ISSN (Print Version) 0975-3931
ISSN (On Line Version) 2278-1277

Inequality in consumption expenditure at household level in Uttar Pradesh: Regression Based Inequality Decomposition (RBID)

Alok Kumar Pandey[1]

The particular method of decomposition depends on the nature of the inequality index and the way in which it is decomposed since the decomposability of the indices differs from measure to measure. The issue relating to subgroup inequality levels to overall inequality has been discussed in a number of recent studies like Cowell 1980, Cowell and Kuga 1981, Bourguignon, 1979, Shorrocks 1980 and 1984, Shorrocks and Mukherjee, 1982, Das and Parikh 1982, Mishra and Parikh 1992. If the total inequality can be expressed as a function of sub-group inequality values, when the sub-groups are mutually exclusive and exhaustive, then variety of ways are found to decompose the total inequality. Many researchers use most popular method for decomposition i.e. additive decomposability. Additively decomposable can be expressed as the sum of a "between-group" term and a "within-group" term. Between-group component is defined as "the value of the inequality index when all the within-group inequalities are assumed to be non-existent by a hypothetical assignment of the group average income to each member of the same group".

The common inequality indicators can be used to assess the major contributors to inequality, by different subgroups of the population and by region. For example, average income may vary from region to region, and this alone implies some inequality "between groups." Income may vary inside each region, adding a "within-group" component to total inequality. For policy purposes, it is useful to be able to decompose these sources of inequality: if most inequality is due to disparities across regions, for instance, then the focus of policy may need to be on regional economic development, with special attention to helping the poorer regions. Inequality decomposition is important for understanding the main determinants of inequality and for policy analysis. The traditional approach to the subject was based purely on the analysis of the mathematical properties of inequality indices and is open to the criticism that the formal requirements for exact decomposition are perhaps too

[1] Assistant Professor, Department of Economics, DAV PG College, Varanasi, India. Email: pandey.alok@rediffmail.com

demanding for some practical applications (Fiorio and Jenkins 2007). It allows inequality accounting but not a causal analysis. Recent applied work in the field of inequality decomposition is focused on the use of regression-based (RB) approaches to avoid some of the restrictions of the traditional methods.

A range of different applications of the regression-based income inequality decomposition literature exist. Among these, note can be made of the extension by Morduch and Sicular (2002) where the composition of income by the different sources is observed, and the case where different income sources are accounted for by different income regimes in Farm-household income (Bardham and Boucher 1998, Yuko et al. 2006 and Kimhi 2007). While encountering a constant as a source of income inequality in empirical analysis of income distribution is possible but rare, the presence of a constant is a rule rather than an exception in regression equations.

It is important to note that expenditures are not similar with income, which may even be a better indicator of well-being, for various reasons. Among them there is the possibility of consumption without expenditures at least within the same period. Rightly observed by the Atkinson (1998), "Expenditures are supposed to better reflect "long-term" or "permanent" income and are from this point of view considered to be a better measure of economic well-being and respective inequalities".

Consumption expenditure data are collected in many developing countries including India, the process is time taking, costly and needs adjustment for household size, composition and for price level. Owing to these difficulties, the economic proxies (consumer durables, housing quality and household amenities) are collected to measure the economic status of the households in both small and large-scale population based surveys. The traditional approach to the subject was based purely on the analysis of the mathematical properties of inequality indices and is open to the criticism that the formal requirements for exact decomposition are perhaps too demanding for some practical applications. It allows inequality accounting but not a causal analysis. Although decompositions of the inequality index by income components appeared in the literature in 1970s, Recent applied work has reawakened interest in inequality decomposition by focusing on the use of regression-based (RB) approaches to avoid some of the restrictions of the traditional methods.

The present study focuses on the regression based inequality decomposition based on household level consumption expenditure for the year 2005-06, 2006-07 and 2007-08. In the light of above the section II deals with the Review of Literature. Data and methodology have been presented in section III. Section IV consists of Empirical results and section V represents conclusion of the study.

Review of Literature:

This issue was also addressed by Houthakker (1957) as early as in the 1950s. The issues related to household expenditures and consumption have been disregarded in

sociology and particularly empirical sociological research to a large degree, although family and household budget data frequently used for empirical study in early days.

Although there is a long history of research on patterns of household expenditures and their changes across time, which goes back to the 19th century and the famous work by Ernest Engel and others, these questions have attracted surprisingly little attention in recent years. Atkinson (1970), Sen (1973), Kakwani (1980) and Shorrocks (1982) proposed several concepts and measures to characterise inequalities in income distribution.

Oaxaca (1973) and Blinder (1973) first use Regression-based inequality decomposition measurement for between-group differences in the mean of income. Juhn et al. (1993) applied this approach to allow for the decomposition of between-group differences in the full wage distribution.

DiNardo et al. (1996) and Deaton (1997) proposed semi-parametric and nonparametric techniques that sought to model and compare the whole distribution of income in terms of density functions. Fields (2003) developed regression-based decomposition to analyze labor earnings inequality in the United States for variables such as gender, race, experience, schooling, occupation, industry, and region.

Morduch and Sicular (2002) in their study adopted the regression-based approach to decompose the inequality of the average income of 259 farm households of 16 villages in Zouping County (Republic of China,) during 1990-1993. The variables were grouped into regional segmentations, human capital accumulation, and political variables. This study decomposed four inequality measures, namely, Theil-T, squared CV/variance, alternative CV, and the Gini coefficient, to quantify the sources of the inequality. They found that the contribution of political variables was relatively small whereas the contributions of spatial characteristics were large in all decompositions.

However, the relative contribution of each factor was highly sensitive to the decomposition rule. This sensitivity became a limitation of the Morduch and Sicular (2002) method. For example, while Theil-T indicated that human capital and demographic variables were strongly inequality-reducing; the Gini decomposition stated that these factors contributed moderately to the inequality.

