(M.Phil Thesis)

A Spatial Econometrics Analysis of Educational Distribution and Regional Income Disparities in Pakistan

By

Haleema Bibi

PIDE2018FMPHILETS03

Supervised by

Dr. Hafsa Hina

Department of Econometrics

Pakistan Institute of Development Economics Islamabad

2020

Pakistan Institute of Development Economics

CERTIFICATE

This is to certify that this thesis entitled: **"A Spatial Econometrics Analysis of Educational Disparities and Regional Income Disparities in Pakistan"** submitted by Ms. Haleema Bibi is accepted in its present form by the Department of Economics & Econometrics, Pakistan Institute of Development Economics (PIDE), Islamabad as satisfying the requirements for partial fulfillment of the degree of **Master of Philosophy in Econometrics**.

External Examiner:

Dr. Saima Nawaz
Assistant Professor,
Department of Economics, COMSATS,
Islamabad

Supervisor:

Dr. Hafsa Hina
Assistant Professor
PIDE, Islamabad

Head, Department of Economics & Econometrics:

Dr. Karim Khan
Associate Professor/Head
Department of Economics &Econometrics
PIDE, Islamabad

Dedication

This thesis is dedicated to:

My beloved parents without whom I could not be, what I am today and all those people who are suffering from regional income Inequalities due to unequal distribution of education.

ACKNOWLEDGEMENTS

"Sometimes our light goes out but is blown into flame by another human being. Each of us owes deepest thanks to those who have rekindled this light."

Foremost I want to offer this endeavor to our Almighty Allah on whom we ultimately depend for sustenance and guidance. I am sure this work would have never become truth without His guidance.

I revere the patronage and moral support extended with love by my parents whose financial support and passionate encouragement made it possible for me to complete this project.

I am deeply thankful to my Supervisor, Dr. Hafsa Hina for her constant supervision, sincere guidance, cooperation, support and motivation throughout this research work.

Besides my advisor, I am highly obliged to my teachers, classmates, friends and Siblings for their encouragement and valuable suggestions. Their contribution has been very crucial in carrying out this research work.

I owe a debt of gratitude to Pakistan Institute of Development Economics for giving me an opportunity to complete this work.

Haleema Bibi

Abstract

Spatial regression models provide the opportunity to analyze spatial data and spatial Dependence. Yet, several model specifications can be used, all assuming different types of spatial dependence. This interactional Multi dimensionality lead to the problem of model identification in spatial modeling. This study is use to identify the best suited spatial model by using classical hypothesis testing approach. While Cross-sectional spatial econometric approach is use to check the spatial dependence between income and educational inequality at district level in Pakistan. Results of the analysis specify the SLM as an optimal spatial model comparative to SEM through classical approach while spatial dependence has found significant in the regional income and educational inequality data.

Educational inequality significantly and positively contributed in reducing the regional income inequality due to spatial interactional effect. Although educational attainment also contributed significantly but educational distribution outer performed than educational attainment in the model. It is concluded that investing for equitable distribution of education will be very effective policy strategy to overcome the problem of regional income disparity in Pakistan.

Keywords: Income inequality, educational inequality, spatial effect

Table of Contents

List of Tables

List of Figures

List of abbreviations

GDP	Gross domestic product
GPI	General Parity Index
REM	Random Effect Model
FEM	Fixed Effect Model
PSLM	Pakistan Social and Living Standard Measurement
ARDL	Autoregressive Distributive lagged Model
TFP	Total Factor Productivity
FDI	Foreign Direct Investment
ML	Maximum Likelihood
GSEM	General Spatial Error Model
DSAR	Durbin-Spatial Autoregressive
SEM	Spatial Error Model
SLM	Spatial lag Model
HIES	Household Integrated Economic Survey
GMM	Generalized method of moment
IV	Instrumental Variable
SWM	Spatial Weight Matrix
OLS	Ordinary Least Square

CHAPTER 1

1.1 Introduction

Spatial effect analysis (i.e. role of space, effect of neighborhood, density and distance) of socio-economic phenomena become extremely important due to the varying types of geographical locations and their interdependence on the basis of strategies and economic planning (Krugman, 1991; Krugman and Venebles, 1995; Quah, 1996; Baldwin *et al*, 2003; van Oort, 2004; World Development Report, 2009). Classical econometric techniques like OLS, used to analyze cross section and panel data, cannot capture the spatial effect as it is assumed that observations/regions are independent to one another. Ignoring spatial dependency could lead to inefficient and biased estimates, invalid inference procedures and as a result wrong conclusions (Lesage 2008).

To capture the spatial effect different spatial econometric models introduced in the literature of spatial econometric however SLM and SEM as proposed by (Anselin 1988), are commonly used spatial models. In many disciplines, as in regional sciences (Goodchild and Haining 2004), economics (Anselin 2002), and marketing (Bronnenberg 2005) applying spatial econometric models provided accurate estimation results in the presence of spatial autocorrelation in the data. However, researchers must consider the nature and pattern of the spatial dependence when applying spatial models to know which model is best fitted to the data. Florax & Nijkamp (2003) demonstrated that the use of appropriate spatial econometric model among the different spatial econometric models is crucial while analyzing social and regional studies having spatially characterized data. Autocorrelation is a common problem in a spatial data. It is often produced spuriously by model misspecification (omitted explanatory variables that are correlated over space and misspecified spatial interaction), McMillem.D (2003).

Misspefication spatial interactions in linear regression models cause serious problems such as meaningless inferences and invalid predictions. As ignoring spatial interactional effect of dependent variable causes OLS estimators to become biased and inconsistent while ignoring spatial error structures OLS remain unbiased, but it becomes inefficient (Cliff and Ord 1981). Recently the econometricians start focusing on the Problem of diagnostic checking, specification testing and models selection because quality of inferences highly depend on the correct specification of the model under consideration; Hossain M.Z (2000). So far different approaches are introduced in the literature of spatial econometric to tackle the model identification problem. But the most commonly used approaches are classical hypothesis testing base procedure and Herdery approach. Florax *et al*, (2003) demonstrated that the classical approach (start with simple OLS, perform LM test and select the optimal model) outperforms than the Hendry approach (follow general to specific methodology). Therefore, we used the classical approach in this analysis.

Most of the existing research analysis of spatial socio-economics characteristics has been focused on provincial level in Pakistan while the role of social interactions among districts is neglected. District level research has become even more important after enactment of the 18th Constitutional Amendment that allowing the distribution of resources from center to regional level, Ahmad.S (2011). Pakistan is characterized with spatial disparities between its key socio-economic characteristics such as education, health, physical infrastructure. Significant spatial effect (spillover effect) of regional education and economic development is found in the neighboring regions confirming that economic geography matters for Pakistan (Burki *et al*, 2010).

In Pakistan there is worsening and persistence level of inter and intra-provincial income and educational inequality along various dimensions of disparity (Oxfam study, March 2015). Inequality in providing basic education facilities lead to many socio-economic disparities i.e income, region ,gender, and social economic indicators which accelerate the vicious cycle of inequality and poverty (Hamid *et al* 2013). In spite of giving the right of free and compulsory education to all by the Government of Pakistan, the school-going children are facing multiple inequalities and disparities in getting primary and secondary education with respect to gender, location, and income and ethnicity profiles. Pakistan is the second largest country in number of out-of-school students in the world, (UNESCO, 2010). GPI for entire Pakistan is calculated 0.86, which shows that only 86 females comparative to every 100 boys are getting primary or lower secondary education. The GPI was estimated at 0.94 for Punjab, 0.79 for Sindh, 0.71 for KPK and 0.69 Baluchistan, UNICEF (2013).

Countries with high income disparity have low literacy rate (Nolan *et al* 2014). Farooq (2010) demonstrated that countries where education is available for every one as a basic necessity, in those countries income inequality is very low. Education is pre-condition for economic development of a nation. It cause to increases the productivity of nation on one side and reduces the poverty on the other. Therefore, in such countries where the distribution of education is equal, the poor peoples of those countries enjoyed an outsized share of country's GDP. As a result, income disparity in such countries is very low, Raja (2000).

Income inequality and regional economic disparity in Pakistan remained a popular topic for researchers in Past. Many researchers such as Siddiqi (1891), Hussain (1993), Tariq *et al* (2003), Akhter (2008), Shaheen *et al* (2016) and Amjad (2016) worked on inter and intra-regional inequalities and disparities .However in Past literature ,the role of geography(space) in the study of regional disparities has been largely disregarded in Pakistan. Many research studies i.e Sergio Rey (2001), Li *et al* (2014), David *et al* (2018), Gonul and Erkut (2019) provided evidences on the existence and significance of spatial dependence in their research study.

The spillover of economic activities across regions create a spatial interdependence. As education spillover across regions could income convergence towards equality Pede *et al* (2012). Tselios (2008) and Umer *et al* (2014) explored the relationship between educational distribution and income inequalities in their regional research studies and demonstrated that geographical location (space) has significant impact on educational and income inequalities. In Pakistan a lot of research work has done on the education's significance in overcoming the problem of regional income inequalities. But no research study focused on capturing the spatial effect (spillover effect) of educational distribution on regional income inequalities. So this study will focus on the implications of spatial effect on educational and income inequalities in Pakistan along with the identification of best suited spatial model Between SLM and SEM Specifications.

