Subgroup decomposability of Theil entropy measure of inequality: Evidence from India

Debasish Mondal
Professor of Economics,
Department of Economics,
Vidyasagar University, Midnapore

Nilmadhab Das
Research Scholar,
Department of Economics,
Vidyasagar University, Midnapore

Abstract

Henry Theil published a seminal book 'Economics and Information Theory' and introduced the essential functional forms for modelling and understanding inequality, known as the Theil Entropy Measure. The measure has a particular advantage as it is subgroup decomposable and satisfies the additive decomposability criterion developed by Bourguignon. This criterion is intended to establish that we can neatly decompose the inequality measure into within-group inequality and between-group inequality. Nevertheless, researchers like Sudhir Anand, Udo Ebert, and Mishra & Parikh have suggested an alternative method of subgroup decomposition. This method exhibits that the entropy measure fails to be neatly decomposed into within-group inequality and between-group inequality; a residual part is also present. Indeed, we observe that this residual part is an 'interaction' of within-group inequality and between-group inequality. For empirical evidence, we decompose the combined (combining rural and urban regions) consumer expenditure inequality into within-region inequality and between-region inequality for all India and its major states from 1983 to 2011-12.

Keywords: Theil Entropy Measure, Subgroup decomposability, Within-group inequality, Between-group inequality, Interaction

1. Introduction:

Henry Theil (1976) published a seminal book, 'Economics and Information Theory,' which provides a landmark in the inequality measurement analysis by introducing an inequality measure based on the information theory, known as Theil Entropy Measure. Scholars have appreciated the significance of the milestone of the Theil entropy measure for some time, but standard references on income inequality studies now acknowledge its usefulness. According to Cowell (2003), "Theil's measure provided a framework for considering the meaning of inequality and an introduction to an essential functional form for modelling and understanding inequality." Moreover, the Theil measure is popular for subgroup decomposition since the structure of the Theil entropy is well-defined for this purpose. It also satisfies the subgroup decomposability criterion developed by Bourguignon (1979). The method developed by Bourguignon (1979) helps us to explain how the Theil entropy measure can be neatly decomposed into within-group inequality and between-group inequality. The within-group inequality is found as a weighted average of subgroup inequality, and the between-group inequality is found as a function of subgroup means. Following Bourguignon, Shorrock (1980) and Cowell (1980) have also suggested that the Theil entropy measure is a neatly decomposable inequality measure. Nonetheless, Sudhir Anand (1983) has proposed an alternative method to

DOI: http://dx.doi.org/10.62424/VUJE.2024.18.00.11

estimate the within-group inequality. He states, "The within-group component is defined as the inequality index value after all between-group income inequalities have been suppressed. Thus, a hypothetical income distribution is created in which the group mean incomes are equalised to the overall mean by an equiproportional change in the income of each member of a group." Likewise, He estimates the between-group inequality by hypothetically suppressing all the within-group inequality. In addition, Mishra & Parikh (1992) have endorsed that the reduction of between-group inequality by equalising all subgroup means does not always reduce the same amount of between-group inequality from overall inequality. It is only possible for strongly decomposable inequality measures. The Theil entropy measure fails to satisfy this strong decomposability criterion, or the sum of within-group and between-group inequality is not exactly equal to overall inequality. Thus, we observe that these two methods show some different results. The primary objective of this paper is to comparatively analyse these two methods in relation to the decomposition of the Theil entropy measure.

In the first section of this paper, we explain the approach of the measure of the Theil index, while the second section discusses the prerequisite of an inequality measure. In the following section, we critically examine the subgroup decomposability of the Theil entropy measure. The next section provides the empirical evidence of the decomposition of combined (combining rural and urban sectors) consumer expenditure inequality into within-group inequality and between-group inequality for all India and its states, followed by the conclusion.

2. The approach of Theil entropy measure:

The inception of income inequality measures is rooted in the concept of the Lorenz curve, introduced by Lorenz (1905). This curve represents the cumulative proportions (or percentages) of income against cumulative proportions (or percentages) of populations. It aligns with the egalitarian (diagonal) line under conditions of equal income distribution but deviates as inequality increases. The popular income inequality measure, the Gini coefficient, is derived from the Lorenz curve. Another approach to income inequality is the welfare-based approach. Dalton (1920) initially proposed a welfare-based inequality measure, while Atkinson (1970) modified and popularised it. This approach examines the inequality through the welfare of the income distribution instead of the actual income distribution. Nonetheless, Theil introduced an inequality measure based on information theory. In general, the information theory provides us information related to distinct events. An event with a higher probability contains less information, and vice versa. Suppose y is the probability that a particular event will occur. Then, the information function h(y) contains less information and vice versa. Hence, h(y) is a decreasing function of y. In a possible way, h(y) can be expressed as a function of the logarithmic reciprocal of y, that is, h(y) = log(1/y). For n number of events, he introduces the concept of entropy to evaluate the expected information for all events. The entropy is the value of the sum of the information contents of all events multiplied by their corresponding probabilities, that is, H $(y_i) = \sum_{i=1}^{n} y_i \log \frac{1}{y_i}$. However, if we consider y_i as the income share of i-th individual, then the H (y_i) could be an inequality of the income distribution. The H (y_i) will be the maximum when one individual has all income, that is, H (y_i) = log (n). Thus, subtracting H (y_i) from its maximum value, we can define the Theil index or Theil entropy measure.

