DECOMPOSITION OF INCOME INEQUALITY IN PAKISTAN:

A DISTRICT LEVEL ANALYSIS

Submitted by

Muhammad Arsalan Ali

Supervised by

Dr. Rehana Siddiqui

Professor/Head

Department of Environmental Economics

A dissertation presented to Pakistan Institute of Development Economics, Islamabad

In fulfullment of the requirment for the degree of Masters of Philosophy (M.Phil) in Economics

DEPARTMENT OF ECONOMICS

PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMIC ISLAMABAD

2018

Pakistan Institute of Development Economics

CERTIFICATE

This is to certify that this thesis entitled: **"Decomposition of Income Inequality in Pakistan: A District Level Analysis"** submitted by Mr. Muhammad Arsalan Ali is accepted in its present form by the Department of Economics, Pakistan Institute of Development Economics (PIDE), Islamabad as satisfying the requirements for partial fulfillment of the degree of **Master of Philosophy in Economics**.

External Examiner:

Dr. Abdul Sattar
Senior Assistant Professor
Department of Management Sciences
Bahria University, Islamabad

Supervisor:

Dr. Rehana Siddiqui
Head
Department of Environmental Economics
PIDE, Islamabad

Head, Department of Economics:

Dr. Attiya Y. Javid
Head/Professor
Department of Economics
PIDE, Islamabad

ALLAH
The Most Beneficent
The Most Merciful

"To Allah belongs whatever is in the heavens and whatever is in the earth. Whether you show what is within yourselves or conceal it, Allah will bring you to account for it. Then He will forgive whom He wills and punish whom He wills, and Allah is over all things competent."

(Al-Baqarah, 2:284)

ACKNOWLEDGEMENTS

Everything in the World starts with the name of ALLAH Almighty. All praises for Him, whatever I am today and what I will be in future is all due to His blessings on me. I thank ALLAH for the strength and guidance that helped me to complete the thesis.

I thank all my teachers, parents, siblings, class fellows, university fellows and the staff at PIDE (Muhammad Shahzad Afsar) who helped and supported me. I was a raw person before this thesis, never had done research. It was the kind generosity and hard work of my supervisor Madam Rehana Siddiqui that made me a researcher. Any question that I had in mind, I asked Dr. Rehana Siddiqui and he readily gave me answers. In the first and second semester, the brilliance of Dr. Wasim Shahid Malik and Dr. Musleh-ud-din Dr. Athar Maqsood and Dr. Rehana Siddiqui motivated and accelerated our enthusiasm in economics.

My class fellows and university fellows were a great help. They never let me down and whenever I felt down they pushed me to raise myself. Sidra Saddiq, Muhammad Taimoor, Qasim Raza, Muhammad Afzal Khan were a constant and uneding moral support. I thank all my class fellows and may ALLAH shower His blessings on them.

The care, the affection and the support that family provides plays a vital role. I am honred to have such parents who always supported my education both financially and morally.

M ARSALAN ALI

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

In this study, an attempt has been made to decompose income inequality in districts of four provinces of Pakistan with respect to rural and urban segments. For this purpose, Gini coefficient is estimated at district level by using data of Pakistan Social and Living Standard Measurement Survey (PSLM) for the years 2004-05 and 2014-15. The findings reveal that there is significant overtime change in income inequality across districts especially in Swat, Bonair, Gwader and Jhal Magsi. Income inequality in Swat, Bonair,and Gwader has decreased while in Jhal Magsi income inequality has been increased significantly. A regression based decomposition methodology, proposed by Field (2003) has been used to examine the contribution of variables like household size, gender of household head, age of household head, education of household head, marital status of household head, employment status of household head and region. The findings reveal that district level income inequality among all four provinces of Pakistan has not changed significantly except few districts which have been mentioned above. In addition high concentration of income inequality is observed in urban areas as compared to rural areas. While within group, inequality is higher as compared to inequality between the groups in all provinces. It has been also observed that, from 2004-05 to 2014-15 inequality within the groups has been increased whereas between the groups inequality has been decreased. It shows that inequality gap between rural and urban areas has decreased while inequality within the groups of urban and rural areas has increased. The variables which contributed more in inequality are education of household head followed by household size and region.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Income and asset inequality has increased both in developed and developing economies and Pakistan is no exception. Issue is not confined to some specific coutries as pointed out by Milanovic (2005) that income inequality has increased in the world. Study of inequality is important to understand welfare and well being of a nation. Seminal work on inequality measurment was done by Atkinson (1970) who has linked it with welfare distribution. Associating issue of subgroup inequality to overall inequality has been discussed by Bourguignon (1979), Cowell (1980), Shorrocks (1980) and Shorrocks (1984).

Pakistan has low income per capita with hight income inequalities (Jamal, 2014). To understand this inequality we have to decompose it to identify the key factor leading to disparity. Decomposition analysis can be divided into two approaches. The first approach relates to the decomposition of income between subgroups of the population. It shows that total inequality is sum of "within groups" and "between groups" inequalities (Shorrocks, 1980; Bourguignon, 1979). The second approach relates to the decomposition of income of households or individuals into different factors contributing to total inequality (Shorrocks, 1982).

As mentioned earlier, inequality is major concern in developed and developing countries and Pakistan is no exception. Studies show that income inequality has continuously changed for both urban and rural areas and overall in Pakistan. Income distribution improved in Pakistan during the period of sixties and eighties but has

worsened in the seventies and nineties [GOP, (2000); Anwar (2004-05); Burki *et al*, (2010); Kemal, (2007)]. Gini coefficient for urban, rural, and overall is 0.40, 0.30 and 0.35 respectively for the year 1987-88 which has increased up to 0.41, 0.38 and 0.41 for the year 2010-11 [Jamal, (2014); Jamal, (2009)].

Most of the existing research on income inequality is conducted either for Pakistan or at provincial level. These studies neglect the distribution of income at district level.[1] The present study is an attempt to fill this gap as it explores the distribution of income across the districts of Pakistan. However because of socioecoomic diversity across districts it is important to examine inequality at district level. In this study, we shall examine the income inequality across districts of Pakistan by using two data sets of Pakistan Social and Living Standard Meassurment Survey survey and check either income inequality across districts of Pakistan has increased or decreased. We restrict our analysis to those districts included in PSLM survey 2004-05. Siddiqui, (2008) claimed that public provision of social services like education play a significant role in development of capability that helps to reduce income inequality. The present study used Gini coefficient to estimate income inequality. Sine inequality is not expectd to change significanlty every year we are using the year of 2004-05 and 2014-15. [2]

1.2. Problem Statement

The existing literature relating to income inequality in Pakistan does not study the problem of income inequality across the districts of Pakistan. Because in literature most of the studies have used data of HIES where a limited number districts are covered.

[1] Exclusion include Jamal and Khan (2003a, 2003b), Naqvi (2007), Siddiqui (2008), Jamal and Khan (2008a, 2008b) , Arif et al (2010), Ahmed (2011) and Burqi et al,(2010).

[2] Inequality may cahnge rapidly din case of shock/disasters like earthquake and floods.

To study income distribution across districts is important because of great variation in standard of living can be observed across different districts and provinces of Pakistan. All the provinces comprise districts of high and low income per capita districts in Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan (Ahmed, 2011). The gap varies across provinces. So the purpose of this study is to identify the factor which cause income inequality across the districts. This will enable us to identify the districts level information needed and also the type of intervention required.

1.3. Objectives of Study

Broadly, the study has four objectives.

1 Examine the income inequlaity, across the districts of Pakistan.

2 Examine household income inequality between and within group (urban and rural areas).

3 Estimate earning function for each province of Pakistan to identify factors responsible for income inequality.

4 Examine how the income inequality has changed overtime and which factors are responsible for the change in income inequality overtime across provinces?

1.4. Significance of the Study

Based on data of Pakistan Social and Living Standard Measurment Survey 2004-05 and 2014-15 this study provides valuable information about the districts of Pakistan. We also examine the income inequality between and within the group of each province. The study is expected to help the course of policy for reduction of inequality. Information of factors which affect income inequality will help policy makers to make suitable policies to reduce inequality across provinces of Pakistan.

1.5. Organization of the Study

This study is organized in five chapters. Chapter one contains background, problem statement, objectives and significance of the study. Chapter two consists of review of existing literature on inequality. Chapter three deals with empirical methodology and data used for this study. Results are discussed in chapter four. Conclusions and Policy recommendations are in the final chapter.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews some relevant studies available on decomposition of income inequality. This section is divided in two subsections. First subsection briefly explains decomposition by population subgroup. Second subsection briefly summarizes the decomposition of inequality by factor component.

Pioneers work done by Oaxaca (1973) and Blinder (1973) measuring the difference in mean income between group by using Regression based inequality decomposition. Juhn et al. (1993) utilize this approach to permit decomposing differences in the full wage distribution between groups rather than mean income. DiNardo, Fortin and Lemiux (1996) & Deaton (1997) proposed semi-parametric and non-parametric proficiency that sought to model and compare the entire income distribution in terms of density functions. Fields (2003) developed regression based decomposition to examine inequality in labor earnings by using variable such as gender, experience, race, schooling, industry, occupation, and region in United States. This technique is based on semi logarithmic income generating function. Field demonstrates further relative factor inequality weight. Further percentage contribution of variable would be same either any measure of inequality used. Finding of this study showed that contribution of schooling is more in inequality followed by occupation, experience and gender.

In Pakistan, most of the studies have used Lorenz Curve, Gini Coefficient, Theil index and coefficient of variation to measure inequality. Kruijk (1987), Ahmed and Ludlow (1989), Adam (1994),), Haq (1998), Ali and Tahir (1999), Ahmad (2001), Naschold (2009), Idrees and Ahmad (2010), Farooq (2010), used Gini-coefficient for

their estimates and some studies also reported Theil's index for Pakistan. Most of the studies have used Household Integrated Economic Survey data while some of other used Household Income and Expenditure Survey for their analysis. A considerable amount of literature has been published to analyze consumption expenditure and income inequalities. Some studies discussed separate income distribution for urban and rural areas while others considered inequality at district level and Pakistan as a whole. Some of these studies described inequality across individuals while others accounted income inequality across household.

2.1. Decomposition of Income Inequality by Subgroups

In this section we are explaining the existing literature regarding inequality decomposition by subgroup. Decomposition by population subgroup provides the information of which groups are affected more by inequality. This approach starts by separating a sample into discrete categories such as urban and rural, individual having primary vs higher education, male or female. After that it calculates the inequality that exist between sub samples and within subsample. This approach is applied by Shorrocks (1983), Silber (1989), Jenkins (1995), and Cowell Jenkins (1995).General thought about inequality is that it may be more noticeable in urban areas as compared to rural areas. Urban areas are generally developed due to more sectoral diversity. Different sectors acquire different skills which cause difference in wages among workers.

Anwar (2004-05); Jamal (2009) and Burki *et al* (2010) claimed that inequality increased in early nineties than declined afterwards. Farooq, (2010); Tahir *et al*, (1999), Anwar (2004-05) highlighted that inequality in urban areas is high as compared to rural areas. Using HIES survey data, Idress and Ahmad (2010) & Haq (1998) estimated the

inequaity of household consumption expenditure in rural and urban areas of Pakistan and found that inequality in non-food consumption is more than in food consumption. Ahmed and Ludlow (1989) estimated household income and expenditure inequality in Pakistan by using Micro Nutrient Survey 1967 and Household Income and Expenditure Survey (HIES) for 1979 and 1984. The study used five measure of inequality i.e. Gini coefficient, coefficient of variation, Atkinson index, logarithmic variance, and Lorenz curve. The Gini coefficient for households per capita increased from 0.363 in 1979 to 0.371 in 1884-85. The estimated Coefficient of Variation reflected the presence of very high incomes in the rural areas of NWFP, Sindh and Baluchistan provinces.

Ahmad (2001) explored the income inequality by occupational groups using household per capita income. The study found that inequality among skilled workers is higher relative to other professionals. Siddiqui and Siddiqui (1998), Ali and Akhter (2014-15) estimated the earning gap between male and female workers using earning function approach. They found that return to schooling is higher for female as compared to male. Individual characteristic like education, marital status, employment status, industrial choice, and occupation are the main factor that contribute to earning differential between male and female workers in Pakistan. Age groups belonging to the two extremes (the youngest and the oldest) have the highest influence in explaining Inequality. Inequality and education not only increase income but also have a major contribution in personal earning inequality [Nasir and Mahmood (1998)].

In the context of district level of inequality, Ahmed (2011) performed an exploratory analysis of socio economic disparities in Pakistan. The study found that Punjab comprises the largest group of high income per capita districts; district wise income inequality level showed weak but education level expose high spatial link.

Azam and Bhatt (2016) explored spatial income inequality at district level in India. Using Field (2003) methodology it was found that differences across district in rural India are observed due to difference in mean income while in urban India differences are observed due to within state factor component.

Skoufias and Olivieri (2013) investigated the difference in living standard among 471 districts in Indonesia by using household consumption survey data for the year 2009. This study simply used welfare ratio to measure the welfare among districts of Indonesia. Primary explanation of differences in welfare among district is observed due to difference in return (marginal welfare gain) of household mobility characteristic. In contrast, similar analysis is done by Skoufias and Olivieri (2013) in Thailand found that difference in welfare in urban and rural areas observed due to higher level of endowment (education and health facilities, access to infrastructure) in urban areas.