1.2 Motivation of the study

The spillover of economic activities across regions create a spatial interdependence. Ignorance the effect of spatial dependence could lead to biased and inconsistent parameter estimates (LeSage, 1998). To capture the spatial interactional effects, different spatial econometric models are introduced in the existing econometric literature. But very little work has done on optimal model selection among different spatial econometric models. Model identification is much debated issue because misspecification of models could lead to inaccurate results and wrong policies. There is need to do a lot of work in this context. Therefore this research study is use to analyzed the spatial effect of education inequality on regional income disparity along with econometric implication of optimal spatial model selection.

1.3 Objectives of the Study

The objectives of this research study are

1: To detect the spatial effect of educational inequality on regional income disparity in Pakistan.

2: To specify the optimal spatial model between SLM and SEM

1.4 Research Questions

1: Is there exist any significant spatial dependence between educational inequality and income disparity at district level in Pakistan?

2: Which specification of the spatial econometric model is optimal to best describe the data?

1.5 Significance of the Study

This research study will contribute in the existing literature of spatial econometrics in term of model specification criteria for regional cross sectional studies. It will benefit the policy makers and regional development planners in making regional policies about the distribution of education to reduce regional income inequalities and disparities because capturing the spatial dependence will help to know which spatial entities (geographical regions /areas) are most affected and require special attention to solve the problem under consideration

1.6 Organization of the Study

The remaining study is organized as follows: chapter two reviews empirical literature on spatial econometric models and spatial effect of educational inequality on regional income disparity. Chapter three discusses the theoretical framework and econometric methodology used in the analysis. Chapter four contains analysis of this research study. Chapter five gives results and policy implications.

CHAPTER 2
LITERATURE RIVIEW

2.1 Introduction

This chapter explains the research gap in available literature and the nature of relationship between educational inequality and income disparity and their spillover effects due to spatial dependence.

Section 2.2 deals with the empirical literature to develop the relationship between educational and income inequality. Critical debates on different spatial econometric techniques has been discussed in section 2.3. Section 2.4 discussed the existing literature on the spatial effect of educational inequality on regional income disparity. While literature gap is reported in the last section of the chapter.

2.2 Empirical Literature on Educational and Income Inequality Relationship

Kanwal and Munir (2015) determined the impact of educational inequality and gender inequality in education on income inequality in South Asian countries for the time period of 1980 to 2010. REM and FEM were used in estimation. Using the education Gini index, it is found that there exist a positive relationship between educational and income inequality. The results also indicated that gender inequality in education at primary and tertiary level has positive and significant impact on income inequality but has negative impact on percapita income while gender inequality at secondary level has negative and significant impact on income inequality but has positive on per capita income. Fair and equal distribution of education is essential along with the expansion of education to remove all these inequalities.

Farooq (2010), analyzed the impact of education on income inequality and estimated the ratios of inequality between regions and male and female labor force through Gini-Coefficient using data from PSLM survey of 2004-05. The results of the study indicated that the income was unequally distributed between male and female labor force. The ratio of inequality was comparatively higher in males than females. The estimates of Gini index was found to be higher in urban areas as compared to rural areas. It is revealed that the distribution of income can be more equal through more educated people. Therefore, it is implied that equal opportunity of education and employment should be given to all without any discrimination of gender and regions.

All developed nations are developed due to educational development. Education playing important role not only in the economic growth but also in reducing the income disparity and poverty. Khan *et al* (2013) explored the various factors affecting the education and its impact on economic growth in Pakistan. This research is based on the interpretive phenomenological approach. It is explored by the study that there are many important factors including differences in income, gender disparity, geographical inequality, feudalism and system of education that affecting the education setup in Pakistan. But the differences in income is most important. It is also found by the study that attainment of education is necessary reducing the level of income inequality in Pakistan. There is need to developed a standardized education system in all over the country through effective policies and government intervention.

Jamal (2016) researched on inter and intra provisional inequalities on socio-economic development indicators related to human resources and standard of living in the dimensions of income, education, health and housing. Micro economic survey data of households (PSLM

2012-13) is used in the analysis. District development rank orders and Provincial multidimensional Gini coefficients are estimated to understand the nature of regional socioeconomic development. The magnitudes of estimated Gini coefficients reflected low level of district per capita income inequality in KPK and Baluchistan as compared Punjab and Sindh provinces while Provincial Gini coefficients just by aggregating household's incomes indicated highest inequality in Baluchistan and lowest in Punjab. This research study also worked to find out the development rank order of districts for both unadjusted and adjusted intra-district inequality. This adjustment significantly affects the development rank orders of districts.

Gragorio and Lee (2002) investigated the role of education distribution on regional income imparity by using the panel data set covering a big range of countries from 1960- 90. The results of the analysis showed that the educational determinants like higher level of schooling and more equitable provision of education contribute significantly to reduce the income inequalities. Education is one of the most important determinant of income equality. The effects of social expenditure also analyzed in the research study and found significant impact of public social security to have equitable division of income.

The empirical literature provided mixed results of inequality-growth nexus. Some empirical studies found positive and significant while others found negative impact of inequality on growth. Majeed (2016) empirically analyzed the impact of income inequality on economic growth in Pakistan using annual time series data from 1975 to 2013. Analysis of the study based on the ARDL approach to Co-integration. The results showed positive significant impact of inequality on annual economic growth of Pakistan. In spite of the fact that inequality has positive impact on growth yet such kind of growth cannot be fruitful for long run as the poor are not include in growth process.

Idrees and Shah (2018) analyzed the educational inequality for urban and rural regions of all provinces and capital region of Pakistan using micro data from PSLM (2014-2015). Educational inequality is measured over the entire and working populace aged of 15 years and above and those who not enrolled in any school. It is found that educational disparities are comparatively lower across working population as compared to entire population and relatively better in urban regions as compared to rural regions but more acute among females than males. The provinces wise analysis found that the intensity of educational disparities are low in Islamabad but high in Baluchistan and Sindh.

Abdullah *et al* (2011) revisited the existing research studies that investigated the impacts of education on disparity. Particularly, it provided an extensive summery of the 64 econometrics studies through a meta-regression analysis that unanimously reported 868 results of the impact of education on disparity. It is explored that education diminishes the income part of high earners and increases the part of the low earners, yet has no impact on the part of the middle class earners. Educational inequality increases income inequality. Education has more negative impact on inequality in Africa than Asia.

Rodriguez-pose and Teslios (2010) examined the impact of income inequalities and educational disparities on regional economic development in Western Europe. They used cross sectional and panel data analysis based on microeconomic survey data for the time span 1994–2001. To measure the income and educational inequality different inequality indices are used.

However due to high correlation among indices only Theil index is represented. The responses of the growth model indicated significant positive correlation between educational and income inequalities and regional economic development. Existing level of income and education inequality creates socioeconomic incentives and considered growth-enhancing. Generally, existing income and education inequality are probably going to expand growth, however the size of their effect is very little. Nonetheless, expanding inequality cannot adopt as a strategy for the development of regions in Western Europe.

Castello-Climent and Domenech (2008) explored the relationship between human assets inequality, longevity and economic development. In this analysis they proposed a model which is based on the assumption that individual's investments in the accumulation of human assets depend upon their longevity which inturn depends upon the human capital and socio-economic status of their parents. Individuals belong to rich families, their parents have higher level of education and life expectancy relative to those who belong poor families. The high life expectancy of their parents enforce them to devote more no of years to their education because they have much time to get benefit from their investments while poor children not do so. So they work as non-educated workers and contribute less in the economic welfare of the country. Result of the study found inverse relationship between human assets inequality and economic development. Study implied that government should give free education to poor and provide them better health facilities to improve their life expectancy.

2.2 Critical Review on spatial econometric models (SLM & SEM)

Spatial dependence can be expected in a data set with observations that are collected from different locations. Cross sectional studies that involve micro level data of households or firms having more chance of the problem of spatial autocorrelation than studies that involve macro data. This problem arises due the large numbers of observations contain by micro level studies which are characterized by spatial relationships (Bell and Bockstael, 2000). Many research studies i.e Burki *et al* (2010), Ahmad.S (2011), Li *et al* (2014), Umer *at el* (2014) , Iqbal & Nawaz (2017), Gonul &Erkut (2019) confirmed the existence of spatial dependence and highlighted the importance of economic geography in their regional research studies.

Regions are not isolated but are a part of a core-periphery system. Disregarding of interactions between regions in the system can lead to adverse impact on regional policies Ahmad.S (2011). As traditional econometric techniques could not capture the spatial dependence and lead to biased and inconsistent results. Therefore, economists have turned their attention to spatial modelling which can capture the spatial effect (spillover effect) due to spatial autocorrelation (Yang and Zheng, 2010). However, it is important to find the nature and pattern of the spatial dependence when applying spatial models to know which model is appropriate for data.

In the literature of spatial econometrics many models have developed which treat three different type of spatial dependence, Endogenous spatial dependence among the dependent variable, exogenous spatial dependence among the independent variable and spatial interactions among stochastic error term (Elhrost, 2013). These three types of dependencies can captured through Spatial Lag Model (SLM), Spatial Lag of X (SLX) and spatial error Model (SEM) which are commonly used in existing literature of spatial econometrics. As research studies of li *et al* (2014), Umer *et al* (2014), Kivi. L.H (2019), Raza &Hina (2016), Nawaz & Mangla, provided evidences on application of these spatial models in their research studies.

Specification of models has been largely neglected, even though it leads to serious inference problem. Seemingly small changes to model specification have major impact on the spatial effect estimates (Numayer and Plumper 2010). Osland.L(2010) analyzed that misspecifications of spatial models (i.e. miss specified model functional form, missing spatial variables and spatial heterogeneity) lead to spurious results. To overcome the problem of spatial model identification Anslin *et al* (1996) proposed Lagrange Multiplier tests (Robust LM Lag and Robust LM Error) which are used to test the Spatial Lag and Spatial Error Specifications. SLM is use to capture the lag dependency in the dependent variable while SEM is use to capture shock transmitted through error term, Anslin (1988).