3. Axioms of an inequality measure:

In this section, we introduce the notation and axioms of the inequality measure. The phrase 'inequality in the distribution of income' (or consumer expenditure or such valued things) is very commonly used in economics and other social sciences. The term 'inequality' in the phrase 'inequality in the distribution of income' means the absence of equality or deviation from

DOI: http://dx.doi.org/10.62424/VUJE.2024.18.00.11

equality in the distribution of income among the persons/households of a community or a geographical region and others. By the term distribution of a total income Y among n individuals in the form y1, y2, …, yn, such that yi ≥ 0 for all i and where μ is the AM (Arithmetic Mean) of y1, y2, …, yn. A function I (y1, y2, …, yn) qualifies a measure of income inequality if it satisfies some axioms. These axioms are,

I. **Axiom I (Symmetry):** I (y1, y2, …, yn) remains unchanged from any permutation of income among the individuals.
II. **Axiom II (Normalisation):** I (μ, μ, …, μ) = 0, where μ is the mean income.
III. **Axiom III (Aggregativity):** Overall inequality is the function of subgroup inequality, that is I (y1, y2, …, yn) = $F_m\{I_{n_1}(y_{11}, y_{12}, \dots, y_{1n_1})$, I_{n_2} $(y_{21}, y_{22}, \dots, y_{2n_2})$, …, I_{n_k} $(y_{k1}, y_{k2}, \dots, y_{kn_k})$; $Y_1, Y_2, \dots, Y_k$; $n_1, n_2, \dots, n_k\}$, with $Y_i = \sum_{i=1}^{n_i} y_{ij}$, for all partitions (k; $n_1, n_2, \dots, n_k$) of the population.
IV. **Axiom IV (Pigou-Dalton income transfer criterion):** Transfer of one unit of income from an individual having income yi to an individual having income yi +a increases inequality by a less amount than the transfer of one unit of income from an individual having income yj to an individual having income yi +a given that yi > yj.
V. **Axiom V (Income homogeneous of degree zero):** I (λy1, λy2, …, λyn) = I (y1, y2, …, yn). This axiom implies that inequality remains unchanged when we proportionately change all individuals' income.

The entropy measure satisfies all these necessary axioms. In addition, Shorrock (1980) has derived the general class of additively decomposable inequality measures for subgroup decomposition. Theil entropy measure is explained to satisfy this subgroup decomposability. As proved by Shorrock, the general class of additively decomposable inequality measures are,

$$I_c(y) = \frac{1}{n}\frac{1}{c(c-1)}\sum_i [(\frac{y_i}{\mu})^c - 1], \qquad c \neq 0,1 \tag{1}$$

$$I_c(y) = \frac{1}{n}\sum_i \log\frac{\mu}{y_i}, \qquad c = 0 \tag{2}$$

$$I_c(y) = \frac{1}{n}\sum_i \frac{y_i}{\mu}\log\frac{y_i}{\mu}, \qquad c = 1 \tag{3}$$

For c = 0, we get the log mean deviation or Theil's second measure or L (Theil, 1967), and for c = 1, we have Theil's first inequality measure or Theil entropy measure or T (Theil, 1967). For c = 2, it becomes the half of the squared coefficient of variation.

4. Methods of Subgroup Decomposition:

An inequality measure is subgroup decomposable when it is expressed as the sum of within-group and between-group inequality. Bourguignon (1979) proposed this subgroup decomposability criterion. According to Bourguignon (1979), "A decomposable inequality measure is defined as a measure such that the total inequality of a population can be broken down into a weighted average of the inequality existing within subgroups (within-group inequality) of the population and the inequality existing between them (between-group inequality)." His large number of followers, including Shorrock (1980) and Cowell (1980), have suggested that the Theil entropy measure follows this criterion. So, the Theil entropy measure is neatly decomposable into within-group inequality and between-group inequality. Nevertheless, researchers like Sudhir Anand (1983), Udo Ebert (1988), Mishra and Parikh

(1992), and Foster and Shneyerav (2000) have proposed another method of subgroup decomposability. They have made different adjustments to the estimation of within-group inequality, which shows different results. We designate these two methods as Method 1 and Method 2, respectively. We discuss these two methods in the following section.

a. Method 1:

Suppose a society has n individuals with income $y_1, y_2, \dots, y_n$ respectively, and I(y) is a relative measure of inequality, which is unit free, then I $(y_1\lambda, y_2\lambda, \dots, y_n\lambda)$ = I $(y_1, y_2, \dots, y_n)$. If the society consists of 'k' groups with the above income vector grouped as $((y_{11}, y_{12}, \dots, y_{1n_1})$, $(y_{21}, y_{22}, \dots, y_{2n_2})$, …, $(y_{k1}, y_{k2}, \dots, y_{kn_k}))$ with $\mu_1, \mu_2, \dots, \mu_k$ as the group means and $\sum_{j=1}^{k} n_j$ = n, then $I_1(y_{11}, y_{12}, \dots, y_{1n_1})$, $I_2(y_{21}, y_{22}, \dots, y_{2n_2})$, …, $I_k(y_{k1}, y_{k2}, \dots, y_{kn_k})$ are called within-group inequalities. The overall inequality is a function of these subgroup inequalities, that is, I $(y_1, y_2, \dots, y_n) = F_m\{I_{n_1}(y_{11}, y_{12}, \dots, y_{1n_1}), I_{n_2}(y_{21}, y_{22}, \dots, y_{2n_2}), \dots, I_{n_k}(y_{k1}, y_{k2}, \dots, y_{kn_k}); Y_1, Y_2, \dots, Y_k; n_1, n_2, \dots, n_k\}$. On the other hand, if through transfers of income of individuals in all subgroups, the income of individuals is equalised with the respective subgroup means, if the inequality function is aggregative and I(μ, μ, …, μ) = 0, then inequality of the population after all such transfers becomes $I(\mu_1, \mu_1, \dots, \mu_1, \mu_2, \mu_2, \dots, \mu_2, \dots, \mu_k, \mu_k, \dots, \mu_k)$, this is the 'between-group' inequality [Bourguignon (1979)], denoted by I_B. Further, he defined 'within-group' inequality, denoted by I_W, as $I - I_B$, leading to the fact that $I = I_B + I_W$. Bourguignon (1979) and his followers have used this definition of I_W for subgroup decomposition. However, after defining I_W as $I - I_B$, Bourguignon (1979) has used a beautiful method to evaluate the contribution of a particular subgroup in I_w. He has considered a distribution obtained through the transfer of income of individuals in the j^{-th} subgroup so that the income of individuals in that subgroup is equal to their subgroup mean, thereby eliminating the inequality within the jth subgroup. Inequality for the new distribution now consists of 'between group' inequality I_B and contributions of all except the j^{-th} subgroup in 'within group' inequality, denoted by $I_{W(-j)}$. Then, he subtracted the inequality of this new distribution from that of the original distribution to evaluate the contribution of the j^{th} subgroup in the 'within-group' inequality, denoted by I_{Wj}. In this way, he has evaluated the contributions of all m subgroups in 'within-group' inequality. He has then examined whether the sum of all these I_{Wj}, for j = 1, 2, …, k, is becoming equal to the value of I_W obtained earlier as $I - I_B$. He has defined the inequality measure as additively decomposable if the above sum is equal to I_W, $I_W = \sum_{j=1}^{m} I_{W_j}$.

b. Method 2:

The researchers in this group also accept the fundamental concept of subgroup decomposability criterion; the within-group inequality is the weighted average of subgroup inequality, and the between-group inequality is the function of subgroup means. Nonetheless, Sudhir Anand (1983) has suggested that within-group inequality can be defined by eliminating between-group inequality from overall inequality. To eliminate all between-group inequalities, he has suggested equalising subgroup means to the overall mean by implementing a proportional change in the income of every unit within each subgroup.

In our previous setup, I (y)= I $(y_1, y_2, y_3 \dots, y_n)$ consist of both within-group and between-group inequality. The within-group inequality can be obtained by eliminating between-group inequality through between-group transfers. If equal amount multiplication is made for the j-th subgroup with amount, $\frac{\mu}{\mu_j}$ then subgroup means will be the overall mean, that is $\mu_j = \mu$, and it is said that there will be no between-group inequality. As a result, we get the new income distribution that only exists within-group inequality. Then, we estimate the between-group inequality by subtracting this within-group from the overall inequality. Moreover, Mishra and Parikh (1992) mentioned that between-group inequality and within-group inequality should be independent. Sometimes, the decomposition coefficients in the within-group inequality can be affected by the change in the group means when the income shares are the coefficients of within-group inequality. In such case, the within-group inequality depends on between-group inequality. Eliminating between-group inequality from overall inequality through equalising all subgroups means not necessarily the same as the amount of within-group inequality. Further, they mentioned that an inequality measure is strongly decomposable when within-group inequality is independent of between-group inequality. The Theil entropy measure is not strongly decomposable; instead, it is a weakly decomposable inequality measure. In the next section, we employ these two methods to decompose the Theil entropy measure.

5. Decomposition of Theil entropy measure:

From equation 3, the Theil entropy measure is

$$T = \frac{1}{n}\sum_{i}\frac{y_i}{\mu}\log\frac{y_i}{\mu} \tag{2}$$

The typical arrangement for decomposition of the Theil entropy measure for k-subgroups, that is,

$$T = \sum_{j=1}^{k}\frac{n_j}{n} * \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu}\log(\frac{y_{ji}}{\mu}) \tag{5}$$

where y_{ji} is income of i-th individual in j-th subgroup.

Or $$T = \sum_{j=1}^{k}\frac{n_j}{n} * \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu_j} * \frac{\mu_j}{\mu}\log(\frac{y_{ji}}{\mu_j} * \frac{\mu_j}{\mu}) \tag{6}$$

Or $$= \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu} * \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu_j}\log(\frac{y_{ji}}{\mu_j}) + \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu} * \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu_j}\log(\frac{\mu_j}{\mu}) \tag{7}$$

Or $$= \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}T_j + \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\log\frac{\mu_j}{\mu} \tag{8}$$

Or $$= \sum_{j=1}^{k}S_jT_j + \sum_{j=1}^{k}S_j\log\frac{\mu_j}{\mu} \tag{9}$$

In equation 9, S_j and T_j imply income share and subgroup inequality for j-th subgroups.