2.2. Determinants of Income Inequality.

In this section we explain the existing literature regarding inequality decomposition by factor component. A second approach used in literature is inequality decomposition by factor components to know the contribution of each factor into overall inequality. Shorrocks (1983) established what proportion of total inequality is attributable to various sources of income in United States. This study has used data of distribution of net family income between 1968 and 1977. This study decomposed the following sources of income, taxes, transfer income, wage earning, and capital earning. Finding of this study showed that labor income had highest contribution to inequality followed by capital earning. Transfer income and tax payment contributed negatively to inequality. For instance incomes generated from non-farm activities have an equalizing effect in following studies by Adam (1994) for Pakistan, Adam (2001) for

Egypt, Luo and Zhu (2006) for China, El-Osta et al. (1995) for the United States, Leones and Feldman (1998) for the Philippines. On the other hand, income generated from non-farm activities contributed positively to inequality in Ecuador by Elbers and Lanjouw (2001). On the contrary, study conducted by Canararajah et.al. (2001) in Uganda and Ghana found that non-farm self-employment income has contributed more in disequalizing the inequality than non-farm wages.

However, there are several studies that have attempted to identify factors that cause income inequality. For instance, Kruijk (1987) observed that occupation choice of labor is major source of income inequality. In contrast, Kruijk and Leeuwen (1985) concluded that shift of household from rural to urban, increase the labor market participation of household increase earning in rural areas, is major source of income inequality in rural and urban areas of Pakistan. Tripathi (2016) examined that household size is the major factor of consumption expenditure in rural and urban areas of India, but the finding of this study contrast with the study of Pandey (2014-15) that education of household head is the main factor that contribute inequality in household consumption expenditure.

Okatch et al, (2013) examined the determinants of income inequality in Botswana. The study used data of HIES 2002-03 by using Field methodology (2003). To examine the income inequality author comprised the variable into three types. Variable which effect more in income inequality are secondary school of education, number of paid employees ,value added tax, and dependency ratio of number of children in the household. Morduch and Sicular (2002) examined the income inequality in rural china by using period of four year. This study used regression based method of inequality and further decompose indices of Gini and Thiel. Theil index shows that

demographic variables and human capital reduce inequality. Alternative demographic and human capital shows the positive contribution to inequality which shows by Gini index. Sologon et al (2016) decomposed cross national difference in welfare inequality in Egypt and Tunisia by using Tunisia as a reference country. Result of this study showed that inequality among household expenditure exists due to education, sector of employment and working status. The study also showed that counter factual distribution of demographics and expenditure structure between Tunisia and Egypt are inequality decreasing while labor market has limited effect on inequality.

Ali et al. (2013), Su and Heshmati (2013) inspected that education and occupation are major determinants of income inequality. In addition Nazli, (2004-05); Farooq, (2010); Karunaratne, (2000); (Epo, 2011) observed that education positively affect the distribution of income. Nazli (2004-05) also found that interactive term of education and experience has positive impact on earnings and effect of education on earning is greater than effect of experience. Shabbir (1994) estimated the Mincerian earnings function to address the issue of inequality in the personal earning by using Population, Labor Force, and Migration survey data for year 1979 and found that human capital explained more variation in personal earning as compared to schooling.

Furthermore there are variety of studies conducted at rural areas which showed that land owner ship is major factor that explain income inequality, [Naschold, (2009); Arayama et al. (2006); Anwar et al. (2004-05); Jamal, (2009)]. Wan and Zhou, (2004-05) examined income inequality in rural china finding of this study showed that geographical condition is important variable that contribute in income inequality. Heshmanti (2004-05 and 2006) acknowledged that inequality have many dimensions.

Economists have long been concerned specifically with monetary dimensions related to household or individual incomes.

2.3 Conclusions

We have shown that the literature on inequality in Pakistan can be divided in two parts i.e decomposition by population subgroups (male and female, household head with primary vs secondary schooling or higher education, rural and urban household etc.) and decomposition of income sources (occupations i.e. senior official, manager, professional, technician, self-employment and wage income etc.). Some studies revealed that education and occupation types are major factors of inequality while other showed that household size is main feature of income inequality. Moreover inequality is higher in urban areas as compared to rural areas. In rural areas physical asset i.e. land ownership is major source of income inequality and similarly high inequality observed in non-food consumption expenditure as compared to food consumption expenditure. Some studies showed that return to education is higher for female as compared to males and other showed that individual characteristic like education, marital-status, employment status, industrial choice and occupation are the main factors that contribute to earning gap between males and females. However the literature lack dissegregated distric level analysis for Pakistan. Given the devolution plans the district specific intervention will be helpful to improve the welfare of population.

This study would fill the gap in existing literature in the following ways. First this study using districts as a geographical unit, we calculate the change in income inequality across the districts of Pakistan for two year. We also examine overtime which factors contribute more in inequality. Second we also provide the assessment of components which contribute in inequality by using Field (2003) methodology. Third we further, decompose the population by group to examine the inequality between and within group.

CHAPTER 3

DATA AND METHODOLOGY

3.1 Data

To estimate income inequality at household level data has been taken from PSLM survey for two years i.e 2004-05 and 2014-15. The survey has been conducted by Pakistan Bureau of Statistics (PBS). Information of various social and economic indicators at district level is available in PSLM surveys. We are using social and economic indicators for 74420 households for the year 2004-05 and 78635 households for the year 2014-15. PSLM survey covers information of households on education, occupation, employment status, income, ownership of assets, household details, marital status, immunization, facilities and services.

Each city town is subdivided into enumeration blocks under the framework of PSLM. Urban areas comprised of 26698 blocks and rural areas consist of 50588 blocks. Enumeration blocks in urban and rural areas have been taken as Primary Sampling Unit (PSU). Sample size of PSU has been selected form each strata/ sub-strata by using probability proportional to size sampling technique. Households in which each PSU is selected by using sample technique of random start, have been taken as Secondary Sampling Unit (SSU). Data from rural and urban areas of Primary Sample Units (PSU) and Secondary Sampling Units (SSU) of each province are discussed in Table 3.1.

Table 3.1 provides us complete information of Primary Sampling Units (PSUs) and Secondary Sampling Units (SSUs). We have shown that there are 5204 Primary Sampling Units (PSUs) in 2004-05 out of which 2993 are rural and 2211 are urban. Similarly for year 2014-15 there are 5326 Primary Sampling Units (PSUs) out of which 4116 are rural and 1210 are urban. Household total sample is 74420 in 2004-05 and

78635 for the year 2014-15. This sample size has provided sufficient information of different household to calculate income inequality at district level.

Urban areas which have population of 500,000 and more are considered as independent stratum. These metropolitan cities consist of Islamabad, Faisalabad, Lahore, Multan, Sargodha, Rawalpindi, Bahawalpur, Sialkot, Gujranwala, Karachi, Hyderabad, Sukkur, Peshawar and Quetta. Each metropolitan city was sub divided into PSU comprises of 200 and 250 households. Each urban city has been classified into low middle and high income groups. In the Table 3.1 we show that total sample of the year 2004-05, 35 percent is urban population and remaining is rural population. Similarly for the year 2014-15 total sample consists of 17 percent of urban population and 83 percent of rural population. We have used 34 districts of Punjab, 16 districts of Sindh, 24 of Khyber-Pakhtunkhwa, and 26 for Baluchistan for our analysis. Detail of districts is given in table A1 in appendix.

Table 3.1: PSLM Profile Survey 2004-05 and 2014-15.

Province	Sample PSUs			Sample SSUs		
	Urban	Rural	Total	Urban	Rural	Total
			2014-15			
Punjab	594	1860	2454	6814	29188	36002
Sindh	375	901	1276	4399	14336	18735
Khyber Pakhtunkhwa	104	764	868	1184	11898	13082
Baluchistan	110	572	682	1276	8971	10247
Islamabad	27	19	46	292	277	569
Total	**1210**	**4116**	**5326**	**13965**	**64670**	**78635**
			2004-05			
Punjab	1086	1182	2268	13032	18912	31944
Sindh	642	684	1326	7704	10944	18648
Khyber Pakhtunkhwa	258	591	849	3096	9456	12552
Baluchistan	195	521	716	2340	8336	10676
Islamabad	30	15	45	360	240	600
Total	**2211**	**2993**	**5204**	**26532**	**47888**	**74420**

Source: Pakistan Bureau of Statistic (2004-05 and 2014-15).

We have selected data of two PSLM surveys, which provide the information of different indicators in household roster which includes, age, gender, marital status, educational status, family detail and per capita income. Our unit of analysis is household. Households who did not provide information of income have been dropped. Total sample of household covered is 74420 in year 2004-05 and for the year 2014-15 is 78635.

After careful analysis we observed that 25% of households have not provided the information of income. When we have removed these households from total sample then our total sample is 54702 household for the year 2004-05. Similarly this same procedure has been done for the data of 2014-15 which has total sample of 78635 household. We observed that 32% of households have not provided the information of their income. After removing these household form total sample our remaining total sample is 53000 households. This total sample is sufficient to estimate the income inequality across districts of Pakistan.

3.2 Methodological Framework

3.2.1 Theoretical Model

This study uses the framework provided by human capital approach which states that acquisition and productive capacity of individuals are enhanced when people invest more in human capital in the desire of earning more in the future. This study relies on analytical framework acquaint by Becker (1962) for investment in education and analysis of different panorama of human capital. According to this approach individuals and household invest in education in order to build more human capital with the objective of obtaining higher benefit in the form of higher earnings. The basic theme of the human capital approach is that productivity of an individual's increases when investing in human capital through education. Apart from education others factors such

as occupation, job status, industry, literacy, region, marital status and province may also have a direct colligate with individual earnings. Wages of labor are affiliated with labor productivity when productivity increases rewards are pay to labor in the form of high wages. In human capital individuals continue to invest until marginal reverts equal the marginal cost. The criterion human capital model is stipulated as:

$$lny_j = \beta_o + \sum_{k=1}^{K} \beta_k X_{kj} + \varepsilon_j \ldots \quad 6)$$

Where, lny_j is denotes the household j natural logarithm of per capita income, and independent variable X_{kj} is a row matrix refers to specific k characteristic of jth household which includes the household demographic characteristic, characteristic of household head such as age, gender, occupation, education, household main income source, job status, and indicator variable for districts. Where ε_i is an error term which distributed normally having property of zero mean, constant variance and no autocorrelation.

3.2.2 Measurement and Decomposition of Gini Index

A number of techniques on measurement of inequality have been proposed in literature. But each has some pros and cons. Apparently a technique would be attributed to be a good if following properties are fulfilled: a) decomposability (decomposition between and within the group is possible), b) Pigou Dalton transfer principal (transfer of income from some poor to rich will increase inequality), c) symmetry (changing the place in distribution will not affect inequality), d) Population homogeneity (inequality will not change if there is same increase or decrease in income of each group of population e) income independence (same increase or decrease in income would not affect income inequality.

Among all of the measures Gini coefficient is the one that satisfies all of the properties mentioned above. Gini coefficient will be used in this study as it is easily decomposable and interpretable. Its value lies between 0 and 1. Value of 1 depicts the perfect inequality in society while 0 shows perfect equality in society. However there is one problem with Gini coefficient that it assigns more weights to income transfer affecting middle income classes as compared to extreme income classes. Geometric approach is the most common approach to define Gini coefficient. According to this approach this can be define as the ratio between the area between the line of perfect equality and Lorenz curve to the total area under the perfect equality line. Rao (1969) presented the formula to calculate Gini coefficient based on geometric approach is:

$$G = \sum_{i=1}^{n-1} P_i Q_{i+1} - P_{i+1} Q_i \qquad 1)$$

Where P_i is the cumulative population share and Q_i is the cumulative income share corresponding to the ith household when all households are arranged in ascending order of income. Decomposition of inequality index examine the structure of inequality. By doing this we will calculate the contribution of each group to total inequality. After decomposing inequality index a consistent relation between overall inequality and its attributes is needed. When dealing with decomposability we must be able to differentiate between and within group inequality. Within group inequality capture the inequality that exists due to income variability. On the other hand between group inequality capture the inequality that exist due to income variability across different groups. Decomposition of inequality index can be written as within and between element.

$$I = I_{within} + I_{between} \qquad 1)$$

We decompose further Gini index into between and within group. Before decomposition of Gini index it is worth noting that it is not perfectly decomposable. Besides between and within group inequality it has a zero residual R.[3] Let's assume we have two groups of individuals living in urban and rural areas. Gini index G_{within} is given by the formula:

$$G_{within} = \left[\frac{n_R}{n}\frac{y_R}{y}\right] G_R + \left[\frac{n_U}{n}\frac{y_U}{y}\right] G_U \qquad 2)$$

Where G_R and G_U are the Gini indexes which are measured on the basis of rural and urban incomes. Whereas the square brackets shows the weights which are given to each group. The weight in the square bracket is the product of two terms. First term shows the population share and second terms shows the income share of each group. Gini indexes will be calculated separately for urban and rural areas. After that these indices multiply with the corresponding weights of each urban and rural areas, to estimate within group inequality.

