may suffer from the heterogeneity bias induced by the restrictive assumptions on which they have been based. As a remedy to the heterogeneity bias, studies on convergence began to adopt a time series analysis to investigate common trends of convergence between regions over time. In such models convergence would occur if the long run forecasts of the differences in the per capita income levels would go to zero (Rey and Montouri 1999; Carlino and Mills, 1996a, 1996b, and Bernard and Durlauf, 1995). This is often not possible because of shocks to individual regional economies. As noted by Bernard and Darlauf (1996), with the occurrence of a shock, the "time series would contain unit roots", and the due to the stationarity requirement, such a form of convergence has been labelled 'stochastic convergence' (Rey and Montouri, 1999 p.145). However, this approach is dependent on the availability of long series of data over time at a national and provincial level, often which either does not exist, is difficult to obtain, or has been collected at irregular intervals over the years (Yildirim et al, 2009).

Most of the above mentioned approaches consider the regions which they study as isolated entities ignoring any role of spatial interaction. Exceptions include studies such as

[54] "Two things are analysed from this analysis, namely the speed of convergence and the so called half life. The speed of convergence refers to the speed at which the economy is converging towards the steady state, while half life is the time necessary for half of the initial gap in per capita output to be eliminated "(Arbia 2005:5)

Quah (1996b), Fingleton (1999) and Rey (1999) which combine Markov chain approach (to study the entire income distribution) with spatial methods. Their findings show that spatial effects are more influential than national factors for explaining regional convergence across Europe (Quah, 1996b). Rey (1999) has also demonstrated similar results for US states. Studies that neglect the spatial features of their data, fail to take into account the fact that spatial data represents the aggregation of individuals in arbitrary geographical borders which portray different political and historical meanings. The level of spatial aggregation is hence crucial because different levels (countries, provinces, districts, states, towns, municipalities, villages etc) can lead to different estimates of a certain phenomenon (Arbia, 1988).

Therefore another branch of research emerged to explicitly take into account spatial association amongst observations and modelled it to examine regional convergence. Such analysis is based on the idea that inter-regional interactions induce dependence between regions. For this reason regional data cannot be considered as independently generated since their exist similarities amongst neighbouring regions (Anselin 1988; Anselin and Bera, 1988). While technological and knowledge spillovers have been identified as important instruments that may contribute towards convergence (Krugman, 1987; Jones, 1997), regional specialization in various economic sectors could even lead to divergence instead of convergence at the aggregate level (Bernard and Jones, 1996b). Such arguments demand an explicit incorporation of *spatial dependence* between the regions under consideration in the model, so that the spatial pattern (if any) of economic growth can be identified and quantified (Anselin and Rey, 1991, Rey and Montouri, 1999). Spatial dependence may also arise due to the "...mismatch between boundaries of the market process...and the administrative boundaries used to organize the data." (Rey and Montouri, 1999:145). Such spatial dependence has been referred to as '*nuisance dependence*' because it is reflected by the spatially autocorrelated error term (Anselin and Rey, 1991).

Another type of spatial effect relevant to studying convergence is the presence of *spatial heterogeneity* i.e. the "...instability of a behavioural relationship across the observational units (Rey and Montouri, 1999)." Spatial heterogeneity challenges the assumption of identical rate of convergence for all regions, on which the traditional cross sectional convergence models are based. Recent developments in spatial econometrics have made it possible to incorporate the spatial effects mentioned above in convergence studies by suggesting new estimators for models that explicitly take spatial effects in account.

3. Spatial Analysis of β-Convergence

As noted in our discussion above, spatial effects can significantly affect any possible regional income convergence estimates. To analyze regional growth and income inequality in a spatial econometric framework, the unconditional model as seen in equation (1) is utilized as a starting point:

$$growth_{i,T-t} = \alpha + \beta \log y_{i,t} + u_{i,t} \quad (5)$$

However as previously argued, ignoring spatial effects when they actually exist may lead to serious econometric misspecification reflected in auto-correlated residuals which violate the OLS assumption of uncorrelated error terms. Hence specification (4) is modified to incorporate spatial effects in order to examine whether spatial autocorrelation takes the form of spatial dependence or of nuisance dependence.

3.1 The spatial lag model (SLM)

The following spatial lag model has been estimated in this chapter:

$$growth_{i,T-t} = \alpha + \beta \log(y_{i,t}) + \rho W growth_{i,T-t} + \gamma X_{i,t} + \varepsilon_{i,t} \quad (6)$$

where $W growth_{i,T-t}$ is the spatially lagged dependent variable for a given spatial weights matrix W, ρ denotes the spatial autoregressive parameter which reflects the spatial dependence in the data being used (Le Sage, 1999; Yildirim et al.,2009), and $\varepsilon_{i,t}$ is a vector of error terms. The spatially lagged dependent variable is always correlated with ε_i, not just at location i but also with the error terms at other locations.

The SLM model also known as spatial auroregressive model (SAR) can be interpreted in various ways (see Rey and Montouri, 1999; Anselin and Bera, 1998). From a strictly technical perspective and in the context of convergence, it can be used to investigate whether spatial dependence of regional growth is a by product of convergence or spatial clustering of initial income (AnneKatrin, 2001). Hence the SAR model can highlight whether the negative relationship between growth and initial income remains robust after controlling for spatial dependence or not. Another interpretation of the SAR model emphasizes the spatial interaction in the data generating process.

The spatial lag model in our case, quantifies how the growth rate in a region is affected by the growth rate of its neighbouring regions. For instance, a ρ of 0.5 would means that GDP growth rate in a given district would increase by 1 percent for every 2 percent increase in the GDP growth rate of its neighbouring districts regardless of the values of the independent variables contained in the X vector (Bernatt, 1996).

When the ignored spatial effects take the form of spatial dependence, the OLS estimation of a spatial lag model would yield biased and inconsistent estimates for the coefficients due to the simultaneity introduced through the spatial lag. Moreover, inference based on the estimated coefficients will become inaccurate. As a remedy, estimators based on maximum likelihood (ML) and instrumental variables (IV) have been suggested as consistent estimators (Anselin 1998; Kelejian and Robinson, 1993; Anselin and Bera, 1998; Kelejian and Prucha 1998; Conley, 1999; Canadas, 2008).

3.2 The spatial error model (SEM)

The spatial error specification—also referred to as nuisance spatial autocorrelation—assumes that spatial dependence operates through the error process. The error term in equation (5) in this case may reveal spatial covariance which can be expressed as:

$$growth_{i,T-t} = \alpha + \beta \log(y_{i,t}) + \gamma X_{i,t} + \varepsilon_{t,i} \tag{7}$$

$$\varepsilon_{t,i} = \tau W \varepsilon_{t,i} + \mu_i$$

$$\varepsilon_{i,t} = (I - \tau W)^{-1} \mu_{i,t} \tag{8}$$

where τ is a scalar spatial error coefficient or a nuisance parameter in the sense that it is not assigned any meaningful economic interpretation; $\mu \sim N(0, \sigma^2 I)$, and W is a spatial weights matrix.

Under such circumstances, the original error term has a non-spherical covariance matrix, and in the presence of non-spherical errors, the ordinary least squares (OLS) estimates of the parameters are inefficient (yet unbiased) estimates of the parameters but give biased estimates of the parameter variances. As a solution, spatial econometric literature suggests the estimation of this model by Maximum Likelihood (ML) or by Generalized Method of Moments (GMM) (Anselin 1998; Kelejian and Robinson 1993, Anselin and Bera 1998; Kelejian and Prucha 1998; Conley 1999; Rey and Montouri, 1999; Canadas,2008). As a spatial process, when a spatial error specification is applied to test for convergence equation (7) takes the following form:

$$growth_{i,T-t} = \alpha + \beta_1 \, lny_{i,t} + \beta_2 \, X_{i,t} + (I - \tau W)^{-1}\mu_{i,t} \tag{9}$$

From the above model it can be observed that through the spatial transformation $(I - \tau W)^{-1}$, a random shock introduced in one district will not only affect its own growth rate but will also influence the growth rates of other districts. This model implies that a districts' growth rate is affected by the growth rate of its neighbouring districts "only to the extent that the neighbours have above or below 'normal' growth rate, where normal is defined as growth rate predicted by equation" (Bernatt, 1996:7).

3.3 Spatial specification tests

Several tests can be utilized as robustness measures for our specifications and in order to endorse which model should be used. The first used in this study is the Moran's I test which is applied to the residuals of the OLS specification (see chapter 2 for details):

$$I = (N/S_0)[e'We/(e'e/N)] \tag{10}$$

where e is the nx1 vector of residuals in the OLS specification, W is the weights matrix, and $S_0 = \sum_i \sum_j w_{i,j}$ is a standardisation factor. Moran's I statistic is sensitive to both types of spatial autocorrelation, and does not assist in distinguishing between which specification is more appropriate. Being able to differentiate between the two forms of spatial dependence is crucial because of different interpretations of the nature of spatial effects.

As a remedy the simple and robust versions of the Lagrange Multiplier (LM) test have been employed (Anselin and Rey, 1991). The LM test is based on Maximum Likelihood function, under which the unrestricted and the restricted models are compared. In the restricted model, the spatial parameters rho and lamda are set equal to 0, while in the unrestricted version rho and lamda are equal to their respective coefficients. For the test to be carried out, it is enough to estimate the restricted model and verify whether the slope of the likelihood function at that point equals 0. The restriction can be considered to be valid if the slope is equal to 0, since it implies that the value of the function will approximately be close to the maximum likelihood (Costa *et al*, 2009).

Following Anselin (1991:21), the LM test can be specified as:

$$LM_{ERROR} = [e'We/(e'e/N)]^2/[tr(W^2 + W'W)] \tag{11}$$

$$LM_{LAG} = [e'Wy/(e'e/N)]^2/D \tag{12}$$

where y is the logarithm of real per capita district GDP, and D = $[(WX\beta)'\,(I - X(X'X)^{-1}(WX\beta)/\sigma^2] + \; tr(\,W^2W'W)$. Under the lag and error versions, the test statistic is asymptotically distributed as χ^2 with one degree of freedom. Finally, the details on the construction of the weights matrix can be found in Section 4.3 of Chapter 3.

Table 1. Decision rule for the LM tests

LM-lag LM-error	Do not reject Ho Do not reject Ho	Reject Ho Do not reject Ho	Do not reject Ho Reject Ho	Reject Ho Reject Ho	
LM-lag robust	--	--	--	Do not reject Ho	Reject Ho
LM-error robust	--	--	--	Reject Ho	Do not reject Ho
Decision	No spatial effects	Lag	Error	Error	Lag
Source: Costa *et al* (2003)					

4. Data and Variable Construction

Real GDP data is not available at a district level in Pakistan. The limited amount of research that has investigated socio-economic issues requiring the estimation of district incomes have utilized district wise cash value of agriculture produce and manufacturing value added as proxies to compute district incomes (National Human Development Report (NHDR), 2003; Jamal and Khan, 2007)[55]. Income data used in this chapter has been obtained from Jamal and Khan (2007). Further technical estimation details can be found in NHDR, 2003: Annex 1(b) and in Jamal and Khan (2007). All income data from 2005-06 was deflated using the Pakistani Consumer Price Index (CPI) of 1998.

[55] District wise crop statistics and Census of manufacturing Industries (CMI) of relevant years close to 1998 and 2005 have been utilized by these studies to estimate district wise cash value of agriculture produce and manufacturing value added as proxies to compute district incomes (Jamal and Khan, 2007).

In order to investigate conditional convergence following Barro (1992) and Trivedi (2002), this chapter has selected the following variables as proxies for human and physical capital that can affect growth: district wise population density, district wise percentage of vaccinated children below 10, district wise percentage of adults having completed at least 10 years of education and district wise percentage of households that consume electricity as the main source of energy. The data on these variables has been obtained from four provincial reports (for the provinces of Punjab, Sindh, Khyber Pahktoonkhwa, and Balochistan) known as ‘Socio-economic Indicators at District Level’ and District Census Reports (1998), both of which have been published by the Federal Bureau of Statistics, Government of Pakistan.

Finally in order to incorporate spatial effects into the growth models estimated in this chapter, two weights matrices have been utilized. The first matrix is a simple binary contiguity W matrix (BC matrix from now onwards) based on the concept of Queen Contiguity i.e. if a district *i* shares a border *or* a vertex with another district *j*, they are considered as neighbours, and $w_{i,j}$ takes the value 1 and 0 otherwise. The second weights matrix is based on inverse average road distance from the centroid of a district to the centroid of the closest district which has a ‘large city’ in it. Details on the constriction of these two matrices have been provided in Chapter 3, Section 3.

5. Empirical Results

5.1 Absolute convergence analysis

As the first step, an unconditional beta-convergence model for 98 Pakistani districts between 1998 and 2005 is estimated by means of Ordinary Least Squares (OLS):

$$growth_{i,05-98} = \alpha + \beta \log(y_{i,98}) + u_i \qquad (13)$$

$$u_i \sim i.i.d\ (0, \sigma_u^2)$$

Growth—the dependent variable is the district’s growth rate of real per capita GDP for the period. It is measured as:

$$growth_i = \log(\frac{GDP2005}{GDP1998}) \qquad (14)$$

If convergence holds, a negative and a significant coefficient for β (our variable reflecting initial income conditions) is expected. As illustrated in Table 2, the regression yields a negative coefficient of the initial level of income (log of real per capita GDP in 1998). However, it is not significant indicting an absence of absolute convergence amongst Pakistani districts between 1998 and 2005.

Table 2 also displays diagnostic results carried out to detect any misspecifications. Three diagnostic tests for spatial dependence for this regression—Moran's I test, and two robust Lagrange multiplier tests, have been utilized. Since the Moran's I does not distinguish between the two forms of misspecification (Rey and Montouri, 1999), the robust versions of Lagrange Multiplier error (RLME) and lag (RLML) tests have been reported to indicate whether the source of spatial dependence is a by-product of measurement errors, or of regional interaction[56]. The RLME and RLML, both reject the null hypothesis of spatial randomness (specifically no spatially autocorrelated error terms in the case of robust LM error test (RLME) and no omitted spatial lag in the case of robust LM lag test (RLML)). However, the robust LM (error) test is slightly more significant than the robust LM (error) test i.e the p-value of RLML test is lower than that of the RLME test.