Adams (2002) followed the regression-based decomposition to examine the income inequality in rural Egypt based on 1,301 rural households in 1997. He found that although agricultural income was the second leading income source, it accounted for the highest share of the inequality, 40.2 percent, whereas nonfarm income, the largest income source in the rural Egypt, contributed 29.7 percent of the inequality.

Wan (2002 and 2004) in their study outlined some shortcomings. They observed that the main difficulty is associated with the failure to carry out an exact decomposition of the estimated sources including the error term, putting aside the

functional form or the inequality index adopted. Heshmanti (2004 and 2006) pays some attention to the interrelationship between income inequality and the non-income inequality dimensions admitting that while inequality can have many dimensions, economists have long been concerned primarily with the monetary dimension.

Similar results were found by Arayama, Kim, and Kimhi (2006) who also extended the regression-based decomposition method suggested by Morduch and Sicular (2002) to estimate regime-specific income-generating functions of Korean farm household with a micro dataset collected in 2003. They found that family size and its composition as well as land ownership were the main determinants of the inequality in Korean farm household.

There are many studies concerned with the inequality in rural China specifically with its high economic growth and high inequality since the late 1970s economic reform. Wan and Zhou (2005) combined both Shapley value framework of Shorrocks (1999) and the regression-based decomposition proposed by Morduch and Sicular (2002) to examine the determinants and the changes of income inequality in rural China with household-level datasets during 1995-2002. To represent the level of development, they selected three villages from each of three provinces; Guangdong, Hubei, and Yunnan. The share of cropping patterns accounted for larger impacts than those of labour and human capital inputs in the rural areas, the Chinese government should improve rural credit services and raise returns for grain-cropping in order to reduce the income inequality in this region. They also suggested that even the impact of education on inequality was rather small but its role would be higher in the near future because of more skill-demanding labour. As observed by Wan and Zhou conventional decompositions by factor components or by population subgroups only provide limited information on the determinants of income inequality.

In this context, the regression-based decomposition (RBD) analysis can shed more light to our understanding of factors that determine income inequality (Oyekale et al, 2007). Cain, Hasan, Magsombol, and Tandon (2008) also adopted regression-based decomposition proposed by Fields (2003) on consumer expenditure surveys to investigate the poverty and the inequality of 17 major states of India in 1983, 1993, and 2004. The explanatory variables were age, gender, social group, production sector, occupation, level of education, and state of residence. Epo et al, (2010) explores determinants of economic welfare econometrically and applies the parameter estimates to compute income sources that explain measured inequality in the distribution of household income. Bay and Epo (2011) in their paper used regression-based inequality decomposition approach to explore determinants of income inequality in Cameroon using the 2007 Cameroon household consumption survey. They used a control function approach that tests for potential endogeneity and unobserved heterogeneity of synthetic variables for education and health, while controlling for other correlates of household consumption to generate contributions of regressed-sources to measured

income inequality. In this paper regressed income sources attributable to education, health, urban residency, household size, fraction of active household members, working in the formal sector and farmland ownership are the main determinants of household income inequality in that order.

In the light of above review of literature it is observed that no significant study is available with respect to the regression based inequality decomposition based on age of the head of household, gender of the head of household, marital status of the head of household, education level of the head of household, household size, household type (activities for the source of income), religion, social group, land possessed, dwelling unit, type of structure, primary source of for cooking, primary source of energy for lighting and sector. It is an area which deserved greater attention to be paid as inequality decomposition based on regression in terms of various characteristics of the head of household and socio demographic characteristics.

The broad aim of the study is to identify the inequality decomposition by focusing on the use of regression-based (RB) approach at households' level. The study also describes the level of inequality in terms of expenditure on various characteristics of the head of households like age, gender, marital status and education level and socio demographic characteristics like household size household type religion social group land possessed dwelling unit type of structure primary source of energy for cooking primary source of energy for lighting and sector.

Uttar Pradesh: State Economy (At a Glance)

Uttar Pradesh is India`s fourth largest and the most populated state in India situated between 23o 52'N and 31o 28 N latitudes and 77o 3' and 84o 39'E longitudes with an area of 93,933 sq mi (243,286 square km), in upper Gangetic plain. There is an average population density of 828 persons per km² i.e. 2,146 per sq meters. In north, Uttar Pradesh shares an international border with Nepal. Other states along Uttar Pradesh's border include Uttarakhand, Haryana and Delhi to the north and northwest, Rajasthan on the west, Madhya Pradesh on the south, Chhattisgarh and Jharkhand on the south east and Bihar on the east. The administrative and legislative capital of Uttar Pradesh is Lucknow.

Geography and Climate:

The climate of the state is predominantly subtropical, however, weather conditions change significantly with location and season. The average temperatures vary from between 12.5– 17.5 °C (55–64 °F) in January to 27.5–32.5 °C (82–91 °F) in May and June. Rainfall in the state ranges from between 1,000–2,000 mm (39–79 in) in the east to 600–1,000 mm (24–39 in) in the west. About 90 percent of the rainfall occurs during the southwest monsoon, lasting from approximately June to September. Uttar Pradesh can be divided into three distinct hypsographical regions i.e. (i) The

Himalayan region in the North, (ii) The Gangetic plain in the centre and (iii) The Vindya hills and plateau in the south

Population and Education:

The state of Uttar Pradesh is expanded in an area of 240,928 sq. km. There are 18 Mandals, 75 Districts, 312 Tehsils 813 Blocks and 106704 Villages. The population of Uttar Pradesh was recorded 199581477 persons during the census 2011 out of which 155.11 million (77.72 percent) are rural and 44.47 million (22.28 percent) are urban. There is an addition of 33383556 persons from census 2001 to census 2011. The population in rural areas declined 1.50 percent during 2001-11. The state shares largest rural percentage i.e. 18.62 percent of the country's rural population during 2001-2011. The population growth rate of Uttar Pradesh was recorded 20.09 percent during 2001 to 2011. During 2001 to 2011 the decadal growth rates in rural and urban areas were recorded as 17.81 percent and 28.75 percent respectively. As per Census 2011, Uttar Pradesh has got 908 females per 1000 males and child sex ratio has been recorded as 899.