Gini index between $G_{between}$ is calculated in the instance of variance. Where we replaced actual incomes by subgroup mean incomes. $G_{between}$ is calculated by using the formula of covariance:

$$G_{between} = \frac{2}{\bar{y}} Cov\left(\bar{y}, F(\bar{y})\right) \qquad 3)$$

Where $\bar{y}$ is the subgroup mean income, $F(\bar{y})$ shows the fractional rank of subgroup mean income. The residual term R adds to general decomposition of Gini index. Overall Gini index is given by the formula:

$$G = G_{within} + G_{between} + R \qquad 4)$$

[3] Residual term explain the overlap characteristic of household (income) in urban and rural areas because this study measure inequality on the basis of income. If income of urban and rural areas does not overlap it mean that residual term equal to zero.

$$G = \left[\frac{n_R}{n}\frac{y_R}{y}\right]G_R + \left[\frac{n_U}{n}\frac{y_U}{y}\right]G_U + \frac{2}{\bar{y}}Cov\left(\bar{y}, F(\bar{y})\right) + R \qquad 5)$$

The residual term R is not very intuitive. However understanding of residual term is necessary while decomposing Gini index. If residual term is equal to zero then Gini index is perfectly decomposable.[4]

Step 1 Regression based decomposition of Income Inequality

Regression based decomposition analysis is used to measure the contribution of characteristics of household in inequality. In addition, we can evaluate the magnitude of these characteristics have contribution to the overtime change in inequality. Cowell and Fiorio (2011) showed that there is some sort of association between the sum of the estimated inequality shares of indicator variable for geographical units used in the regression framework and between regional component of decomposition addressed in section 3.1. Gini index is used to measure inequality that permit us to check the robustness of our finding of inequality in Pakistan.

By using Field (2003) methodology we first estimate the log linear income generating process by using Ordinary Least Square. Because factors which we have used to examine the income inequality are linearly related with the income of household

$$lny_j = \beta_o + \sum_{k=1}^{K}\beta_k X_{kj} + \varepsilon_j.. \qquad 6)$$

[4] when ranking incomees of subgroup from poorest to the richest individual then income do not overlap i.e. if the total income distribution of each individual in relative position is same. When ranking by subgroup incomes overlaps then residual term R is positive i.e. when the position of individual in total income distribution is different from the relative position of a given individual in the subgroup income distribution.

Where, lny_j denotes the natural logarithm of household jth per capita income, and independent variable X_{kj} is a row matrix refers to specific k characteristic of j[th] household, which includes the household demographic characteristic, such as age, gender, occupation, education, household main income source, job status, and indicator variable for districts. Where ε_j is an error term which can be explained the part of variation in income by those variable which we have not incorporated in earning equation. Error term is normally distributed having property of zero mean, constant variance and no autocorrelation. The income generating equation is of semi-log specification and instigate by the outcome that the income variable can be approached fit by using a log-normal distribution (Shorrocks and Wan, 2004-05).

Step 2 Factor Influence Inequality

In the second step we find the percentage contribution of these variable mentioned earlier to level of inequality. For this purpose we have used the estimated coefficient from Ordinary Least Square regression. These coefficients are also known as factor inequality weights, $s_k(lny)$:

$$s_k(lny) = \ cov\ (\ \beta_k X_k, lny\ )/\sigma^2\ (lny)$$

$$s_k(lny) = \ \beta_k * \sigma(X_k) * cor\ (X_k, lny)/\sigma(lny)\ ^5 \qquad (7)$$

Where β_k characterizes the estimated coefficient of k^{th} characteristic of household obtained from OLS regression, and X_k characterizes the value taken on by the k^{th} characteristic. $\sigma(lny)$ and $\sigma(X_k)$ are the standard deviation of lny andX_k, respectively and $cor\ (X_k, lny)$ is the correlation between factor k and lny. Therefore,

[5] This formula permit us to decompose any inequality measure, while the decomposition rules are comply (Shorrocks, 1982) if model is log linear, as the participation rate of the variable remain the same to the inequality. Versatile all inequality indices including the Gini coefficient, coefficient of variation, generalize entropy family, Atkinson index, and the various centile measure are the inequality that satisfying these condition.

$s_k(lny)$ shows the share of k^{th} characteristic in inequality index (Gini coefficient), due to the fact that X_k is distributed unequally among households.

The positive s_k indicates that factor k increases inequality. Whereas negative s_k implies that k is an inequality decreasing factor. When s_k and s_ε are added together $\sum s_k + s_\varepsilon = 1$. Where s_ε includes those variable which are not included in regression equation ,s_ε is inequality that rises when variables are omitted from the regression equation. $\sum s_k = R^2$ explains the explanatory power of k[th] household characteristics that explained proportion of income inequality. Equation (8) clarifies that the factor inequality weight will be large if i) there is high correlation between per capita income and household k^{th} characteristic; ii) β_k is large i.e. there is a large return of household X_k characteristic; iii) Household X_k characteristics fluctuate highly relative to yearly income ; iv) Household characteristic X_k fluctuate due to large provision of social services.

Step 3: Change in Inequality overtime 2004-05 and 2014-15.

In the final stage, we used factor inequality weights to observe the overtime change in Gini coefficient due to modification of houseohld characteristics. For this purpose we required at least two comparable household survey that are done in different time periods. We have used Pakistan Social and Living Standard Measurement Survey for two time period of 2004-05 and 2014-15. Earning equation is estimated separately for two time period. Similarly inequality index is calculated separately for two time periods. Let $I_{2004-05}$ and $I_{2014-15}$ represent any measure of income inequality , $s_k^{2004-05}$and $s_k^{2014-15}$ represent the relative factor inequality weight of k^{th} characteristic of households in the time period 2004-05 and 2014-15. Then the change

in inequality can be explained in terms of factor inequality weight and inequality index of each period.

$$I_{2014} - I_{2004} = \sum_{k=1}^{K} \left(s_k^{2014} I_{2014} - s_k^{2004} I_{2004}\right) + \left(s_\varepsilon^{2014} I_{2014} - s_\varepsilon^{2004} I_{2004}\right) \qquad 8)$$

Therefore, to derive percentage contribution household k^{th} characteristic to the difference in any measure of income inequality between 2004-05 and 2014-15 is as follows:[6]

$$\Delta I_k = \frac{s_k^{2014-15} * I_{2014-15} - s_k^{2004-05} * I_{2004-05}}{I_{2014-15} - I_{2004-05}}, where \sum_k \Delta I_k + \Delta I_\varepsilon = 1 \qquad 9)$$

The decomposition of change in inequality depends on the measure of inequality which we used in the analysis. Therefore the weight of factor inequality in the decomposition of changes in inequality also depends on choice of measure of income inequality that we used. Field (2003) technique has some advantages such as contribution of each individual in inequality remains the same. It does not matter which measure of inequality we use. Moreover this technique also manages the endogeneity problem.

3.3 Econometric Methodology

Following log-linear model is used to estimate which factors contribute more in income inequality.

$$lny_i = \beta_0 + \beta_1 gen + \beta_2 age + \beta_4 m\,stat + \beta_5 edu + \beta_6\, emp\ status + \beta_7 hhsize + \beta_8 region + \varepsilon_i \qquad 10)$$

[6] Field (2003) contends that inequality index I(Y) should be symmetric, continuous and $I(\mu, \mu, \mu, .., \mu) = 0$ where μ is mean income. Virtually all inequality indices including the Atkinson index, coefficient of variation, generalize entropy family, Gini coefficient, and the various centile measures satisfy these condition.

Where lny_i represents the natural logarithmic of per capita income of ith household. The independent variables gender, age, age^2, education, employment status represents the characteristic of household. By using economic theory and knowledge variable affecting income inequality can be identified [Wan and Zhou (2004-05)]. The income generating equation is of semi-log specification and instigate by the outcome that the income variable can be approached fit by using a log-normal distribution (Shorrocks and Wan, 2004).

The vriable used in this study, at the household to analyze the income inequality at household level are presented in Table 3.2.come inequality at household level.

Household head information: (gender, marital status, age, education level, employment status, household size and region)

Representative of individual household is considered as household head either he/she is male or female. In this study we consider head of household is the person who is a sole earner for the household. If there are more than one earners in the household then we consider elder person as head of household. Marital status variable indicates that weather household head was married; unmarried, widowed, divorced or only Nikha has solemnized but not living with his spouse. To make our analysis simple we have introduced a dummy for married is equal to one other wise zero. Variable age has also included which informed us the age of household head in years. Further we used age square as proxy of experience to analyze its effect on per capita income. Information of all households which we have used for analysis is taken from all four provinces of Pakistan. Education of household head included to capture the effect of human capital on earning. Household head with higher education is expected to earn more as compared to head of household with low education.household. Individuals and

household invest in education in order to build more human capital and enjoy higher benefit in the form of higher earnings in the future (Becker, 1962). According to these studies, conducted by [Su and Heshmati (2013); Pandey (2014-15); Sologon et al (2016); Bay and Epo (2011)] education positively affect the distribution of income.

Table 3.2 : Information of Variables

Dependent Variable	**Defination/ Unit of Measurment**	
Lny	**Natural Logarithim of monthly per capita income in Rs.**	
Independent Variable	**Defination/ Unit of Measurment**	**Expected Sign**
Gender	Dummy variable (1 for male, 0 otherwise)	Positive
Age	Continuous Variable(years)	Positive
Education	Household head years of Schooling	Positive
Marital Status	Dummy variable (1 for married, 0 otherwise)	Negative
Household Size	Member of household live and eat together	Negative
Employment Status	Head of household Employment Status Dummy variable (1 for Employer and self employee, 0 otherwise)	Positive
Region	Dummy Variable (1 for urban , 0 otherwise)	Positive

Household head employment status is also an important variable which affects the income of household. If household head is working as an employer or self-employed then he may earn more income. PSLM survey has categorized employment status into 8 eight categorized such as, paid employee, owner cultivator, self-employed, contract cultivator, unpaid family helper, share cropper, employer and livestock only. We have used a dummy for employer and self-employed is equal to one for all others factors is equal to zero.Household size variable is used as an explanatory variable in the analysis it provides information of member of household that live and eat together and have no

other place to live. Household size is a major factor that contributes to income inequality (Tripathi, 2016). If there is more than one earner in household than its contribution to income is positive. If household size is large and there is only one sole earner in household then contribution to income is negative. Then income of that household is less as comapared to those household in which there is more than one earner in the household. Instead of using monthly income we used per capita income of household. If we divide household monthly income with household size we get per capita income of each households. This study used region as an independent variable to analyze either there exist a gap in income between urban and rural households. Because it is generally believed that houshold living in urban areas earn more as compared to rural households.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Income Inequality in Pakistan

For the districts of Punjab highest percapita income inequality is observed in Islamabad and lowest in Mandi Bahuddin and Rajanpur.[7] This study shows that over time income inequality remain stagnant in Islamabad, Mandi Bahuddin ,Rajanpur and across all districts of Pakistan. Chakwal and Faislabad are those districts in which per capita income inequality has changed. Income ienquality in Chakwal has increased from 0.38 to 0.49, while for Faislabad it has decreased from 0.44 to 0.37. For districts of Sindh, highest per capita income inequality is observed in Karachi and lowest in Thatta. Income inquality for Karachi is 0.46 for the year 2004-05 and 0.48 for the year 2014-15. Nowshero Feroze, Ghotki, Jaccobabad, and Badin are those districts in which income inequaity has changed a lot. For Nowshero Feroze inequality has decreased form 0.47 to 0.36, for Ghotki inequality has increased from 0.36 to 0.47, for Jaccobabad inequaity has increaed from 0.37 to 0.46, for Badin inequality has increased form 0.54 to 0.41. Other than those districts inequality is remained stagnant over the time period. For the districts of Khyber Pakhtunkhwa highest inequality is observed in Lakki Marwat and lowest observed in Shangla. Per capita income inequality for Lakki Marwat is 0.52 for the year 2004-05 and 0.44 for the year 0.44. Singnifincat decrease in inequality has observed in Lakki Marwat. Inequalty for Shangla is 0.37 for the year 2004-05 and 0.33 for the year 2014-15. District in which inequality is decreased over time period is Swat, Upper Dir, Lower Dir, Malakand, Bonair, Charsada and Tank. Per

[7] The list of all districts used shown in Appendix of this study. Table A and table B in appendix show the percapita income of districts of Punjab, Sindh, Khyber Pakhtunkhwa and Baluchistan. The unit of analysis of this study is household instead of individuals.

capita income inequality for Swat has decreased from 0.55 in 2004-05 to 0.35 in 2014-15, for Upper Dir inequality has decreased form 0.47 to 0.37, for Lower Dir inequality has decreased form 0.44 to 0.37, for Malakand inequality has increased from 0.38 to 0.49, for Bonair inequality has decreased form 0.50 to 0.38, for Charsada inequality has decreased from 0.46 to 0.39, for Tank inequality has decreased from 0.48 to 0.38.

For districts of Balochistan highest percapita income inequality observed in Ziarat and lowest in Deara Bughti. 0.48 per capita income inequality is recorded in Ziarat for the year 2004-05 and 0.41 for the year 2014-15. Per capita income inequality for Dheara Bughti is 0.23 for the year 2004-05 and 0.35 for the year 2014-15. Increase in percapita income inequality is observed in Deara Bughti. Districts in which percapita income inequality has changed significant is Pashin, Deara Bughti, Gwadar, Jaffarabad, and Jahal Magsi.

For Pashin per capita income inequality has decreased form 0.41 to 0.30 for the year 2004-05 to 2014-15, for Gwader inequality has decreased from 0.50 to 0.28, for Jaffarabad inequality has increased form 0.25 to 0.38, for Jhal Magsi inequality has increased from 0.32 to 0.5. It is broadly viewed that inequality is more noticeable in urban areas. Urban areas are usually more developed, with diverse sectors acquiring different skills and technical knowhow. This can lead to wage differences and gaps that cause greater disparities in urban areas. Table 4.1 provides us the decomposition of income among four provinces of Pakistan which include Punjab, Sindh, KhyberPakhtunkhwa and Balochistan.