Table 2: Absolute ß-convergence of per capita income in 98 Pakistani districts (1998-2005) Dependent variable: Growth rate of real district per capita income	
Log GDP98	-0.0159
	(0.0106)
Constant	0.148*
	(0.0777)
Goodness of Fit	
R-squared	0.0230
Akaike Information Criterion (AIC)	-254.0670
Observations	98
Regression Diagnostics	
Breusch-Pagan Test	38.79***
Jarque-Bera Test	59.39***
Moran's I (error)	-1.3620
Robust Lagrange Multiplier (lag)	4.386**
Robust Lagrange Multiplier (error)	5.167**
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations	

[56] "While the robust LM test (error) tests for spatially autocorrelated error terms and the possible presence of a missing spatially lagged variable, the robust LM test (lag) does the opposite"(Feldkircher, 2006).

The robust tests reveal adequate evidence of spatial autocorrelation implying rejection of the null hypothesis of spatial randomness for values of growth of real per capita GDP between Pakistani districts. However, extra caution must be exercised in interpreting the overall result of divergence because we fail to reject the null hypothesis of homoskedasticity and normality amongst residuals. Long and Ervin (2000) demonstrates via various simulations that there exists a very high possibility for the failure of heteroskedasticity detection tests. As a result it recommends that "a "test for heteroskedasticity should not be used to determine whether (an HC estimator) should be used". Even on the suspicion of the slightest heteroskedasticity, it suggests using an HC estimator. Hence, Col 2 in Table 3, uses the HC1 estimator and provides the results with bootstrapped standard errors.

There still remain other econometric issues that need to be taken into consideration. For example, cross-sectional regression results that are based on relatively smaller samples can be highly sensitive to outliers. As noted in Trevedi (2003), few influential but atypical outliers need to be appropriately accommodated to prevent them from having a distortionary effect on the parameter estimates. Under such circumstances, Temple (2000) suggests on adopting robust estimation techniques along with OLS estimates for which we compute Re-weighted Least Squares (RWLS) estimates as well (Table 3, Col 2). RWLS estimates are less sensitive to outliers than OLS estimates. To compute RWLS estimates, first OLS regression is carried out and the Cooks distance (D) is calculated. This distance (D) can be perceived as an index which is influenced by the size of the residuals, outliers, and the size of the levearage of each observation[57]. It measures the distance between "the coefficient estimates when the *i*th observation is omitted and when it is not" (Trivedi, 2003: 8). An observation for which D is greater than 1 is considered as an outlier and is omitted. This initial screening process is followed by a series of iteratively performed weighted regressions which simultaneously use two types of weight functions—Huber and biweights—to derive the weights. Iterations seize when the maximum amount of change in weights drops below a level of predetermined level of *tolerance.* However, even the RWLS estimate of initial income (Table 3, Col 2) continues to remain negative and insignificant supporting the previous finding of lack of convergence.

[57] Large residual and leverage values raise the value of D.

It has been noted in Verardi and Croux (2010) that diagnostics such as the Cook's Distance used in RWLS, evaluate each observation individually and there might not be adequate pairs or groups of outliers to exert excessive influence. They may conceal the influence of each other when testing for a single one, and may not be able to guarantee the identification of all leverage points. Hence to bolster our results, Iterative Re-weighted Least Squares (IRWLS) technique has been utilized as another robustness measure[58]. The estimates from this technique again support the claim of lack of unconditional convergence (Col 3, Table 3).

Since our measure of GDP per capita has been indirectly calculated by adding the two values of value addition (in Rs million) for agriculture and manufacturing for every district , separate convergence regressions for the two components have also been estimated to confirm the results from Table 1. In both the regressions (see Table 1 and Table 2 in the appendix), the coefficient of β is no longer negative and is highly significant suggesting absolute divergence of district-wise manufacturing and agriculture income values between 1998 and 2005. The results after obtaining bootstrapped robust standard errors, and after estimating RWLS and IRWLS estimates for the components, still remain unchanged i.e. they reject absolute convergence (Tables 5 and 6 in the appendix).

Table 3: Absolute β-convergence of per capita income in 98 Pakistani districts (1998-2005)
Dependent variable: Growth rate of real district per capita income

	(1)	(2)	(3)
	OLS (Robust SE)	Robust Reg RWLS	IRWLS
Log GDP98	-0.0159	-0.0215	-0.0213
	(0.0189)	(0.0212)	(0 .0195)
Constant	0.1479	0.1959	0.1988
	(0.1423)	(0.1608)	(0.1473)
R-squared	0.0227	0.0611	-
RMSE	0.0655	0.0531	-
P-subsets			20
Observations	98	98	98

Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations

[58] The IRWLS estimates are obtained using the **mmregress** command in STATA 10. The initial values for the iteratively reweighted least squares algorithm are monotone M-estimators that are not robust to bad leverage points and may lead the algorithm to converge to a local instead of a global minimum.

5.2 Conditional convergence analysis

For a more accurate idea of which model is more appropriate for the data being utilized, it is necessary to investigate the causes that have led to increasing divergence amongst districts between 1998 and 2005. In this section, a conditional convergence analysis is carried out to take into account the role of some district-specific characteristics and geographical interactions in regional growth. The choice of these explanatory variables has been based on studies such as Barro (1992) and Trivedi (2003) in such a way so that regional differences in social, economic and physical differences can be taken into account. First the following growth model has been estimated through least squares:

$$growth_{i,05-98} = \alpha + \beta \ln GDP98 + \gamma X_{i,98} + \varepsilon_i \qquad (15)$$

where *growth* is an (nx1) vector of district *i*'s growth rates of real per capita GDP between 1998 and 2005. The variable $X_{i,98}$ represents a vector of control variables in 1998. In equation (15), initial values of district specific explanatory variables have been added to the log of real per capita GDP for 1998 (*initial income* of a district) to also investigate their impact on the growth rate of district wise real per capita GDP between 1998 and 2005.

As argued in Partridge (2005), there exist two schools of thought on whether to consider the level of initial income as a proxy for the initial level of development in growth regressions or not. In the neoclassical growth models, initial levels of income are considered as proxies for the initial level of development and inversely related to subsequent growth (Canadas, 2008; Barro, 2000)[59]. These models confirm gradual convergence amongst state economies during the 1800s and the 1990s (Barro and Sala-i-Martin, 1991a). This is because neoclassical convergence models are based on diminishing marginal returns and differentials in returns to capital, implying that factors of production will shift to developing countries where returns to them are higher (Ventura, 1997; Partridge, 2005).

However, the second of school of thought argues that initial income should not be included since eventually –as also argued by endogenous and neoclassical growth models—

[59] Alesina and Perotti (1994), Alesian and Rodrick (1994), Persson and Tabellini (1994), Clarke (1995), Perotti (1996), Li and Zou (1998), and Partridge (1997) are examples of other studies which employ the level of income. Barro (2000), Rey and Montouri (1999), and Catello and Domenech (2002) are studies which employ the log for initial income level.

the initial income term drops out when economies approach their steady state growth paths (Darlauf and Quah, 1999). As claimed in Partridge (2005), initial income term can be omitted when states are close to their steady state levels, because then deviations mainly reflect transitory cyclical and structural shocks instead of neoclassical convergence. When cyclical conditions are dominant, the inclusion of initial income requires instrumental variable techniques to be utilized. In such cases a negative transitory shock at the beginning of the time period being considered would reduce the initial level of income and produce a faster subsequent growth rate as the economy recovers. Moreover Durlauf and Quah (1999), suggests that the convergence effect could just be an example of Galton's fallacy which would just make this effect spurious[60]. For these reasons Scully (2002) states that the inclusion of the initial income would lead to severely biased results. However following Canadas (2008), since districts of Pakistan are not close to reaching their steady state levels, I include initial income as a proxy for the level of initial development in 1998.

Furthermore as the size of population is an important determinant of an economy's growth, population density is included as an explanatory variable. *Popdens*, is the number of inhabitants who lived in each district per squared kilometre in 1998[61]. Due to lack of data on infant mortality rates in Pakistani districts for the years 1998 and 2005, the variable *vaccine* has been utilized as a proxy for non-educational aspect of human capital. Vaccine is measured as the percentage of children 10 years or below having been vaccinated against all major diseases in each district. The variable *matric* has been included as a measure of human capital in each district. Adults having 10 years of completed education are considered as matriculates. It is the minimum level of education required for most entry level jobs in the tertiary sector in Pakistan. *Matric* is calculated as the percentage of adults with matric degrees for each district in 1998.

Another important variable that should be controlled for while studying steady state incomes is physical capital. Data for capital formation only exists for a few main districts and that too for some selective years. As a result *electricity*, the percentage of houses pre district which have electricity as their main source of energy consumption has been used as a proxy. The electricity consumption per district also highlights how much industry and manufacturing activity can it attract.

[60] Galton's Fallacy is used to denote problems encountered during the testing of the neoclassical convergence model (see Bliss, 1999 for details). "Generally speaking ... (it could be the case that) ...a negative association between growth rates and initial development conditions may be consistent with declining, stationary, and rising cross-section income dispersion" (Arbia, 2003:15).

[61] All of the explanatory variables used also appear in the Barro (1992) model and Trivedi (2003).

Table 7 (see appendix) provides the results of the long run OLS growth model. The coefficient for initial income still remains negative (although insignificant) implying lack of conditional convergence in this case. The rest of the additional variables are statistically insignificant as well. However, with the inclusion of explanatory variables the R-squared improves to 20 percent. While the regression diagnostic measures fail to reject the hypothesis of homoskedasticity, the null of hypothesis of spatial randomness is rejected since both the robust LM (error and lag) tests are significant, with the LM (error) test being more significant than the LM(lag) test. Figure 1 (in the Appendix) illustrates the relation between growth and initial income level as implied by the regression in Table 7 (in Appendix). It exhibits a lack of a clear convergence pattern and also seems to be influenced by a few outliers. As a robustness measure, we again produce bootstrapped standard error estimates for the OLS model along with the RWLS and IRWLS estimates for the conditional growth convergence model to take into account the influence of outliers (see Table 8 in Appendix). Under all these three specifications we fail to obtain any solid evidence of conditional convergence across districts between 1998 and 2005.

5.3 Spatial econometric analysis of convergence across districts

As previously discussed, two reasons motivate the inclusion of spatial effects explicitly into the OLS convergence model in this chapter:

1) Assuming spatial independence amongst states in a cross-sectional study of regional growth analysis is a highly restrictive assumption.
2) Ignoring spatial autocorrelation would yield inefficient (yet unbiased) OLS estimates for the convergence analysis.

It should be noted here that this chapter does not address the issue of heteroskedasticity encountered in Section 4.1 and assumes that spatial dependence is the only possible source [similar to the approach of Anselin and Griffith (1988)]. Since OLS regression diagnostics in Table 2 illustrate the presence of spatial interdependence, we incorporate it explicitly in our spatial econometric models which are estimated through maximum likelihood estimation.

5.3.1 Absolute spatial convergence model

The β-coefficient under spatial specifications is negative for both the SAR and SEM models. However since its level of significance in both the cases is very low (about10 percent), these results cannot be confidently treated as an indication of convergence across districts (Table 3 in Appendix). Moreover, caution must be exercised in their interpretation since the Jarque Bera test indicates that the residuals are non-normally distributed. Finally, the spatial diagnostics under both the models are insignificant, hence implying spatial randomness for the absolute convergence process across Pakistani districts.

5.3.2 Conditional spatial convergence model: A comparison of spatial lag vs. spatial error model for convergence analysis in Pakistan

Under conditional spatial specification, we obtain completely opposite and statistically significant results. This section estimates and compares conditional growth models with their spatial counterparts. In Table 4, columns 1 and 2 represent the spatial autoregressive (SAR) and spatial error models (SER) respectively. Both the models indicate significant convergence. The spatial diagnostics for the OLS model indicate towards the presence of spatial autocorrelation as shown by statistically significant Moran's *I* and Lagrange Multiplier tests. While Rho is not significant, lamda is significant at a 5 percent level indicating that the spatial error process may enhance our understanding of the convergence process across Pakistani districts between 1998 and 2005. Finally, the diagnostic tests on the spatial coefficients indicate a high significance level of the error coefficient. The LM lag and LM err tests demonstrate that the omitted spatial dependence is of the nuisance form, since LM err test has a higher level of significance. Moreover, the SEM model also has the lowest AIC value[62]. These results confirm that the spatial error model may be more suitable for analysing convergence / growth between 1998 and 2005[63].

5.3.3 Robustness of spatial results

The results of spatial models are sensitive to how spatial proximity has been defined in the spatial weight matrix. Hence to confirm whether the results above based on the binary contiguity matrix for 98 districts of Pakistan, are sensitive to modified definitions of weights,

---

[62] Since the R-squared measure is not applicable to spatial regression models that are estimated via maximum likelihood, the fit of the model can only be compared by the Akaike Information Criterion.
[63] Since the aim of this study is to detect convergence, I have not delved into the discussion on the effects of the explanatory variables on growth. However I must mention here that the negative coefficients of education and health variables are in line with the empirical estimates and political economy discussions in many similar studies for example, Canadas (2008) for Argentina, and Trivedi (2002) for India.

I also employ a non binary inverse distance matrix (as discussed in chapter 2) and re-estimate the three models contained in Tables 2 and 4. In all three cases, the tests provide the same conclusion i.e. spatial autocorrelation mainly takes the nuisance form.

Table 4. Conditional ß-convergence of per capita income in 98 Pakistani districts (1998-2005) Spatial Lag and Error Models- Maximum Likelihood Estimates		
	Model 1	Model 2
	Spatial Lag Model	Spatial Error Model
Log GDP98	-0.0212**	-0.0199***
	(0.0098)	(0.0073)
Matric	-0.0085***	-0.009***
	(0.0021)	(0.002)
Vacine	-0.00072*	-0.0009**
	(0.0004)	(0.0004)
Popden	0.00002	0.00002
	(0.0001)	(0.0001)
Electric	0.0004	0.0004*
	(0.0002)	(0.0002)
Constant	0.3579***	0.36587***
	(0.0899)	(0.0749)
Tau	--	0.3191**
Rho (W_growth)	0.1517	--
Goodness of Fit		
Akaike Information Criterion	-263.1373	-266.0148
Variance ratio	0.207	0.253
Squared Corr	0.216	0.198
Log likelihood	139.5687	141.0074
Observations	98	98
Regression Diagnostics		
Wald test (rho/tau=0)	1.420	4.320**
Likelihood ratio test (rho/tau=0)	1.413	4.290**
Lagrange Multiplier test (rho/tau)	1.401	3.260*
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations		

In summary, the diagnostic tests for the OLS models in Tables 2 and 7 (see Appendix) point towards a spatial dependence of the nuisance form. This is confirmed by the diagnostic results of the spatial error model. These results imply that there exists spatial dependence which is not being explained. Hence spatial effects may not be of a substantive form but a by-product of convergence. These results are in line with the findings of Rey and Montouri (1999) that demonstrate how the spatial error model is more appropriate specification for explaining the growth process across of the US states.