Zhu *et al* 2017 compared the spatial econometric models and Random Forest for modelling the fire occurrence in regional cross-sectional study of China. They explored that spatial econometric models have better predictive ability than RF and among them Spatial autocorrelation model also known Spatial lag model is best to described the data. To capture the Spatial clustering and identifying the optimal spatial model Higazi *et al (*2013) applied Exploratory spatial data analysis (ESDA) and compared the OLS, SLM and SEM on the basis of Likelihood ratio test .The SEM proved to be the better model than SLM.

2.3 Education inequality, income inequality and spatial econometrics

The spatial analysis of socio-economic characteristics has extremely important in the regional growth and development studies as regions of an economy are interconnected with each others through boarders and policies. To contribute in this context Amad.S (2011) analyzed the spatial aspects of income and education inequalities in Pakistan in three interrelated studies. The first study investigated extend of changes in earning inequalities due to changing in returns to human capital. The second study analyzed extend of spatial clustering of economic inequalities, growth and development across Pakistani districts between 1998 and 2008 by applying exploratory data analysis. The third study income convergence across Pakistani districts between 1998 and 2005 by using spatial and non-spatial econometric techniques. The results of the first study revealed that returns to the low level of education have declined while the returns to the higher education level have increased, remained much higher for females as compared to males, and higher within provinces as compared to between provinces. The results from the second study demonstrated that dependency of neighboring districts on eachothers's development, confirmed the importance of economic geography for regional inequalities (income & education), growth and development across Pakistan. The finding of the last study observed conditional convergence across Pakistan once spatial effects have been taken into account.

The operation of human capital and knowledge spillover play an important role in generating dependencies and disparities. Umar *et al* (2014) formulated regional production model that account educational inequity as a key factor of Nigerian regional inequality of income. Spatial econometric techniques (SLM, SEM) confirmed the presence of spatial interactions of educational inequality on Nigerian regional income disparity. It is explored that equitable provision of education has positive and beneficent impacts on regional income level and it outperformed than educational attainment in the model. To minimize the regional income disparities and improve the regional economic performance, allocation of resources for fair distribution of education will be very a good strategy for Nigeria.

LV *et al* (2017) analyzed the spatial impact of education on economic development using the spatial modeling across 31 chine's regions from the period of 1996-2010. Results revealed that education is a key determinant of economic growth and educational factors are more beneficial for economic growth than labor force and capital investment. It is also found that educational factors are spatially auto correlated and have spatial spillover effect on neighboring regions. Education sector can benefits and upgrade non education sectors due to significant spatial spillover effect. So it is suggested that spatial effect of regions located in neighborhood not be overlooked while making educational strategies and government should make such polices that enhance education development and reduce the educational distribution disparity among regions.

Takahashi (2007) explored the factors of regional income inequality in Vietnam focusing mainly on the role of human capital and land endowments. Human capital found as a leading indicator of regional income disparity rather land endowments and other assets because there is no correlation between these assets inequality and regional income disparity except human capital. The researcher divided the country into two regions. It is found that growth of returns to human assets in one region become an indicator to increase the returns to human assets in another region. Surprisingly differences in land contribution do not firmly correspond with regional income inequality because lower revenue were found by better provision to land in the region. There is need to develop the human capital to get benefits from all type of assets.

Perugini and Martino (2008) determined the factors of income inequality within European regions along with correlation between economic disparity and regional economic development. They explored different economic, social, demographic and institutional factors of income inequalities. It is suggested by the study that detection of spatial patterns are very important and necessary to examine the regional inequality determinants. By employing the spatial descriptive tools, spatial effects of inequality were noticed over the regions by employing spatial autoregressive models through ML method. The significance of spatial effects is explored while measuring income inequality. Variegated determinants of income inequality had found through both data sets. As respects the effect of economic disparity on development, the outcomes demonstrated a positive relationship

Umar *et al* (2013) worked to measure the educational inequalities between and within regions of Nigeria. They conducted a cross- sectional study on northern and western states of Nigeria. Theil index is used to measure the educational inequality. Results of the study indicated that Theil index is a robust measure of inequality than other measures because of its inequality decomposability property into between and within regions. It is found that educational inequality is higher in northern region comparatively southern regions as the magnitude of Theil index is higher for 17 out of 19 regions than estimates of state-level Theil index. Intra-regional educational disparity is more than interregional disparity and is the big reason of income disparity in Nigeria. It is also found that there is a negative correlation between level of education and its disparity which indicates that regions having more educational achievement are capable to attain more equal distribution of education and lesser regional disparity.

Tirado *et al* (2016) analyzed the work of Jeffrey G. Williamson about an inverted U-shaped pattern of regional income inequality across Spanish provinces (NUST 3) by using the novel data set of 150 years from 1860 to 2010.They worked on the different dimensions of regional income disparity by examine the mobility (poor/rich provinces retrain their position or not) and spatial clustering effects. The results of analysis affirmed the presences spatial clustering and U-shaped pattern which shows a little group of well off regions and big share of bad off ones. Regional income inequality was moderately low in the early stages unlikely to the early many years of twentieth century. It has steadily declined since the 1930s and followed the convergence until the 1980s, yet impact of movability rather low and spatial gathering significantly expanded.

Fischert and stribock (2006) analyzed the club convergence hypothesis about regional income growth in the frame work of spatial econometric. This research study provided a more practical and complete picture about cross regional development in 256 European regions using data from 1995-2000. This research study testing two club convergence hypotheses and focused on the heterogeneous pattern and spatial error dependence. Results are divided into threefold. First, standard Barro-style regression model is rejected .Second, heterogeneous pattern has found in the pan-European convergence. Third, spatial error correlation introduced an imperative inclination in the recognition of the club-convergence but study showed that neglecting this inclination leads to deceiving results. So it is implied that spatial dependence shouldn’t be ignore to get unbiased results.

2.5 Literature Gap

Detection of Spatial dependence at micro level cross-sectional studies is extremely important because regional governments are interlinked on basis of strategies and boarders. The act of one government have feedback effect (spillover effect) to another. If spatial dependence and spillover effect not take into account then it could lead to biased and inconsistent parameter estimates (LeSage, 1998). And to capture these spatial interactions, different spatial econometric models have put to use in the existing literature. But very little work has done on model selection which is best suited to the data among different spatial econometric models. There is need to do more work in this context. Therefore this research study is use to analyzed the spatial effect of education inequality on regional income disparity along with econometric implication of optimal spatial model selection.

CHAPTER 3
DATA AND METHODOLOGY

3.1 Introduction

Section 3.2 and 3.3 explain the theoretical framework of regional income and educational inequality and econometric approach of model respectively. The rest of the sections describe the econometric methodology used in the analysis of this research study.

3.2 Theoretical Framework

In existing economic literature, Human capital had great importance as a key factor of economic growth as Human capital increases the productivity of both labor and physical capital (Lucas, 1988 and Barro, 1991. However human capital stock depends upon the level of educational attainment of a nation. It is the level of education that boost the economic activities by developing ability to take initiatives or adapt existing technologies. Investment in human capital is very beneficent that it increases the individual's income and rate of return of all assets (Becker, 2009). To check the impact of education on regional growth, Polasek, Schwarzbauer and Sellner, (2010) promote hypothesis that skilled labor plays a key role in regional growth.

Many research studies explored the determinants of income inequality across regions. As Naschold (2009) explored the factors affecting the income inequality and impacts of changes in household age, gender, education, wealth and location on it. He found that income inequality reduced over time due to changes in demographic characteristics of the households and attainment of higher education along with the provision of higher education more equally. By analyzing the determinants of income inequity in the Portuguese region of Lisboa , Crespo *et al* (2012) found that inequality increases because of differences in the socioeconomic and demographic characteristics of households such as age, gender, household size, average education, income and employment status. In the comparison of all the variables, the larger contribution to income inequity is due to differences in average education levels.

Similarly some other research studies also found these socio-economic and demographic characteristics as the key determinants of regional income inequality. As Nebebe and Rao (2016) explored that household size, household head income sources, general economic condition and residential place are the main factors of income disparity for all type of income earners. Whereas, the household educational attainment and housing residency are the root cause of income inequality at the upper and lower quantiles distribution. Umar *at el* (2014) proposed a regional growth model that account education disparity as a key factor of regional income inequality in Nigeria implied that investing on equitable distribution of education will be very good strategy, to minimize the regional income inequalities in Nigeria.

3.3 Econometric Approach

To determine the role of educational distribution on income inequality across the regions of Pakistan, regional inequality model is developed on different economic and demographic variables. Regional income inequality (yInq) is hypothesized to be the function of the following regional features; the demographic characteristics of the region (Demo) such as Age and household size, the regional level of Educational attainment (Edu), the regional educational inequality (EduIneq), the regional industry structure (Indstry) which is measured by the proportion of working population that is in agriculture and proportion of working population that is living in urban area and the general economic condition of the region (Ecr), measured by the regional per capita GD.

Hence the model is specified as:

$$YInq_{ij} = \beta_{\circ} + \beta_1 Demo_{ij} + \beta_4 Indstry_{ij} + \beta_5 Ecr_{ij} + \beta_4 Edu_{ij} + \beta_4 EduInq_{ij} + \mu_{ij} \qquad (1)$$

3.4 Methodology

The section of methodology contains all the components of spatial econometric methodology used in the analysis i.e. spatial econometric models, model section technique and data & variables.