Table 4.1: Decomposition of Income Inequality by Provinces

Provinces	Population share		Income share		Mean Income		Gini Coefficient	
	2004-05	2014-15	2004-05	2014-15	2004-05	2014-15	2004-05	2014-15
Punjab	0.46	0.46	0.46	0.47	1321	3358	0.44	0.41
Sindh	0.25	0.24	0.29	0.23	1592	3150	0.47	0.40
KhyberPakhtunkhwa	0.17	0.16	0.13	0.15	1076	3104	0.44	0.40
Baluchistan	0.13	0.14	0.12	0.15	1225	3650	0.40	0.36

Source: Author's own estimation based on PSLM 2004-05 and 2014-15.

In both survey years Punjab province has highest share in population in overall sample while Baluchistan has lowest one. The population share of Punjab and Baluchistan province is 0.46 and 0.13 for the year 2004-05 respectively and 0.46 and 0.14 for the year 2014-15 respectively. Similarly the share of Sindh and KhyberPakhtunkhwa is 0.25 and 0.17 for the year 2004-05 respectively and for the year 2014-15 is o.24 and 0.16. Punjab has highest income share while Baluchistan has lowest income share in both survey year. For the year 2004-05 income share of Punjab, KhyberPakhtunkhwa , Baluchistan has increased while the share of Sind in overall income has decreased. Income share of Punjab, KhyberPakhtunkhwa and Baluchistan increased from 0.46, 0.13 and 0.12 percent to 0.47, 0.15 and 0.15 percent respectively. Income share of KhyberPakhtunkhwa has declined from 0.29 to 0.23 percent. However Sindh average income is higher in 2004-05 which is Rs1592 and lowest observed in KhyberPakhtunkhwa which is Rs1072. The mean income of all provinces has increased in the survey year of 2014-15. Highest per capita income for the year 2014-15 is observed in Baluchistan which is Rs3650 and lowest observed in Khyber Pakhtunkhwa which is Rs3104.

According to the Gini in 2004-05 percapita income inequality is higher in Sindh and lowest in Baluchistan. Using the Gini coefficient the inequality for Punjab, Sindh, KhyberPakhtunkhwa and Balochistan for the year 2004-05 were 0.44, 0.47, 0.44 and 0.40 respectively. However, these four provinces experienced a decrease in inequality between 2004-05 and 2014-15. Gini index decreased from 0.44 to 0.41 for the Punjab, and from 0.47 to 0.40 for the Sindh, and from 0.44 to 0.40 for KhyberPakhtunkhwa , and 0.40 to 0.36 for the Balochistan.

Table 4.2: Decomposition of Income Inequality by Region 2004-05

Province	Region	Income share	Mean income	Gini
Punjab	Urban	0.63	1758	0.44
	Rural	0.37	931	0.39
Sindh	Urban	0.7	2080	0.46
	Rural	0.3	1037	0.42
KhyberPakhtunkhwa	Urban	0.43	1573	0.47
	Rural	0.57	866	0.4
Balochistan	Urban	0.41	1644	0.41
	Rural	0.59	1041	0.35

Source: Author's own estimation

Table 4.2 and 4.3 provide the decomposition of income inequality by region among all four provinces of Pakistan for the year 2004-05 and 2014-15. The two categories under consideration are urban and rural areas. The income share of urban region of Punjab and Sindh is higher in survey period of 2004-05. The urban region share of income is less for KhyberPakhtunkhwa and Baluchistan in survey period of 2004-05.

Table 4.3: Decomposition of income inequaity by Region 2014-15

Province	Region	Income share	Mean income	Gini
Punjab	Urban	0.35	4467	0.42
	Rural	0.65	2908	0.39
Sindh	Urban	0.45	4411	0.4
	Rural	0.55	2553	0.37
KhyberPakhtunkhwa	Urban	0.16	4467	0.42
	Rural	0.84	2929	0.41
Balochistan	Urban	0.3	5563	0.47
	Rural	0.7	3173	0.35

Source : Author's own estimation

In survey period of 2014-15 the urban share of all region is less as compared to all rural regions. It is worth mentioning that mean income of all urban regions is higher between two survey periods that is why higher income inequality observed more in urban areas. However inequality using Gini coefficient has registered a significant decrease for both rural and urban areas.

Table 4.4:Decomposition of income inequality between and within group (2004-05 and 2014-15)

Provinces	Within	Between	Residual	Within	Between	Residual
		2004-05			2014-15	
Punjab	0.21	0.16	0.08	0.23	0.10	0.08
Sindh	0.23	0.16	0.08	0.20	0.13	0.08
KhyberPakhtunkhwa	0.22	0.14	0.08	0.31	0.03	0.05
Balochistan	0.20	0.10	0.10	0.23	0.04	0.09

Source : Author own estimation

Further table 4.4 provides the decomposition of income inequality between and within group. We have shown that highest income inequality is observed more in within group as compared to between group inequalities. Over the time with in group inequality has increased for Punjab, KhyberPakhtunkhwa and Balochistan. For Punjab it has been increased from 0.21 to 0.23 and for KhyberPakhtunkhwa it has increased from 0.22 to 0.31, for Balochistan it has increased from 0.20 to 0.23. Over the time the

between group inequality is decreased for Punjab as 0.16 to 0.08 ,for Sindh 0.16 to 0.13, for KhyberPakhtunkhwa 0.14 to 0.03 and for Balochistan 0.10 to 0.04. The residual term shows the income overlap for urban and rural region.

In the next section we find the factors which contribute more in inequality. For this purpose we used Field (2003) methodology. In this methodology we first estimate the earning function by using Ordinary Least Square (OLS). After that we use the cofficient obtain from OLS to examine the contribution of houshold factor in inequality. We find the contribution of household factor in inequality for each provinces.

4.2Punjab

4.2.1 Descriptive Data Analysis for Punjab

Table 4.5: Descriptive Analysis of Punjab

	2004-05		2014-15	
Variables	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Age	44.1	13	45.43	13
Education	8.33	3	8.06	3
Household Size	6.62	3	6.4	3
Per Capita Income	1638	3057	3973	4296

Source : Author own estimate based on PSLM 2004-05 and 2014-15.

For the year 2004-05 average age for household head remained at 44.10 years while education of household head is recorded at 8.33 on average. Household size for Punjab province is 6.62 households per house on average and per capita income stood at 1638 rupees. For the year 2014-15 average age for house hold remained at 45.43 years while education is recorded at 8 standard on average. House hold size for Punjab province is 6.40 household per house on average and per capita income stood at 3973 rupees. Comparing this descriptive for both years we can see that average age has been

increased after ten years gap. This shows that medical facilities has properly been provided or not less than earlier in the year 2004-05. Education standard is as on 8th rank as it was in 2004-05 reflecting that no improvement has been made in this area. Household average size is however decreased in the forthcoming years which is a good sign while considering resource constraints and per capita income. Per capita income has also been increased in a greater proportion as compare to decrease in household size. It has been doubled in year 2014-15.

4.2.2 Estimation of Earning Equation for Punjab.

Estimates of earning equation are provided by using Ordinary Least Square (OLS) for both study periods 2004-05 and 2014-15. The value of R^2 indicates the variation in dependent variable explained by independent variables. The value of R^2 is 0.39 and 0.36 for the study period of 2004-05 and 2014-15 which is enough for crossectional data set. Analyzing the variables separately, the result shows that most of the estimated coefficients are significant at 1 percent level and few are significant at 5 percent. The estimates show that male household head earns 20 percent more than female household head in 2004-05 and 40 percent in 2014-15. It is clear that main differences of income between male and female workers exists due to labor market discrimination. The variable age has a positive sign for both study periods. It means that if age of the household head increase its income is also increases in both survey period. The variable education has a positive sign for both study periods. If household gets more education its contribution is positive towards income.

Household head which has higher education earns more as compare to those household which has lowest education. The variable household size has a negative contribution in both study periods. It shows that if household size increases then percapita income of household decline. The contribution of household size variable is

more in 2014-15 as compared to 2004-05. This study considers marital status as an independent variable which shows that weither household of household is currently married or not. The variable marital status has a negative sign for two study periods which shows that household who is married has low income. The household head who is married earned 23 percent less in 2004-05 and 13 percent less in 2014-15.

Table 4.6: Estimated Earning Function for Punjab

	2004-05	2014-15
Variable	**Coefficient**	**Coefficient**
Gender	0.196(5.04)*	0.394(10.48)*
Age	0.0106(4.01)*	0.013(4.68)*
Education	0.082(48.48)*	0.071(38.21)*
Employmnt Status	0.205(17.38)*	0.132(10.80)*
Household Size	-0.085(41.12)*	-0.087(39.51)*
Marital Status	-0.227(10.05)*	-0.134(5.41)*
Region	0.36(30.78)*	0.293(22.24)*
Constant	6.183(89.98)*	7.039(93.71)*
R^2	0.39	0.36
F-Statistic	994.9012	653.2126
N	14627	13810

Note: * denotes that variable is significant at 1%.

Employment status is used as an independent variable which shows positive sign for both study periods. This study considers three categories of employment status i.e employer, paid employee and self-employed is equal to 1 and otherwise 0. Results indicate that estimated coefficients are positive for both years. Estimate coefficients shows that employer, self-employee and paid employee earn 20 percent more in 2004-05 and 13 percent more in 2014-15 than base categories (Owner cultivator, contract cultivaotr, unpaid family helper, share cropper and live stock). Estimated coefficient of region is positive for both study periods which shows that household living in urban areas earn more than household living in rural areas. Household living in urban areas earn 36 percent more in 2004-05 and 29 more percent in 2014-15. The coefficient of

region shows that over time period gap between income of household living in urban and rural areas has decreased.

4.2.3 Estimation of Factor Inequality Weight

Field (2003) proposed regression based decomposition methodology which enable the current study to measure how much the inequality in per capita income is explained by various characteristics of household. Factor inequality weight s_j attribute to each household j^{th} characteristic is calculated by using the coefficient estimate reported in table 4.6.Table 4.7 and 4.8 show how the inequality factor weight of each variable is calculated. If value of s_j is positive it means that variable is inequality increasing and negative value means that variable is inequality decreasing. Factor inequality weight in current section shows that only a few variables are significant in determining the level of inequality.

4.2.4 Estimation of Factor Inequality Weight for Punjab 2004-05

Factor inequality weight of each variable in 2004-05 is presented in table 4.7 Variables included in earning equation explains 39 percent inequality in percapita income of household which is equivalent to the coefficient of determination. All variables included in the earning equation have a positive s_j affect except the age. The positive value of s_j means that variable has increased inequality whereas negative value shows variable has decreased inequality. This study concludes that the variable which has highest contribution in increasing inequality are education, housheold size and region. The consolidate share of education, region and household size in increasing inequality is 20, 8 and 7 percent. The variable which has lowest contribution in increasing inequality are age square, male, marital status and employment status. The consolidate share of age square, male, marital status and employment status in increasing inequality is 0.1, 0.3, 0.5, and 1 percent respectively. The variable which has

negative contribution in decreasing inequality is age. If age of the household head increases he earns more. The share of age in reducing inequality is 0.2 percent respectively.

Table 4.7:Factor inequality weight of 2004-05 for Punjab

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2004-05
Male	0.2	0.002	0.003
Age	0.01	-0.001	-0.002
Household Size	-0.09	0.057	0.086
Education	0.08	0.133	0.2
Marital Status	-0.22	0.003	0.005
Employment Status	0.21	0.012	0.018
Region	0.36	0.05	0.075

Source : Author own estimate based on PSLM 2004-05.

4.2.5 Estimation of Factor Inequality Weight for Punjab 2014-15

A little increase in unexplained variation is observed in earning equation of 2014-15 when we compared it with earning equation of 2004-05 which shows that the variables included in earning equation do not explain the inequality in per capita income as accurately as before. Variable included in earning equation of 2014-15 explains 36 percent inequality in percapita income. Variables which contribute more in increasing inequality are education, household size and region. The consolidated share of education, household size, employment status and region in increasing inequality is 13, 8, 5 and 4 percent respectively. The variables which have lowest contribution in increasing inequality are male, age,and employment status with corresponding share of 3, 0.2, and 5 percent respectively. The contribution of gender dummy male is 3 percent to the inequality in comparison to 0.3 percent in 2004-05 which shows a significant

increase in inequality. The increase in inequality turns out from the fact that gender earning gap has increased from 2004-05 to 2014-15.

Table 4.8 Factor Inequality Weight 2014-15 for Punjab

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2014-15
Male	0.39	0.020	0.0329
Age	0.01	0.001	0.0021
Household Size	-0.09	0.054	0.0893
Education	0.07	0.085	0.1394
Marital Status	-0.13	-0.011	-0.0184
Employment Status	0.13	0.032	0.0521
Region	0.29	0.025	0.0403

Source : Author own estimate.

4.2.6 Overtime change in Inequality during 2004-05-2014-15

The current study uses regression based decomposition methodology to explain the overtime change in inequality due to household characteristics. Contribution of each household characteristic to change in inequality is measured by using the value of Gini coefficient and equation 9.

$$\Delta I_k = \frac{s_k^{2014-15} * I_{2014-15} - s_k^{2004-05} * I_{2004-05}}{I_{2014-15} - I_{2004-05}} \quad (9)$$

The contribution of variable to inequality is determined by the sign of factor inequality weight. Variable is inequality decreasing if factor inequality weight is negative whereas positive value shows the variable is inequality increasing. However if the factor weight increase as compared to previous year then variable has an ability to increasing the inequality in household percapita income. If factor inequality weight decrease as compared to previous year then variable has an ability to decrease the inequality in household percapita income. First two columns of table 4.9 show the factor inequality weights for the year 2004-05 and 2014-15. Comparison of factor inequality

weight for the years 2004-05 and 2014-15 shows that the variable improving the inequality are education, household size, marital status, age square and region. On the contrary the variable which increase inequality in 2014-15 are male and age. The positive value in the third column shows the variable ability in increasing the Gini coefficient.