The total literate population in Uttar Pradesh was recorded at 118.42 million persons as per census 2011 out of this 88.39 million belong to rural areas and 30.02 million belong to urban areas. Total literacy rate in Uttar Pradesh was recorded 69.72 percent as per census 2011. This shows that there is a significant increase of 45 percent over Census 2001. In state rural and urban literacy rates are 67.55 percent and 77.01 percent respectively. The decadal change indicates that there is an increase in literacy rate by 15.02 percent and 7.26 percent in rural and urban areas respectively.

Male literacy in Uttar Pradesh is recorded as 78.48 percent in rural areas and 81.75 percent in urban areas during the Census 2011. There is a significant increase in male literacy of the state in rural as well as in urban areas i.e. 11.89 percent and 4.99 percent respectively during 2001-11. Female literacy in rural areas of Uttar Pradesh is recorded at 55.61 percent and in urban areas it is recorded at 71.68 percent during the Census 2011. Female literacy rate in rural areas and urban areas increased by 18.71 percent and 9.95 percent respectively during the census 2001 to census 2011.

Poverty and Unemployment:

Poverty prevents people from advancing in economic and social spheres of life. National Sample Survey (NSS) data is used to form the basis for computation of poverty statistics. The poverty rates for rural and urban areas of the State are found to be 30.7 and 32.9 % respectively. Updated poverty lines were used in conjunction with the 2007-08 MPCE distribution of state sample data of 64^{th} round NSS. In Uttar Pradesh, 19.2 percent population (19.2 percent rural, 19.5 percent urban) was found to be below the poverty line in 2007-08. A steeper fall in rural poverty as compared to

urban poverty resulted in the pattern that urban poverty rate in the state now surpasses the rural poverty rate.

The major challenge before economy of Uttar Pradesh is unemployment problem. The different rounds of NSS give the unemployment situation. The unemployment rate in 43rd, 50th, 55th and 61st round was 3.73, 3.70, 4.47 and 4.61 percent respectively.

The above analysis shows that Uttar Pradesh is one of the backward states in India. The slow growth of the state economy can be attributed to various factors. Political willingness is one of the major factors which affected pace of economic development though there was political stability in the state. Low productivity in agriculture and allied sectors has adversely affected employment and income generation. Poor industrial infrastructure along with low level of investment is the major reason for the slow growth of industrial sector in the state. The development of potential sectors like handloom, tourism, bio-technology, IT needs an integrated approach for the speedy development of the state.

Data and Methodology for calculating regression based inequality decomposition:

The National Sample Survey Organisation (NSSO) conducted an all-India survey of households on participation and expenditure on education, employment, unemployment, migration and consumer expenditure on a regular basis for over four decades. The present study utilised unit level data from the 62nd, 63rd and 64th round of NSSO consumer expenditure survey collected in the years 2005-06, 2006-07 and 2007-08 respectively.

Decomposition of Income Inequality: Regression Based Inequality

Morduch and Sicular (2002) and Fields (2003) extended the decomposition of inequality to regression-based inequality decomposition by determinants of income. They suggested expressing household income (or log-income) as:

$$y=X\beta+\varepsilon, \quad (1)$$

Where, X is a (nxk) matrix of explanatory variables (including a constant),

β is a (kx1) vector of coefficients, and

ε is a (nx1) vector of random error terms. Given a vector of consistently estimated coefficients b, income can be expressed as a sum of predicted income and a prediction error as:

$$y = x\hat{b} + \hat{\varepsilon} \quad (2)$$

Per capita income of household is represented as (Cowell and Fiorio 2006):

ISSN (Print Version) 0975-3931
ISSN (On Line Version) 2278-1277

$$y_i = \sum_{m=1}^{M} \hat{b}_m x_i^m + \hat{\varepsilon}_i \qquad (3)$$

Shorrocks (1982) suggested that inequality measures can be written as a weighted sum of incomes i.e.

$$\mathrm{I(y)} = \sum_{i=1}^{n} a_i(y) y_i \qquad (4)$$

where, ai are the weights, yi is the income of household i, and **y** is the vector of household incomes.

Substituting (1) into (4) and dividing by I(y), the share of inequality attributed to explanatory variable m is obtained as

$$\mathrm{s_m} = \mathrm{b_m} \sum_{i=1}^{n} a_i(y) x_i^m / \sum_{i=1}^{n} a_i(y) y_i \qquad (5)$$

Using the regression coefficients, it is possible to compute the "income shares" of the explanatory variables as

$$\alpha_{\mathrm{m}} = \mathrm{b_m} \sum_{i=1}^{n} x_i^m / \sum_{i=1}^{n} y_i \qquad (6)$$

and evaluate the marginal effect on the Gini index of inequality of a uniform increase in an explanatory variable m, as in Lerman and Yitzhaki (1985) by computing s_m-$\alpha_m G(Y)$

In the present study inequality and inequality decomposition of income and household expenditure has been calculated in respect to age of the head of household, gender of the head of household, marital status of the head of household, education level of the head of household, household size, household type, religion, social group, land possessed, dwelling unit, type of structure, primary source of energy for cooking, primary source of energy for lighting and sector.

Fiorio and Jenkins (2007) developed Regression-based inequality decomposition (ineqrbd for STATA), by using Fields (2003) and Shorrocks decomposition rule. According to model, the Y_i and X_i variables based on n observations estimates following relationship as

$$y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots\ldots\ldots\ldots + \beta_k X_k + \mu \qquad \ldots\ldots\ldots\ldots\ldots (7)$$

The model can be rewritten as;

$$Y_i = \beta_0 + Z_1 + Z_2 + Z_3 + \ldots\ldots\ldots\ldots\ldots + Z_k + \mu_1 \qquad \ldots\ldots\ldots\ldots\ldots (8)$$

Z_1, Z_2, Z_3, ...and Z_k are composite variable, product of regression coefficient and variables. For inequality decomposition calculations, the value of β_0 is irrelevant as it is constant for every observation. The predicted value $\hat{Y}$

$$\hat{Y}_i = \beta_0 + Z_1 + Z_2 + Z_3 + \ldots\ldots\ldots\ldots\ldots + Z_k \qquad \ldots\ldots\ldots\ldots\ldots(9)$$

Equations (8) and (9) are of exactly the same form as the Shorrocks (1982) used equation for deriving inequality decomposition by factor components (For example total income is the sum of labour earnings, income from savings and other assets, private and public transfers.