3.4.1 Spatial Econometrics

Spatial econometrics is a sub dimensional field of econometrics which deals with spatial interaction effects among geographical regions. Detection and capturing Spatial autocorrelation (Spatial spillover effects) have main interest in regional sciences. A valuable point of spatial econometrics is that, the magnitude and spatial spillovers can be empirically investigated. Spatial econometrics models are used to explain the behavior of economic agents of different geographical units (Elhorst, 2014).

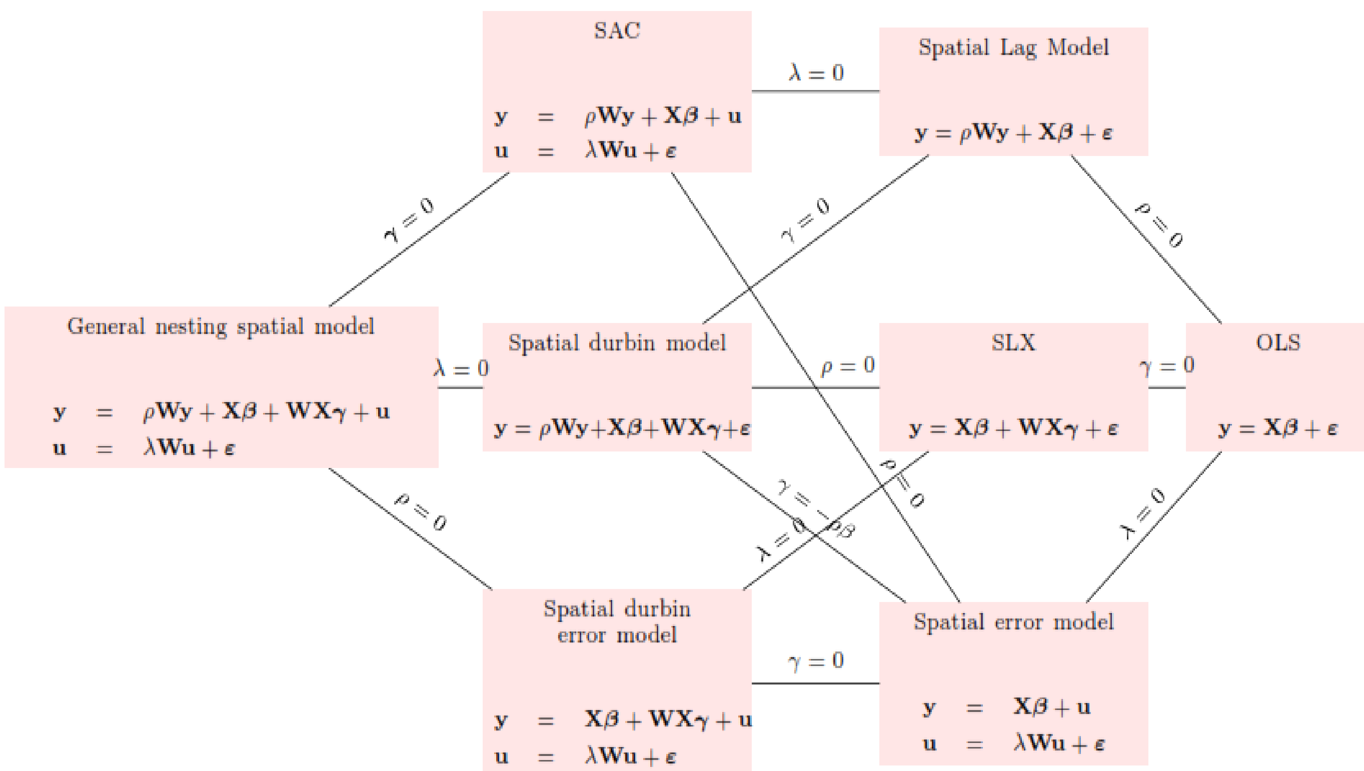
3.4.2 Spatial Autocorrelation

The spatial autocorrelation is the dependence of an observation of a variable in a particular location on the same variable observations in neighboring region (Anselin, 2001). For example, consumption decisions of an area can be influenced by the expenditure decision of the neighboring locations. Detection of the spatial patterns in studies that covers large geographical space is important because of the potential econometric issues emerging from it. Spatial autocorrelation leads to the violation of some statistical assumptions used in the traditional analysis approach i.e. the assumptions of uncorrelated error terms and independent observations (LeSage and Pace, 2009).

3.4.4 Spatial Econometrics Models & Possible Spatial Interaction Effects

It is important to find the nature and pattern of the spatial dependence when applying spatial models to know which model is appropriate for data. Existing literature of spatial econometrics identified two types of spatial dependence; spatial lag and spatial error (Anselin, 1988). However advance literature of spatial econometric introduced some other types of spatial interactions which can be captured by different spatial models.

Figure 3.1: Classification of Linear Spatial Dependence Models



Elhorst (2014)

$$Y = \rho W y + X\beta + WX_{\gamma} + u, \quad u = \lambda Wu + \varepsilon \qquad (2)$$

Model shown in equation 2 is General Nusted Model (GNM) which takes all possible specifications of spatial interactions. Different spatial econometric models can be taken by imposing linear restrictions on its parameters as shown in the diagram. According to Elhrost (2013) three types of spatial interactions are introduced in spatial econometric literature which can be captured by different spatial econometric models. This research study is confined to use and make comparison between the most commonly used models spatial lag model, spatial error model.

3.4.4.1 Spatial Lag Model (SLM)

In Spatial Lag Model, the explained variable in a particular region is influenced by the explanatory variables in both that particular regions and other regions as well. SLM contains spatially lagged dependent variable as shown in equation (2)

$$Y = \rho W y + X\beta + \varepsilon \qquad (3)$$

Where Y is a vector of N observations on the explained variable; W is a $N \times N$ spatial weights matrix; ρ is spatial autoregressive parameters; X is a matrix of observations on the independent variables, with $K \times 1$associated regression coefficient vector β, ε is a vector of $N \times N$ Residuals.

3.4.4.2 Spatial Error Model (SEM)

Error terms over the different spatial entities are dependent on each other's generally due to omitted variables which are themselves spatially correlated. SEM incorporates a spatially autoregressive process in the error term as appeared in equation (4).

$$Y = X\beta + \mu \qquad (4)$$

$$\mu = \gamma W\mu + \varepsilon$$

Here Y is a vector of N observations on the explained variable; W is a $N \times N$ spatial weights matrix; γ is spatial autoregressive parameter; X is a matrix of observations on the independent variables, with $K \times 1$associated regression coefficient vector β, μ is distributed normally and an independently error term with a constant variance.

According to Anselin (2002), ML, GMM or IV estimation techniques require in the estimation of this type of models. In this research study, we applied ML estimation method using STATA software package.

3.4.5 Spatial Weight Matrix

Spatial weight matrix (W) is precondition for the estimation of spatial models. The SWM is a $N \times N$ non negative matrix of binary numbers, in which one is assign for neighbor, and zero is assign to prevent a region to the neighbor of itself (LeSage and Pace, 2009) .Spatial weight matrix describes the closeness of every observation (spatial unit) with the rest of observations that are considered in the sample.

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$$

There are different types of weights matrices based on boundaries and distance that are used in spatial modelling, depending on the nature and phenomenon being studied. According to Bell and Bockstael, (2000), if the units of observation are households or firms the spatial relationship will be best captured by considering distance decay effects between points. Going by the suggestion of Bell and Bockstael, (2000), we used "Inverse Distance Decay". This type of matrix provides greater weighting to observations that are closer to each other than those that are further apart. In the matrix, the weighting between location *i* and *j*, with W*ij* = 1/*dij*, for *dij* <*Dmax*, and W*ij* = 0 otherwise, where *dij* is the distance between the centroids of location *i* and *j* and *Dmax* is a threshold distance.

3.4.6 Model Comparison & Specification

In this research study Classical technique of hypothesis testing is used in which optimal spatial model is selected on the basis of Lagrange Multiplier (LM) tests as proposed by Anslien (1998). Econometric model is estimated through OLS technique then the presence of spatial autocorrelation is checked in the residuals of estimated OLS through Moran's I statistics and Lagrange Multiplier tests (Classical and Robust LM tests). If the hypothesis of no spatial autocorrelation is rejected then it is tested either the spatial lag model or spatial error model is more appropriate to best describe the data. For this LM tests are more powerful than Moran's I statistics because LM tests not use only to capture the spatial dependence in the data but also specify the best fitted model between

SLM and SEM. After specifying the spatial model (either spatial lag or spatial error) that would be used in the research analysis by using maximum likelihood estimation method.

3.4.7 Gini Index

The Gini Index, introduced by the Italian statisticians Carrado Gini in 1921(Gini 1921), has been used in a wide variety of resources allocation contexts to measure inequality including income, wealth, education, health care. The value of this index lies between 0 and 1; a value closer to one meaning higher level of inequality and vice versa. It shows the least fraction of a variable that must be redistributed to get perfect equality. It has following formula:

$$GINI = \frac{2}{N-1}\sum_{i=1}^{N-1}|F_i - Q_i|$$

Where N is the number of regions, $F_i = i/N$, $Q_i = \sum_{j=1}^{i} y_j \Big/ \sum_{i=1}^{n} y_i$ and y_i is the value of variable y (e.g. income, education etc.) in region j when ranked from low (y_1) to high (y_N) among all regions within a country.