Method 1: All y_{ji} are equalised with subgroup means (μ_j) through the within-group income transfer, and this makes all $T_j = 0$. From equation 9, we get $T = 0 + \sum_{j=1}^{k} S_j\log\frac{\mu_j}{\mu}$ or only the between-group inequality. If we subtract this between-group inequality from overall inequality, we get $\sum_{j=1}^{k} S_jT_j$; this indicates the within-group inequality. This method suggested that the Theil entropy measure is neatly decomposable into within-group inequality and between-group inequality.

Method 2: All y_{ji} are multiplied by $\left(\frac{\mu}{\mu_j}\right)$, this makes all $\mu_j = \mu$ and all T_j unchanged; therefore, T (from equation 8) is reduced to $\sum_{j=1}^{k}\frac{n_j}{n}T_j + 0$ or $\sum_{j=1}^{k}P_jT_j$, where P_j is population share. It indicates the within-group inequality, where the weights are population share instead of income share. Now, if we subtract this within-group inequality from overall inequality, we get,

$$\left(\sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}T_j + \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\log\frac{\mu_j}{\mu}\right) - \sum_{j=1}^{k}\frac{n_j}{n}T_j \quad (10)$$

Or $$= \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\log\frac{\mu_j}{\mu} + \left(\sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}T_j - \sum_{j=1}^{k}\frac{n_j}{n}T_j\right) \quad (11)$$

However, Equation 11 does not clearly indicate the between-group inequality when considering it as a function of subgroup means. In addition, a residual part is present in this equation. Therefore, researchers in this group acknowledge that the Theil entropy measure is not strongly decomposable inequality measure (Mishra & Parikh, 1992).

According to Mishra and Parikh (1992), "This is because sometimes the decomposition coefficients in the within-group term can be affected by the change in the group means. This happens when the income shares are the coefficients in the within-group term. In such a case, if the between-group inequality is eliminated by equalising all the group means, the reduction in total inequality will not necessarily be the amount of between-group inequality. However, when the weights or coefficients of the within-group indices are population shares instead of income shares, the total reduction in the inequality will be exactly by the amount of between-group inequality (because the population shares are not affected by the change in group means)." Theil (1967) himself admitted this issue and pointed out that if income is the weight or coefficient of the within-group inequality, then within-group inequality and between-group inequality are not independent. However, they fail to clarify the between-group inequality, and they adequately do not address the residual part. In contrast, researchers in these two groups explicitly state that between-group inequality is only the function of subgroup means, that is, $\sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\log\frac{\mu_j}{\mu}$ in the Theil entropy measure.

In this paper, we suggest considering both within-group inequality and between-group inequality independently; the within-group inequality must be a weighted average of subgroup inequality, where the weight is population share only and the between-group inequality is the function of subgroup means.

So, the within-group inequality $$= \sum_{j=1}^{k}\frac{n_j}{n}T_j \quad (12)$$

And the between-group inequality $$= \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\log\frac{\mu_j}{\mu} \quad (13)$$

Now, the residual part in equation 11 that remains after subtracting the within-group inequality obtained in Method 2 from that obtained in Method 1 is, $\sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}T_j - \sum_{j=1}^{k}\frac{n_j}{n}T_j$. It is important to note that this residual part is not observable if we exclusively utilize any single method for the decomposition; it only arises when we use both methods simultaneously.

Now, the residual part is, $$\sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}T_j - \sum_{j=1}^{k}\frac{n_j}{n}T_j \quad (14)$$

Or $$= \sum_{j=1}^{k} \frac{n_j}{n} * T_j * \left(\frac{\mu_j - \mu}{\mu}\right) \quad (15)$$

The residual part (in equation 15) is actually the weighted average of the product of within-group inequality and group mean difference or the between-group inequality. In other words, it is an interaction of within-group inequality and between-group inequality. This 'interaction' is neither an exclusive part of within-group inequality nor that of between-group inequality; instead, it falls on the overlapping region. In this case, within-group inequality and between-group inequality are not mutually exclusive; they are overlapping and this arises from their interdependence. The 'interaction' may appear for all other inequality measures where the income share is the coefficient of within-group inequality. We suggest that the Theil entropy measure is decomposable into three parts; first, the 'within-group inequality' which is estimated from within subgroups only; second, the 'between-group inequality' that is estimated from the subgroup mean difference, and third, the 'interaction' which is obtained as an interaction of the above two components.

6. Empirical Evidences

Data Source: The National Sample Survey Office (NSSO) has started publishing data on rural and urban consumer expenditure for all India and its states from the year 1950-51 on an annual basis, but the sample size was small. Large sample survey was first conducted in 1972-73 and subsequently in 1977-78, 1983, 1987-88, 1993-94, 1999-00, 2004-05, 2009-10, 2011-12, 2015-16 and 2022-23. However, the 2015-16 survey data could not be published due to some technical reasons and the 2022-23 survey data are yet to be published. Moreover, the NSSO published consumer expenditure data by distributing monthly per capita consumer expenditure (MPCE) into different classes, but the number of expenditure classes varies across different periods. We utilise data from 1983 to 2011-12 for the subgroup decomposition of combined inequality (combining rural and urban inequality) for all India and its fifteen major states.