Although the SEM model seems to be the best among the alternatives considered to examine convergence, it is not enough to eliminate completely the spatial dependence. Further possible improvements could include a spatial panel data analysis coupled with a stochastic simulation analysis to observe how a shock to a district operates via the error term throughout the geographic system. Finally note that there is a possibility for spatial heterogeneity in the sample under consideration, however I have not taken it into consideration since this chapter focuses on the effects of spatial autocorrelation only.

6. Non-parametric Analysis of Growth and Convergence

Intra-distribution dynamics and spatial effects

As noted in Trivedi (2002), conditional convergence does not necessarily imply that districts are actually coming closer in terms of income levels. ß-convergence approach for convergence analysis has been criticized for being unable to distinguish between convergence, divergence, and stationarity. This is because “generally speaking a negative association between growth rates and initial development conditions may be consistent with a declining, stationary, and rising cross-section income dispersion” (Arbia, 2003:15). This failure—as previously mentioned—is referred to as Galton’s Fallacy (Quah, 1993).

Due to this limitation of the ß-convergence approach along with the caveats of the σ-convergence, there stemmed a new approach of analysing convergence and growth which focuses on estimating the entire income dynamics rather than “just fitting the first two moments and thus revealing the evolution of income distribution” (Arbia 2003:15). As noted by Quah (1993, 1996a-b, 1997), convergence analysis should be carried out by evaluating the shape and the dynamics of the distribution under study. Under such circumstances, non parametric techniques for the estimation of univariate density function can be used.

The standard deviation of the cross-district income distribution increased by 14.8 percent between 1998 and 2005, indicating an increase in the overall cross-district income dispersion. To examine closely the districts that have actually driven the increased dispersion and whether they are situated at the core or periphery of the cross district income distribution, Tukey Box Plots for a sample of *normalized* GDP of 98 districts have been utilized[64]. The inter-quartile range (middle 50% of the income distribution) slightly increased between the two years as the size of the lower quartile increased by 2005 (see Figure 2 in Appendix). This indicates that the districts located at the extremes of the distribution (in this case at the very bottom of the distribution) have the highest influence on the distribution (examples include districts of Khushab (80), Okara (54), Dadu (90)). These states have driven the increase between district income inequality between 1998 and 2005.

As a next step the chapter analyzes the distribution itself rather than just examining the moments of the cross-district income distribution. Figure 1 illustrates kernel-smoothed densities of log-relative per capita incomes between 98 districts of Pakistan. The kernel densities were estimated for the years 1998 and 2005 to verify the behaviour of per capita income. The density of the natural logarithm of ratio between GDP values for each district and the national average has been produced. The kernel-smoothed estimates have been obtained by using a Gaussian Kernel with normal optimal bandwidths that minimize the mean integrated squared error[65].

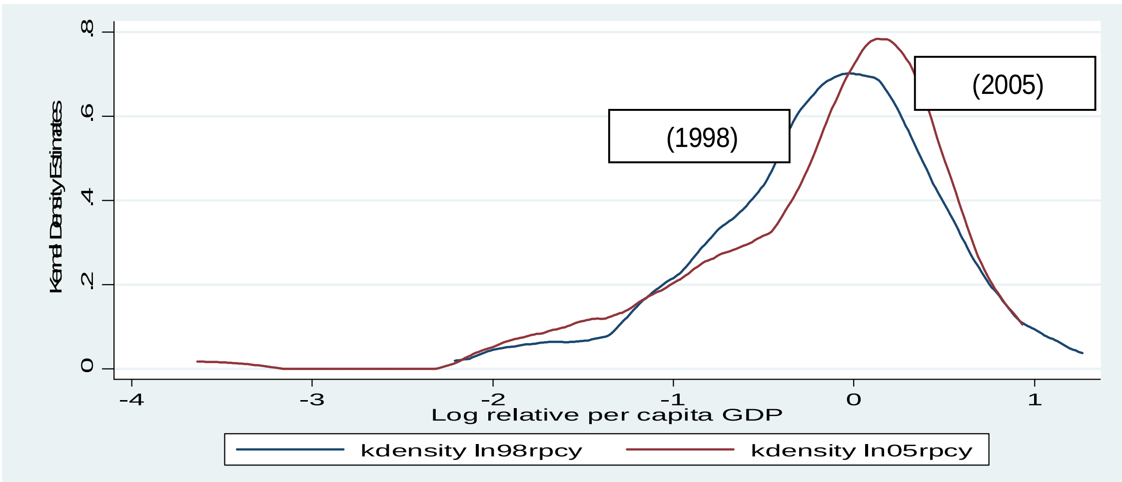
Another way is to examine the mobility dynamics by quantifying the intra-distribution dynamics. This is carried out using bivariate kernel which estimates the joint density of the income distribution at two points in time (Figure 3 in Appendix). From any point on the axis marked 1998 extending parallel to the axis marked 2005, the stochastic kernel is a probability density function. It describes how the cross sectional distribution of income at one point (1998) in time evolves into that in a future time period (2005). In figure 1 and figure 3 (see Appendix), the shifting of the kernel density estimate of income distribution has shifted towards the right for 2005, implying that on average district incomes have increased over time. Moreover 2005 income distribution became more peaked which may indicate a decreased dispersion within cross-state income distribution. The mode (peak point) of the distribution became larger in 2005 which can again be taken as an evidence of increased dispersion or polarization within cross-district income distribution. A longer and a thinner left

[64] where 'normalized' refers to the log of the ratio between s district's GDP and the national average.

[65] The kernel estimator is a non parametric technique and a smoothed version of the histogram used to estimate the probability density function f of a random variable X (e.g income). The estimator can be written as: f(x)= $\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$, where n is the number of finite observations in x; h is the smoothing parameter called the bandwidth; and K is the kernel function of the variable x which can adopt various functions.

tail of the distribution can be taken as an indication that districts at the bottom end of the income distribution are pulling away from the rest.

Figure 1. Kernel density estimates of log relative cross-district per capita income distribution



7. Conclusions

This chapter is the first detailed study on Pakistan which has taken into account the role of spatial effects in order to empirically assess the convergence process across Pakistani districts between 1998 and 2005. At the theoretical level it utilized the convergence hypothesis based on the neoclassical growth model. For empirical analysis, the chapter applied Least Squares, Re-weighted Least squares (RWLS), and spatial econometric techniques (in particular the spatial lag and spatial error models) to investigate the catch-up process between Pakistani districts.

At least three important conclusions emerge from the empirical analysis presented in Section 4. First, I find evidence for lack of σ-convergence over time as illustrated by increasing standard deviation of real per capita district GDPs, and other measures of dispersion (e.g the shape of the cross-district income distribution). This finding is corroborated by evidence from Trivedi (2002) for Indian states and Young, Higgins, and Levy (2007) for US States.

Second, there is no evidence of absolute β-convergence (between 1998 and 2005) implying that there is no tendency for initially poorer districts to growth faster. As a robustness measure various estimators were used to confirm this finding which is corroborated by findings from similar studies that have analysed absolute β-convergence in emerging market economies post 1970s (Trivedi, 2002). Some recent examples include evidence from Mexcio (Juan-Ramon and Rivera-Batiz, 1996), China (Jian, Sachs and Warner, 1996), India (Bajpai and Sachs, 1996; Trivedi, 2002), and Russia (Buccellato, 2007).

Third, adding spatial effects increases the precision of our estimates. Without their inclusion the detection of spatial autocorrelation in the unconditional and conditional convergence models would result in inefficient OLS estimators ad unreliable statistical inferences (Magrini, 2004). Most recent studies that have investigated intra-regional convergence using spatial econometric techniques have also demonstrated this (for example see Baumont, Ertur, and Gallo, 2001; Basile, Nardis, and Mantuan, 2003) for convergence estimates across European regions, Paraguas and Dey, 2006 for evidence from Indian states, and Elias and Rey, 2011 for evidence from Peruvian regions.

Moreover while standard econometric estimation of conditional β-convergence points towards lack of convergence, spatial estimates report otherwise. The spatial estimates show the inconsistency of the OLS technique since they point towards possible convergence. The spatial dimension emerges as non negligible through the error process, highlighting the importance of incorporating the role of geographic data while analyzing convergence and growth. The robustness tests for OLS (absolute and conditional) and spatial models also portray the spatial error model as a more appropriate model for investigating convergence. Even with a different criterion for contiguity, the results remained unchanged. The spatial error model has also appeared as an appropriate model specification in various other studies that have analysed convergence using spatial econometrics (examples include Rey and Montouri, 1999 for US states; Basile, Nardis, and Mantuan, 2003 for European regions). The most important economic implication of this result is that the presence of significance spatial error dependence implies that a random shock to a specific district would diffuse throughout the system' (Rey and Montouri, 1999; Baumont, Ertur, and Gallo, 2001), however the strength of the shock would depend on the distance of districts from the district in which the shock is introduced. Moreover, in terms of obtaining policy implications, this model could highlight a spatial spillover effect if it is re-estimated using a Spatial Durbin Model (SDM) specification (Baumont, Ertur, and Gallo, 2001; Ertur and Koch, 2006). This is because SDM is the reduced form of a model with cross-sectional dependence in the errors, yet at the same time it can also be perceived as the nesting model in a more ‘general approach of model selection’ (Mur and Angulo, 2005). The inclusion of lagged spatial variables (WX) may model/capture the spatial correlation and therefore enhance explanatory power. It could even be possible that a spatial correlation in the error term no longer exists which would enhance our existing interpretations. Since the estimation of a SDM was beyond the scope of this chapter, it will be an included in the forthcoming extension of this analysis. Moreover as previously noted further possible improvements will also include a stochastic simulation analysis to observe how a shock to a district operates via the error term throughout the geographic system since the spatial error model emerged to be model with the best fit in this chapter.

APPENDIX

Figure 1. Growth rate versus initial income

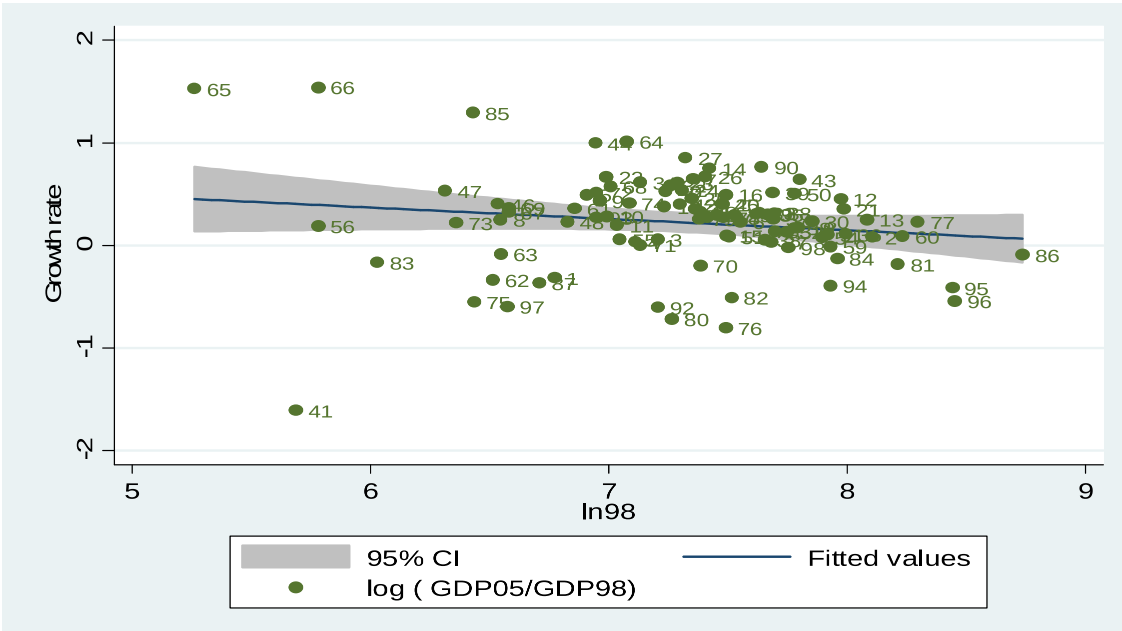


Table 1. Absolute ß-convergence of per capita agriculture income in 98 Pakistani districts (1998-2005) OLS- Estimates	
Log GDP Agri 98	0.01574***
	(0.0059)
Constant	-0.0880*
	(0.0457)
Goodness of Fit	
R-squared	0.068
Akaike Information Criterion (AIC)	-258.7537
Observations	98
Regression Diagnostics	
Breusch -Pagan test	2.20
Jarque-Bera normality test of residuals	61.95***
Moran's I (error)	0.111
Robust Lagrange Multiplier (lag)	1.990*
Robust Lagrange Multiplier (error)	1.843*
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations	

Table 2. Absolute ß -convergence of per capita manufacturing income in 98 Pakistani districts (1998-2005) OLS- Estimates	
Log GDP Manuf 98	0.0041**
	(0.0019)
Constant	0.0122
	(0.0111)
Goodness of Fit	
R-squared	0.049
Akaike Information Criterion (AIC)	-256.7383
Observations	98
Regression Diagnostics	
Breusch -Pagan test	0.80
Jarque-Bera normality test of residuals	34.99*
Moran's I (error)	0.108
Robust Lagrange Multiplier (lag)	0.993*
Robust Lagrange Multiplier (error)	0.961*
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations	

Table 3. Absolute ß -convergence of per capita income in 98 Pakistani districts (1998-2005) Spatial Lag and Error Models- Maximum Likelihood Estimates		
	Model 1	Model 2
	Spatial Lag Model	Spatial Error Model
Log GDP98	-0.0171*	-0.0160*
	(0.0106)	(0.0097)
Constant	0.1599**	0.1488**
	(0.0789)	(0.0708)
Tau	--	0.0912
Rho (W_growth)	0.0871	--
Goodness of Fit		
Akaike Information Criterion	-250.4902	-250.5331
Variance ratio	0.025	0.023
Squared Corr	0.029	0.023
Log likelihood	129.2451	129.2665
Observations	98	98
Regression Diagnostics		
Wald test (rho/tau=0)	0.422	0.465
Likelihood ratio test (rho/tau=0)	0.423	0.466
Lagrange Multiplier test (rho/tau)	0.424	0.471
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations		

Table 4. Absolute β-convergence of per capita income in 98 Pakistani districts (1998-2005) Dependent variable: Growth rate of real district per capita income			
	(1)	(2)	(3)
	OLS (Robust SE)	Robust Reg RWLS	IRWLS (mmreg)
Log GDP98	-0.0159	-0.0215	-0.0213
	(0.0189)	(0.0212)	(0 .0195)
Constant	0.1479	0.1959	0.1988
	(0.1423)	(0.1608)	(0.1473)
R-squared	0.0227	0.0611	-
RMSE	0.0655	0.0531	-
P-subsets			20
Observations	98	98	98
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations			