Alternatively, one may apply the decomposition rule to the inequality of $\hat{Y}$ itself, in which case there is also a decomposition term corresponding to the residual (Cowell and Fiorio, 2006.In STATA, ineqrbd provides a regression-based Shorrocks-type decomposition of a variable labelled "Total", where Total is defined as $\hat{Y}$, unless the fields option is used in which case Total refers to predicted $\hat{Y}$. In either case, the contribution to inequality in Total of each term is labelled "s_f" in the output (From help for ineqrbd in STATA , Carlo V. Fiorio and Stephen P. Jenkins, December 2008) .

In ineqrbd modules provides the means, standard deviations, and correlations, of Total, the residual and the composite variables Z_1, Z_2, ... Z_k. Results for the composite variables are ordered in the same order as the underlying variables are ordered in Z_1, Z_2, ... Z_k..

Also I_2 summarises inequality using half the squared coefficient of variation (the Generalized Entropy measure I2), rather than the coefficient of variation (CV). Based on various empirical studies it is observe that inequality may be negative, e.g. because the mean of a composite variable may be negative. In this paper right hand side variables are as follows:

X_1= Head of household (h_age)

X_2= Gender of the head of household (h_gen)

X_3= Marital status of the head of household (h_edu)

X_4= Education level of the head of household (h_maritulstatus)

X_5= Household size (hhsize_c)

X_6= Household type (Source of Earning in rural and urban areas, i.e. htype_c)

X_7= Religion (religion_c)

X_8= Social group (Socialgroup_c)

X_9= Land possessed (land_code)

X_{10}= Dwelling unit (dwelling_c)

X_{11}= Type of structure (structure_c)

X_{12}= Primary source of energy for cooking primary (cook_c)

X_{13}= Source of energy for lighting (light_c) and

X_{14}= Sector (Sector)

In the paper results are presented as (http://www.stata.com/stb/stb48/sg104/ineqfac.hlp):

The decomposition rule is the proportionate contribution of factor f to total inequality (for f=1, 2, ..., 14) , s_f: s_f = rho_f * sd(factor_f) / sd(totvar)

Where, rho_f is the correlation between factor_f and total variable, and sd(.) is the standard deviation. (Equivalently, s_f is the slope coefficient from the regression of factor_f on totvar.) For each observation, $\sum_{1}^{F}(s\_f)=1$, and S_f = s_f*I2(Total), Mean: m_f = mean(f); Standard Deviation: sd(f) = std.dev. of f. The member of the Generalised Entropy class of inequality measures, I2_f = 0.5*[sd(f)/m_f]^2

Findings of the study:

Maximum percentage of the head of households belong to 30-39 years of age group and minimum percentage of head of households belong to 30 years of age. But the Average MPCE is found the maximum for the head of households having below 30 years of age of the head of household and minimum for the head of households who belong to 30-39 years of age group for the three rounds of NSSO unit level data. Out of the total head of households 89.39 percent were male and only 10.61 percent were female. In the year 2005-06 and average MPCE was maximum i.e. ` 683 for female head of households. The same trend was observed for the year 2007-08.

43.16 percent head of households were illiterate in the year 2005-06, 42.28 percent head of households were not literate in the year 2006-07 and 43.83 percent were not literate in the year 2007-08. Average monthly per capita consumption expenditure for the illiterate head of households was minimum for the year 2005-06 and 2007-08, while AMPCE was minimum for the head of households having literate without formal schooling for the year 2006-07. The minimum percentage of the head of households were having diploma/certificate course for all the three rounds of NSSO survey but the AMPCE of such households having high. This shows that technical knowledge provides better opportunity of earnings. Thus, the state government focus

on such technical courses at diploma or certificate level which provide immediate better job opportunities or entrepreneur's in the economy. This will obviously increase the income level of households as well as state as whole. In the state, for the three rounds of survey the majority of head of households are married but their AMPCE is minimum among all the categories like unmarried, divorcee/separated and widow. Hindu households were having maximum percentage while Muslims households were having minimum AMPCE among all the religions in states

In 2005-06, 39.79 percent households belonged to Other Backward Class, 20.10 percent to Schedule Caste and 9.17 percent to Schedule Tribe. The same trend was found in the year 2006-07. In 2007-08, 41.28 households belonged to other backward classes and 19.77 percent to schedule cast and 8.97 belonged to schedule tribe. AMPCE was maximum for the others categories social groups and minimum for ST groups in the state for the three rounds under study.

The majority of the households were self-employed in agriculture activities in rural areas for all the rounds of survey under study. The expenditure of households involved in other activities especially in urban areas for all the three rounds of NSSO survey reported maximum AMPCE i.e. ` 1137, ` 1259 and ` 1416 respectively. As per 62$^{nd}$ rounds of survey, 13.40 percent households were having two or less than two members in the family and their AMPCE was ` 1108. Majority of the families i.e. 33.56 percent had seven or above members in the family and their average AMPCE was ` 541. As per 63$^{rd}$ round of the survey, majority i.e. 30.30 percent families had seven or above seven members in their families with an average AMPCE of ` 631. 12.07 percent families had two and less than two family members with an AMPCE of ` 1197 for the year 2006-07. As per 64$^{th}$ round of survey majority of the households i.e. 29.70 percent had seven or above seven members in their families with an AMPCE of ` 650. As per 62nd round of survey majority i.e. 93.14 percent households had own houses and their average MPCE was ` 607 and 5.49 percent households had hired houses and their AMPCE was ` 1301. The same trend was found for 63rd and 64th rounds of survey. Land possession situation shows that approx one percent household posses land greater than 4 hectares in the states and majority of the households (85 %) posses one hectare or less than one hectare land in the state. The same trend can be seen for 63rd and 64th round of survey. Majority of the households were having firewood and chips as major source of energy for cooking i.e. 56.63 percent and their AMPCE was ` 544 in the year 2005-06. LPG as a source of energy for cooking in the household is contributing 16.86 percent and AMPCE was ` 1052 in the 62nd round of survey. As for as source of energy for lighting is concerned 58.66 percent households depended on kerosene and AMPCE for those household was ` 525 in the year 2005-06. The same trend was noticed for 63rd and 64th rounds of survey.