3.4.8 Data &Variables

To examine the spatial effect of educational inequality on regional income disparity in Pakistan micro economic cross-sectional data of households is used which is derived from the HIES survey of 2018-19 provided by the Pakistan Bureau of Statistics. Data is based upon 1, 59,000 of households from Punjab, Sindh, KPK and Baluchistan. 90 districts from all four provinces are included in this research study. Different Spatial econometric techniques are used in the Analysis. Educational and income inequality is calculated by Gini Index. Variables used in the regression analysis are described in the following table.

Variable	Definition
Dependent variable (yInq)	Regional Income Inequality
Demographic Characteristics (Age, Household size)	Average age of household At regional level, Regional average Household size
Schooling	Regional level of educational attainment
Educational inequality(EduInq)	Regional level of educational inequality
Regional industrial structural(Indstry)	Proportion of working population that is in agriculture
Sector	Proportion of working population that is living in urban area
GDP per capita	General economic condition of the region

CHAPTER 4

REGRESSION AND ESTIMATES

In the first part of regression analysis, province wise income and educational inequalities pattern are explained through bar chart across all the districts KPK, Punjab, Sindh and Baluchistan. Income and educational inequalities are calculated through Gini index. Its values lies between 0 and 1; a value closer to 1 means perfect inequality and vice versa for closer to 0.

Figure 4.1: KPK Income & Educational Inequalities

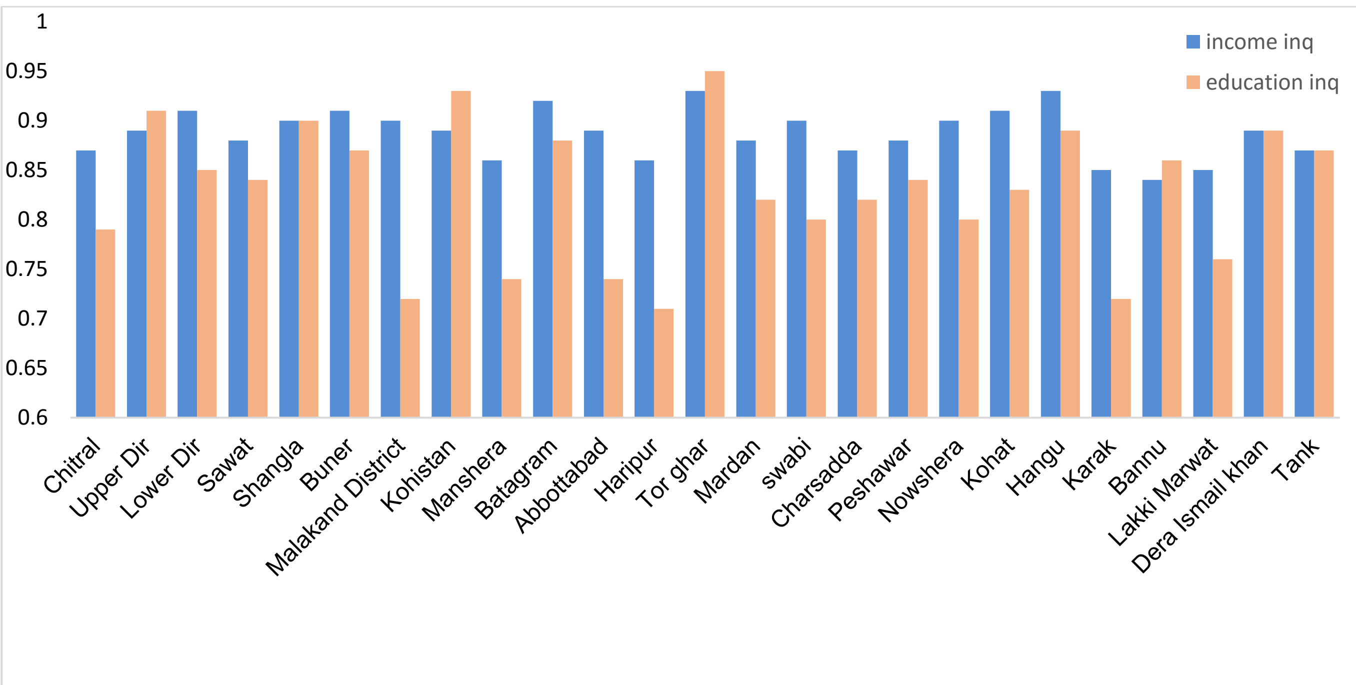


Chart 4.1 displays income and educational inequality patterns across all the districts of KPK. The estimated values of Gini coefficient remained between 0.7 to 1.0 for both income and education inequality estimates which depicts severe inequality among all the districts. In some districts (i.e. Shangla, Dera Ismail khan and Tank) the ratio of income and educational

inequality remained equal while in the others districts was unequal. The level of income inequality is high than education inequality in all the districts except Upper Dir, Kohistan, Torghar and Bannu districts.

Figure 4.2: Punjab Income & Educational Inequalities

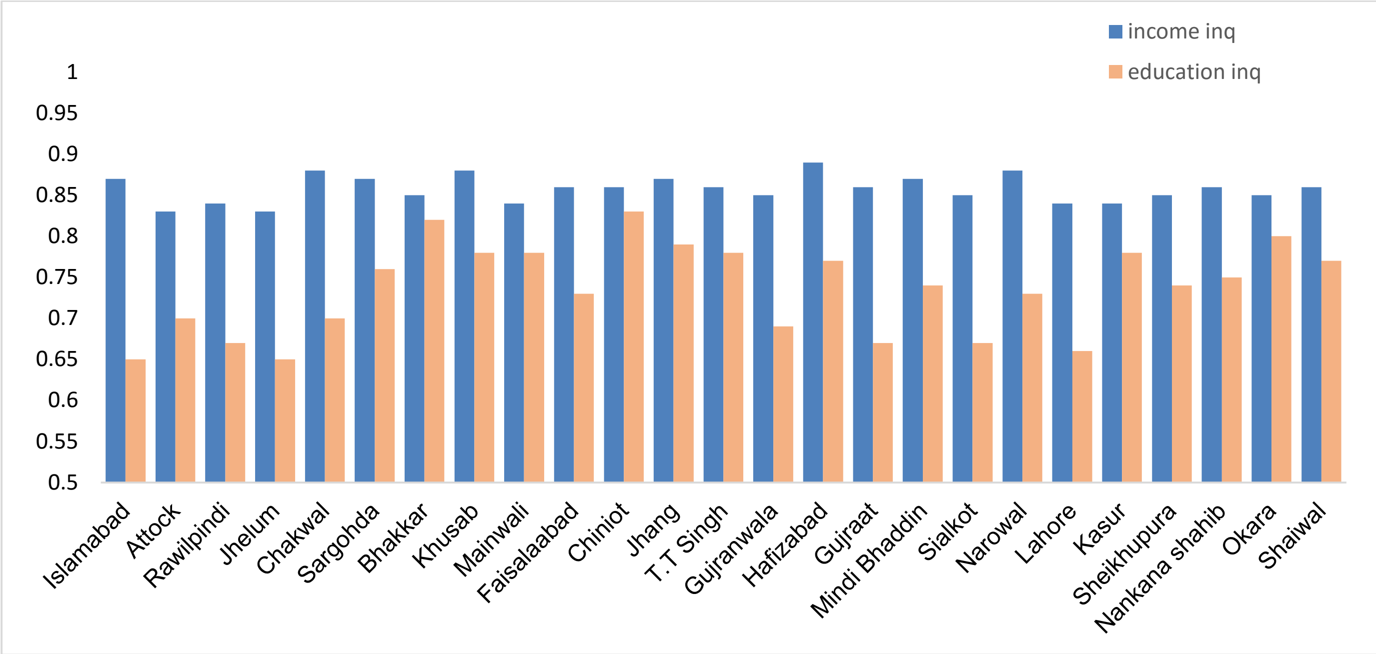


Chart 4.2 displays income and educational inequality patterns across all the districts of Punjab. Values of Gini coefficient lies in 0.65 to 0.9 for both inequality estimates which show high income and educational inequality among all the districts. High level of income inequality pertained in Islamabad, Chakwal, Sargohda, Khusab, Hafizabad, Mindi Bhaddin and Norowal districts while level of education inequality remained high in Bhakkar, Chiniot, Kasur, Okara and Shaiwal districts. Income inequality is found to be high comparative to the educational inequality across all districts of Punjab.