Table 5. Absolute β-convergence of agriculture income (value added) in 98 Pakistani districts (1998-2005) Dependent variable: Growth rate of real district per capita income			
	(1)	(2)	(3)
	OLS (Robust SE)	Robust Reg RWLS	IRWLS (mmreg)
Log Agri 98	0.0157**	0.0239***	0.0211***
	(0.008)	(0.005)	(0.008)
Constant	-0.0880	-0.1533***	-0.1307**
	(0.065)	(0.046)	(0.067)
R-squared	0.0684	0.2426	-
RMSE	0.0640	0.0464	-
P-subsets	-	-	20
Observations	98	98	98
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations			

Table 6. Absolute β-convergence of manufacturing income (value added) in 98 Pakistani districts (1998-2005) Dependent variable: Growth rate of real district per capita income			
	(1)	(2)	(3)
	OLS (Robust SE)	Robust Reg RWLS	IRWLS (mmreg)
Log Manuf 98	0.0041**	0.0041**	0.0018
	(0.004)	(0.0018)	(0.0022)
Constant	0.0122	0.0116*	0.02922*
	(0.0139)	(0.0124)	(0.0173)
R-squared	0.0490	0.0679	-
RMSE	0.0646	0.0540	-
P-subsets	-	-	20
Observations	98	98	98
Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1% Source: Author's calculations			

Table 7. Conditional ß-convergence of per capita income in 98 Pakistani districts (1998-2005) Dependent variable: Growth rate of real district per capita income	
Log GDP98	-0.0188
	(0.0998)
Matric	-0.0083***
	(0.0022)
Vacine	-0.0006***
	(0.0005)
Popden	0.0004***
	(0.0002)
Electric	0.00002***
	(0.0001)
Constant	0.3297*
	(0.0904)
Goodness of Fit	
R-squared	0.2003
RMSE	.06055
Akaike Information Criterion (AIC)	-265.7243
Observations	98
Regression Diagnostics	
Breusch -Pagan Test	7.65
Jarque-Bera Test	46.25*
Moran’s I (error)	-1.601
Robust Lagrange Multiplier (lag)	2.921*
Robust Lagrange Multiplier (error)	4.781**

Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author’s calculations

Table 8. Conditional ß -convergence of per capita income in 98 Pakistani districts (1998-2005)			
	(1)	(2)	(3)
	OLS (Robust SE)	Robust Reg RWLS	IRWLS (mmreg)
		Non boot	
Log GDP98	-0.0188	-0.0136	-0.0164
	(0.0155)	(0.0084)	(0.0172)
Matric	-0.0083***	-0.008***	-0.0029***
	(0.002)	(0.002)	(0.003)
Vacine	-0.0006***	-0.0005***	0.0001***
	(0.0004)	(0.0004)	(0.0009)
Popden	0.00002***	0.00001***	5.93e-06
	(0.00001)	(0.00009)	(0.00001)
Electric	0.0004***	0.0004***	-0.0002***
	(0.0003)	(0.0002)	(0.0006)
Constant	0.3297	0.2783*	0.2178*
	(0.1215)	(0.0757)	(0.0917)
R-squared	0.2003	--	--
RMSE	0.0606	--	--
P-subsets	na	na	20
Observations	98	98	98

Note: Standard errors in parenthesis, * significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author’s calculations

Figure 2. Tukey box plot: log of district income relative to sample average[66]

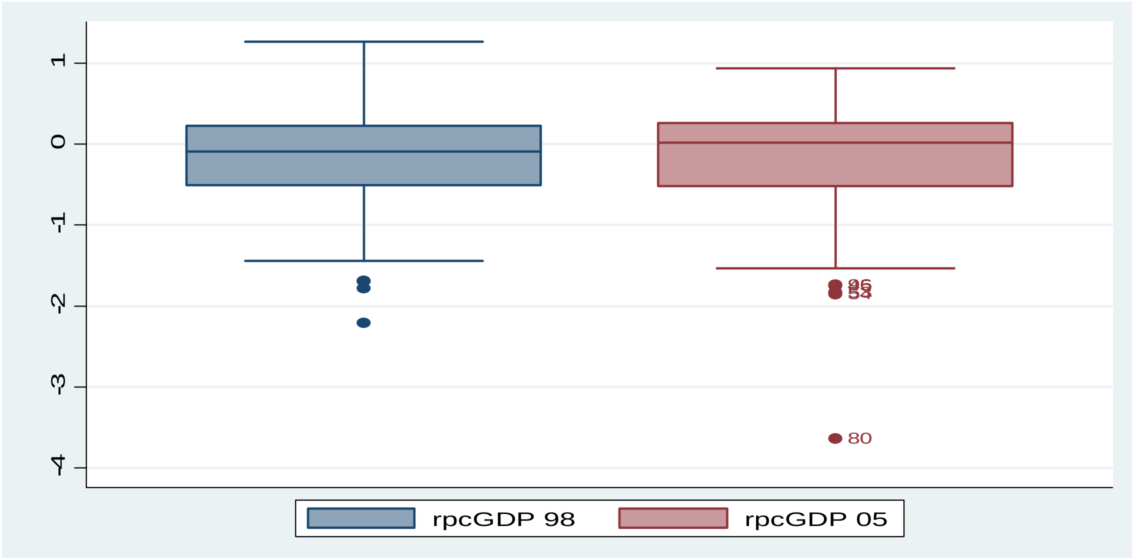
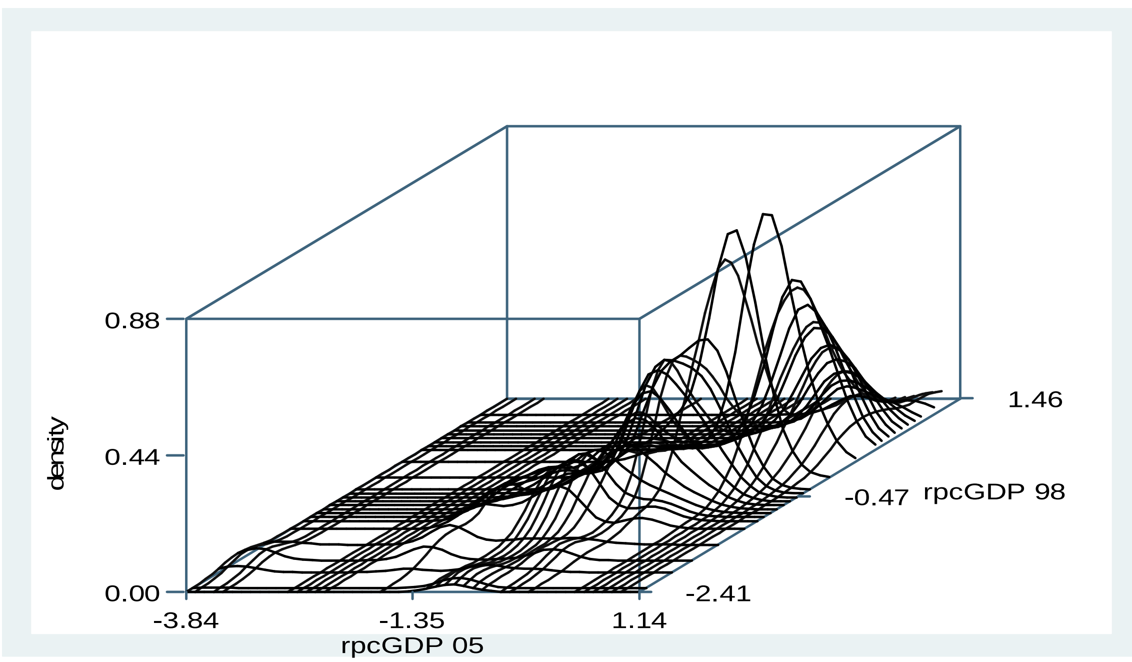


Figure 3: Stochastic Kernel Plot for relative per capita GDP in 2005 and 1998



66 Rpc stands for relative per capita GDP

Chapter 5

CONCLUSIONS

Socio-economic inequalities in Pakistan have been consistently high since the 1960s. The country is characterised by widespread internal socio-economic disparities as its core areas have experienced rapid development while the peripheral areas continue to lag behind. These spatial inequalities have become an urgent political issue reflected through public dissatisfaction and frequent anti-government demonstrations in lagging areas, and requires focused efforts if spatially balanced economic growth is to be achieved.

Furthermore, Pakistan is also experiencing a demographic transition according to which it has been estimated that the number of its potential workforce (15-64 years age group) will increase to approximately 221 million by 2050. However, currently out of a population of over 170 million, 25 million children do not go to schools. Moreover due to a weak education system, the educated class is embracing the phenomenon famously referred to as the *brain drain*, while the uneducated tier of the society is confronting frictional unemployment due to changes labour demand patterns (Ahmed, 2005). If Pakistan is unable to tackle these issues on a priority basis by increasing investment in its human capital and providing employment to its emerging workforce, it will lose its opportunity to take advantage of its demographic dividend (Nayab, 2007; Cohen, 2008; 2011).

Finally, the recently enacted 18th Constitutional Amendment in April 2011 and the 7th National Finance Commission Award have allowed the transfer of more funds from the federation to the provinces which will now have more authority over the provision of health, educational and physical infrastructure facilities[67]. This fundamental shift towards the division of power between the centre and the provinces has significant implications for the country’s long term policy planning, management and implementation. Specifically, it has made district level research even more important as education and other public services have

[67] The 18th Amendment to the Constitution of Pakistan (8th April 2010) has removed the power of the President of Pakistan to unilaterally dissolve the Parliament, removed the limit on a Prime Minister serving more than two terms, renamed the North West Frontier Province to Khyber Pakhtoonkhwa in accordance with the wishes of the Pashtun majority of the province, and increased provincial autonomy.

become the sole domain of the provinces. There is a need for increased research and policy prescriptions on how will the provinces build their capacities and generate resources for efficient public service delivery in social sectors across their districts.

This dissertation has addressed the above mentioned issues in three studies based on the fields of income and educational inequalities, and regional growth in Pakistan. By utilizing recent micro and macro data sets for Pakistan, it re-examined and updated estimates of the returns to education and income inequality estimates for Pakistani provinces, shed light on the effect of spatial clustering on the income and education inequality in Pakistani districts, and finally investigated for the first time the extent to which Pakistani districts are catching up in terms of their real per capita GDP (convergence). The next section describes and discusses the key findings from the three studies in this dissertation. It is followed by Section 3 which highlights the policy implications. Section 4 discusses the key empirical contributions of this dissertation followed by a short agenda for future research in this area in Section 5. Finally, Section 6 discusses how the output of this research assignment will be disseminated.

2. Summary and Conclusions

2.1 Can education explain changes in earnings inequality? Decomposition of earnings inequality in Pakistan (1993-2006)

Between 1993 and 2006 earnings inequality in Pakistan rose by approximately 34 percent as the proportion of wage earners in its working population increased from 40 to 62 percent[68]. This steady rise in earnings income inequality now dominates the overall inequality pattern amongst its total employed population. Since nearly a quarter (22 percent) of the total increase in earnings inequality between 1993 and 2006 was due to differences in education levels amongst workers, this chapter investigated the extent to which can this observed change in earnings inequality be attributed to changes in returns to human capital characteristics across Pakistan? It also demonstrated how the dynamics of earnings inequality have changed between 1993 and 2006.

[68] When this chapter was written, the latest available household data set (PSLM) was used to carry out the analysis for the fiscal year 2005-06. The next PSLM survey was carried out in 2007-08 but was not made available officially until 2010.

In order to do so this chapter employed various statistical and econometric inequality decomposition techniques. First, it analyzed the dynamics of earnings inequality between 1993 and 2006 by calculating, comparing and decomposing some General Entropy Indices and the Gini coefficient of earnings inequality within and across provinces, and over time. This descriptive analysis was extended and complemented with the regression based approach to inequality decomposition for which this study utilized a relatively recent empirical framework similar to that used in Bourguignon et al (2005) to identify the how changes in the structure of earnings and socio-demographic characteristics affect the overall earnings inequality. It simulated the effect of the observed changes in: returns to education and the demographic structure on the distribution of earnings. This decomposition allowed us to investigate the effect of unequal human capital characteristics of individuals on their earnings over time in Pakistan.

Empirical results from this chapter have demonstrated that earnings inequality within rural and urban regions has increased much more than earnings inequality between rural and urban regions. Moreover, it remained much higher amongst females as compared to earnings inequality amongst males. However for the case of educational returns in Pakistan, inequality between individuals with different education levels has remained much higher than inequality amongst individuals with the same amount of education during 1998 and 2005. With regards to the private returns to education in Pakistan, except for primary education, the returns to each successive level of education (i.e. secondary, high school, graduation, post graduation) have increased between 1993 and 2006 with females having higher returns as compared to males throughout. Although this would imply that parents should favour education amongst girls as compared to boys, it is not actually so in reality. Furthermore, the falling returns to primary education in Pakistan implies that if earnings-education profile continues to disadvantage workers with lower levels of education (i.e. remain convex), increased distribution of basic education might not even contribute much towards poverty reduction. Finally, the counterfactual analysis highlights the importance of education, in influencing the labour market returns and the distribution of earnings income in Pakistan. Most importantly, the micro simulations have demonstrated how females in general and rural females in particular (due to the high returns to their skills) are most sensitive to changes in policies that can affect inequality and labour market outcomes.

2.2 *Does economic geography matter for Pakistan? A spatial exploratory analysis*

Generally, econometric studies on income and education distribution consider regions and provinces as independent entities, ignoring the likely possibility of spatial interaction particularly within a country. This interaction may cause spatial dependency or clustering, which is referred to as spatial autocorrelation in a spatial econometric framework. Chapter 3 discussed why there is a need to re-examine previous findings on income and education inequalities in the light of recent developments in the field of spatial econometrics for Pakistan. It also highlighted why there is a need for increased amount of research on geographical levels smaller than the commonly researched provincial level for example districts.

Specifically, it has exploited the geographical characteristics of the latest available micro data sets to analyze the spatial clustering of income and education distribution in the districts of Pakistan by employing spatial exploratory data analysis (ESDA) techniques. Global and local measures of spatial autocorrelation were computed using the Moran's *I* and Geary's *C* index to obtain estimates of the spatial autocorrelation of spatial disparities across 98 districts between 1998 and 2005. The overall finding is that the distribution of district wise income inequality, income, education attainment, growth, and development levels, exhibits a significant tendency for inequality and similar levels of education to cluster in Pakistan (i.e. the presence of spatial autocorrelation is confirmed).