Motivation for Decomposition: Using the NSSO data, Mondal and Kayet (2018) have published combined inequality data for all India and its major states from 1983 to 2011-12 using the Gini index and other inequality measures. However, here, we estimate the combined inequality measured by the Theil index in the same way as they have calculated the combined Gini index. We calculate the combined consumer expenditure inequality for all India and its states by combining all expenditure classes of rural and urban regions in an ascending order. Inherently, the combined within inequality, as explained by Bourguignon (1979) is the weighted average of rural and urban consumer expenditure inequality, and the weights are the income share of rural and urban regions.

In addition, the income share is almost the same for rural and urban areas because the income share is the multiplication of the ratio of rural and urban Monthly Per Capita Consumer Expenditure (MPCE) to combined MPCE with their respective population share. The population share is always higher in rural regions than in urban regions for all India and its states. On the other hand, MPCE for the urban population is higher compared to the rural population. As a result, by multiplying the rural population share with the rural MPCE / combined MPCE and the urban population share with the urban MPCE / combined MPCE, we observe that the income shares of both rural and urban regions are more or less the same. The combined within inequality is expected to be close to the simple average of rural and urban inequality. On the other hand, the MPCE gap between urban and rural regions creates a between inequality and this is added to the within inequality to arrive at combined inequality. Automatically, depending on the conditions, combined inequality may be slightly less than the simple average of rural and urban inequalities, may be greater than the simple average of rural and urban inequalities but less than the larger of rural and urban inequalities or may be greater than both of rural and urban inequalities. For illustrations, we estimate the rural, urban, and combined inequality of all India and its states for 2011-12 and those are presented in Table-1. Table 1 illustrates that the combined inequality in the state of Bihar is lower than the simple average of rural and urban inequality (indicated by horizontal shades). This may be because of relatively high rural income share and low MPCE gap in this state.

Table 1: Combined, Rural and Urban Inequality (Theil entropy measure) in all India and its Major States for the year 2011-12

State	Combined Inequality	Rural Inequality	Urban Inequality
Andhra Pradesh	0.088	0.063	0.089
Assam	0.067	0.037	0.095
Bihar	0.047	0.040	0.066
Gujarat	0.078	0.068	0.063
Haryana	0.106	0.056	0.128
Karnataka	0.162	0.096	0.159
Kerala	0.176	0.183	0.165
Madhya Pradesh	0.115	0.061	0.142
Maharashtra	0.115	0.070	0.109
Odisha	0.082	0.049	0.095
Punjab	0.082	0.072	0.088
Rajasthan	0.072	0.051	0.087
Tamil Nadu	0.092	0.073	0.086
Uttar Pradesh	0.103	0.055	0.150
West Bengal	0.122	0.048	0.135
All India	0.116	0.076	0.118

Source: Authors' calculation using NSSO data.

On the contrary, some states exhibit combined inequality higher than both rural and urban inequality resulting from high urban income share and high MPCE gap or high between inequality (indicated by diagonal shades). This is observed for Gujarat, Karnataka, Maharashtra, and Tamil Nadu. Other states exhibit combined inequality higher than the simple average of rural and urban inequality but lower than the larger of rural and urban inequality indicating a low MPCE gap or low between inequality. Out of these states, for states like Kerala and Punjab, the combined inequality is very close to the simple average of rural and urban inequality, while it is far for West Bengal, Madhya Pradesh, and all India.

The location of combined inequality in relation to rural and urban inequality is governed mainly by the MPCE gap or the between inequality along with rural and urban income shares, but as is observed in Table 2, it is also governed by inequality gap.

Table 2: Combined Inequality (Theil entropy measure), MPCE gap and Inequality gap in all India and its major states for 2011-12

State	Combined Inequality	Inequality Gap*	MPCE Gap**
Andhra Pradesh	0.088	34.00	47.92
Assam	0.067	88.95	63.98
Bihar	0.047	47.63	33.59
Gujarat	0.078	-7.82	47.54
Haryana	0.106	78.89	53.17
Karnataka	0.162	49.80	71.00
Kerala	0.176	-10.25	19.92
Madhya Pradesh	0.115	79.61	61.29
Maharashtra	0.115	43.44	61.56
Odisha	0.082	64.17	66.83
Punjab	0.082	20.61	23.45
Rajasthan	0.072	52.12	42.22
Tamil Nadu	0.092	15.91	45.03
Uttar Pradesh	0.103	92.81	57.24
West Bengal	0.122	95.45	72.50
All India	0.116	43.95	60.81

Source: Authors' calculation using NSSO data.