Figure 4.3: Sindh Income & Educational Inequalities

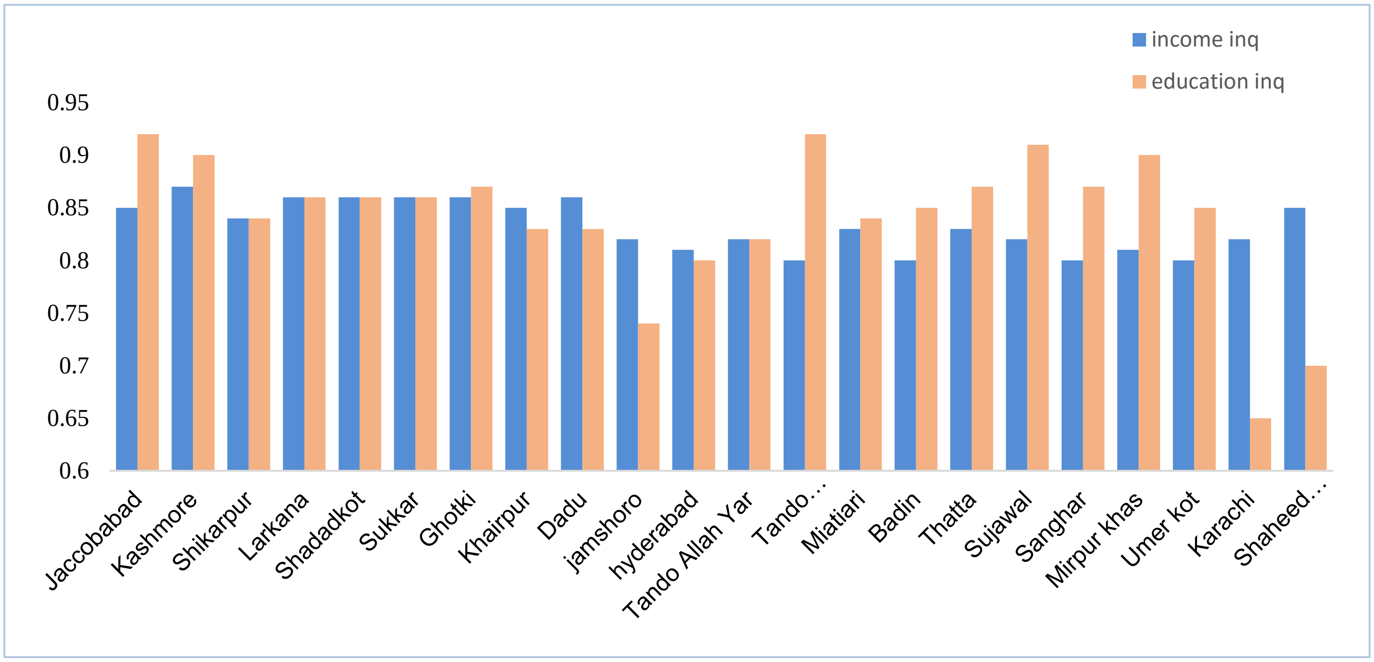


Chart 4.3 depicts the picture of income and educational inequality in the districts of Sindh. The estimated values of Gini coefficient show high level of income and educational inequality but the level of income and educational inequality either remained equal in some districts (i.e. Shikarpur, Larkana, Shadadkot, Sukkar and Tando Allah Yar) or the level of educational inequality is found higher than income inequality in most of the districts (i.e. Miatiari, Badin, Thatta, Sujawal, Sanghar, Mirpurkhas, Umerkot). It shows that there is high educational inequality in Sind province comparative to the income inequality.

Figure 4.4: Baluchistan Income & Educational Inequalities

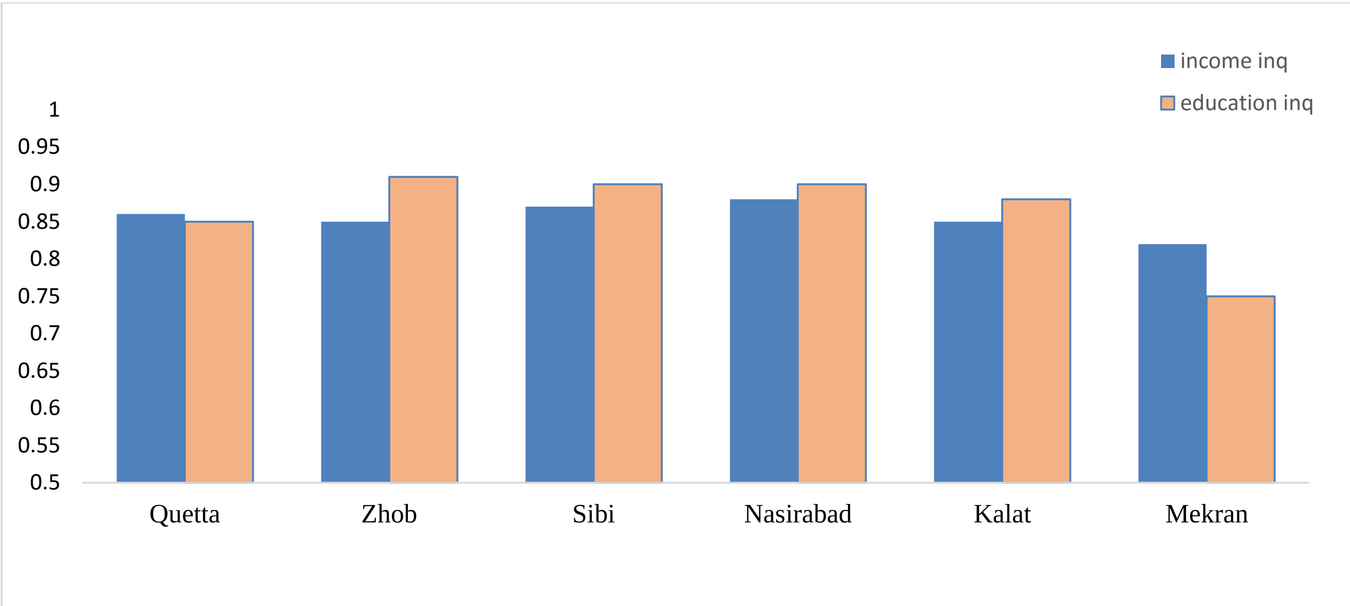


Chart 4.4 shows income and educational inequality estimates across the districts of Baluchistan. High level of income and educational inequality is found through Gini coefficient in all the districts. The ratio of educational inequality is higher than the income inequality in Zhob, Sibi, Nasirabad and Kalat districts. It is concluded that severe income and educational inequality exist in all the provinces however the intensity of income inequality is more in some districts and intensity of educational inequality in others districts. Income inequality is comparatively higher in the districts of KPK and Punjab while educational inequality is higher in Sind and Baluchistan relatively. In the second part of regression analysis regional income inequality model is estimated. The first analytical step includes in the estimation of the regional inequality model is applying OLS technique.

The second step of the analysis is detection of spatial autocorrelation in the residuals of the estimated OLS equation, and if the hypothesis of no spatial autocorrelation is rejected, then, in the next step it is tested that either the SLM or the SEM is more suited to describe the data. The results of summery statistics of the variables used in the model and OLS diagnostic test are shown in table 1 and 2 respectively.

Table 4.1: Summery Statistics

Provinces	Variables	Observation	Mean	S.D	Min	Max
Punjab	Inc Inq	37	0.85	0.01	0.83	0.89
	Edu Inq	37	0.76	0.06	0.65	0.9
	HHage	37	25.07	1.61	21.9	29.2
	HHsize	37	7.05	0.57	6	8
	Ind dum	37	0.04	0.03	0	0.13
	Sector	37	0.04	0.17	0	1
	GDPPC	37	7314.3	2679.14	3570.16	15589.1
	Schooling	37	2.71	0.89	1.07	4.99
KPK	Inc Inq	25	0.88	0.02	0.84	0.93
	Edu Inq	25	0.82	0.06	0.71	0.95
	HHage	25	22.9	1.94	19.8	27.3
	HHsize	25	9.16	1.21	7	11
	Ind dum	25	0.02	0.02	0	0.09
	Sector	25	0	0	0	0
	GDPPC	25	4243.5	1740.5	1354.2	8319.9
	Schooling	25	2.11	0.88	0.57	3.64
Sindh	Inc Inq	22	0.83	0.02	0.8	0.87
	Edu Inq	22	0.84	0.06	0.65	0.92
	HHage	22	23.1	1.61	20.63	26.18
	HHsize	22	7.8	1.56	6	11
	Ind dum	22	0.11	0.07	0	0.25
	Sector	22	0.32	0.47	0	1
	GDPPC	22	5367.3	2047.5	2708.15	11835.2
	Schooling	22	1.87	0.93	0.89	4.63
Baluchistan	Inc Inq	6	0.85	0.02	0.82	0.88
	Edu Inq	6	0.86	0.06	0.75	0.91
	HHage	6	21.95	1.47	20.7	24.6
	HHsize	6	10	2.44	6	12
	Ind dum	6	0.02	0.01	0.01	0.04
	Sector	6	0.83	0.41	0	1
	GDPPC	6	4812.4	2379.9	2837.14	9368.4
	Schooling	6	1.68	0.75	1.12	3.13

The table 1 shows the province wise summery statistics of dependent and independent variables included in the regional inequality model. Dependent and independent variables are kept in the 1st column of the table. Income inequality is the dependent variable while remaining all are independent variables. No of observations shows the no of districts included in the estimation from all the provinces of Pakistan i.e 37 from Punjab, 27 from KPK, 22 from Sindh and 6 from Baluchistan. Descriptive statistics shows the actual condition of each province. GDP percaptia is high in Punjab and Sindh. The ratio of income inequality is high in KPK then in Punjab and Baluchistan and lowest in Sindh while educational inequality is more in Baluchistan and Sindh relative to Punjab and KPK and educational attainment is more in Punjab and KPK than others.

Table 4.2: OLS Estimates

Income Inequality	Coefficients	T	P>|t|
Intercept	0.526	2.49	0.015
HHage	0.001	0.27	0.792
HHsize	0.003	1.33	0.188
Indusdmy	-0.251	-5.31	0.000
Sector	-0.024	-3.81	0.000
GDP Percapita	$-8.91e^{-07}$	-0.54	0.590
Schooling	-0.0241	2.01	0.044
Education inq	0.4600	2.53	0.013
Number of obs	90	R Square	0.573
F(9,80)	15.73	Prob>F	0.000

Table 2 displays the estimates of OLS regression. Except regional household age, household size and GDP percapita all the remaining variables are significantly affect the regional income inequality. schooling (regional educational attainment), Indusdmy (regional proportion of agricultural working population), sector (proportion of working population living in urban areas) are negatively and significantly affect the regional income inequality meaning that lower values of these variables associated with lower regional income inequality. While Eduinq (regional educational inequality) is positively and significantly affect the regional income inequality. Regional educational inequality is the main variable which is used to determine the regional income inequality. The high positive value of its coefficient depicted that it is the key determinant of regional income inequality.