Particular findings include the fact that there exists clustering of high and low income districts in 1998 and in 2005, with Punjab having the largest cluster of high income districts in both the years. District education levels also revealed high spatial associations especially when neighbourhood was defined in terms of sharing borders (contiguity). This implies that districts that share a border may actually influence each other's education attainment levels. When individual education levels were analysed, the chapter concludes that there exists positive spatial autocorrelation dependence for education levels below bachelors (i.e. primary, matric i.e. grade 10, and inter i.e. grade 12). As for higher levels of education (Bachelors and above), sharing a border with a district with a high percentage of graduates, is more influential than the distance from the provincial capital or the nearest large city. Finally, an analysis of spatial association of a set of district wise Human Development Indicators confirmed that a district's development levels are associated on the development levels prevailing in its neighbouring districts in Pakistan. In summary, the detection of significant spatial autocorrelation in income inequality, education and human development levels across districts implies that districts (and regions in general) should not be viewed as independent observations in econometric analysis of socio-economic phenomenon.

2.3 *Are districts across Pakistan catching up? A spatial econometric analysis of regional convergence*

Do initially poorer countries grow faster in terms of per capita income and catch up to those that were initially richer? This issue has been referred to as regional' convergence' in economic theory and has remained an important field of research for economists. While Pakistan has been a part of various cross-country convergence studies, a limited amount of research has been carried out to study this phenomenon within the country across its districts. Building on the results obtained in Chapter 3, this chapter has contributed towards the literature on the spatial analysis of regional convergence, by investigating the extent to which Pakistani regions are catching up in terms of their income. It has explored convergence of GDP across 98 Pakistani districts between 1998 and 2005 using both spatial and non spatial econometric techniques. It is the first study of its kind to carry out a convergence analysis based on a district level and that too from a spatial econometric perspective for Pakistan.

The chapter first investigated absolute convergence across Pakistani districts and then proceeded to a conditional convergence analysis to take into account the role of some district specific socio-economic/ demographic characteristics and geographical interactions in regional growth. In order to do so it utilized Ordinary Least Squares, Reweighted Least Squares, and Iterative Re-Weighted Least Squares techniques. The empirical results from these analyses indicated towards the lack of both, absolute and conditional convergence. However, results of district income spatial autocorrelation from chapter 3 and the diagnostic tests carried out in this chapter to detect for model misspecifications revealed evidence of positive spatial autocorrelation implying that the values for growth of real per capita GDP across Pakistani districts are spatially interdependent. This finding implies that assuming spatial independence in regional econometric studies for Pakistan would be a highly restrictive assumption.

Hence, after explicitly incorporating spatial interdependence in the model by utilizing spatial econometric models, the conditional spatial model revealed the possibility of convergence across Pakistani districts between 1998 and 2005. The diagnostics of the models suggested that spatial dependence between the districts emerges through the error process, highlighting that that there exists unexplained spatial dependence which could just be a by-product of the convergence process instead of actually being of a substantive form (i.e. the actual existence of spatial interaction effects between districts). Even with a different criterion for contiguity, the results remained unchanged.

Non parametric techniques were also utilized to examine the shape and the dynamics of the distribution of GDP per capita in the two years under study. Results revealed that the standard deviation of the cross-district income distribution increased by 14.8 percent between 1998 and 2005, indicating an overall increase in the cross-district income dispersion. In conclusion, empirical findings from Chapter 4 have demonstrated how adding spatial effects to econometric analysis can enhance our understanding of the regional convergence process across Pakistan.

3. Policy Recommendations

First, policies seeking to reduce overall income inequality in Pakistan should focus more on achieving greater equality in the distribution of the characteristics of paid employees. If earnings-education profile continues to disadvantage workers with lower levels of education (i.e. be convex), increased distribution of basic education might not even contribute much towards poverty reduction as presumed by the Millennium Development Goals. The policy implication for this finding lies in targeted government interventions that can facilitate in distributing the gains of education across different groups in the Pakistani labour market. In other words the government should facilitate the education of more people instead of making arrangements to pay less educated people more.

Second, about 68% of Pakistan's current population is below 30 years of age and as most of them are entering the labour force, it has been estimated that the total labour force is increasing by 3% annually (Planning Commission of Pakistan, 2011). This increasing proportion of the working age group implies a decreased dependency ratio which can have an extremely beneficial impact on economic growth if Pakistan can capitalize its demographic dividend. However when compared to other emerging market economies, Pakistan has a relatively large proportion of uneducated youth (32%) with very little vocational training. This group either ends up in elementary occupations or remains unemployed. If unattended this youth becomes disconnected with the political problems and may become extremely vulnerable to rising extremism in the country. If Pakistan is to take advantage of its increasing youth by making them beneficial participants of the economy, it must provide for their education, health, and employment. Focused efforts must be carried out such as the formulation of a National Youth Service Policy as recently suggested by the Planning Commission of Pakistan's Growth Strategy, 2011.

Third, the education emergency in Pakistan requires immediate attention and remedial measures. If Pakistan has to capitalize its demographic dividend, focused effort must be directed towards improving its human capital. Policies that promote education attainment, creativity, innovation and job creation need to be implemented. Educational reforms will have to cater to both public and private educational institutions in rural and urban areas of Pakistan. Moreover while the results (in Chapter2) demonstrate the returns to education have been higher for females throughout, it remains true that males continue to earn more in both, rural and urban areas. In this background, education can play a crucial role in obtaining gender equality in the Pakistani labour market as it has been noted that gender gap in earnings in Pakistan is much smaller for those who have obtained higher education (see Aslam, 2009).

Fourth, the empirical results and graphical evidence provided in Chapters 3 and 4 provide new insights for urban planning and regional development policies. Their results have provided sufficient evidence of the fact that districts with similar levels of education and other development characteristics, do tend to cluster in Pakistan. From a policy perspective, these results have demonstrated how 'neighbourhoods matter' for improving local human development conditions since a district's development levels depend on the development levels prevailing in its neighbouring districts in Pakistan. If the existing clusters of commerce and development are allowed to flourish and the creation of new ones in other provinces is facilitated, they can become suppliers of human capital and centres of trade, recreation, knowledge and domestic commerce, and a means of reducing rural urban migration.

` Finally, chapters 3 and 4 have highlighted the importance of additional data collection and research (that takes spatial effects into account) at lower levels of geographical aggregation in Pakistan because Pakistani provinces are characterised by extreme within diversity in terms of their socio-demographic characteristics such as language and culture, and economic conditions.

4. Contributions

This dissertation has contributed towards the literatures and policy debate on economic inequalities, convergence and growth in emerging market economies and in particular, in Pakistan. Its overall goal has been to reconsider exiting research on socio-economic inequalities and regional growth in Pakistan, by taking spatial effects into account to produce new insights on how these issues should be tackled while modelling them. The methodologies utilized in each of the three main studies of this dissertation have been implemented for the first time on Pakistani data. Chapter 2 has utilized regression based counterfactual analysis via Bourguignon et al (2005) to identify and analyze how changes in the structure of earnings and socio-demographic characteristics affect the overall earnings inequality in Pakistan. It has also re-examined the findings of previous studies on human capital inequalities in Pakistan and has presented the latest estimates for returns to education across Pakistan, taking into account gender and spatial differences.

Although preliminary in nature, the empirical results from Chapters 3 and 4 have provided the first detailed evidence on why spatial effects should be taken into consideration for analyzing socio-economic issues in Pakistan by utilizing spatial exploratory and spatial econometric techniques. Furthermore this analysis has been carried out at a district level which is often neglected in research. While most studies present results on a provincial level in Pakistan, the few that have considered districts, only analyse few districts with similar characteristics (e.g analyse districts belonging to the same province). Chapter 4, not only uses the spatial econometric methodology for the first time on Pakistani data, it is also the first study on intra-regional convergence in Pakistan in order to do so.

Given the recent shift towards increased provincial autonomy in Pakistan, and the launch of the ‘Education Emergency Programme’ in March 2011 to tackle the Pakistan’s education crisis, the contributions of this dissertation on educational and income inequalities make it a timely input in research for enhanced policy making in Pakistan and South Asia.

5. Future possibilities and Next steps

Testing issues such as socio-economic inequalities and regional income convergence involves a number of data issues. Organizing data on a district level aggregation and over time was the most demanding task encountered during the compilation of this dissertation. This was because, despite its need, there is a lack of data which is representative at the district level.

However as of 2007-08, the Federal Bureau of Statistics of Pakistan has regularly started to collect district representative data for the Pakistan Living Standards Measurement Survey. The official data was not made publically available until the end of 2008, which is why this dissertation had to utilize the micro data sets for 1998 and 2005, since they were the *only* two years for which district representative data existed when the chapters of this dissertation were being written. The plausible initiative of FBS to collect data at district level now, provides new opportunities to richness and detail to the existing analysis by providing various opportunities to exploit spatial exploratory and econometric techniques for Pakistani data. It also makes it possible to carry out a spatial panel data analysis for socio-economic inequalities and regional convergence and offers the possibility to test the effect various other explanatory variables besides the ones already used in the convergence analysis carried out in Chapter 4.

Moreover, I plan to extend the results for earnings income inequality in Chapter 2 by analyzing overall household income inequality across Pakistan. As for my chapters on spatial analysis of socio-economic issues, I first plan to estimate a Spatial Durbin Model for analysing regional convergence and carry out a stochastic simulation analysis to observe how a shock to a district operates via the error term throughout the geographic system since the spatial error model emerged to be model with the best fit in this chapter.

Finally, another future research possibility involves the investigation of the club convergence hypothesis (also known as spatial regimes analysis) using spatial techniques for the case of Pakistan. Club convergence occurs when regions with similar socio-economic characteristics converge only when their initial conditions are also similar (Cappelen, 2004). Increasingly literature in growth economics has demonstrated this tendency for economies to ‘evolve within groups and not in isolation’ called convergence clubs (Bandyopadhyay, 2003:3; Bernaud and Durlauf, 1994; Esteban and Ray, 1994). In this context I would like to identify and investigate whether high income districts in Pakistan have formed a *club* and to what extent are the low income districts lagging behind.

BIBLIOGRAPHY

A

Abbas, Q. (2000). "The Role of Human Capital in Economic Growth: A Comparative Study of India and Pakistan." *The Pakistan Development Review*, 39 (4): 451-473.

Abbas, Q. (2001). "Endogenous Growth and Human Capital: A Comparative Study of Pakistan and Sri Lanka." *The Pakistan Development Review*, 40(4): 987- 1007.

Acemoglu, D. (1998). “Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality.” *Quarterly Journal of Economics*, 113:1055-1090.

Acemoglu, D. (2001). “Good Jobs versus Bad Jobs.” *Journal of Labor Economics*, 19(1): 1-21.

Adams, R. H. J. (1992). Sources of Income Inequality in Rural Pakistan: A Decomposition Analysis, Agriculture and Rural Development Department, World Bank, WPS0836.

Adams, R. H. J. (1994). "Non-farm Income and Inequality in Rural Pakistan: A Decomposition Analysis." *Journal of Development Studies* 31(1): 110-133.

Adams, R. H. J. (1995). "Sources of Income Inequality and Poverty in Rural Pakistan." IFPRI Research Report (No. 102).

Ahmad, E & Stephen Ludlow (1989) Poverty, Inequality and Growth in Pakistan. The Pakistan Development Review 28:4, 831–850.

Ahmad, M. (2000). "Estimation of Distribution of Income in Pakistan, Using Micro Data." *Pakistan Development Review*, 39(4): 807-824.

Ahmed, V. (2005). “Welfare Impact of Trade and Taxation Reforms: A CGE-Microsimulation Analysis.” Unpublished PhD Dissertation, Department of Economics. Galway, National University of Ireland.

Aghion, P., E. Caroli & C. García-Peñalosa. (1999). “Inequality and Economic Growth: The Perspective of the New Growth Theory.” *Journal of Economic Literature, 37*:1615–1660.

Aghion, P. (2002). “Schumpeterian Growth Theory and the Dynamics of Income Inequality.” *Econometrica*, 70: 855-82.

Alesina, A & R. Perotti. (1994). “The Political Economy of Economic Growth: A Critical Survey of the Literature.” *World Bank Economic Review, 3*:351–371.

Alesina, A & D. Rodrik. (1994). Distributive Politics and Economic Growth. *The Quarterly Journal of Economics, 109*(2): 465-490.

Alesina, A & R. Perotti. (1996). “Income Distribution, Political Instability, and Investment.” *European Economic Review, 40*: 1203-1228.

Amos, O. M. Jr. (1988). “Unbalanced Regional Growth and Regional Income Inequality in the Latter Stages of Development,” *Regional Science and Urban Economics*, 18: 554-66.

Anselin, L., Varga, A, & Z, Acs. (1997). "Local Geographical Spillovers between University Research and High Technology Innovations." *Journal of Urban Economics*, 42: 422-448.

Anselin, L. (1988a). “Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeniety,” *Geographical Analysis,* 20:1-23.

Anselin, L. (1988b). *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer Academic Press.

Anselin, L. (1995a). SpaceStat. A Software Program for the Analysis of Spatial Data (version 1.80), Morgantown: Regional Research Institute, West Virginia University

Anselin, L. (1995b). “Local Indicators of Spatial Association—LISA”, *Geographical Analysis* 27: 93-115.

Anselin, L.(1996). “The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association”, in M.Fisher, H.J Scholten and D. Unwin (eds.), Spatial Analytical Perspectives on GIS, London: Taylor and Francis.

Anslein, L & A.K. Bera. (1998). “Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics” in Hullah A., and Gelis, D.E.A. (Eds.) *Handbook of Applied Economic Statistics*, Marcel Deker, New York, pp.237-290.

Anslein, L., & S.J Rey. (1991). “Properties of Test for Spatial Dependence,” *Geographical Analysis*, 23(2): 112-131.

Anselin, L. (2003a). "Spatial Externalities, Spatial Multipliers and Spatial Econometrics." *International Regional Science Review,* 26: 147-152.

Anwar, T. (2003). “Trends in Inequality in Pakistan between 1998-99 and 2001-02.” *The Pakistan Development Review,* 42(4): 809–821.

Arbia, G. (1988). *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, Kluwer Academic Publisher.

Arbia, G., R. Benedetti, & G. Espa. (1996). ‘‘Effects of the MAUP on Image Classification.’’ *Geographical Systems,* 3: 123–41.

Arbia, G. (2001). “The role of Spatial Effects in the Empirical Analysis of Regional Concentration,” *Journal of Geographical Systems*, 3: 271-281.