*Inequality Gap = $\frac{urban\ inequality - rural\ inequality}{(rural\ inequality + urban\ inequality)/2} * 100$

**MPCE Gap = $\frac{urban\ MPCE - rural\ MPCE}{(rural\ MPCE + urban\ MPCE)/2} * 100$

As explained earlier, rural and urban income shares determine whether the combined within inequality of Method 1 lies below or above the simple average of rural and urban inequalities. However, the substantial MPCE gap or substantial between inequality can make the combined inequality lying not only above the simple average of rural and urban inequalities, but also above both these inequalities. It implies that combined inequality not only contains within-group inequality but also contains inequality between rural and urban regions (known as 'between inequality'), and a high 'between inequality' may lead to such results. Figures of Table 2 show that in 2011-12 MPCE gap was highest in West Bengal at 72.50%, followed by Karnataka (71.00%); but for West Bengal combined inequality is observed to lie well between rural and urban inequalities, whereas for Karnataka the same is observed to surpass both rural and urban inequalities. The main reason behind this phenomenon can be traced in the values of inequality gap presented in Table 2. For West Bengal MPCE gap is highest at 72.50% leading to a high between inequality, but at the same time it has the highest inequality gap at 95.45% and this inequality gap is significantly greater than the MPCE gap and this leads to the fact that high between inequality fails to make combined inequality surpass both rural and urban inequalities. On the other hand, for Karnataka inequality gap is only 49.80% and this inequality gap is significantly less than the MPCE gap and this leads to the fact that high between inequality makes combined inequality surpass both rural and urban inequalities. For states like Gujarat and Tamil Nadu even if the MPCE gaps are not very high, combined inequality surpasses both rural and urban inequalities as inequality gaps are significantly less than the respective MPCE gaps. Finally for Maharashtra, combined inequality surpasses both rural and urban inequalities for the same reason even if the MPCE gap is moderate. For other states, either inequality gap is greater than MPCE gap (Assam, Bihar, Haryana, Madhya Pradesh,

DOI: http://dx.doi.org/10.62424/VUJE.2024.18.00.11

Rajasthan and Uttar Pradesh) or inequality gap is slightly less than MPCE gap (Andhra Pradesh, Kerala, Odisha, Punjab and All India).

Based on the above discussion, we conclude that income share, inequality gap, and MPCE gap (between inequality) are the three main factors in determining the position of combined inequality in relation to rural and urban inequalities. Notably, the MPCE gap, compared to the inequality gap plays the most crucial role. In Table 3 we present the values of combined inequality in 15 major states and all India for all the years with horizontal and diagonal shades as explained in Table 1. It shows that in 1983, 10 out of 16 cells contain diagonal shades. This means that in that year, for those states, between inequalities were very high, or inequality gaps were very low or both. With the passage of time this has reduced and from 199-00 it has become stagnant. All these results motivate us to decompose combined consumer expenditure inequality into within-group and between-group inequalities for all India and its major states.

Table 3: Combined Inequality (Theil entropy measure) of all India and its major states from 1983 to 2011-12

State	1983	1987-88	1993-94	1999-00	2004-05	2009-10	2011-12
Andhra Pradesh	0.077	0.096	0.079	0.072	0.099	0.107	0.088
Assam	0.038	0.065	0.045	0.055	0.053	0.063	0.067
Bihar	0.062	0.065	0.053	0.047	0.047	0.050	0.047
Gujarat	0.064	0.064	0.066	0.066	0.081	0.088	0.078
Haryana	0.064	0.065	0.067	0.052	0.091	0.083	0.106
Karnataka	0.087	0.091	0.079	0.083	0.106	0.093	0.162
Kerala	0.099	0.096	0.077	0.062	0.103	0.195	0.176
Madhya Pradesh	0.080	0.089	0.088	0.076	0.108	0.097	0.115
Maharashtra	0.096	0.120	0.114	0.102	0.117	0.136	0.115
Odisha	0.068	0.077	0.069	0.060	0.084	0.091	0.082
Punjab	0.067	0.064	0.057	0.052	0.102	0.085	0.082
Rajasthan	0.091	0.090	0.061	0.049	0.080	0.078	0.072
Tamil Nadu	0.104	0.108	0.099	0.125	0.109	0.085	0.092
Uttar Pradesh	0.076	0.076	0.073	0.065	0.089	0.078	0.103
West Bengal	0.091	0.088	0.085	0.087	0.104	0.102	0.122
All India	0.087	0.096	0.089	0.086	0.105	0.108	0.116

Source: Authors' calculation using NSSO data.

However, our theoretical discussion shows that Theil entropy measure is not perfectly decomposable into within-group and between-group inequalities. This is because for this measure, the two components are not mutually exclusive or they are not independent; there is an interaction between within-group and between-group inequalities. Therefore, we shall make this decomposition into three parts.

Decomposition of combined consumer expenditure inequality into within-group inequality and between-group inequality:

In this section, we analyse the within-group inequality, between-group inequality, and interaction of Theil entropy measure for all India and its states as explained in the methodology section. Table 4 presents the within-group (rural and urban) inequality values for all India and its major states. The within-group inequalities are high (lying in the range of 59.16 to 99.27%), implying that inequality within the rural and urban sectors is more than the inequality between them. The within-group inequality was highest in the state of Rajasthan in 1983, followed by Punjab in 1987-88, Haryana in 1993-94, and Kerala from 1999-00 to 2011-12. On the other hand, the within-group inequality was lowest in the state of Maharashtra in 1983, followed by Assam in 1987-88, 1993-94 and 2004-05, West Bengal in 1999-00 and 2011-12 and Karnataka in 2009-10.

Table 4: The within-group inequality (Theil entropy measure) in all India and its states from 1983 to 2011-12.