Regression based inequality decomposition regression results for the 62nd, 63rd and 64th rounds of NSSO unit level data are presented in Table 1. In the regression equations explanatory variables are h_age, h_gen, h_edu, h_maritulstatus, religion_c,

Socialgroup_c, dwelling_c, structure_c, land_code, hhsize_c, htype_c, cook_c, light_c and Sector for age of the head of households, gender of the head of households, marital status of the head of households, education level of the head of households, household size, household type, religion, social group, land possession, dwelling unit, type of structure, primary source of energy for cooking, primary source of energy for lighting and sector respectively and explained variable is Average Monthly per capita expenditure of the households. In the Table estimates of coefficients, their standard error and t statistics are given in the table. For the 62nd round, education level of the head of household, household size, household type, religion, social group, land possessed and primary source of energy for lighting are significant. Excluding household size above variables are positively affecting the average monthly per capita consumption expenditure. For 63rd round age of the head of household, marital status of the head of household, education level of the head of household, household size, household type, religion, social group, land possessed, type of structure and primary source of energy for lighting are significant. Marital status of the head of household, type of structure and household size are negatively affecting the average monthly per capita consumption expenditure. For 64th round, age of the head of household, education level of the head of household, household size, social group, land possessed, type of structure, primary source of energy for cooking, primary source of energy for lighting and sector are significant. Type of structure and household size are negatively affecting the average monthly per capita consumption expenditure. In this Table, R^2, Adjusted R^2, F statistics and total sample observations are also given.

Decomposition of inequality in AMPCE and predicted AMPCE for the Year 2005-06, is presented for Uttar Pradesh in Table 2. For inequality decomposition for AMPCE maximum value of s_f (= rho_f * sd(f) / sd(total) is for education of the head of household. This trend is also followed for predicted AMPCE. The size of households contributing 8.595 percent and 26.438 percent and type of household contributing 6.282 and 19.326 for average monthly per capita expenditure and predicted average monthly per capita expenditure in total inequality respectively. Social group and energy for lighting contributing 3.586 and 3.392 respectively for AMPCE and 11.032 and 11.009 respectively for predicted AMPCE. Structure of household for S_f (= s_f*I2(Total)) gives negative value in inequality decomposition of AMPCE and predicted AMPCE. The ratio S_f and I2_f for total is 0.218 in AMPCE and 0.071 for predicted AMPCE.

In the year 2006-07 the regression based inequality decomposition for Uttar Pradesh in terms of AMPCE and predicted AMPCE is given in Table 3. The maximum contribution in inequality is again education of the head of households followed by household size and type of household, social group and energy for light. Similar trend follows for the other measure of inequality decomposition. The ratio S_f and I2_f for total is 0.230 in AMPCE and 0.075 for predicted AMPCE.

ISSN (Print Version) 0975-3931
ISSN (On Line Version) 2278-1277

Table 4 presents regression based inequality decomposition for the Uttar Pradesh in the year 2007-08 in terms of AMPCE and predicted AMPCE. For the year 2007-08, Education of the head of households, size of households, rural and urban region, social groups and energy for lighting are contributing 10.859, 9.516, 4.877, 3.925 and 2.037 respectively in terms of AMPCE and 30.644, 26.854, 13.763, 11.077 and 5.747 respectively in terms of predicted AMPCE in inequality. The ratio S_f and I2_f for total is 0.246 in AMPCE and 0.087 for predicted AMPCE.

In Appendix 1, summary statistics like mean, standard deviation, minimum and maximum and in Appendix 2, correlation coefficient presented for Total, residual, and composite RHS variables for regression model for AMPCE and predicted AMPCE (equation 8 and equation 9) for the year 2005-06, 2006-07 and 2007-08.

Table 1: Regression Based Inequality decomposition: Regression Results for the 62nd, 63rd and 64th rounds of NSSO unit level data

	62nd Round			**63rd Round**			**64th Round**		
Variables	**Coef.**	**se**	**t***	**Coef.**	**se**	**t***	**Coef.**	**se**	**t***
h_age	11.48	5.08	2.26	40.86	5.14	7.94*	37.46	5.43	6.90*
h_gen	5.90	24.44	0.24	-59.13	25.73	-2.30	30.55	26.92	1.13
h_edu	41.45	2.20	18.81*	40.65	2.39	16.96*	53.74	2.62	20.45*
h_maritulsus	-2.56	19.27	-0.13	-57.80	18.30	-3.16*	-40.81	21.16	-1.93
religion	39.65	8.93	4.44*	37.75	9.00	4.19*	21.75	9.34	2.33
Socialgroup	19.01	2.35	8.08*	20.37	2.29	8.89*	26.80	2.55	10.49*
dwelling	-1.62	8.47	-0.19	2.09	7.86	0.27	23.32	9.18	2.54
Structure_hh	13.42	7.60	1.76	-43.69	7.89	-5.53*	-55.74	8.84	-6.31*
land	17.50	3.29	5.31*	27.10	3.17	8.54*	18.52	3.71	4.99*
hhsize	-75.52	3.74	-20.15*	-89.24	3.91	-22.79*	-106.90	4.25	-25.11*
htype	19.85	3.27	6.07*	20.09	3.09	6.50*	4.86	3.32	1.46
Cookenergy	7.00	3.78	1.85	-1.15	4.17	-0.28	22.29	4.58	4.86*
lightenergy	24.94	3.04	8.19*	26.58	3.03	8.75*	18.54	3.38	5.48*
Sector	-0.17	34.43	-0.01	11.36	31.50	0.36	197.50	34.47	5.73*
constant	480.47	64.21	7.48*	792.18	64.71	12.24	611.43	67.71	9.03*
R^2	0.325			0.326			0.354		
Adj R^2	0.323			0.324			0.353		
F(14, 39241)	130.2			178.0			197.1		
Observation	39256			5173			5043		

*statistically significant at the one percent level.