Table: 4.3 Diagnostic tests for spatial dependence

Test	Stat. value	P-value
Moran's I (error)	4.786	0.000
LM lag	24.554	0.000
Robust LM lag	18.517	0.000
LM error	6.498	0.011
Robust LM error	0.460	0.498

Five diagnostic test statistics that used to detect the presence of spatial autocorrelation are shown in table 3. The first test is Moran's I statistics. As discussed earlier, the Moran's I statistic is considered a powerful diagnostic test for detection of spatial dependence, but it is less helpful in suggesting which alternative spatial model should be used. For this purpose, the LM test statistics are used as a best alternative model selection technique. The next two (LM Lag and Robust LM-Lag) test statistics in the table are refer to the SLM as an alternative, if these are more significant over their alternative error tests. The next two tests statistics pertain to the SEM as an alternative specification if they are found to be significant alternatively. The lower P-value of Moran's I statistics supported to the rejection of null hypothesis of no spatial Autocorrelation and indicated spatial dependence in the data. Here, P-values of both LM Lag and LM Error are significant but the p-value of LM Lag is comparatively smaller than the LM Error tests and P-value of robust LM Error is not significant relative to robust LM lag. So it is suggested that a spatial lag specification is best suited to the data and should be used in estimation.

Table: 4.4 Estimates of Spatial Lag Model & Spatial Error Model

	SLM			SEM		
Income Inq	**Coefficients**	**Z**	**P>|z|**	**Coefficients**	**z**	**P>|z|**
Intercept	-0.0520	-0.26	0.79	0.5210	2.77	0.006
HHage	0.0006	0.39	0.699	0.006	0.40	0.690
HHsize	0.0028	1.32	0.187	0.003	0.002	1.73
Indusdmy	-0.1448	-3.27	0.001	-0.117	-2.44	0.015
Sector	-0.0113	-1.92	0.055	-0.010	-1.70	0.088
GDPPC	-4.84e^{-07}	-0.35	0.728	-8.24 e^{-07}	-0.56	0.575
Schooling	-0.0197	2.01	0.04	-0.016	1.90	0.152
Edu Inq	0.3937	2.41	0.016	0.3214	1.69	0.091
Rho/lamda	0.7777	5.54	0.000	0.915	11.57	0.000
Observation	90	R-Square	0.73	R-Square	0.46	

Table 4 displays the results of SLM and SEM estimated with ML method. First column contains all the independent variables while coefficients values of these variables for SLM, their z and P-values are shown in second, third and fourth columns and same estimated results of SEM are shown in the last three columns respectively. The estimated results of SLM and SEM are almost same as OLS. All variables display with their expected signs of coefficient. Here regional educational inequality displayed as a key determinant of regional income inequality with high positive value in both models. Rho and Lamda are coefficients of spatial autocorrelation in the SLM and SEM both are positive with high coefficient value and significant with lower p-value exhibited the positive autocorrelation (spatial dependence) in the model. 74 percent cross regional variation is explained in regional income inequality through

SLM while only 46 percent through SEM. As the regional educational inequality and regional educational attainment both variables are significantly affected the regional income inequality with their expected signs. However the relative sizes of coefficients are very important. As shown in the table 2 and 4, the measure of educational inequality out-performed that of educational attainment in spatial and non-spatial models. Thus, distribution of education is more important in reducing regional income inequality, than a skewed educational attainment for the few of the populace.

Table: 4.5 Diagnostic tests for model specification

	Spatial lag Model		***Spatial Error Model***	
Null Hypothesis;	**Rho=0**		**Lambda=0**	
Test	***Stat. value***	**P-value**	***Stat. value***	**P-value**
Wald test	30.69	0.000	133.7	0.000
LR test	18.79	0.000	15.19	0.002
LM test	24.55	0.000	6.49	0.011

Three diagnostic tests; Wald test, Likelihood ratio test and Lagrange multiplier test are performed for model specification between SLM and SEM. The null hypothesis rho=0 is tested for Spatial lag dependence and Lambda=0 is tested for Spatial error dependence. The stats values and p-values of all three diagnostic tests are sufficient for the rejection of null hypothesis for both model specification but LR and LM test statistics are more significant with lower p-values for spatial lag specification against the spatial error specification while Wald test is equally significant for both model specification. In the light of above test statistics SLM is best fitted model for the spatial effect analysis of regional income and educational inequality.

CHAPTER 5
CONCLUSION AND RECOMMANDATIONS

5.1 Conclusion

To determine the spillover of educational inequality on regional income disparity in Pakistan, regional inequality model is developed in which dependent variable regional income inequality set to be equal regional household age, household size, educational attainment, educational inequality, general economic condition and industrial structural. Microeconomic cross-sectional data of 1, 59,000 of households from 90 districts of all provinces of Pakistan is included in the analysis. Spatial econometric technique is used to check the role of space in the model. While educational and income inequality is calculated by Gini index.

In the first analytical step, inequality model is estimated through OLS technique then for the detection of spatial dependence Moran'I and LM tests are applied on the estimated residuals of OLS. Both the diagnostic tests statistics confirmed the spatial dependence (spillover effect) in the model. SLM is specified over SEM due to more significant p-value of LM lag test comparative to LM error and insignificant p-value of Robust LM Error relative to Robust LM lag. For correct model specification further three diagnostic tests; Wald test, Likelihood ratio test and Lagrange Multiplier test are performed. LR and LM test statistics are more significant with lower p-values for spatial lag specification against the spatial error specification while Wald test is equally significant for both model specification. So it is concluded that SLM is best fitted model for the spatial effect analysis of regional income and educational inequality.

Results of the analysis showed that almost all the independent variables are significantly affect the regional income inequality with their expected positive or negative signs. The both main variables regional educational inequality and its attainment significantly affect the income inequality which means that higher level of educational inequality is associated with higher level of income inequality and higher educational attainment is associated with lower level of income inequality. But the relative sizes of their coefficients showed that educational inequality outer performed than the educational attainment in the model. It means fair and equitable distribution of education across the regions is more helpful in reducing regional income inequality than the higher educational attainment of few of populace.

The positive spatial effect demonstrated that the main cause of regional income inequality is the unequal distribution of education among the regions .If educational opportunities are not provided in a region, not only the people of that particular region remain uneducated and unskilled but the people of the linked regions which based on that region for their education also deprived to get education. These regions produce very low human capital that contribute less in the economy and get unsatisfactory portion of country's GDP, due to which these are faces the problem of regional income inequality.

5.2 Policy Implications

Misspecification of model may led to biased and inefficient estimators which cause the wrong policy decision and inaccurate future predictions. This research study has policy implication for researchers and econometricians that they must take into account the problem of identification when specifying best suited model to the data. The second implication of this study is for regional policy makers and development planners that they must consider the role of geography (spatial effect) especially in regional economic cross-sectional studies because ignorance of spatial interactions from neighboring regions also led to inaccurate estimates. And it is suggested that to narrowing down the regional income disparity government should focus on the more equitable distribution of education across all the regions along with its attainment at higher level.

References

Abdullah, A. J., Doucouliagos, H., & Manning, E. (2011). Education and income inequality: A meta-regression analysis. *Unpublished manuscript, Deakin University*.

Akhtar, S. (2008). Trends in regional inequalities in Pakistan: Evidence since 1998. *The Lahore Journal of Economics*, *13*, 205-220.

Amjad, Z. (2015). Trade and Income Distribution in Pakistan. *Global Journal of Management*

And Business Research.

Anselin, L. (2001). Spatial econometrics. *A companion to theoretical econometrics*, *310330*.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, *27*(3), 247-267.

Anselin, L. (2013). *Spatial econometrics: methods and models* (Vol. 4). Springer Science & Business Media.

Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food policy*, *26*(4), 315-331.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407-443.

Barufi, A. M., Haddad, E., & Paez, A. (2012). Infant mortality in Brazil, 1980-2000: A spatial panel data analysis. *BMC public health*, *12*(1), 181.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago press.

Bell, K. P., & Bockstael, N. E. (2000). Applying the generalized-moments estimation approach to spatial problems involving micro-level data. *Review of Economics and Statistics*, *82*(1), 72-82.

Bonet, J. (2006). Fiscal decentralization and regional income disparities: evidence from the Colombian experience. *The Annals of Regional Science*, *40*(3), 661-676.

Burki, A. A. (2011). *Exploring the links between inequality, polarization and poverty: empirical evidence from Pakistan* (No. 11-04). SANEI Working Paper Series.

Castelló‐Climent, A., & Doménech, R. (2008). Human capital inequality, life expectancy and economic growth. *The Economic Journal*, *118*(528), 653-677.

Chen, C. H., & Wu, H. L. (2005). Determinants of regional growth disparity in China's transitional economy. *Journal of Economic Studies*.

Crespo, N., Simões, N., & Diogo, A. P. (2012). Determinant factors of income inequality:

evidence from a Portuguese region. *Economics Bulletin*, (3), 2056-2064.

David, A., Guilbert, N., Hamaguchi, N., Higashi, Y., Hino, H., Leibbrandt, M., & Shifa, M.

(2018). Spatial poverty and inequality in South Africa: A municipality level analysis.