State	1983	1988-87	1993-94	1999-00	2004-05	2009-10	2011-12
Andhra Pradesh	0.067 (86.06)	0.082 (86.05)	0.068 (85.96)	0.050 (70.34)	0.081 (81.05)	0.076 (70.41)	0.072 (82.26)
Assam	0.031 (80.75)	0.047 (71.82)	0.027 (60.45)	0.035 (63.07)	0.035 (66.35)	0.049 (77.98)	0.045 (67.72)
Bihar	0.052 (84.13)	0.060 (93.00)	0.041 (77.79)	0.038 (80.70)	0.038 (81.06)	0.041 (82.89)	0.043 (91.61)
Gujarat	0.052 (81.20)	0.049 (77.08)	0.049 (75.15)	0.048 (72.56)	0.060 (74.23)	0.062 (70.48)	0.065 (84.39)
Haryana	0.057 (88.33)	0.062 (95.08)	0.064 (94.29)	0.046 (89.23)	0.087 (95.52)	0.076 (91.71)	0.081 (76.19)
Karnataka	0.071 (81.16)	0.076 (84.20)	0.060 (76.36)	0.055 (66.38)	0.076 (71.83)	0.055 (59.16)	0.120 (74.47)
Kerala	0.092 (92.73)	0.089 (93.38)	0.070 (90.13)	0.058 (93.67)	0.099 (96.30)	0.182 (93.15)	0.174 (99.27)
Madhya Pradesh	0.066 (82.96)	0.071 (79.57)	0.068 (77.10)	0.054 (71.09)	0.074 (68.56)	0.074 (75.71)	0.083 (72.46)
Maharashtra	0.067 (69.49)	0.100 (82.95)	0.078 (68.94)	0.067 (65.17)	0.088 (75.30)	0.088 (64.81)	0.088 (76.03)
Odisha	0.056 (81.19)	0.061 (79.29)	0.049 (71.00)	0.047 (78.90)	0.067 (80.66)	0.061 (66.36)	0.057 (69.61)
Punjab	0.063 (93.22)	0.063 (98.13)	0.053 (93.18)	0.048 (91.81)	0.085 (83.38)	0.078 (90.97)	0.078 (95.27)
Rajasthan	0.085 (93.73)	0.082 (90.80)	0.054 (88.31)	0.039 (78.94)	0.064 (81.01)	0.057 (73.23)	0.060 (83.78)
Tamil Nadu	0.086 (82.43)	0.089 (81.99)	0.083 (83.66)	0.089 (70.80)	0.092 (84.28)	0.065 (76.81)	0.079 (86.27)
Uttar Pradesh	0.067 (88.74)	0.065 (86.37)	0.062 (85.85)	0.054 (82.82)	0.077 (87.12)	0.062 (80.13)	0.076 (73.82)
West Bengal	0.065 (71.68)	0.065 (73.73)	0.061 (71.28)	0.053 (61.59)	0.075 (72.02)	0.064 (62.77)	0.076 (61.93)
All India	0.071 (81.37)	0.079 (81.99)	0.069 (76.99)	0.061 (71.67)	0.083 (78.68)	0.078 (72.56)	0.089 (76.79)

Source: Author's calculation using NSSO data.

(N.B. Figures in the parentheses indicate percentage of within inequality in combined inequality.)

The between-group inequality is much lower than the within-group inequality (Table 5). The low between-group inequality indicates the inequality of consumer expenditure between rural and urban sector is less in comparison to inequality among these sectors. The low between-group inequality (presented in Table 5) is significant for a society, but if it exceeds a specific limit then it might be harmful and may affect social harmony. The between-group inequality was highest in the state of Maharashtra in 1983, followed by West Bengal in 1987-88, Assam in from 1993-94 to 2004-05, West Bengal in 2011-12. Therefore, the between-group inequality requires control in these states. Moreover, after the liberalization most of states between-group inequality increased except Tamil Nadu.

Table 5: The between-group inequality (Theil entropy measure) of all India and its states from 1983 to 2011-12.