Table 4.9: Contribution of factor to change in inequality 2004-05 to 2014-15.

Variables	Factor inequality weight 2004-05	Factor inequality weight 2014-15	Sj*G2015-Sj*G2005	Contribution to change in Gini coefficient
Male	0.003	0.033	0.012	-0.381
Age	-0.002	0.002	0.001	-0.062
Household Size	0.086	0.089	-0.001	0.037
Education	0.2	0.139	-0.031	0.993
Marital Status	0.005	-0.018	-0.01	0.302
Employment Status	0.018	0.052	0.013	-0.425
Region	0.075	0.04	-0.017	0.532

Source : Author own estimate.

If third column has a negative value then it work in opposite direction. Education of household head, marital status, household size, and region are the variable which improve the Gini coefficient. Last column of table 4.9 shows how the change in Gini coefficient is explained by each variable. Gini coefficient for Punjab has improved to 0.41 in 2014-15 as compared to 0.44 in 2004-05. The positive value in last column indicates that variable that account for decreasing the Gini coefficient and vice versa

4.3 KhyberPakhtunkhwa

4.3.1 Descriptive Data Analysis for Khyber Pakhtunkhwa .

For the year 2004-05 average age for house hold head remained at 46.75 years while education of household head is recorded at nine on average. Household size for Khyber Pakhtunkhwa province is 8.17 households per house on average and per capita

income stood at 1076 rupees. For the year 2014-15 average age for house hold remained at 46.17 years while the education of household head is recorded at nine standard on average. Household size is six households per house on average and per capita income stood at 3103 rupees. Comparing this descriptive for both years we can see that average age has been decreased in both years even after ten years gap. Education standard is as on 9th rank as it was in 2004-05 reflecting that no improvement has been made in this area. Household average size is however decreased in the forthcoming years which is a good sign while considering resource constraints and per capita income. Per capita income has also been increased in a greater proportion as compare to decrease in household size.

Tablem 4.10: Descriptive analysis for Khyber Pakhtunkhwa

	2004-05		2014-15	
Variables	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Age	46.75	14	46.17	13
Education	8.67	4	8.89	3
Household Size	8.17	4	7.69	4
Per Capita Income	1076	2919	3103	3893

Source : Author own estimate based on PSLM data 2004-05 and 2014-15.

4.3.2 Estimation of Earning Equation for Khyber Pakhtunkhwa

Estimates of earning equation are provided by using Ordinary Least Square (OLS) for both study periods 2004-05 and 2014-15. The value of R^2 indicates the variation in dependent variable explained by independent variable. The value of R^2 is 0.42 and 0.39 for the study period of 2004-05 and 2014-15 which is enough for crossectional data set. Analyzing the variables separately, the result show that most of the estimated coefficients are significant at 1 percent level and few are significant at 5

percent. The estimates show that male household head earn 20 percent more than female household head in 2004-05 and 17 percent in 2014-15. It is clear that main differences of income between male and female workers exists due to labor market discrimination. The discrimination in labor market has decreased but there is enough gap in earning of male and female workers. The variable age has a positive sign for both study periods.

Table 4.11: Earning function for Khyber Pakhtunkhwa .

Variable	2004-05 Coefficient	2014-15 Coefficient
Gender	-0.203276(2.31)*	0.170(1.98)*
Age	0.023(4.41)*	0.030(5.87)*
Education	0.062(20.98)*	0.071(23.01)*
Employment Status	0.254(10.75)*	0.088(4)*
Household Size	-0.059(20.70)*	-0.079(24.35)*
Marital Status	-0.130(2.79)*	-0.163(3.05)*
Region	0.342(15.15)*	0.201(6.63)*
Constant	6.141(45.12)*	6.939(49.41)*
$\boldsymbol{R^2}$	0.42	0.39
F-Statistic	209.554	176.985
N	4530	4309

Note: * denotes that variable is significant at 1%.

It means that if age of the household head increases its income is also increases. The variable education has a positive sign for both study periods. If household head gets more education its contribution is positive in income. Household head which has higher education earn more as compare to those household which has lowest education. Return to education is more in 2014-15 as compared to 2004-05. The variable household size has a negative contribution in 2004-05 a both study period. It shows that if household size increases by one unit then per capita income of that household decline. The contribution of household size variable is more in 2014-15 as compared to 2004-

05. This study consider marital status as an independent variable which shows weither head of household is currently married or not. The variable marital status has a negative sign for two study periods which shows that household who is married has low income. The household head who is married earned 13 percent less in 2004-05 and 16 percent less in 2014-15.

Employment status is used as an independent variable which shows positive sign for both study periods. This study considers three employment status categories i.e employer, paid employee and self-employee is equal to 1 and otherwise 0. Results indicate that estimated coefficients are positive for both years. Estimated coefficients show that employer, self-employee and paid employee earn 25 percent more in 2004-05 and 8 percent more in 2014-15 than base categories (Owner cultivator, contract cultivaotr, unpaid family helper, share cropper and live stock). Estimated coefficient of region is positive for both study periods which shows that household living in urban areas earn more than household living in rural areas. Household living in urban areas earn 42 percent more in 2004-05 and 39 more percent in 2014-15. The gap between percapita income of households living in urban and rural areas has significantly decreased during last decade.

4.3.3 Estimation of Inequality Factor Weight.

Field (2003) proposed regression based decomposition methodology which enables the current study to measure how much the inequality in per capita income is explained by various characteristics of households. Factor inequality weight s_j attributed to each household j^{th} characteristic is calculated by using the coefficient estimates reported in 4.11. Table 4.12 and 4.13 show how the inequality factor weight of each variable is calculated. If value of s_j is positive it means that variable is inequality increasing and negative value means that variable is inequality decreasing. Factor

inequality weight in current section shows that only a few variables are significant in determining the level of inequality.

4.3.4 Estimation of Factor Inequaity Weight for KPK 2004-05

Factor inequality weights of each variable for 2004-05 are presented in table 4.12. Variables included in earning equation explains 42 percent inequality in percapita income of households which is equivalent to the coefficient of determination. All variables included in the earning equation have a positive s_j except the age. The positive value of s_j means that variable cause an increase in inequality whereas negative value shows variable decreases the inequality. This study concludes that the variables which have highest contribution in increasing inequality are education, household size, and region. The consolidated share of education, household size, and region in increasing inequality is 12, 7, 5 percent respectvely. The reason for positive impact can be due to lower level of human capital resulting in higher premia on education leading to inequality. The variables which have lowest contribution in increasing inequality are age square and marital status. The consolidate share of age square and marital status in increasing inequality is 1 and 2 percent respectively. The variable which has negative contribution in decreasing inequality is gender and age. If age of the household head increases he earns more but at a diminishing rate. The share of gender and age in reducing inequality is 0.5 and 1 percent respectively.

Table 4.12: Factor Inequality Weight of 2004-05

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2004-05
Male	-0.20	-0.003	-0.005
Age	0.02	-0.013	-0.019
Household Size	-0.06	0.053	0.079
Education	0.06	0.086	0.129
Marital Status	-0.13	0.001	0.002
Employment Status	0.25	0.014	0.021
Region	0.34	0.033	0.050

Source: Author own estimate.

4.3.5 Estimation of Factor Inequality Weight for KPK 2014-15

When we compared the earning equation of 2004-05 with 2014-15 a little increase in unexplained variation is observed in earning equation of 2014-15 which shows that the variables included in earning equation do not explain the inequality in per capita income as accurately as before. In 2014-15, variables included in earning equation explain 39 percent of the inequality. This study concludes that the variables which contribute more in increasing inequality are educationa and household size. The consolidate share of education and household size is 15 and 12 percent respectively. The variables which have lowest share in increasing inequality are gender, age square, marital status, employment status and region with corresponding share of 0.3, 2, 0.1, and 0.2 percent respectively. Variable age has a negative contribution in decreasing inequality. The share of age in reducing inequality is 2 percent. The contribution of gender dummy is 0.3 percent to the inequality in comparison to -0.5 percent in 2004-05 which shows a significant increase in inequality. The increase in inequality turn out from the fact that gender earning gap between male and female worker has increased from 2004-05 to 2014-15.

Table 4.13: Factor Inequality weight 2014-15 for KPK

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2014-15
Male	0.17	0.0020	0.003
Age	0.03	-0.015	-0.026
Household Size	-0.08	0.072	0.123
Education	0.07	0.089	0.152
Marital Status	-0.16	0.0005	0.001
Employment Status	0.09	0.0013	0.0023
Region	0.20	0.008	0.013

Source : Author own estimate based on PSLM 2014-15

4.3.6 Overtime Change in Inequality during 2004-05-2014-15

The current study uses regression based decomposition methodology to explain the overtime change in inequality due to household characteristics. Contribution of each household characteristic to change in inequality is measured by using the value of Gini coefficient and equation 9.

$$\Delta I_k = \frac{s_k^{2014-15} * I_{2014-15} - s_k^{2004-05} * I_{2004-05}}{I_{2014-15} - I_{2004-05}} \quad (9)$$

The contribution of variable to inequality is determined by the sign of factor inequality weight. Variable is inequality decreasing if factor inequality weight is negative and inequality increasing if factor inequality weight is positive. However if the factor weight in current year increases as compared to previous year then variable has an ability to increase the inequality in percapita income of household. If factor weight in current year decreases as compared to previous year then variable has an ability to decrease the inequality in percapita income of household. First two columns of table 4.14 show the factor inequality weights for the year 2004-05 and 2014-15. Comparison of factor inequality weight for the year 2004-05 and 2014-15 shows that the variables improving the inequality are gender, age, marital status, employment

status and region. On the other hand the variable which contribute positively in increasing inequality in 2014-15 is household size and education. The positive value in the third column shows the variable has an ability in increasing the Gini coefficient. If third column comprises of negative value then it works in opposite direction. Age, marital status, employment status and region are the variables which improve the Gini coefficient and household size, education of household head is a variable which worsen the inequality. Last column of table 11 shows how the change in Gini coefficient is explained by each variable.

Table 4.14 : Contribution to change in inequality 2004-05 to 2014-15

Variables	Factor inequality weight 2004-05	Factor inequality weight 2014-15	Sj*G2015-Sj*G2005	Contribution to change in Gini coefficient
Male	-0.005	0.003	0.004	-0.082
Age	-0.019	-0.026	-0.002	0.048
Household Size	0.079	0.123	0.014	-0.334
Education	0.129	0.152	0.004	-0.097
Marital Status	0.002	0.001	-0.001	0.012
Employment Status	0.021	0.002	-0.008	0.193
Region	0.05	0.013	-0.017	0.392

Source : Author own estimate.

Gini coefficient for khyber pakhtunkhwa has improved to 0.4 in 2014-15 as compared to 0.44 in 2004-05. The positive value in last column indicates that variable that accounts for decreasing the Gini coefficient and vice versa.

4.4 Balochistan

4.4.1 Descriptive Data Analysis for Balochistan

Average age for household head remained at 43.95 years for the year 2004-05 while education of household head is recorded at 8.83 on average. House hold size for

Baluchistan is 7.1 households per house on average and per capita income stood at 1225 rupees. For the year 2014-15 average age for household remained at 42.92 years while the education of household head is recorded at 8.88 standard on average. Household size is 7.19 household per house on average and per capita income stood at 3650 rupees. Comparing this descriptive analysis for both years we can see that average age has been decreased in both years even after ten years gap. Education standard is as on 9th rank as it was in 2004-05 reflecting that no improvement has been made in this area. Household average size is remained the same in both study periods. Per capita income has also been increased more than double in 2014-15.

Table 4.15: Descriptive analysis of Balochistan

	2004-05		2014-15	
Variables	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Age	43.95	13	42.92	12
Education	8.83	4	8.88	4
Household Size	7.1	3	7.19	3
Per Capita Income	1225	1494	3650	7460

Source : Author own estimate based on PSLM data 2004-05 and 2014-15.

4.4.2 Estimation of Earning Equation for Balochistan

Estimates of earning equation are provided by using Ordinary Least Square (OLS) for both study periods 2004-05 and 2014-15. The value of R^2 indicates the variation in dependent variable explained by independent variables. The value of R^2 is 0.38 and 0.36 for the study period of 2004-05 and 2014-15 respectively which is enough for crossectional data set. Analyzing the variables separately, the resuts show that most of the estimated coefficients are significant at 1 percent level and few are significant at 5 percent. The estimates show that male household head earns 17 percent less than

female household head in 2004-05 and 27 less than in 2014-15. But the result of coefficient of gender is statistically insignificant for both study period. The variable age has a positive sign for both study periods. It means that if age of the household head increases its income is also increases.The variable education has a positive sign for both study periods. If household head gets more education its contribution is positive in income. Household head which has higher education earn more as compare to those household which has lowest education. Return to education is remain the same for both study period.

Table 4.16 : Estimated Earning Function for Balochistan.