Note: Results are based on STATA 11.0 "ineqrbd" developed by Fiorio and Jenkins (2007); NSSO 62nd, 63rd , and 64th rounds unit level data has been used.

ISSN (Print Version) 0975-3931
ISSN (On Line Version) 2278-1277

Table 2: Regression-based decomposition of inequality in AMPCE and predicted AMPCE for the Year 2005-06

Inequality Decomposition	For AMPCE					For Predicted AMPCE				
	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)
residual	67.492	0.147	0.000	9.E+30	4.E+31					
h_age	-0.062	0.000	5.939	0.074	0.339	-0.191	0.000	5.939	0.074	1.043
h_gen	0.015	0.000	1.016	0.030	0.139	0.047	0.000	1.016	0.030	0.428
h_edu	10.598	0.023	23.768	0.304	1.392	32.601	0.023	23.768	0.304	4.282
h_maritulsus	-0.001	0.000	-0.851	0.012	0.056	-0.004	0.000	-0.851	0.012	0.171
religion	0.394	0.001	7.649	0.140	0.640	1.212	0.001	7.649	0.140	1.969
Social group	3.586	0.008	12.084	0.215	0.987	11.032	0.008	12.084	0.215	3.037
dwelling	-0.022	0.000	-0.283	0.192	0.882	-0.069	0.000	-0.283	0.192	2.712
Structure_hh	-0.427	-0.001	3.386	0.130	0.598	-1.313	-0.001	3.386	0.130	1.839
land	0.047	0.000	10.706	0.137	0.629	0.143	0.000	10.706	0.137	1.936
hhsize	8.595	0.019	-70.362	0.037	0.168	26.438	0.019	-70.362	0.037	0.516
htype	6.282	0.014	15.774	0.306	1.403	19.326	0.014	15.774	0.306	4.317
Cookenergy	0.116	0.000	3.379	0.120	0.551	0.357	0.000	3.379	0.120	1.694
lightenergy	3.392	0.007	11.009	0.291	1.334	10.435	0.007	11.009	0.291	4.105
Sector	-0.005	0.000	-0.033	0.054	0.245	-0.014	0.000	-0.033	0.054	0.754
Total	100.000	0.218	100.000	0.218	1.000	100.000	0.071	100.000	0.071	1.000

Note: Results are based on STATA 11.0 "ineqrbd" developed by Fiorio and Jenkins (2007); proportionate contribution of composite var f to inequality of Total, s_f = rho_f*sd(f)/sd(Total); S_f = s_f*I2(Total); m_f = mean(f); sd(f) = std.dev. of ;. I2_f = 0.5*[sd(f)/m_f]^2; NSSO 62nd rounds unit level data has been used. More details of various estimates visit http://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

Table 3: Regression-based decomposition of inequality in AMPCE and predicted AMPCE for the Year 2006-07

Inequality Decomposition	For AMPCE					For Predicted AMPCE				
	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)
residual	67.428	0.155	0.000	2.E+32	7.E+32					
h_age	0.242	0.001	18.053	0.071	0.310	0.742	0.001	18.053	0.071	0.952
h_gen	0.019	0.000	-8.702	0.027	0.116	0.059	0.000	-8.702	0.027	0.357
h_edu	9.227	0.021	20.428	0.294	1.280	28.327	0.021	20.428	0.294	3.930
h_maritulstatus	0.183	0.000	-16.534	0.013	0.058	0.560	0.000	-16.534	0.013	0.179
religion_c	0.396	0.001	6.253	0.144	0.625	1.217	0.001	6.253	0.144	1.919

Socialgroup_c	3.641	0.008	11.717	0.219	0.954	11.177	0.008	11.717	0.219	2.928
dwelling_c	0.021	0.000	0.320	0.220	0.959	0.065	0.000	0.320	0.220	2.945
structure_c	1.747	0.004	-9.645	0.133	0.577	5.362	0.004	-9.645	0.133	1.772
land_code	0.514	0.001	13.485	0.161	0.701	1.578	0.001	13.485	0.161	2.152
hhsize_c	7.251	0.017	-70.557	0.039	0.168	22.260	0.017	-70.557	0.039	0.516
htype_c	5.536	0.013	14.132	0.313	1.365	16.996	0.013	14.132	0.313	4.189
cook_c	-0.020	0.000	-0.440	0.123	0.537	-0.062	0.000	-0.440	0.123	1.650
light_c	3.554	0.008	10.104	0.309	1.346	10.911	0.008	10.104	0.309	4.131
Sector	0.264	0.001	1.890	0.056	0.244	0.809	0.001	1.890	0.056	0.749
Total	100.000	0.230	100.000	0.230	1.000	100.000	0.075	100.000	0.075	1.000

Note: Results are based on STATA 11.0 "ineqrbd" developed by Fiorio and Jenkins (2007); proportionate contribution of composite var f to inequality of Total, s_f = rho_f*sd(f)/sd(Total); S_f = s_f*I2(Total); m_f = mean(f); sd(f) = std.dev. of ;. I2_f = 0.5*[sd(f)/m_f]^2; NSSO 62nd rounds unit level data has been used. More details of various estimates visit http://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

Table 4: Regression-based decomposition of inequality in AMPCE and predicted AMPCE for the year 2007-08