Arbia, G., R, Basile & M. Salvatore. (2003). “Measuring Spatial Effects in Parametric and Non-parametric Modelling of Regional Growth and Convergence.” Paper for UNU/ WIDER Project Meeting on Spatial Inequality in Development. Available at: http://www.rimisp.org/FCKeditor/UserFiles/File/documentos/docs/pdf/arbia_et_al_2003_spatial_effects_regional_growth.pdf (accessed on 22 November 2010).

Arbia, G., Dominicis, L & G, Piras. (2005). "The relationship between Regional Growth and Regional Inequality in EU and transition countries - a Spatial Econometric Approach," *ERSA conference papers* No. 05p168, European Regional Science Association.

Arbia, Guiseppe. (2006). Spatial Econometrics. Statistical Foundations and Applications to Regional Convergence. Berlin, Heidelberg, Springer-Verlag.

Arbia, G., G, Espa., D, Giuliani & A, Mazzitelli. (2009). "Clusters of firms in space and time," *Department of Economics Working Papers* No. 0902, Department of Economics, University of Trento, Italy.

Arby, M.F., M.J. Malik & M.N. Hanif. (2010). “The Size of Informal Economy in Pakistan.” *SBP Working Paper Series*, No. 33.

Ashraf, J & B. Ashraf. (1993). “Analysis of the Male-Female Earnings Differential in Pakistan.” *The Pakistan Development Review*, 32(4): 895–904.

Aslam, M. (2007). "Rates of Return to Education by Gender in Pakistan." Global Poverty Research Group Working Paper Series No. 64.

Aslam, M & G. G. Kingdon. (2008). “Gender and Household Education Expenditure in Pakistan.” *Applied Economics,* 40: 2573–91.

Autor, D., L, Katz & A.B, Krueger. (1999). “Computing Inequality: Have Computers Changed the Labour Market?” *Quarterly Journal of Economics*, 113:1169-1214.

B

Baldwin, R.E. & R. Forslid. (2000a). “The Core-periphery Model and Endogenous Growth: Stabilizing and Destabilizing Integration.” *Economica,* 67: 307–324.

Baldwin, R., R. Forslid., P. Martin., G. Ottaviano & F. Robert-Nicoud. (2003a). *Economic Geography and Public Policy*, Princeton University Press.

Baldwin, R. & P. Krugman. (2004). “Agglomeration, Integration and Tax Harmonization.” *European Economic Review*, 48(1): 1–23.

Balisacan, A. M & F, Nobuhiko. (2004). "Changes in Spatial Income Inequality in the Philippines: An Exploratory Analysis," *Working Papers UNU-WIDER Research Paper* , World Institute for Development Economic Research (UNU-WIDER).

Bandyopadhyay, S. (2003). "Convergence Club Empirics : Some Dynamics and Explanations of Inequal Growth Across Indian States", *Discussion Paper Distributional Analysis Research Programme* No. *DARP* 69, The Toyota Centre.

Basile, R., Nardis, S de & M. Mantuan. (2011). “Multiple regimes in cross region growth regression with spatial dependence: A Parametric and semi parametric approach.” ISAE, Institute of Studies and Economic Analysis, Rome, Available at: http://www.jyu.fi/ersa2003/cdrom/papers/4.pdf (accessed on January 13, 2011)

Basile, R. (2008). “Regional economic growth in Europe: A semiparametric spatial dependence approach”, *Papers in Regional Science*, 87(4): 527-544.

Banerjee, A., & A. Newman. (1993). “Occupational Choice and the Process of Development.” *Journal of Political Economy*, *101*(2): 274–298.

Banerjee, A & E. Duflo. (2003). “Inequality and Growth: What Can the Data Say?” *Journal of Economic Growth*, *8*(3): 267-299.

Barro, R & X. Sala-i-Martin. (1991a). “Economic Growth in a Cross Section of Countries.” *Quarterly Journal of Economics*, *106*: 407-443.

Barro, R & X. Sala-i-Martin. (1991b). “Convergence Across States and Regions.” *Brookings Papers on Economic Activity*, *1*: 107-182.

Barro, R & X. Sala-i-Martin. (1992). “Convergence.” *Journal of Political Economy*, *100*:223-251.

Barro, R.J & Sala-i-Martin, X., (1995). *Economic Growth*. McGraw-Hill, New York.

Barro, R. J. (2000). “Inequality and Growth in a Panel of Countries.” *Journal of Economic Growth*, *5*(1): 5-32.

Barro, R., & X. Sala-i-Martin. (2004). *Economic Growth*. Cambridge, MA: MIT Press.

Baumont, C., C. Ertur & J. Le Gallo. (2001). "A Spatial Econometric Analysis of Geographic Spillovers and Growth for European Regions, 1980-1995," LATEC - Document de travail Economie (1991-2003) 2001-04, LATEC, *Laboratoire d'Analyse et des Techniques EConomiques*, CNRS UMR 5118, Université de Bourgo.

Beck, T., A. Demirgüç-Kunt & R. Levine. (2004). “Finance, Inequality and Poverty: Cross Country Evidence.” *World Bank Policy Research*, Working Paper No. 3338.

Beck, N., K. Gleditsch, & K. Beardsley. (2006). "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly,* 50: 27-44.

Bell, L & R. Freeman. (2001). "The Incentive for Working Hard: Explaining Hours Worked Differences in the US and Germany." *Labour Economics, 8*: 181–202.

Benabou, R. (1996). "Inequality and Growth. National Bureau of Economic Research." Working Paper no.5658.

Benhabib, Jess & M. M. Spiegal. (1994). "The Role of Human capital in Economic Development: Evidence from Aggregate Cross-Country Data." *Journal of Monetary Economics* 34: 143-173.

Benhabib, J., & A. Rustichini. (1996). Social Conflict and Growth. *Journal of Economic Growth, 1*: 129-146.

Benneworth, P & G. J. Hospers. (2007). "The New Economic Geography of Old Industrial Regions: Universities as Global–local Pipelines." Environment and Planning C: *Government and Policy*, 25(6): 779–80.

Bernaud, A & S. Durlauf. (1994). "Interpreting Tests of the Convergence Hypothesis", *NBER Technical Working Papers*, No. 159.

Berman E., J. Bound, & S. Machin. (1998). "Implications of Skill-Biased Technological Change: International Evidence." *Quarterly Journal of Economics*, 113(4): 1245-79.

Berman, E., R. Somanathan, & H. Tan (2003). "Is Skill-Biased Technological Change here yet? Evidence from Indian Manufacturing in the 1990." *NBER* Working Paper No. 5110.

Bernat, G.A. Jr. (1996). "Does Manufacturing Matter? A spatial econometric view of Kaldor's Laws." *Journal of Regional Science*, 36(3): 463-477.

Bernard, A. & C. Jones. (1996a). "Productivity and Convergence Across US States and Inductries," *Empirical Economics*, 21: 113-35.

Bernard, A. & C. Jones. (1996b). "Productivity across Industries and Countries: Time Series Theory and Evidence." *Review of Economics and Statistics*, 78: 135-46.

Berry ,Christopher. & E. L. Glaeser. (2005). "The Divergence of Human Capital Levels Across Cities." *Papers in Regional Science,* 84(3): 407-444.

Bertola, G. (1993). "Factor Shares and Savings in Endogenous Growth". *American Economic Review,* 83: 1184-1199.

Beyer, H., P. Roja. & R. Vergara. (1999) "Trade Liberalisation and Wage Inequality." *Journal of Development Economic*s, 59(1): 103-123.

Bils, M & P. Klenow. (2000). “Does Schooling Cause Growth?”*American Economic Review,* 90: 1160—1183.

Birdsall, N & J. L. Londono. (1997). "Asset Inequality Matters: An Assessment of the World Bank's Approach to Poverty Reduction." *American Economic Review,* 87(2): 32-37.

Blinder, A. S. (1973). “Wage Discrimination: Reduced Form and Structural Variables.” *Journal of Human Resources,* 8: 436–455.

Bliss, C. (1999). “Galtons Fallcy and Economic Convergence.” *Oxford Economic Papers* No. 51.

Bloom, D, D. Canning, & J. Sevilla. (2002). "The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change." Santa Monica: *RAND Corporation Report* No. MR-1274.

Bound, J & G. Johnson. (1992). “Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations.” *American Economic Review,* 83: 371-392.

Bourguignon, F., & C. Morrisson (2002). “Inequality among World Citizens: 1820-1992.” *The American Economic Review, 92*(4): 727-744.

Bourguignon, F., F. Ferreira, & M. Menéndez. (2003). *Inequality of Outcomes, Inequality of Opportunities, and Intergenerational Educational Mobility in Brazil.* Washington, DC: The World Bank.

Bourguignon, F. (2003). “The Growth Elasticity of Poverty Reduction; Explaining Heterogeneity across Countries and Time Periods.” In T. Eichler and S. Turnovsky (ed.), *Inequality and growth: Theory and Policy Implications*. Cambridge: The MIT Press.

Bourguignon, F. (2004). “The Poverty-Growth-Inequality Triangle.” *The World Bank Policy Research,* Working Paper No. 28102.

Bourgiugnon, F and Ferreira, F.H.G (2005), “Decomposing Changes in the Distribution of Houshold Incomes: Methodological Aspects.” En Bougiognon, F, and Ferriera, F.H.G, The Microeconomics of Income Distribution Dynamics, Oxford University Press.

Bowles, S & H. Gintis. (2002). “Social Capital and Community Governance.” *Economic Journal, 112*: 419–436.

Buccellato, T. (2007). “Convergence across Russian Regions: A Spatial Econometrics Approach.”, Centre for Financial and Management Studies, Discussion Paper 70.

Burki, A., A. K. Munir., M. Khan., U. Khan., A. Faheem., A. Khalid & S.T. Hussain. (2010). “Industrial Policy, Its Spatial Aspects and Cluster Development in Pakistan.” Analysis Report to the Industrial Policy 2010, Lahore University of Management Sciences.

C

Canadas, A. (2008a). “The Effect of Spatial Clustering on Economic Growth.” Mount St. Mary’s University. Mimeo.

Canadas, A. (2008b). “Does Spatial Clustering help Explaining Differences in the Inequality of Income Distribution.” Mount St. Mary’s University. Mimeo.

Cappelen, A. (1999). “Convergence, Divergence and the Kuznets Curve.” European Commission, TSER Programme, http://www.merit.unimaas.nl/tser/teis018.pdf (accessed on 8 December , 2010).

Caramuta, D. (2005). “A Dynamic Approach to the Relationship between Inequality, Social Capital and Institutions.” Paper presented at the CONICET, Universidad Nacional del Sur, Argentina.

Carlino, G & L. Mills. (1996a). “Testing Neoclassical Convergence in Regional Incomes and Earnings.” *Regional Science and Urban Economics*, 26: 565-90.

Carlino, G & L. Mills. (1996b). “Convergence and the US States: A Time Series Analysis.” *Journal of Regional Science*, 36: 597-616.

Castello, A & R. Domenech. (2002). “Human Capital Inequality and Economic Growth: Some New Evidence.” *Economic Journal*, 112:187–200.

Celebioglu, F & S. Dall'erba. (2010). "Spatial Disparities across the Regions of Turkey: An Exploratory Spatial Data Analysis," *The Annals of Regional Science*, 45(2): 379-400.

Charles, D. (2003). “Universities and Territorial Development: Reshaping the Regional Role of UK Universities.” *Local Economy*, 18: 7-20.

Clarke, G. (1995). “More Evidence on Income Distribution and Growth.” *Journal of Development Economics*, 47: 403–427.

Cliff, A.D., and J.K. Ord. (1973). *Spatial Autocorrelation*. London: Pion.

Cliff, A.D., and J.K. Ord. (1981). *Spatial Processes*. London: Pion.

Cohen, S. P. (2011) “Coping with a Failing Pakistan.” Norwegian Peace Building Centre, *NOREF Policy Brief* No 1. Available at:

http://www.peacebuilding.no/eng/Publications/Noref-Policy-Briefs/Coping-with-a-failing-Pakistan, (accessed on March 7, 2011).

Cohen, C. (2008). "Pakistan 2020: The Policy Imperatives of Pakistani Demographics." *NPEC Working Papers Series*. Available at: http://www.npec-web.org/Essays/20080928-Cohen-Pakistan2020.pdf

Colcough, C., G. Kingdon, & H. Patrinos. (2009). "The Pattern of Returns of Education and Its Implications." *RECOUP Policy Brief*, No 4.

Conley, T. (1999). "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1-45.

Costa, Alan A.B., B. P. Schettini & S. C. Diniz. (2009). "Changing the Focus: A Spatial Econometric Assessment of the Convergence Process in Minas Gerais (1991-2000)", http://www.bnb.gov.br/content/aplicacao/eventos/forumbnb2009/docs/changing.pdf (accessed on 27 Dec 2010).

Currie, J. & A. Harrison. (1997). "Sharing Costs: The Impact of Trade Reform on Capital and Labour in Morocco." *Journal of Labor Economics*, 15: 44-71.

D

Dollar, D & A. Kraay. (2002). "Growth is Good for the Poor." *Journal of Economic Growth, 7*(3): 195-225.

Dominicis, L., G, Arbia & H. L.F, de Groot. (2007). "The Spatial Distribution of Economic Activities in Italy." *Tinbergen Institute Discussion Papers* 07-094/3, Tinbergen Institute.

Darlauf, S.N & D.T., Quah. (1999). "The New Empirics of Economic Growth," in Taylor, J.B. and Woodford, M. (Eds), *Handbook of Macroeconomics*, Vol. IA, Chap 4, North-Holland, Amsterdam.

E

Easterly, W. (2001). "The Middle Class Consensus and Economic Development." *Journal of Development Economics, 6*: 317–335.

Easterly, W. (2001). "The Political Economy of Growth without Development: A Case Study of Pakistan." Paper for the Analytical Narratives of Growth Project, Kennedy School of Government, Harvard University. http://www.nyu.edu/fas/institute/dri/Easterly/File/Pakistan.pdf

Easterlin, R. (1960b). “Regional Growth of Income,” in Kuznets, S., Miller, A., and Easterlin, R., (Eds) *Population Redistribution and Economic Growth in the United States*, 1870-1950.American Philosophical Society, Philadelphia.

Elias, M & S.J. Rey. (2011). “Educational Performance and Spatial Convergence in Peru.” *Région et Développement*, 33.

Ertur, C & W. Koch. (2006). "Regional disparities in the European Union and the enlargement process: an exploratory spatial data analysis, 1995–2000." *The Annals of Regional Science*, 40(4): 723-765.

Estaban, J & D. Ray. (1994). “On the Measurement of Polarization.” *Econometrica*, 62:819-852.