	2004-05	**2014-15**
Variable	**Coefficient**	**Coefficient**
Gender	-0.175(1.02)***	-0.272(1.58)***
Age	0.018(3.24)*	0.024(4.15)*
Education	0.052(17.49)*	0.056(19.74)*
Employment Status	-0.096(4.09)*	-0.009(0.411)***
Household Size	-0.092(25.16)*	-0.093(26.01)*
Marital Status	-0.178(3.07)*	-0.261(3.93)*
Region	0.267(12.19)*	0.191(8.03)*
Constant	6.968(33.96)*	7.937(38.48)*
$\boldsymbol{R^2}$	0.38	0.36
F-Statistic	175.014	183.47
N	3062	3918

Note: * denotes that variable is significant at 1%, *** denotes varialbe is significant at 10% t statistic is reported in parenthesis.

The variable household size has negative contribution in both study periods. It shows that if household size increases by one unit its percapita income of household declines.This study consider marital status as an independent variable which shows

weither head of household is married or not. The variable marital status has a negative sign for both study periods which shows that household who is married has low income. The household head who is married earned 17 percent less in 2004-05 and 26 less percent less in 2014-15 as compared to unmarried household heads.

Employment status is used as an independent variable which shows negative sign for both study periods. This study considers three employment status categories i.e. employer, paid employee and self-employee is equal to 1 and 0 otherwise. Results indicate that estimated coefficients are negative for both study years. Estimated coefficients show that employer, self-employee and paid employee earn 9 percent less in 2004-05 and 0.9 percent less in 2014-15 than base categories (Owner cultivator, contract cultivaotr, unpaid family helper, share cropper and live stock). Estimated coefficient of region is positive for both study periods which shows that household living in urban areas earn more than household living in rural areas. Households living in urban areas earn 38 percent more in 2004-05 and 36 more percent in 2014-15. The gap between percapita income of household living in urban and rural areas has significantly declined during last decade.

4.4.3 Estimation of Inequality Factor Weight.

Field (2003) proposed regression based decomposition methodology which enabled the current study to measure how much the inequality in per capita income is explained by various characteristics of household. Factor inequality weight s_j attributed to each household j^{th} characteristic is calculated by using the coefficient estimated reported in table 4.16. Table 4.17 and 4.18 show how the inequality factor weight of each variable is calculated. If value of s_j is positive it means that variable is inequality increasing whereas negative value means that variable is inequality decreasing.Factor

inequality weight in current section shows that only a few variables are significant in determining the level of inequality.

4.4.4 Estimation of Factor Inequality Weight for Balochistan 2004-05

Table 4.17 : Factor Inequality Weight 2004-05 for Balochistan

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2004-05
Male	-0.1751	-0.0002	-0.0004
Age	0.0181	-0.0088	-0.0176
Age square	-0.0001	0.0045	0.0090
Household Size	-0.0922	0.0670	0.1346
Education	0.0524	0.0599	0.1203
Marital Status	-0.1784	0.0010	0.0021
Employment Status	-0.0966	-0.0013	-0.0026
Region	0.2674	0.0234	0.0470

Source : Author own estimate by using PSLM 2004-05 data.

Factor inequality weight of each variable in 2004-05 is presented in table 4.17. Variables included in earning equation explain 38 percent inequality in percapita income of household which is equivalent to the coefficient of determination. All variables included in the earning equation have positive s_j except the age. The positive value of s_j means that variable increases the inequality whereas negative value shows variable decreases the inequality. This study concludes that the variables which have highest contribution in increasing inequality are education and household size. The reason for positive impact can be due to lower level of human capital resulting in higher premium on education leading to inequality. The consolidated share of education and household size increasing inequality is 12, 13 percent respectively. The variables which have lowest contribution in increasing inequality are age square, marital status and region. The consolidate share of age square, marital status and region in increasing inequality is 0.9, 0.2, and 4 percent respectively. The variables which have negative

contribution in increasing inequality are gender and age. If age of the household head is increased he earns more but at a diminishing rate. The share of age in reducing inequality is 1 percent.

4.4.5 Estimation of Factor Inequality Weight for Balochistan 2014-15.

When we compared the earning equation of 2004-05 with the earning equation of 2014-15 a little increase in unexplained variation is observed in earning equation of 2014-15 which shows that the variables included in earning equation do not explain the inequality in per capita income as accurately as before. In 2014-15, variables included in earning equation explained 36 percent of the inequality.

Table 4.18: Factor Inequality weight 2014-15.

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2014-15
Male	-0.27	0.0005	0.001
Age	0.02	-0.007	-0.012
Household Size	-0.09	0.065	0.120
Education	0.06	0.076	0.141
Marital Status	-0.26	0.002	0.004
Employment Status	-0.01	-0.0002	-0.0003
Region	0.19	0.011	0.020

Source : Author own estimate based on PSLM 2014-15 data.

This study concludes that the variables which contribute more in increasing inequality are education and household size. The consolidated share of education, and household size in increasing inequality is 14 and 12 percent respectively. The variables which has lowest share in increasing inequality is age square, marital status and region with corresponding share of 0.6, 0.4 and 2 percent. Variable age and employment status has a negative contribution in decreasing inequality. The share of age, and employment status is 1, and 3 percent respectively.

4.4.6 Overtime Change in Inequality during 2004-05-2014-15

The current study uses regression based decomposition methodology to explain the overtime change in inequality due to household characteristics. Contribution of each household characteristic to change in inequality is measured by using the value of Gini coefficient and equation 9.

$$\Delta I_k = \frac{s_k^{2014-15} * I_{2014-15} - s_k^{2004-05} * I_{2004-05}}{I_{2014-15} - I_{2004-05}} \quad (9)$$

The contribution of variable to inequality is determined by the sign of factor inequality weight. Variable is inequality decreasing if factor inequality weight is negative whereas positive sign describes the variable is inequality increasing. However if the factor weight in current year increases as compared to previous year then variable has an ability to increase the inequality in percapita income of household. If factor weight in current year decreases as compared to previous year then variable has an ability to decrease the inequality in percapita income of household. First two columns of table 11 show the factor inequality weights for the year 2004-05 and 2014-15.

Table 4.19 : Contribution of Factor to Change in Inequality 2004-05- to 2014-15

Variables	Factor inequality weight	Factor inequality weight	Sj*G2015-Sj*G2005	Contribution to change in Gini coefficient
Male	-0.0004	0.001	0.0005	-0.0105
Age	-0.018	-0.012	0.0026	-0.0596
Age square	0.009	0.006	-0.0016	0.0361
Household Size	0.135	0.120	-0.0112	0.2522
Education	0.120	0.141	0.0022	-0.0486
Marital Status	0.002	0.004	0.0007	-0.0150
Employment Status	-0.003	-0.0003	0.0009	-0.0212
Region	0.047	0.020	-0.0119	0.2689

Source : Author own estimate.

Comparison of factor inequality weight for the year 2004-05 and 2014-15 shows that the variables improving the inequality are gender, age, education, household size employment status and region. On the other hand the variables which contribute positively in increasing inequality in 2014-15 are age and marital status. The positive value in the third column shows the variable ability in increasing the Gini coefficient. If third column has a negative value then it works in opposite direction. Gender, age, education of household head, employment status are the variables which improve the Gini coefficient, age square and region are the variable which worsen the inequality. Last column of table 4.19 shows how the change in Gini coefficient is explained by each variable. Gini coefficient for Balochistan has improved to 0.36 in 2014-15 as compared to 0.40 in 2004-05. The positive value in last column indicates that variable that account for decreasing the Gini coefficient and vice versa.

4.5 Sindh

4.5.1 Descriptive Data analysis for Sindh

For the year 2004-05 average age for house hold head remained at 42 years while education of household head is recorded at nine on average. House hold size for Sindh province is seven household per house on average and per capita income stood at 1936 rupees. For the year 2014-15 average age for house hold remained at 41 years while the education of household head is recorded at nine standard on average. House hold size is six household per house on average and per capita income stood at 3818 rupees. Comparing this descriptive analysis for both years we can see that average age has been decreased in both years even after ten years gap. Education standard is as on nine rank as it was in 2004-05 reflecting that no improvement has been made in this area. Household average size is however decreased in the forthcoming years which is a good sign while considering resource constraints and per capita income. Per capita

income has also been increased in a greater proportion as compare to decrease in household size. Per capita income of Sindh has been doubled in 2014-15. For the year 2004-05 average age for house hold head remained at 43.6 years while education of household head is recorded at nine standard on average. House hold size for Sindh province is 6.74 household per house on average and per capita income stood at 1592 rupees. For the year 2014-15 average age for house hold remained at 42.36 years while the education of household head is recorded at nine standard on average. House hold size is 6.41 household per house on average and per capita income stood at 3149 rupees.

Table 4.20: Descriptive analysis for Sindh

	2004-05		2014-15	
Variables	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Age	43.6	13	42.36	12
Education	8.95	4	8.91	4
Household Size	6.74	3	6.41	3
Per Capita Income	1592	4381	3149	3846

Source : Author own estimate based on PSLM data 2004-05 and 2014-15

Comparing this descriptive analysis for both years we can see that average age has been decreased in both years even after ten years gap. Education standard is as on nine rank as it was in 2004-05 reflecting that no improvement has been made in this area. Household average size is however decreased in the forthcoming years which is a good sign while considering resource constraints and per capita income. Per capita income has also been increased in a greater proportion as compare to decrease in household size. Per capita income of Sindh has been doubled in 2014-15.

4.5.2 Estimation of Earning Equation for Sindh

Estimates of earning equation are provided by using Ordinary Least Square (OLS) for both study periods 2004-05 and 2014-15. The value of R^2 indicates the variation in dependent variable explained by independent variable. The value of R^2 is 0.41 and 0.38 for the study period of 2004-05 and 2014-15 which is enough for crossectional data set. Analyzing the variables separately, the results show that most of the estimated coefficients are significant at 1 percent level and few are significant at 5 percent. The estimates show that male household head earns 9 percent less than female household head in 2004-05 but in 2014-15 male household head earn 2 percent more than female household in 2014-15. It is clear that main differences of income between male and female workers exists due to labor market discrimination. The discrimination in labor market has been increased in Sindh. The variable age has a positive sign for both study periods. It means that if age of the household head increases its income is also increases

Table 4.21: Estimated Earning Function for Sindh.

	2004-05	**2014-15**
Variable	**Coefficient**	**Coefficient**
Gender	-0.095(1.36)***	0.025(1.54)***
Age	0.011(3.29)*	0.029(7.50)*
Education	0.081(39.82)*	0.077(35.40)*
Employment Status	0.225(14.06)*	0.088(0.89)***
Household Size	-0.106(40.76)*	-0.093(34.50)*
Marital Status	-0.172(5.32)*	-0.148(3.73)*
Region	0.435(27.70)*	0.332(20.79)*
Constant	6.427(63.39)*	6.855(53.99)*
$\boldsymbol{R^2}$	0.41	0.38
F-Statistic	850.9293	500.8461
N	8724	7216

Note: * denotes that variable is significant at 1%, *** varialbe is significant at 10% t statistic is reported in parenthesis

The variable education has a positive sign for both study periods. If household head gets more education its contribution is positive in income. Household head which has higher education earn more as compare to those household which has lowest education. Return to education is less in 2014-15 as compared to 2004-05. The variable household size has a negative contribution in both study period. It shows that if household size increases by one unit then percapita income of household decreases. This study consider marital status as an independent variable which shows that weither head of household is married or not. The variable marital status has a negative sign for study period which shows that houseold who is married has low income as compared to those household who is unmarried. The houshehold head who is married earns of 2004-05 while positive for the period of 2014-15. In study period 2014-15 married houseolds earn more than unmarried household. The household head who is married earned 17 percent less in 2004-05 and 14 percent less in 2014-15. Employment status is used as an independent variable which shows positive sign for both study periods. This study considers three employment status categories i.e. employer, paid employee and self-employee is equal to 1 and otherwise 0. Results indicate that estimated coefficients are negative for both study years. Estimated coefficients show that employer, self-employee and paid employee earn 22 percent less in 2004-05 and 8 percent less in 2014-15 than base categories (Owner cultivator, contract cultivaotr, unpaid family helper, share cropper and live stock).

Estimated coefficient of region is positive for both study periods which shows that households living in urban areas earn more than households living in rural areas. Household living in urban areas earn 43 percent more in 2004-05 and 33 more percent in 2014-15. The gap between percapita income of household living in urban and rural areas has significantly decreased during last decade.

4.5.3 Estimation of Inequality Factor Weight

Field (2003) proposed regression based decomposition methodology which enabled the current study to measure how much the inequality in per capita income is explained by various characteristics of household. Factor inequality weight s_j attributed to each household j^{th} characteristic is calculated by using the coefficient estimated reported in table 4.21. Table 4.22 and 4.23 show how the inequality factor weight of each variable is calculated. If value of s_j is positive it means that variable is inequality increasing and negative value means that variable is inequality decreasing. Factor inequality weight in current section shows that only a few variables are significant in determining the level of inequality.

4.5.4 Estimation of Factor Inequality Weight for Sindh 2004-05

Factor inequality weight of each variable in 2004-05 is presented in table 4.22. Variables included in earning equation explain 41 percent inequality in percapita income of household which is equivalent to the coefficient of determination. All variables included in the earning equation have a positive s_j except the age. The positive value of s_j means that variable increases inequality whereas negative value shows variable cause a decrease in inequality. This study concludes that the variables which have highest contribution in increasing inequality are education, household size, and region. The consolidated share of education, household size and region in increasing inequality is 20, 12 and 9 percent respectively. The reason for positive impact can be due to lower level of human capital resulting in higher premia on education leading to inequality. The variables which have lowest contribution in increasing inequality are age, age square, marital status, and employment status. The consolidate share of age, age square, marital status and employment status in increasing inequality is0.9, 0.2, 0.4

and 1 percent respectivley. The variable which has negative contribution in decreasing inequality is gender. The share of gender in reducing inequality is 0.3 percent.