Farooq, M. (2010). Education and Income Inequality in Pakistan. *Dialogue (1819-6462)*, *5*(3).
Fingleton, B. (2000). Spatial econometrics, economic geography, dynamics and equilibrium: a 'third way'?. *Environment and planning A*, *32*(8), 1481-1498.

Fischer, Manfred M., and Claudia Stirböck. "Pan-European regional income growth and clubconvergence." *The Annals of Regional Science* 40, no. 4 (2006): 693-721.

Fleisher, B., Li, H., & Zhao, M. Q. (2010). Human capital, economic growth, and regional inequality in China. *Journal of development economics*, *92*(2), 215-231.

Gonul, D., & Erkut, G. (2019). Why do skilled people migrate to cities? A spatial econometric analysis for understanding the impact of the social environment on the attraction of human capital to cities in Turkey. *European Spatial Research and Policy*, *26*(1), 127-148.

Gonul, D., & Erkut, G. (2019). Why do skilled people migrate to cities? A spatial econometric analysis for understanding the impact of the social environment on the attraction of human capital to cities in Turkey. *European Spatial Research and Policy*, *26*(1), 127-148.

Gregorio, J. D., & Lee, J. W. (2002). Education and income inequality: new evidence from cross‐country data. *Review of income and wealth*, *48*(3), 395-416.

Hamid, A., Akram, N., & Shafiq, M. (2013). Inter and Intra Provincial Educational Disparities in Pakistan. *Pakistan Journal of Social Sciences (PJSS)*, *33*(2).

Hussain, A. (1993, March). Regional economic disparity in Pakistan and a framework for regional policy. In *Wilton Park Conference at Wiston House, Sussex, England (8–12 March). Accessed February* (Vol. 12, p. 2019).

Idrees, M. U. H. A. M. M. A. D., & Shah, A. N. W. A. R. (2018). An empirical analysis of educational inequalities in Pakistan. *Pakistan Economic and Social Review*, *56*(2), 313-324.

Jamal, H. (2016). Spatial disparities in socioeconomic development: the case of Pakistan. *Pakistan Development Review*, *55*(4), 421-435.

Kanwal, A., & Munir, K. (2015). The impact of educational and gender inequality on income

inequality in South Asia.
Kayam, S. S., Hisarciklilar, M., & Yabrukov, A. (2012). What causes the regional disparity of FDI in Russia?‖. *MPRA Paper*, (39151).

KHAN, M. T. Y., & SASAKI, K. (2003). Regional disparity in Pakistan's economy: Regional econometric analysis of causes and remedies. *Interdisciplinary information sciences*, *9*(2), 293308.

Khan, M. Z., Rahman, S., & Chaudhry, A. R. (2015). Education and income inequality in Pakistan'. *Management & Administrative Science Review*, *4*(1), 134-45.

Kyriacou, A. P., & Roca-Sagalés, O. (2012). The impact of EU structural funds on regional disparities within member states. *Environment and Planning C: Government and Policy*, *30*(2), 267-281.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
Li, Q., Song, J., Wang, E., Hu, H., Zhang, J., & Wang, Y. (2014). Economic growth and pollutant emissions in China: a spatial econometric analysis. *Stochastic environmental research and risk assessment*, *28*(2), 429-442.

Lucas, R. E. (1989). On the mechanics of economic development. *NBER Working Paper*, (R1176).

Lv, K., Yu, A., Gong, S., Wu, M., & Xu, X. (2017). Impacts of educational factors on economic growth in regions of China: a spatial econometric approach. *Technological and Economic Development of Economy*, *23*(6), 827-847.

Majeed, M. T. (2016). Economic growth and income inequality nexus: An empirical analysis for

Pakistan.

Ministry of Finance: *Pakistan* economic survey 2018-19, *10,193-194*
Naschold, F. (2009). Microeconomic determinants of income inequality in rural Pakistan. *The Journal of Development Studies*, *45*(5), 746-768.

Nawaz, S., & Iqbal, N. (2016). Education poverty in Pakistan: A spatial analysis at district level. *Indian Journal of Human Development*, *10*(2), 270-287.

Nebebe, A. F., & Rao, C. A. (2016). Determinants of income inequality in urban Ethiopia: A study of South Wollo Administrative Zone, Amhara National Regional State. *International Journal of Applied Research*, *2*(1), 550-563.

Nolan, B., Salverda, W., Checchi, D., Marx, I., McKnight, A., Tóth, I. G., & van de Werfhorst, H. G. (Eds.). (2014). *Changing inequalities and societal impacts in rich countries: thirty countries' experiences*. OUP Oxford.

Raza, Q., & Hina, H. (2016). Fiscal decentralisation, provincial economic growth and spillover effects: a spatial panel data analysis. *Pakistan Development Review*, *55*(4), 743-760.

Rodríguez-Pose, A., & Tselios, V. (2010). Inequalities in income and education and regional economic growth in western Europe. *The annals of regional science*, *44*(2), 349-375.

Shaheen, S., Awan, M. S., & Cheema, A. R. (2016). MEASURING INEQUALITY OF

OPPORTUNITY IN PAKISTAN. *Pakistan Economic and Social Review*, *54*(2), 165-190.
Siddiqi, A. H. (1981). Regional inequality in the development of Pakistan. *GeoJournal*, *5*(1), 17-

32.

Takahashi, K. (2007). Sources of regional income disparity in rural Vietnam: Oaxaca-Blinder decomposition. *IDE Discussion Papers*, *95*.

Tirado, D. A., Díez-Minguela, A., & Martinez-Galarraga, J. (2016). Regional inequality and economic development in Spain, 1860–2010. *Journal of Historical Geography*, *54*, 87-98.

Umar, H. M., Ismail, R., & Abdul-Hakim, R. (2013). Measuring the Regional Variations in

Educational Attainment and Inequality in Nigeria. *Abstract of Economic, Finance and Management Outlook*, *1*, 1-52.

Umar, H. M., Ismail, R., & Eam, L. H. (2014). A spatial econometrics analysis of educational distribution and regional income disparities in Nigeria. *Proceedings Book of ICETSR 2014, Malaysia Handbook on the Emerging Trends in Scientific Research*, 722-731.

UNDP: United Nation Development Report on Pakistan 2019.

UNESCO (2010). *Education for All – EFA Global Monitoring Report 2010: Reaching the Marginalized*, Oxford University Press, Oxford.

UNICEF (2013). *Out-of-School Children in the Balochistan, Khyber Pakhtunkhwa, Punjab and Sindh Provinces of Pakistan*, UNICEF Pakistan Country Office, Islamabad.

Ahmed, S. (2011). Does economic geography matter for Pakistan? A spatial exploratory analysis of income and education inequalities. *The Pakistan Development Review*, 929-952.

Ahmed, S. (2011). *Essays on Spatial Inequalities in Income and Education: Econometric Evidence from Pakistan* (Doctoral dissertation, University of Trento).

Iqbal, N., & Nawaz, S. (2017). Spatial Differences and Socioeconomic Determinants of Health Poverty. *The Pakistan Development Review*, *56*(3), 221-248.

Nawaz, S., & Mangla, I. U. The Economic Geography of Infrastructure in Asia: The Role of Institutions and Regional Integration.

Guliyev, H. (2020). Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial statistics*, 100443.

Wang, S., Chen, Y., Huang, J., Chen, N., & Lu, Y. (2019). Macrolevel Traffic Crash Analysis: A Spatial Econometric Model Approach. *Mathematical Problems in Engineering*, *2019*.

Anselin, L., & Bera, A. K. (1998). Introduction to spatial econometrics. *Handbook of applied economic statistics*, *237*.

Hossain, M. Ä Note on Model selection in statistics and Econometrics.

McMillen, D. P. (2003). Spatial autocorrelation or model misspecification?. *International Regional Science Review*, *26*(2), 208-217.

LeSage, J. P. (2008). An introduction to spatial econometrics. *Revue d'économie industrielle*, (123), 19-44.

Plümper, T., & Neumayer, E. (2010). Model specification in the analysis of spatial dependence. *European Journal of Political Research*, *49*(3), 418-442.

Osland, L. (2010). An application of spatial econometrics in relation to hedonic house price modeling. *Journal of Real Estate Research*, *32*(3), 289-320.

Song, C., Kwan, M. P., Song, W., & Zhu, J. (2017). A comparison between spatial econometric models and random forest for modeling fire occurrence. *Sustainability*, *9*(5), 819.

Raza, Q., & Hina, H. (2016). Fiscal decentralisation, provincial economic growth and spillover effects: a spatial panel data analysis. *The Pakistan Development Review*, 743-760.

Kivi, L. H. (2019). Spatial interactions of regional labour markets in Europe. *Available at SSRN 3330778*.

Elhorst, J. P. (2014). *Spatial econometrics: from cross-sectional data to spatial panels* (Vol. 479, p. 480). Heidelberg: Springer. Vega, S. H., & Elhorst, J. P. (2013, August). On spatial econometric models, spillover effects, and W. In *53rd ERSA Congress, Palermo, Italy*.

Goodchild, M. F., & Haining, R. P. (2004). GIS and spatial data analysis: Converging perspectives. *Papers in Regional Science*, *83*(1), 363-385.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, *27*(3), 247-267.

J. Bronnenberg, B. (2005). Spatial models in marketing research and practice. *Applied Stochastic Models in Business and Industry*, *21*(4-5), 335-343.

Florax, R. J., & Nijkamp, P. (2003). Misspecification in linear spatial regression models. *Tinbergen Institute Discussion Papers*, (2003-081), 3.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: models & applications*. Taylor & Francis.

Quah, D. (1996). "Regional convergence clusters in Europe." *European Economic Review*, 35: 951-958.

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

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