F

Fallah, B & M. Partridge. (2007). “The Elusive Inequality-Economic Growth Relationship: Are there Differences between Cities and the Countryside?” *The Annals of Regional Science, 41*(2): 375-400.

Fan, C.C & E. Casetti. (1994). “The Spatial and Temporal Dynamics of US Regional Income, 1950-1989.” *The Annals of Regional Science*, 28: 177-96.

Feldman, Maryann P. (1994). *The Geography of Innovation*. Boston, Kluwer.

Feldkircher, M. (2006): 'Regional Convergence within the EU-25: A Spatial Econometric Analysis.' In: New Regional Economics in Central European Economies: The Future of CENTROPE, *Proceedings of OeNB Workshops* No. 9, Oesterreichische Nationalbank, 2006, pp. 101-119.

Ferreira, F. (2001). “Education for the Masses? The Interaction between Wealth, Educational and Political Inequalities.” *Economics of Transition, 9*(2): 533–552.

Ferreira, H.G.F., R. Paes de Barros. (2005). “The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976-96.” In: F. Bourguignon, F.H.G. Ferreira, N. Lustig, eds. (2005). *The Microeconomics of Income Distribution Dynamics*. World Bank and Oxford University Press.

Fields, G. S. (2003). “Accounting for Income Inequality and its Change: A New Method with Application to Distribution of Earnings in the United States.” In Polachek, S. W., Editor, *Research in Labour Economics*, Vol 22.

Fingleton, B & E. López-Bazo. (2006). “Empirical Growth Models with Spatial Effects.” *Papers in Regional Science*, 85 (2): 177-219.

Florida, R. (2002a). "Bohemia and Economic Geography." *Journal of Economic Geography*, 2: 55-71.

Forbes, K. J. (2000). “A Reassessment of the Relationship Between Inequality and Growth.” *American Economic Review, 90*: 869–887.

Fotheringham A, Bunsden C and Charlton M. (2000). *Quantitative Geography*. London: Sage.

FraBdorf, A., M. M. Grabka & J. Schwarz. (2008). "The Impact of Household Capital Income on Income Inequality: A Factor Decomposition Analysis for Great Britain, Germany and the USA." *IZA Discussion Paper* No. 3492.

Franzese, R. & J. Hays. (2007b). “Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data,” *Political Analysis*, 15(2):140-64.

Fujita, M., P, Krugman., & A. J. Venebles. (1999). *The Spatial Economy: Cities, Regions and International Trade.* Cambridge, Massachusetts., MIT Press.

G

Galor, O & J. Zeira. (1993). “Income Distribution and Macroeconomics.” *Review of Economic Studies,* 60: 35-52.

Galor, O., & D. Tsiddon. (1997). “Technological Progress, Mobility, and Economic Growth.” *American Economic Review, 87*: 363–382.

Gehring, K & K. Kulkarani. (2006). “Economic Growth and Income Inequality in India.” Available at: http://kulkarnibooks.com/assets/downloads/kishore_papers/Growth_and_Inequality_in_India_with_keith.pdf (accessed on November 2010).

Glaeser, E., J. Scheinkman, & A. Shleifer. (2002). “The Injustice of Inequality.” *National Bureau of Economic Research,* Working Paper No. 9150.

Glomm, Gerhard & B. Ravikumar. (1992). "Public Versus Private Investment in Human capital: Endogenous Growth and Income Inequality." *Journal of Political Economy,* 100(4): 818-834.

Government of Pakistan (2001). Socio-Economic Indicators at District Level—Sindh, Federal Bureau of Statistics, Statistical Division.

Government of Pakistan (2001). Socio-Economic Indicators at District Level—NWFP, Federal Bureau of Statistics, Statistical Division.

Government of Pakistan (2002). Socio-Economic Indicators at District Level—Punjab, Federal Bureau of Statistics, Statistical Division.

Government of Pakistan (2002). Socio-Economic Indicators at District Level—Balochistan, Federal Bureau of Statistics, Statistical Division.

Government of Pakistan (2010). Economic Survey 2005-06, Government of Pakistan, Finance Division, Islamabad, Pakistan.

Green, W. H. (2003). *Econometric Analysis.* (5th ed.) Englewood Cliffs, NJ: Prentice Hall.

Griffin, K & A.R. Khan. (1972). K. Griffin: Financing Development Plans in Pakistan, in *Growth and Inequality in Pakistan*, Macmillan, London.

H

Haining, R. (2003). Spatial Data Analysis. Theory and Practice. Cambridge, Cambridge University Press

Halvorsen, R., & R. Palmquis. (1980). "The Interpretation of Dummy Variables in Semilogarithmic equations." *American Economic Review*, 70(4): 474—475.

Haq, Rashida (1998). Inequality in Income, Expenditure and Welfare in Pakistan. Unpublished M.Phil Dissertation, Quaid-i-Azam University, Islamabad.

Haque, N-ul- (1977). "An Economic Analysis of Personal Earnings in Rawalpindi City." *Pakistan Development Review*, 16: 4: 353–382.

Haque, N-ul- (2005). "Brain Drain or Human Capital Flight." *PIDE, Lectures in Development Economics*, No. 11.

Haque, N-ul-. (2006). "Beyond Planning and Mercantilism—An Evaluation of Pakistan's Growth Strategy." Inaugural Lecture, *PIDE Lectures in Economics,* Pakistan Institute of development Economics.

Haque, N-ul- (2006). "Awake the Sleeper Within: Releasing the Energy of Domestic Commerce." *PIDE- Working Papers*, 2006:11.

Heckman, J. (1979). “Sample Selection Bias as a Specification Error.” *Econometrica*, 47: 153-161.

Henry, Nick & S, Pinch. (2006). “Knowledge and Clusters,” in C. Pitelis, R. Sugden and J.R. Wilson (eds), *Clusters and Globalisation: The Development of Urban and Regional Economies, Cheltenham*: Edward Elgar, pp. 114-132.

Heshmati, A. (2004). "A Review of Decomposition of Inequality." IZA Discussion Paper No 1221.

Hornstein, A., P. Krusell & G.L. Violante. (2005). The Effects of technical Change on Labour Market Inequalities. In *Handbook of Economic Growth* 1B (Elsevier Press: Amsterdam), 1275-370.

Hospers, G-J. (2006). “Silicon somewhere? Assessing the Usefulness of Best Practices in Regional Policy.” *Policy Studies*, 27(1): 1–15.

Hussain, F & MQasim, M. A. (2005). "Inequality in the Literacy Levels in Pakistan: Existence and Changes Overtime." *South Asia Economic Journal* 26: 251-264.

Hussain, A. (1988) *Strategic Issues in Pakistan’s Economic Policy: Technical Change and Social Polarization in Rural Punjab*. Lahore: Progressive Publishers.

Hussain, A. (2003). “Pakistan National Human Development Report 2003—Poverty, growth and Governance.” Islamabad: United Nations development Program.

Hussain, A. (2008). “Institutional Imperatives of Poverty Reduction.” Research Paper Series. Beaconhouse, National University.

Hussain, I. (1999). “*The* Political Economy of Reforms—A Case Study of Pakistan.” *Lectures in Development Economics,* No.10. PIDE.

Hussain, I. (1999). *The Economy of an Elitist State*. Karachi, Oxford University Press.

Hyder, A. (2007). "Wage Differentials, Rate of Return to Education and Occupational Wage Share in the Labour Market of Pakistan." PIDE Working Paper No 17.

I

Idrees, M. (2006). “An Analysis of Income and Consumption Inequalities in Pakistan.” Unpublished PhD Dissertation, Department of Economics. Quaid -i- Azam ,University, Islamabad.

International Labour Organization (2008). *Income Inequality in the Age of Financial Globalization*, World of Work Report (WWR), International Institution for Labour Studies. Geneva.

J

Jafari, S. Y. and A. Khattak.(1995). “Income Inequality and Poverty in Pakistan.” *Pakistan Economic and Social Review,* 33:1 and 2.

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

Jamal. H, I. Toor & F. Khan. (2003). "Private Returns to Education: Evidence from Pakistan." SPDC Research Report No.50.

Jamal, H & A.J. Khan. (2003). "The Changing Profile of Regional Inequality," *The Pakistan Development Review*, Pakistan Institute of Development Economics, 42(2): 113-123.

Jamal, H, A. J. Khan., I. A. Toor & N. Amir. (2003). "Mapping the Spatial Deprivation of Pakistan," *The Pakistan Development Review*, Pakistan Institute of Development Economics, 42(2): 91-11.

Jamal, H & A. J. Khan. (2005). “The Knowledge Divide: Education Inequality in Pakistan” *The Lahore Journal of Economics*, 10(1): 83-104.

Jamal, H & A. J. Khan. (2008). “Trends in Regional Human Development Indices”, Research Report 73.

Jamal, H & A. J. Khan. (2008). “Education Status of Districts: An Exploration of Inter-Temporal Changes”, Research Report 71.

Joekes, S., N. Ahmed, A. Ercelawn, & S. Zaidi. (2000). “Poverty Reduction without Human Development in Pakistan.” *Development Policy Review*, 18: 37-62.

Juhn, C., K. M. Murphy & B. Pierce. (1993). “Wage Inequality and the Rise in Returns to Skill.”
Journal of Political Economy, 101: 410-442.

K

Karlsson, C. & K. Haynes. (2002). “Regional labor markets in transition.” *Papers in Regional Science*, 81(3): 301-304.

Kawachi, I., B. P. Kennedy., K, Lochner & D. Prothrow-Stith . (1997). "Social Capital, Income Inequality and Mortality," *American Journal of Public Health*, 87(9): 1491-1498.

Kelejian, H.H., & D. P. Robinson (1993). "A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, An Application to a County Expenditure Model," *Papers in Regional Science*, 72: 297-312.

Khan, R. I. A. (2003). *Spatial Distribution of Population with Special Refernce to 1998 Population Census*. In A.R.Kemal, M.Irfan and N.Mahmood (eds.*). Population of Pakistan: An Analysis of 1998 Population and Housing Census*. Islamabad: Pakistan Institute of Development Economics.

Khan, F. S & I. A. Toor. (2003). "Changes in Returns to Education in Pakistan: 1990-2002." *The Lahore Journal of Economics,* 8(2): 83-99.

Khan, S. R., & M. Irfan. (1985). "Rate of Returns to Education and Determinants of Earnings in Pakistan." *Pakistan Development Review,* 34(3): 671–80.

Kelejian, H.H & Prucha., I.R. (1998). "A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." Journal of Real Estate Finance and Methods 17, 99-121.

Knack, S & P. Keefer. (1997). "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics,* 112: 1251–1288.

Knowles, S. (2001), "Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data." mimeo.

Kraay, A. (2006). "When Is Growth Pro-Poor? Evidence from a Panel of Countries." *Journal of Development Economics,* 80(1): 198-227.

Krugman, P. (1991). *Geography and Trade*. MIT Press, Cambridge.

Krugman, P & A. J. Venables. (1995). "Globalization and Inequality of Nations." *Quarterly Journal of Economics,* 11: 857-880.

Krugman., P. (1999). "Increasing Returns and Economic Geography." *The Journal of Political Economy,* 99(3): 483-499.

Kucera., D. (2002). "The Effects of Wealth and Gender Inequality on Economic Growth: A Survey of Recent Empirical Evidence." *International Institute for Labour Studies*, Working Paper DP/136/2002.

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review,* 45(1):1-28.

L

Lagendijk., A. (2000). “Learning in Non-core Regions: Towards ‘Intelligent Clusters’; Addressing Business and Regional Needs.” In: F. Boekema, K. Morgan, S. Bakkers and R. Rutten, Editors, *Knowledge, Innovation and Economic Growth*, Edward Elgar, Cheltenham (2000), pp. 165–191.

Larosse, J. (2004). “Towards a third generation industrial policy in Flanders.” *IWT Studies* No. 52, Het Vlaamse Instituut voor Innovatie door Wetenschap en Technologie , Brussels.

Lawson C & E. Lorenz. (1999). “Collective Learning, Tacit Knowledge and Regional Innovative Capacity.” *Regional Studies,* 33: 305-317.

Le Gallo, J & C. Ertur. (2003). An Exploratory Spatial Data Analysis of European Regional Disparties, 1980-1995, in European Regional Growth (Advances in Spatial Sciences) by Bernard Fingleton (Ed), Springer.

Le Gallo, J. & S. Dall'erba. (2004) "Evaluating the temporal and the spatial heterogeneity of the European convergence process, 1980-1999." Urban/Regional 0401003, EconWPA.

Le Gallo, G & C. Ertur. (2003) "Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995." *Journal of Economics*, 82(2): 201.

Legovini, A, C. Bouillon & N. Lustig. (2005). Can Education Explain Changes in Income Inequality in Mexico? In F. Bourguignon, F.H.G. Ferreira, N. Lustig, eds. (2005). *The Microeconomics of Income Distribution Dynamics*. World Bank and Oxford University Press.

Leon, G. (2007). The Military, Wealth and Strategic Redistribution. *Paper presented at the Northeast Universities Development Consortium Conference (NEUDC),* Harvard University.

Leoni, T & W, Pollan. (2003). "The Impact of Inequality on Economic Growth," *WIFO Working Papers* 211, WIFO.

LeSage, J., & R. K. Pace. (2009). Introduction to Spatial Econometrics. Boca Raton, FL: Chapman & Hall.

Levine, R. (2004). “Finance and Growth: Theory and Evidence.” *NBER Working Paper Series* No. 10766.

Li, H & Z. Hengfu. (1998). “Income Inequality is Not Harmful for Growth: Theory and Evidence.” *Review of Development Economics, 2*: 318–334.

Lopez, R., V. Thomas & Y. Wang. (1998). "Addressing the Education Puzzle. The Distribution of Education and Economic Reforms." *World Bank Working Papers, No*. 2031.

Lopez, H & L. Serven. (2006). “A Normal Relationship? Poverty, Growth, and Inequality.” *World Bank Policy Research,* Working Paper, 3814.

Lucas, Robert. E. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics,* 22: 3-42.

M

Magrini, S. (2004). "Regional (di)convergence," Handbook of Regional and Urban Economics, in: J. V. Henderson & J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, Edition 1, Volume 4, Chapter 62, pages 2741-2796, Elsevier.

Mankiw, N.G., Romer, D., & Weil, D. (1992). “A contribution to the empirics of economic growth,” *Quarterly Journal of Economics*, 107: 407-437.

Ministry of Education of Pakistan (MOEP) & The Pakistan Education Task Force (2011). *Education Emergency Pakistan*. Available on: http://educationemergency.com.pk/wp-content/uploads/EE-Killer-Facts-English.pdf (accessed on 10 March, 2011)

Malik., A. (2009). “Quality and Coordination of Official Development Aid in Pakistan.” Brookings Institute, Available at: http://www.brookings.edu/papers/2009/08_pakistan_aid_malik.aspx, (accessed on January 20, 2011).