Inequality Decomposition	For AMPCE					For Predicted AMPCE				
	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)	100*s_f	S_f	100*m_f/m	I2_f	I2_f/I2(total)
residual	64.566	0.159	0.000	4.E+31	2.E+32					
h_age	0.069	0.000	15.739	0.076	0.309	0.195	0.000	15.739	0.076	0.873
h_gen	0.002	0.000	4.259	0.030	0.123	0.005	0.000	4.259	0.030	0.348
h_edu	10.859	0.027	24.107	0.299	1.215	30.644	0.027	24.107	0.299	3.429
h_maritulstatus	0.063	0.000	-11.014	0.013	0.051	0.178	0.000	-11.014	0.013	0.145
religion_c	0.146	0.000	3.468	0.153	0.621	0.411	0.000	3.468	0.153	1.752
Socialgroup_c	3.925	0.010	14.050	0.210	0.852	11.077	0.010	14.050	0.210	2.404
dwelling_c	0.334	0.001	3.315	0.197	0.802	0.942	0.001	3.315	0.197	2.263
structure_c	1.967	0.005	-11.032	0.129	0.525	5.552	0.005	-11.032	0.129	1.482
land_code	-0.113	0.000	8.677	0.147	0.598	-0.318	0.000	8.677	0.147	1.689
hhsize_c	9.516	0.023	-79.648	0.036	0.148	26.854	0.023	-79.648	0.036	0.417

htype_c	1.229	0.003	3.291	0.294	1.195	3.468	0.003	3.291	0.294	3.372
cook_c	0.525	0.001	7.962	0.119	0.483	1.481	0.001	7.962	0.119	1.363
light_c	2.037	0.005	6.711	0.292	1.186	5.747	0.005	6.711	0.292	3.346
Sector	4.877	0.012	30.899	0.056	0.229	13.763	0.012	30.899	0.056	0.647
Total	100.000	0.246	100.000	0.246	1.000	100.000	0.087	100.000	0.087	1.000

Note: Results are based on STATA 11.0 "ineqrbd" developed by Fiorio and Jenkins (2007); proportionate contribution of composite var f to inequality of Total, s_f = rho_f*sd(f)/sd(Total); S_f = s_f*I2(Total); m_f = mean(f); sd(f) = std.dev. of ;. I2_f = 0.5*[sd(f)/m_f]^2; NSSO 62nd rounds unit level data has been used. More details of various estimates visit http://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

Conclusions and Suggestions:

In the present study an attempt has been made to focus on the inequality decomposition based on regression analysis in Uttar Pradesh developed by Morduch and Sicular (2002), Fields (2003) and Fiorio and Jenkins (2007) for the year 2005-06, 2006-07 and 2007-08. Data used in the study is consumption expenditure unit level data collected by National Sample Survey Organisation. The variables used for the decomposition based on regression for AMPCE are age of the head of household, gender of the head of household, marital status of the head of household and education level of the head of household, household size household type religion social group land possessed dwelling unit type of structure primary source of energy for cooking primary source of energy for lighting and sector. Education level of the head of household is contributing as the main factor for inequality followed by the size of household, social groups and energy for lighting in Uttar Pradesh for the years 2005, 2006 and 2007. Results of the study suggest that there is an urgent need for the enhancement in government expenditure on education especially in technical and higher level of education in the state. The results are also revealing the fact that in rural areas still job opportunities are not sufficient and or of low level for the period under study.

References:

Adams, R. H.Jr., (2001), "Non-Farm Income, Inequality and Land in Rural Egypt", Economic Development and Cultural Changes, vol 50, pp 339-363.

Atkinson, A., (1970), "The Economics of Inequality", Oxford University Press, Delhi.

Atkinson, Antony B., (1998), "Poverty in Europe", Oxford University Press, Blackwell.

Bardham, B., and S., Boucher, (1998), "Mitigation, Remittances and inequality: Estimating the Net effects of Migration on Income Distribution", Journal of Economics, vol 55, pp 307-331.

Baye F. M. And Epo B N. (2011), "Inequality Decomposition by Regressed-Income Sources in Cameroon1", Paper Prepared for the Special IARIW-SSA Conference on Measuring National Income, Wealth, Poverty, and Inequality in African Countries Cape Town, South Africa.

Blinder, A.S., (1973), "Wage Discrimination: Reduced Form and Structural Estimates", Journal of Human Resources, vol. 8, pp 436-55.

Bourguignon, F., (1979), "Decomposable income inequality measures", Econometrica vol. 47, pp. 901-920.

Bourguignon, F., M. Fournier and M. Gurgand, (2001), "Fast Development with a stable Income Distribution: Taiwan, 1979-1994", Review of Income and Wealth, vol. 47, no,2, pp 139-63.

Cain, J. Salcedo, Rana Hasan, Rhoda Magsombol, and Aja y Tandon (2008), "Accounting for Inequality in India: Evidence from Household Expenditures", ERD Working Paper No. 117, Asian Development Bank

Cowell, F. A., (1980), "On the structure of additive inequality measures", Review of Economic Studies vol. 47, pp. 521.531.

Cowell, F.A. and Fiorio, C.V., (2006), "Rethinking Inequality Decomposition: Comment" Distributional Analysis Research Programme Working Paper 82, STICERD, London School of Economics, London, http://sticerd.lse.ac.uk/dps/darp/DARP82.pdf

Cowell, F.A. and Kuga,K., (1981), "Additivity and the entropy concept: an axiomatic approach to inequality measurement", Journal of Economic Theory vol. 25, no. 1, pp. 131-43.

Das, T. and Parikh, A., (I982), "Decomposition of inequality measures and a comparative analysis", Empirical Economics, vol. 7, pp. 23-48.

Deaton, A., (1997), "The Analysis of Household Surveys", John Hopkins University Press, Baltimore.

DiNardo, J., N.M. Fortin and T. Lemiux, (1996), "Labour Market Institution and the Distribution of Wages, 1973-1992: A semi-parametric Approach", Econometrica, vol. 64 no 5, pp 1001-44.