Table 4.22: Factor Inequality Weight of 2004-05 for Sindh

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2004-05
Male	-0.1	-0.002	-0.003
Age	0.01	0.007	0.009
Household Size	-0.11	0.091	0.122
Education	0.08	0.149	0.2
Marital Status	-0.17	0.003	0.004
Employment Status	0.23	0.012	0.015
Region	0.44	0.073	0.098

Source : Author own estimate based on PSLM 2004-05.

4.5.5 Estimation of Factor Inequality Weight for Sindh 2014-15

When we compared the earning equation of 2004-05 with the earning equation of 2014-15 a little increase in unexplained variation is observed in earning equation of 2014-15 which shows that the variables included in earning equation do not explain the inequality in per capita income as accurately as before. In 2014-15, variables included in earning equation explained 38 percent of the inequality.

Table 4.23: Factor inequality Weight of 2014-15 for Sindh

Variables	Bj	Covariance of (Bxj, lny)	Factor inequality weight 2014-15
Male	0.09	-2.90E-05	-5.00E-05
Age	0.03	0.009	0.016
Household Size	-0.09	0.066	0.114
Education	0.08	0.117	0.202
Marital Status	-0.15	0.001	0.003
Employment Status	0.03	0.001	0.001
Region	0.33	0.036	0.062

Source : Author own estimate based on PSLM 2014-15.

This study concludes that the variables which contributed more in increasing inequality are education, household size and region. The consolidate share of education, household size and region is 20, 11 and 6 percent respectively. The variables which have lowest share in increasing inequality are age, employment status, marital status with corresponding share of 1, 0.1 and 0.3 percent respectively. Variable gender and age square has a negative contribution in decreasing inequality.

4.5.6 Overtime Change in Inequality during 2004-05-2014-15

The current study uses regression based decomposition methodology to explain the overtime change in inequality due to household characteristics. Contribution of each household characteristic to change in inequality is measured by using the value of value of Gini coefficient and equation 9.

$$\Delta I_k = \frac{s_k^{2014-15} * I_{2014-15} - s_k^{2004-05} * I_{2004-05}}{I_{2014-15} - I_{2004-05}} \quad (9)$$

The contribution of variable to inequality is determined by the sign of factor inequality weight. Variable is inequality decreasing if factor inequality weight is negative whereas positive sign describe the variable is inequality increasing. However if the factor inequality weight in current year increase as compared to previous year then variable has an ability to increasing the inequality in percapita income of household. If factor weight in current year decreases as compared to previous year then variable has an ability to decrease the inequality in percapita income of household.

Table 4.24 : Contribution to Change in Inequality by Factor

Variables	Factor inequality weight 2004-05	Factor inequality weight 2014-15	Sj*G2015-Sj*G2005	Contribution to change in Gini coefficient
Male	-0.003	-5.00E-05	0.0013	-0.02
Age	0.009	0.0156	0.0019	-0.03
Household Size	0.122	0.114	-0.0116	0.17
Education	0.2	0.2025	-0.0129	0.19
Marital Status	0.004	0.0026	-0.0007	0.01
Employment Status	0.015	0.0015	-0.0067	0.1
Region	0.098	0.0623	-0.0209	0.3

Source : Author own estimate.

First two column of table 4.24 show the factor inequality weights for the year 2004-05 and 2014-15. Comparison of factor inequality weight for the year 2004-05 and 2014-15 shows that the variables improving the inequality are age square, marital status, education, household size, employment status, and region. On the other hand the variables which contribute positively in increasing inequality in 2014-15 are education of the household head. The positive value in the third column shows the variable ability in increasing the Gini coefficient. If third column has a negative value then it works in opposite direction. Age square, marital status, employment status, education, household size and region are the variables which improve the Gini coefficient and gender, education and household size are the variables which worsen the inequality. Last column of table 11 shows how the change in Gini coefficient is explained by each variable. Gini coefficient for Baluchistan has improved to 0.40 in 2014-15 as compared to 0.47 in 2004-05. The positive value in last Column indicates the variable that accounts for decreasing the Gini coefficient and vice versa.

CHAPTER 5

CONCLUSIONS AND POLICY RECOMMENDATIONS

5.1 Conclusions

Following study attempts to decompose the income inequality in Pakistan. Gini coefficient has been used for this purpose and inequality is decomposed into between and within group. Further we used regression based decomposition methodology developed by Field (2003) for all four provinces of Pakistan to examine which household factor contributes more inequality for the year 2004-05 and 2014-15. Overtime change in inequality has also been observed either it has been decreased or increased. Data for this study has been taken from Pakistan Social and Living Standard Measurement Survey conducted by Pakistan Bureauof Statistic (PBS). The variables used for regression based decomposition methodology are gender, age of household head, education of household, household size, employment status of household head and region. The main findings emerged from this study are as follows. For the year 2004-05 highest per capita income inequality is observed in Sindh and lowest is observed in Baluchistan. For the year 2014-15 income inequality is decreased among all four provinces of Pakistan. Gini coefficient for Punjab is 0.41 in 2014-15 as compared to 0.44 in 2004-05, for Sindh is 0.40 in 2014-15 as compared to 0.47 in 2004-05, for Khyber Pakhtunkhwa is 0.40 in 2014-15 as compared to 0.44 in 2004-05, for Baluchistan is 0.36 in 2014-15 as compared to 0.40 in 2014-15. Highest income inequality is observed in urban area as compared to rural areas. During the last decade income inequality has been decreased both in urban and rural areas of Pakistan.

Decomposition of Gini coefficient in to between and within group shows that highest inequality is observed in within group as compared to between groups. For the

years 2004-05 to 2014-15 inequality within group has been increased whereas between the groups inequality has decreased among all four provinces of Pakistan. For estimating the contribution of household characteristics into inequality earning function is estimated separately for all provinces as per the requirement of Field (2003) methodology. Result suggests that all estimated coefficient are significant at 1 percent which implies that all variables included in earning function have significant role in explaining household per capita income. Education of household head is contributing the main factor for inequality followed by size of household and region (urban and rural) for all provinces of Pakistan. This study also calculate the income inequality across the districts of Pakistan. For the districts of Punjab highest percapita income inequality is observed in Islamabad and lowest in Mandi Bahuddin and Rajanpur. Significant changed in inequality has observed in Chakwal and Faislabad. For the districts of Sindh highest inequality is observed in Karachi and lowest observed in Thatta. For the district of Khyber Pakhtunkhwa highest income inequality is observed in Lakki Marwat and lowest observed in Shangla. Similarly for the districts of Baluchistan highest per capita income inequality is observed in Ziarat and lowest observed in Daera Bughti.

5.2 Policy Recommendations

Present study suggests the following policy recommendation to reduce income inequality.

- Income markup of poor households would increase if they are intended to get more education. Their per capita income would eventually increase as their years of schooling increase. In the long run the vicious earning gap between rich and poor will also shrink. Attainment of primary and secondary education can be a very effective measure to

reduce inequality. When poor household invest in education then return to education increase due to markup obtain from each additional years of schooling. Education attainment would also ensure increased real wage of poor households as a return to invest in education. Therefore it is recommended that households should consider education inevitable at both primary and secondary level. Poor households should be provided some incentives to send their children to school. That would increase wage premium of poor households if they are provided reasonable education facilities. And that would eventually cause a decline in existing income inequality between rich and poor households.

- Second household size both in rural and urban areas need to be reduced alternatively a reduction in the number of dependent members in household is suggested. People must be provided adequate knowledge about the downsides of large family size. Proper employment opportunities should be made easily available so that dependency decrease.

5.3 Future Research Area

Data for this study has been taken from Pakistan Social and Living Standard Measurement Survey (PSLM) to measure income inequality at household level. However this study provides some recommendations for future research. To confirm/contradict income inequality at household level can be measured these findings by using expenditure data at household level. The result of this study show that education is the major factor that contribute in inequality. However there is a need to study considerable inequality in education of households.

REFERENCES

Abid, Yosr, Cathal O'Donoghue, and Denisa Sologon. (2016) "Decomposing Welfare Inequality in Egypt and Tunisia: An Oaxaca-Blinder Based Approach." Paper presented at the Economic Research Forum Working Papers.

Adams, Richard H Jr. (2001). 'Nonfarm Income, Inequality, and Land in Rural Egypt,'Economic Development and Cultural Change, vol. 50, pp 339-363.

Adams Jr, Richard H. (1994)"Non-Farm Income and Inequality in Rural Pakistan: A Decomposition Analysis." *The Journal of Development Studies* 31, no. 1: 110-133.

Ahmad, Ehtisham, Stephen Ludlow, and Mir Annice Mahmood. (1989) "Poverty, Inequality andGrowth in Pakistan [with Comments]." *The Pakistan Development Review* 28, no. 4 831-850.

Ahmad, Mehboob.(2002) "Income Inequality among Various Occupations/Professions in Pakistan-Estimates Based on Household Income Per Capita. The Lahore Journal of Economics, Vol.7, No.1.

Ali, Liaqat, and Naveed Akhtar.(2014-15) "An Analysis of the Gender Earning Differentials in Pakistan." *Research Journal of Applied Sciences, Engineering and Technology* 7, no. 13: 2772-84.

Ali, Salman Syed, Sayyid Tahir, and GM Arif.(1999) "Dynamics of Growth, Poverty, and Inequality in Pakistan [with Comments]." *The Pakistan Development Review* : 837-858.

Ali, Liaqat, MI Ramay, and Z Nas.(2013) "Analysis of the Determinants of Income and Income Gap between Urban and Rural Pakistan." *Interdisciplinary Journal of Contemporary Research in Business* 5, no. 1: 858-885.

Anwar, Talat. (2006)"Trends in Absolute Poverty and Governance in Pakistan: 1998-99 and 2004-05-05." *The Pakistan Development Review* : 777-793.

Atkinson, A.B. (1970)."On the Measurement of Inequality." Journal of Economic Theory. 2, 244-263.

Azam, Mehtabul, and Vipul Bhatt.(2016) "Spatial Income Inequality in India, 1993-2011: A District Level Decomposition. IZA Discussion Papers, No. 9892.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, *70*(5, Part 2), 9-49.

Bellù, Lorenzo Giovanni, and Paolo Liberati. (2006) "Policy Impacts on Inequality: Decomposition of Income Inequality by Subgroups." *EASYPol Module* 52.

Blinder, Alan S. (1973): Wage discrimination: Reduced Form and Structural Estimates. Journal of Human Resources. 8, 436–455.

Bourguignon, F. (1979): Decomposable Income Inequality Measures. Econometrica 47, 901–920.

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 for the Industrial Policy 2010, Lahore University of Management Sciences.

Canagarajah, Sudharshan., Newman Constance, & Ruchira Bhattamishra (2001). "Non-Farm Income, Gender, and Inequality: Evidence from Rural Ghana and Uganda," Food Policy, vol. 26: pp405-420.

Cowell, Frank A. (1980) "On the Structure of Additive Inequality Measures." *The Review of Economic Studies* 47, no. 3: 521-531.

Cowell, Frank A & Jenkins, Stephen P, (1995), 'How Much Inequality Can We Explain? A Methodology and an Application to the United States', Economic Journal, vol. 105, no. 429, pp. 421-430.

De Kruijk, Hans, and AR Kemal. (1987)"Sources of Income Inequality in Pakistan [with Comments]." *The Pakistan Development Review* 26, no. 4: 659-672.

DiNardo, J., Fortin, N. and T. Lemieux (1996): Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semi-Parametric Approach. Econometrica 64(5), 1001–1044.

Fields, G. (2003). _Accounting for Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States', in Solomon W. Polachek (ed.) Worker Well-Being and Public Policy (Research in Labor Economics, Volume 22) Emerald Group Publishing Limited, pp.1 – 38.

Haq, Rashida, and Nasim Shah Shirazi. (1998)"Trends in Inequality and Welfare in Consumption Expenditure: The Case of Pakistan [with Comments]." *The Pakistan Development Review* (1998): 765-779.

Idrees, Muhammad, and Eatzaz Ahmad. (2010)"Measurement and Decomposition of Consumption Inequality in Pakistan. The Lahore Journal of Economics 15:2: pp. 97-112.

Jamal, Haroon (2009), "Growth and Income Distribution for Poverty Reduction: Pakistan's Experience", Policy Brief . Number 1, UNDP-Pakistan.

Jamal, Haroon.(2014-15) "Growth and Income Inequality Effects on Poverty: The Case of Pakistan (1988-2011). MPRA Paper No. 59897, posted 14 November 2014-15 20:17 UTC.

Kemal, AR. (2006)"Income Inequalities in Pakistan and a Strategy to Reduce Income Inequalities." *Background Paper for PRSP-II, PRSP Secretariat, Government of Pakistan* .

Karunaratne, Hettige Don.(2000) "Age as a Factor Determining Income Inequality in Sri- Lanka." *The Developing Economies* 38, no. 2 (2000): 211-242.

Lambert P.J., Aronson J.R., 1993. Inequality Decomposition Analysis and the Gini Coefficient Revisited, Economic Journal, 103, 1221-1227

Jenkins, Stephen P., (1995). "Accounting for Inequality Trends: Decomposition Analyses for the UK, 1971- 86" Economica, Vol. 62, No. 245 pp. 29-63.

Juhn, C., K. Murphy and Pierce B., (1993), "Wage Inequality and the Rise in Returns to Skill" Journal ofPolitical Economy.

Kuznets, S. (1955) "Economic Growth and Income Inequality", The American Economic Review, Vol XLV, N° 1.