Martin, R & P. Sunley. (1998). "Slow Convergence? The New Endogenous Growth Theory and Regional Development. " *Economic Geography,* 74(3): 201-227.

Maskell, P & A. Malmberg. (1999). "Localised Learning and Industrial Competitiveness." *Cambridge Journal of Economics,* 23(2): 167-185.

Mauro, P. (2002). *The Persistence of Corruption and Slow Economic Growth*. Washington, DC. IMF.

Midelfart-Knarvik, K., H, G. Overman, S. G. Redding & A. G. Venables. (2004). The Location of European Industry, in: A. Dierx, F. Ilzkovitz and K. Sekkat (eds), *European Integration and the Functioning of Product Markets.* Cheltenham, Edward Edgar Publishing.

Mincer, J. (1974). *Schooling, Experience and Earning*s. National Bureau of Economic Research, New York.

Moreno., R & B. Trehan, (1997). "Location and the Growth of Nations," *Working Papers in Applied Economic Theory*, 97-02. Federal Reserve Bank of San Francisco.

Morduch, J. & T. Sicular (2002). "Rethinking Inequality Decomposition, with Evidence from Rural China." *The Economic Journal,* 112: 93-106.

Molho, I. (1995). “Spatial Autocorrelation in British Unemployment.” *Journal of Regional Science*, 35(4): 641-658.

Moulaert, F & F. Sekia. (2003). “Territorial Innovation Models: A Critical Survey.” *Regional Studies*, 37(3): 289-302.

Mur, J & A. Angulo. (2005). “ A Closer look at the Spatial Durbin Model.” *ERSA Conference paper* No. 392.

N

Nan, L. (2000). “Inequality in Social Capital.” *Contemporary Sociology, 29*(6): 785-795.

Naqvi, S. A. A (2008). "Mapping Spatial Inequality in Punjab: Theory and Application of Geographically Weighted Regressions." *University of Massachusetts – New School University Economics Graduate Student Workshop.*

Naqvi, S. A. A (2007). "A Look at Spatial Inequality in Pakistan - Case Study of District Sargodha." *Center for Global, International and Regional Studies. Mapping Global Inequalities - conference papers.* Paper mgi-9.

Naseem, S.M. (1977). Rural Poverty and Landlessness in Asia, ILO Report, Geneva.

Naseem, S.M. (2008). “Political Economy of Structural Reforms in Pakistan.” EABER Working Paper.

Nasir, Zafar M. (1998). “Determinants of Personal Earnings in Pakistan: Findings from the Labour Force Survey 1993-94.” *The Pakistan Development Review*, 37(3): 251–274.

Nasir, Z. M. & H. Nazli (2000) “Education and Earnings in Pakistan.” Pakistan Institute of Development Economics, Islamabad. Research Report No. 177.

Nasir, Z. M. (2002). "Returns to Human Capital in Pakistan: A Gender Disaggregated Analysis." *Pakistan Development Review*, 41(1): 1-28.

Nayab, Durr-e- (2006). "Demographic Dividend or Demographic Threat in Pakistan." *PIDE Working Papers*, 2006:10.

Nelson, J. (ed). (1990). *Economic Crisis and Policy Choice: The Politics of Adjustment in the Third World*. Princeton: Princeton University Press.

Nelson, R & E. Phelps. (1966). "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review: Papers and Proceedings*, 51(2): 69-75.

Nijkamp, P & J. Poot. (1998). "Spatial Perspectives on New Theories of Economic Growth." *Annals of Regional Science*, 32(1): 7-37.

O

O' Donoghue, C. (1999). "Estimating the Rate of Return to Education using Microsimulation." *The Economic and Social Review*, 30(3): 249-265.

Oaxaca, R. L. (1973). “Male-Female Wage Differentials in Urban Labour Markets.” *International Economic Review* 14: 693–709.

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Washington: Brookings Institution.

Olson, M. (1962). *The Logic of Collective Action*. Cambridge: Harvard University Press.

Ottaviano G.I.P. & D. Puga. (1998). "Agglomeration in the Global Economy: A Survey of the New Economic Geography." *World Economy,* 21(6):707-731.

P

Paci, R. & S. Usai. (1999). “Externalities, Knowledge Spillovers and the Spatial Distribution of Innovation.” *GeoJournal*, 49: 381-390.

Paci , R. & S., Usai. (2000). “Technological enclaves and industrial districts. An analysis of the regional distribution of innovative activity in Europe.” *Regional Studies*, 34: 97-114.

Paci, R. & S., Usai. (2001). "Externalities and Local Economic Growth in Manufacturing Industries." Centro Ricerche Economiche Nord Sud, Universita degli Studi di Cagliari, Contributi di Ricerca 01/13.

Panizza, U. (2002). “Income Inequality and Economic Growth: Evidence from American Data.” *Journal of Economic Growth, 7*(1): 25-41

Partridge, M. (1997). “Is Inequality Harmful for Growth? Comment.” *The American Economic Review , 87*: 1019–1032.

Partridge, M. (2005). “Does Income Distribution Affect U.S. State Economic Growth?” *Journal of Regional Science, 45*(2): 363–394.

Pede, Valerien O., Raymond. J. G. M. Florax & Henri L.F. de Groot, (2006). "Technological Leadership, Human Capital and Economic Growth: A Spatial Econometric Analysis for U.S. Counties, 1969-2003." *Purdue University, Department of Agricultural Economics, Staff Paper* No. 06-09.

Perotti, R. (1993). “Political Equilibrium, Income Distribution, and Growth.” *Review of Economic Studies, 60*: 755-776.

Perotti, R. (1996). “Growth, Income Distribution, and Democracy: What the Data Say.” *Journal of Economic Growth, 2*:149–187.

Persson, T., & G. Tabellini. (1994). “Is Inequality Harmful for Growth?” *American Economic Review, 84*:600–621.

Perry, G., O. Arias., H. López., W. Maloney & L. Servén. (2006). *Poverty Reduction and Growth: Virtuous and Vicious Circles.* Washington, D.C.: The World Bank.

Planning Commision of Pakistan. (2011). Framework for Economic Growth Pakistan. Government of Pakistan. Islamabad.

Plotnick, B., E. Smolensky., E. Evenhouse, & S. Reilly. (1996). The Twentieth-Century Record of Inequality and Poverty in the United States. In G. Engerman and R.Gallman (ed.), *The Cambridge Economic History of the United States* (Vol. II). Cambridge, U.K.: Cambridge University Press.

Pritchett, L. (1999). “Where has all the education gone?” World Bank mimeo.

Psacharopoulos, G. (1994). “Return to Investment in Education: A Global Update.” *World Development* 22:9, 1325–1325.

Psacharopoulos, G., and H. Patrinos. (2002). “Returns to Investment in Education.” The World Bank. Policy Research Working Paper No. 2881.

Puga, Diego & A., J. Venables (1997). "Preferential Trading and National Location." *Journal of International Economics* 43(3-4): 347-368.

Pyatt, G. (1976). “On the Interpretation and Disaggregation of Gini Coefficients.” *The Economic Journal*, 86(342):243—255.

Q

Quah, D. (1993a). “Galton’s fallacy and tests of the convergence hypothesis.” *Scandinavian Journal of Economics*, 95: 427-43

Quah, D. (1993b). “Empirical Cross-Section Dynamics in Economic Growth,” *European Economics Review,* 37 (2-3): 426-434.

Quah D. (1996). “Regional convergence clusters in Europe.” *European Economic Review*, 35: 951-958.

Quah, D. (1997). “Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs,” *Journal of Economic Growth*, 2(1): 27-59.

R

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

Rey, Sergio J & Janikas, Mark V (2005). “Regional Convergence, Inequality and Space.” *Journal of Economic Geography, 5*: 155-176.

Redding, S. & P. K. Schott. (2003). "Distance, Skill Deepening and Development: Will Peripheral Countries Ever Get Rich?" *NBER Working Papers* 9447, National Bureau of Economic Research, Inc.

Robertson, R. (2000). “Trade Liberalisation and Wage Inequality: Lessons from the Mexican Experience.” *World Development*, 23 (6): 827-849.

Rodrik, D. (1999). “Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses.” *Journal of Economic Growth, 4*(4): 385-412.

Rodríguez-Pose, A & V, Tselios. (2007). "Mapping the European regional educational distribution: Educational attainment and inequality," Working Papers 2007-18, *Instituto Madrileño de Estudios Avanzados* (IMDEA) Ciencias Sociales.

Rousseeuw, P.J & A.M. Leroy. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.

Rumayya., W.Wardaya & E. A. Landiyanto. (2005). “Growth in East Java: Convergence or Divergence?” available at http://ideas.repec.org/p/wpa/wuwpge/0508005.html, (accessed on 27 October 2010).

S

Sabir, M. & Z. Aftab. (2006). “Province-wise Growth Patterns in Human Capital Accumulation.” *Pakistan Development Review,* 45(4): 873–92.

Sabir, M. & Z. Aftab. (2007). "Dynamism in the Gender Wage Gap: Evidence from Pakistan." *Pakistan Development Review,* 46(4): 865-882.

Saint-Paul, G. & T, Verdier. (1993). "Education, Democracy and Growth." Journal *of Development Economics,* 42(2): 399-407.

Sala-i-Martin, X. (1997). “I Just Ran Two Million Regressions,” *American Economic Review*, 87(2): 178-183.

Sawada, Y. (1997). “Human Capital Investment in Pakistan: Implications of Micro Evidence from Rural Pakistan.” *The Pakistan Development Review,* 36(4): 695–712.

Scarpetta, S. & T. Tressel. (2002). “Productivity and Convergence in a Panel of OECD Industries: Do Regulations and Institutions Matter? *OECD Economics Department Working Papers*, N.342.

Scully, Gerald W. (2002). “Economic Freedom, Government Policy and the Tradeoff between Equity and Economic Growth,” *Public Choice*, 113: 77–96.

Sen, A. (1992). *Inequality Reexamined.* Cambridge, MA: Harvard University Press.

Siebert, H. (1998). “Commentary: Economic Consequences of Income Inequality.” Paper presented at the Symposium of the Federal Reserve Bank of Kansas City on Income Inequality: Issues and Policy Options.

Shabbir, T (1991). “Sheepskin Effects in the Returns to Education in a Developing Country.” *Pakistan Development Review,* 30 (1): 1–19.

Shabbir, T & A. H. Khan (1991). “Mincerian Earnings Functions for Pakistan: A Regional Analysis.” *Pakistan Economic and Social Review,* 29(2): 99–112.

Shorrocks, A. (1980). “The Class of Additively Decomposable Inequality Measures.” *Econometrica*, 48(6): 13—625.

Shorrocks, A. F. (1982). “Inequality Decomposition by Factor Components.” *Econometrica*, 50:193—211.

Shorrocks, A. F. (1983). “The impact of income components on the distribution of family incomes.” *The Quarterly Journal of Economics*, 98(2):311—326

Siddiqui, R., & R. Siddiqui. (1998). “A Decomposition of Male-Female Earnings Differentials.” *Pakistan Development Review*, 37(4):885–98.

Sokal R R., N.L. Oden N L & B.A. Thomson. (1998). “Local Spatial Autocorrelation in a Biological Model.” *Geographical Analysis*, 30: 331- 354.

Solow, R.M. (1956). “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics*, 70: 65-94.

Stanovnik T. (1997). “The Returns to Education”. *Economic of Education Review*, Vol. 16 (4), Pergamon Publications.

Stata Corporation (2001). *Stata Reference Manual: Release 7*. College Station, Stata Press, Texas.

Swan, T.W. (1956). “Economic Growth and Capital Accumulation,” *Economic Record*, (32): 334-361.

T

Trivedi, K. (2002). “Regional Convergence and Catch up in India between 1960 and 1992.” University of Cambridge, Online working paper : http://www.uoit.ca/sas/Economic%20Growth/regional.convergence.pdf

U

United National Development Program (2003). *Pakistan National Human Development Report*. Oxford University Press, Karachi.

V

Van Oort, F. G. (2004). Urban Growth and Innovation. Spatially Bounded Externalities in the Netherlands. Aldershot, Ashgate.

Violante, Giovanni L. (2011) "Skill-biased technical change." The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008. *The New Palgrave Dictionary of Economics Online*. Palgrave Macmillan. 10 Available at: http://www.dictionaryofeconomics.com/article?id=pde2008_S000493> doi:10.1057/9780230 226203.1538 (accessed on January 6, 2011).

W

Wagle, Udaya, R. (2008). “How have the Poor in South Asia Fared between 1980 and 2004? An Assessment of living Conditions.” *South Asia Economic Journal*, 9: 261.

Welch, F. (1999). “In Defense of Inequality.” *American Economic Review, 89*: 1–17.

White, H. (1956). “A Heteroskedasticity – Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, 48: 817-838.

Wooldridge, Jeffrey M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA, MIT Press.

World Bank (2000). World Development Report, 2000-2001: Attacking Poverty. Washington, D. C.

World Bank (2003). Poverty in Pakistan: Vulnerabilities, Social Gaps, and Rural Dynamics. Washington, D.C.

World Bank (2005). Handbook on Poverty and Inequality, World Bank. Available at: http://siteresources.worldbank.org/INTPA/Resources/4299661259774805724/Poverty_Inequality_Handbook_Ch03.pdf

World Development Report (2006). *Equity and Development*. New York: World Bank and Oxford University Press.

World Development Report (2009). *Reshaping Economic Geography*. New York: World Bank and Oxford University Press.

Y

Young, A., M. Higgins, & D. Levy. (2007). “Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data” *MPRA* Paper No. 2714.

You, J & Khagram, S. (2005). “A Comparative Study of Inequality and Corruption.” Paper presented at the annual meeting of the Midwest Political Science Association, Palmer House Hilton, Chicago, Illinois. Online pdf version: http://www.allacademic.com/meta/p_mla_apa_research_citation/0/8/6/5/3/p86534_index.html, (Accessed on 15th Feb, 2011).

Z

Zaidi, S. A., (2005). *Issues in Pakistan’s Economy*, Second Edition, Revised and Expanded, Oxford University Press.

Zakir, N & M. Idrees., (2009). “Trends in Inequality, Welfare and Growth in Pakistan,1963-64 and 2004-2005.” Pakistan Institute of Development Economics. *PIDE working paper* 2009: 53.

Zhang, L. (2008). “Political Economy of Income Distribution Dynamics.” *Journal of Development Economics*, 87: 119 – 13.