Epo, B., N, Baye, F., M and Manga. N., A., T, (2010), "Explaining Spatial and Intertemporal sources of Poverty, Inequality and Gender Disparities in Cameroon: A Regression-Based Decomposition Analysis", Poverty and Economic Policy (PEP) Research Network, Research project No. 11321, June 2010, www.pep-net.org.

Fields, G., (2003), "Accounting for Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States", Research in Labor Economics.

Fiorio, Carlo V., and Jenkins, Stephen P., (2007), "ineqrbd: Regression-based inequality decomposition, following Fields (2003), Motivation Model ineqrbd Example: wage inequality" UKSUG, http://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

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

Heshmati A., (2006), "The World Distribution of Income and Income Inequality: A Review of the Economics Literature", Journal of World Systems Research, vol XII, no. I, pp 61-107.

Houthakker, H.S., (1957), "An International Comparison of Household Expenditure Patterns, Commemoration the Centenary of Engel´s Law" Econometrica, Volume 25, Issue 4, S. 532-551.

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

Kakwani, N.,(1980), "On a Class of Poverty Measures", Econometrica, vol 48, pp 437-446.

Kimhi, A., (2007), "Regression-Based Inequality Decomposition: A Critical review and Application to Farm-Household Income Data", Agricultural Economics Department, The Hebrew University, Israel.

Lerman, Robert I., and Yitzhaki. S., (1985), "Income Inequality Effects by Income Source: A New Approach and Applications to the United States", Review of Economics and Statistics, vol 67, pp. 151-156.

Mishra, Padmaja & Parikh, A., (1992), "Household Consumer Expenditure Inequalities in India: A Decomposition Analysis", Review of Income and Wealth, Blackwell Publishing, vol. 38(2), pages 225-36.

Mookherjee, Dilip & Shorrocks, Anthony F, (1982), "A Decomposition Analysis of the Trend in UK Income Inequality", Economic Journal, Royal Economic Society, vol 92, no368, pp 886-902.

Morduch, J. and Sicular T., (2002), "Rethinking Inequality Decomposition with Evidence from Rural China", Economic Journal, vol 112 (January), pp 93-106.

NSSO, (2001), "Consumption of some important commodities in India", National Sample Survey Organization, Ministry of Statistics and Program Implementation, Govt. of India.

NSSO, (2005), "Household Consumer expenditure in India", National Sample Survey Organization, Ministry of Statistics and Program Implementation, Govt. of India.

NSSO, (2005), "Note on Sample Design and Estimation Procedure: 62nd Round", National Sample Survey Organisation, Survey Design Research Division, March.

NSSO, (2006), "Note on Sample Design and Estimation Procedure: 63rd Round", National Sample Survey Organisation, Survey Design Research Division, March.

NSSO, (2006), "Press not on Household consumer expenditure in India, 2006-07", National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

NSSO, (2007), "Household Consumption of Various Goods and Services in India", 2004/05 National Sample Survey Organization, Ministry of Statistics and Program Implementation, Govt. of India.

NSSO, (2007), "Note on Sample Design and Estimation Procedure: 64th Round", National Sample Survey Organisation, Survey Design Research Division, March.

NSSO, (2008), "Household Consumer expenditure in 2006/07", National Sample Survey Organization, Ministry of Statistics and Program Implementation, Govt. of India.

Oaxaca, R., (1973), "Male-Female Wage Differences in Urban Labour Markets" International Economic Review, vol 14 no 3, pp 693-709.

Oyekale, A.S, A.I., Adeoti and Oyekale, T.O., (2007), "Regression-Based Approach to Income Inequality Decomposition in Rural and Urban Nigeria", International Journal of Economic Perspective, vol 1, no 2, pp 45-54.

Sen, A. K., (1973), "Poverty, Inequality and Unemployment: Some conceptual issues in Measurement", in Poverty and Income Distribution in India, e.d. T. N. Srinivasan and P.K. Bardhan, Statistical Publishing Society, Calcutta.

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

Shorrocks, A. F., (1984), "Inequality decompositions by population subgroups", Econometrica vol. 52 no. 6, pp. 1369-88.

Shorrocks, A.F., (1982), "Inequality Decomposition by Factor Components" Econometrica, vol. 65, pp.177-207.

Shorrocks, A.F., (1999), "Decomposition procedure for Distribution Analysis: A unified Framework Based on Shapely value", University of Essex and Institute for Fiscal studies, Colchester, England.

STATA 11.0, (2011), "ineqrbd help and ado file", http://www.stata.com

Wan, G.H. and Z.Y. Zhou, (2005), "Income Inequality in Rural China: Regression-Based Decomposition Using Household Data", Review of Development Economics, vol. 9, pp 107-20.

Wan, G.H., (2002), "Regression-Based Inequality Decomposition: Pitfalls and a Solution Procedure" UNU-WIDER Discussion Paper No. 2002/101, Helsinki.

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

Yuko, A., J.M. Kim. and Kimhi, A., (2006) "Determinants of Income Inequality among Korean Farm Households" Economic Research Centre, Discussion paper, No. 161, School of Economics, NAGOYA University.

Annual Plan, (2010-11), Planning Department, Government of Uttar Pradesh http://planning.up.nic.in/documents.htm

Annual Plan, (2009-10), Planning Department, Government of Uttar Pradesh http://planning.up.nic.in/documents.htm

Annual Plan, (2008-09), Planning Department, Government of Uttar Pradesh http://planning.up.nic.in/documents.htm

Sankhyikiya Diary, (2011), Economics and Statistics Division, Directorate of Economics and Statistics, Government of Uttar Pradesh, http://updes.up.nic.in.

Uttar Pradesh-Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Uttar_Prades

Annual Report 2011-12, Ministry of Power, Government of India, New Delhi.

Basic Statistical Returns of Scheduled Commercial Banks in India, (2012), Reserve Bank of India, Mumbai.

PHD Research Bureau, (2011), "Uttar Pradesh : The State Profile", PHD Chamber of Commerce and Industry, New Delhi.