Leones, Julie P., and Shelley Feldman (1998). "Nonfarm Activity and Rural Household Income: Evidence from Philippine Microdata." Economic Development and Cultural Change 46: 789-806.

Milanovic, Branko.(2004-05) "Worlds Apart: Measuring Global and International Inequality." *Retrieved March* 8: 2011.

Morduch, J. and T. Sicular (2002). Rethinking inequality decomposition, with evidence from rural China. The Economic Journal 112, 93.106.

Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions *National Bureau of Economics Research*.No. 2.

Naschold, Felix.(2009) "Microeconomic Determinants of Income Inequality in Rural Pakistan." *The Journal of Development Studies* 45, no. 5: 746-768.

Nasir, Zafar Mueen, and Riaz Mahmood.(1998) "Personal Earnings Inequality in Pakistan: Findings from the Hies 1993-94." *The Pakistan Development Review* (1998): 781-792.

Nazli, Hina. (2004-05)"The Effect of Education, Experience and Occupation on Earnings:
Evidence from Pakistan. The Lahore Journal of Economics, Vol.9, No.2.

Okatch, Zelda, Abu Siddique, and Anu Rammohan.(2013) "Determinants of Income Inequality in Botswana discussion paper 13.15.

Pandey, Alok Kumar.(2014-15) "Household Level Inequality Decomposition in the Consumption Expenditure: Regression Approach. Journal of Exclusive Management Science - Vol 4 Issue – ISSN 2277 – 5684.

Shabbir, Tayyeb (1994) Mincerian Earnings Function for Pakistan. The Pakistan Development Review 26:4 687-696.

Siddiqui, R. (2008) . "Income, Public Social Services, and Capability Development: A Crossdistrict Analysis of Pakistan," PIDE-Working Papers 2008:43.

Siddiqui, Rehana, Rizwana Siddiqui, and M Ramzan Akhtar. "A Decomposition of MaleFemale Earnings Differentials [with Comments]." *The Pakistan Development Review* (1998): 885-898.

Shorrocks, A.F. (1980): The Class of Additively Decomposable Inequality Measures. Econometrica 48, 613–625.

Shorrocks, A.F. (1982) "Inequality Decomposition by Factor Components", Econometrica, 50, pp. 19211.

Shorrocks, A.F. (1984), "Inequality Decomposition by Population Subgroups", Econometrica, 52, pp. 1369-1385.

Shorrocks, Anthony F. and Guanghua Wan, (2004-05) "A Method for Generating Income Data from Lorenz Coordinates," unpublished manuscript, WIDER-UNU.

Silber, Jacques.(1998) "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality." *The Review of Economics and Statistics* : 107-115.

Skoufias, E., & Olivieri, S. (2013).Geographic disparities in well-being and fiscal expenditures in Thailand: 2000 vs 2009.Journal of the Asia Pacific Economy,18(August (3)), 359–381.

Skoufias, Emmanuel, and Sergio Olivieri. (2013)"Sources of Spatial Welfare Disparities in Indonesia: Household Endowments or Returns?". *Journal of Asian Economics* 29 : 62-79.

Su, Biwei, and Almas Heshmati. (2013)"Analysis of the Determinants of Income and Income Gap between Urban and Rural China." *China Economic Policy Review* 2, no. 01 1350002.

Tripathi, Sabyasachi.(2016) "Source of Inequality in Consumption Expenditure in India: A Regression Based Inequality Decomposition Analysis. MPRA Paper No. 72117, posted 20 14:10 UTC.

Wan, G.H., (2004-05), "Accounting For Income Inequality in Rural China: A Regression-Based Approach", Journal of Comparative Economics, vol 32, no. 2, pp 348-63.

APPENDIX

Regression Based Decomposition Methodology

Derivation of the Model

Field (2003) proposed a regression based decomposition methodology to explain the level of income inequality computed by an inequality index. In this methodology first income earning fucntion is estimate in which income is expressed as a function of variable such as age, gender , education etc. The decomposition is based on the estimate of the income earning function similar to equation 6)

$$lny_{jt} = \beta_o + \sum_{k=1}^{K} \beta_k X_{kjt} + \ \varepsilon_{jt}$$

Where lny is the percapita income, the subscript j refers to each housheold, t denotes the year,and k denotes the number of varaibles.

The income earning function given in above equation can be rewritten in matrix form as

$$lny_{jt} = \ \beta_t Z_{jt}$$

Where

$$\beta_t = \ [\beta_t \beta_{1t} \beta_{2t} \dots \dots \beta_{kt}\ 1]$$

$$Z_{jt} = \ \left[1\ X_{j1t} X_{j2t} \dots \dots X_{jkt} \varepsilon_{jt}\right]$$

The present regression based decomposition focuses on the decomposition of the logarithim income variance which is a measure of inequality.

Theorem

Let $L_1, \dots . L_P$ and $M_1, \dots . M_Q$ be two set of random variables, and $l_1, \dots l_p$ and $m_1, \dots m_q$ be two set of constant. Mood and Boas proposed a theorem according to this theorem covariance between the two random variable can be written as

$$\boldsymbol{Cov}\left[\sum_{\boldsymbol{p=1}}^{\boldsymbol{P}} \boldsymbol{l_p L_p}, \sum_{\boldsymbol{q=1}}^{\boldsymbol{Q}} \boldsymbol{m_q M_Q}\right] = \sum_{\boldsymbol{p=1}}^{\boldsymbol{P}} \sum_{\boldsymbol{q=1}}^{\boldsymbol{Q}} \boldsymbol{l_p m_q}\ \boldsymbol{Cov}\left[\boldsymbol{A_p B_q}\right]$$

In the context of single random variable y when applying this theorem such that

$$lny = \sum_{k=1}^{K+2} \beta_k Z_k$$

The covariance for the single variable y is obtained as

$$Cov\left[\sum_{k=1}^{K+2} \beta_k Z_k, lny\right] = \sum_{k=1}^{K+2} Cov\left[\beta_k Z_k, lny\right]$$

In equation 6 left hand side shows the covariance between lny and itself, it is simply the variance of lny. Thus equation 6 can be rewritten as

$$\sigma^2(lny) = \sum_{k=1}^{K+2} Cov\left[\beta_k Z_k, lny\right] \qquad A)$$

Equation B is obtained by dividing equation A by $\sigma^2(lny)$.

$$100\% = \sum_{k=1}^{K+2} \frac{Cov[\beta_k Z_k, lny]}{\sigma^2(lny)} = \sum_{k}^{K+2} s_k(lny) \qquad B)$$

Where $s_j\,(lny)$ is the relative factor inequality weight. The relative factor inequality tells us how much inequality in percapita income is explain by variable included in the model such as age, gender, education etc.

$$s_k(lny) = \frac{Cov[\beta_k Z_k, lny]}{\sigma^2(lny)}$$

If the last element of Z is excluded which is the factor inequality weight of the error term, then the remaining relative factor inequality weight can be represented as

$$\frac{\sum_{k=1}^{K+1} Cov\left[\beta_k Z_k, lny\right]}{\sigma^2(lny)}$$

The sum of this equation is exactly equal to goodness of fit of the regression model. Equation 8 shows the ordinary correlation coefficient is related to the covariance

$$Cor\left[\beta_k Z_k, lny\right] = Cov[\beta_k Z_k, lny]/\sigma(\beta_k Z_k)\sigma(lny)$$

let s_j (lny) denotes the share of log variance of income that is explained by the kth explanatory factor and R^2(lny) be the fraction of the log variance that is explain by all of the Z variables taken together except the error term. Then the log variance of income can be decomposed as.

$$s_k(lny) = Cov\,[\beta_k Z_k, lny]/\sigma^2(lny)$$

$$s_k(lny) = \beta_k * \sigma(Z_k) * Cor[Z_k, lny]/\sigma(lny)$$

Where

$$\sum_{k=1}^{K+2} s_j(lny) = 100\%$$

and

$$\sum_{k=1}^{K+1} s_j(lny) = R^2(lny)$$

Table A1 List of Districts

Punjab		KPK	Balochistan
Islamabad	Lodhran	Swat	Quetta
Attock	D.G Khan	Upper Dir	Pashin
Rawalpindi	Rajanpur	Lower Dir	QillahAbdullah
Jehlum	Layyah	Chitral	Chaghi
Chakwal	Muzaffar Garh	Shangla	Sibbi
Sargodha	Bahawalpur	Malakand	Ziarat
Bhakhar	Bahawalnagar	Bonair	Deara Bughti
Khushab	Rahim YarKhan	Peshawar	Kalat
Mianwali	**Sindh**	Charsada	Mastung
Faislabad	Khairpur	Nowshera	Khuzdar
Jhang	Sukkur	Kohat	Awaran
T.T singh	Nawabshah	Karak	Kharan
Gujranwala	NowsheroFeroze	Hangu	Lasbilla
Gujrat	Ghotki	D I Khan	Ketch/ Turbat
Sialkot	Jaccobabad	Tank	Gwadar
Hafizabad	Shikarpur	Manshera	Panjgur
MandiBahuddin	Larkana	Abbotabad	Zohb
Narowal	Dadu	Batagram	Loralai
Lahore	Hyderabad	Kohistan	Barkhan
Kasur	Badin	Haripur	Musa Khel
Okara	Thatta	Bannu	Qilla Saifullah
Sheikhupura	Sanghar	LakkiMarwat	Nasirabad
Vehari	Mirpur Khas	Mardan	Jafarabad
Sahiwal	Tharparker	Swabi	Jhal Magsi
Multan	Karachi		Bolan/ kachhi
Khanewal			
Pakpattan			

Table B Gini of percapita income 2004-05 and 2014-15

Punjab					
District	**Gini of per capita income 2004-05**	**Gini of per capita income 2014-15**	**District**	**Gini of per capita income 2004-05**	**Gini of per capita income 2014-15**
Islamabad	0.5	0.515	Lodhran	0.38	0.393
Attock	0.42	0.429	D.G Khan	0.44	0.43
Rawalpindi	0.43	0.423	Rajanpur	0.35	0.362
Jehlum	0.4	0.398	Layyah	0.41	0.393
Chakwal	0.38	0.494	Muzaffar Garh	0.36	0.393
Sargodha	0.43	0.428	Bahawalpur	0.44	0.391
Bhakhar	0.43	0.405	Bahawalnagar	0.42	0.43
Khushab	0.45	0.409	Rahim Yar Khan	0.44	0.412
Mianwali	0.43	0.421			
Faislabad	0.44	0.372		**Sindh**	
Jhang	0.39	0.409			
T.T singh	0.4	0.418	Khairpur	0.347	0.4
Gujranwala	0.39	0.43	Sukkur	0.391	0.4
Gujrat	0.35	0.379	Nawabshah	0.428	0.38
Sialkot	0.39	0.355	Nowshero Feroze	0.476	0.36
Hafizabad	0.39	0.364	Ghotki	0.364	0.47
Mandi Bahuddin	0.35	0.364	Jaccobabad	0.376	0.46
Narowal	0.45	0.428	Shikarpur	0.347	0.4
Lahore	0.45	0.431	Larkana	0.397	0.43
Kasur	0.37	0.41	Dadu	0.39	0.38
Okara	0.38	0.383	Hyderabad	0.434	0.4
Sheikhupura	0.37	0.393	Badin	0.548	0.41
Vehari	0.35	0.399	Thatta	0.345	0.32
Sahiwal	0.38	0.417	Sanghar	0.394	0.4
Multan	0.42	0.409	Mirpur Khas	0.448	0.44
Khanewal	0.39	0.381	Tharparker	0.353	0.38
Pakpattan	0.43	0.381	Karachi	0.463	0.48

Table C Gini of percapita income 2004-05 and 2014-15.

KPK			Balochistan		
District	Gini of per capita income 2004-05	Gini of per capita income 2014-15	District	Gini of per capita income 2004-05	Gini of per capita income 2014-15
Swat	0.55	0.35	Quetta	0.38	0.38
Upper Dir	0.47	0.37	Pashin	0.41	0.3
Lower Dir	0.44	0.37	Qillah Abdullah	0.31	0.34
Chitral	0.39	0.35	Chaghi	0.35	0.34
Shangla	0.37	0.33	Sibbi	0.36	0.4
Malakand	0.38	0.49	Ziarat	0.48	0.41
Bonair	0.50	0.38	Deara Bughti	0.23	0.35
Peshawar	0.49	0.41	Kalat	0.35	0.39
Charsada	0.46	0.39	Mastung	0.35	0.33
Nowshera	0.36	0.41	Khuzdar	0.41	0.36
Kohat	0.40	0.43	Awaran	0.34	0.31
Karak	0.45	0.42	Kharan	0.33	0.28
Hangu	0.44	0.41	Lasbilla	0.36	0.34
D I Khan	0.44	0.4	Ketch/ Turbat	0.33	
Tank	0.48	0.38	Gwadar	0.50	0.28
Manshera	0.45	0.42	Panjgur	0.40	
Abbotabad	0.42	0.47	Zohb	0.39	0.31
Batagram	0.40	0.41	Loralai	0.39	0.39
Kohistan	0.42	0.39	Barkhan	0.34	0.31
Haripur	0.39	0.38	Musa Khel	0.44	0.43
Bannu	0.42	0.43	Qilla Saifullah	0.36	0.38
Lakki Marwat	0.52	0.44	Nasirabad	0.39	0.39
Mardan	0.38	0.38	Jafarabad	0.25	0.38
Swabi	0.37	0.38	Jhal Magsi	0.32	0.5