State	1983	1988-87	1993-94	1999-00	2004-05	2009-10	2011-12
Andhra Pradesh	0.010 (13.44)	0.009 (09.88)	0.010 (12.45)	0.017 (24.12)	0.013 (13.42)	0.023 (21.81)	0.013 (14.30)
Assam	0.006 (16.64)	0.012 (18.29)	0.014 (32.05)	0.016 (29.74)	0.013 (24.98)	0.011 (18.12)	0.016 (23.36)
Bihar	0.009 (14.25)	0.004 (05.83)	0.009 (17.71)	0.006 (13.70)	0.006 (12.68)	0.006 (11.96)	0.003 (06.32)
Gujarat	0.012 (18.46)	0.012 (18.98)	0.014 (20.83)	0.015 (23.14)	0.019 (22.78)	0.021 (23.97)	0.013 (16.40)
Haryana	0.007 (10.34)	0.003 (04.24)	0.004 (06.56)	0.004 (08.24)	0.003 (03.14)	0.005 (05.91)	0.016 (14.86)
Karnataka	0.015 (17.26)	0.012 (12.87)	0.016 (20.56)	0.023 (27.32)	0.024 (22.63)	0.031 (32.94)	0.030 (18.34)
Kerala	0.006 (06.07)	0.004 (04.44)	0.005 (06.96)	0.003 (04.49)	0.002 (01.96)	0.007 (03.70)	0.002 (01.23)
Madhya Pradesh	0.014 (17.39)	0.016 (18.23)	0.017 (18.88)	0.017 (22.95)	0.023 (21.28)	0.019 (19.31)	0.020 (17.57)
Maharashtra	0.027 (27.72)	0.022 (18.11)	0.032 (28.00)	0.028 (27.59)	0.024 (20.16)	0.033 (24.07)	0.022 (18.67)
Odisha	0.012 (17.48)	0.014 (17.89)	0.018 (25.72)	0.011 (17.92)	0.013 (15.92)	0.022 (23.97)	0.019 (23.58)
Punjab	0.004 (05.26)	0.002 (02.42)	0.004 (06.35)	0.003 (05.77)	0.009 (08.73)	0.005 (05.46)	0.003 (03.59)
Rajasthan	0.008 (08.29)	0.006 (06.57)	0.006 (10.32)	0.008 (16.17)	0.009 (11.56)	0.012 (15.55)	0.008 (11.78)
Tamil Nadu	0.017 (16.40)	0.017 (15.75)	0.014 (13.84)	0.025 (20.26)	0.016 (14.56)	0.016 (18.77)	0.011 (12.17)
Uttar Pradesh	0.008 (10.57)	0.008 (11.19)	0.009 (11.69)	0.008 (12.60)	0.008 (09.08)	0.011 (14.23)	0.016 (15.31)
West Bengal	0.023 (25.12)	0.017 (19.50)	0.020 (23.77)	0.025 (28.58)	0.022 (21.19)	0.026 (26.01)	0.031 (25.12)
All India	0.015 (17.30)	0.014 (14.49)	0.017 (19.15)	0.019 (22.12)	0.018 (16.90)	0.022 (20.79)	0.021 (17.78)

Source: Author’s calculation using NSSO data.

(*N.B. Figures in the parentheses indicate percentage of between inequality in combined inequality.)*

In addition, we incorporate the concept of 'interaction' for the subgroup decomposition of the Theil entropy measure. Table 6 presents the 'interaction' for all India and its states from 1983 to 2011-12. The interaction indicates both negative and positive values. This observation implies that there are only two groups (rural and urban), and the rural MPCE is consistently lower than the urban MPCE across all states and all India. Consequently, the difference between rural MPCE and overall MPCE ($\mu_1 - \mu$) is consistently negative, while the difference between urban MPCE and overall MPCE ($\mu_2 - \mu$) is consistently positive. Therefore, the direction of the interaction — whether it becomes positive or negative depends on the value of $(\mu_1 - \mu)$ or $(\mu_2 - \mu)$. Nevertheless, the strength of this direction depends on the value of subgroup inequality and population share. For example, Kerala and Haryana in 2011-12 have negative interactions, which implies the multiplication of $(\mu_1 - \mu)$ with rural population share $(\frac{n_1}{n})$ and rural inequality (T_1) is more dominating than the multiplication of $(\mu_2 - \mu)$ with urban population share $(\frac{n_2}{n})$ and urban inequality (T_2). In the same way, the interaction is negative for Madhya Pradesh and Maharashtra in 1983, Maharashtra and Punjab in 1987-88, and Haryana in 1993-94.

Table 6: The interaction (Theil entropy measure) of all India and its states from 1983 to 2011-12.

State	1983	1988-87	1993-94	1999-00	2004-05	2009-10	2011-12
Andhra Pradesh	0.0004 (00.50)	0.004 (04.07)	0.001 (01.59)	0.004 (05.54)	0.005 (05.53)	0.008 (07.78)	0.003 (03.44)
Assam	0.001 (02.60)	0.006 (09.89)	0.003 (07.50)	0.004 (07.19)	0.005 (08.67)	0.002 (03.90)	0.006 (08.92)
Bihar	0.001 (01.62)	0.001 (01.16)	0.002 (04.50)	0.003 (05.60)	0.003 (06.26)	0.003 (05.15)	0.001 (02.08)
Gujarat	0.0002 (00.34)	0.003 (03.94)	0.003 (04.02)	0.003 (04.30)	0.002 (02.99)	0.005 (05.55)	-0.001 (-00.79)
Haryana	0.001 (01.32)	0.000 (00.67)	-0.001 (-00.84)	0.001 (02.52)	0.001 (01.34)	0.002 (02.38)	0.010 (08.96)
Karnataka	0.001 (01.58)	0.003 (02.93)	0.002 (03.08)	0.005 (06.30)	0.006 (05.53)	0.007 (07.89)	0.012 (07.19)
Kerala	0.001 (01.20)	0.002 (02.18)	0.002 (02.91)	0.001 (01.84)	0.002 (01.74)	0.006 (03.15)	-0.001 (-00.51)
Madhya Pradesh	-0.0003 (-00.35)	0.002 (02.20)	0.004 (04.02)	0.005 (05.96)	0.011 (10.16)	0.005 (04.98)	0.011 (09.97)
Maharashtra	0.003 (02.79)	-0.001 (-01.06)	0.003 (03.06)	0.007 (07.24)	0.005 (04.54)	0.015 (11.13)	0.006 (05.29)
Odisha	0.001 (01.33)	0.002 (02.82)	0.002 (03.28)	0.002 (03.17)	0.003 (03.42)	0.009 (09.67)	0.006 (06.81)
Punjab	0.001 (01.52)	-0.000 (-00.55)	0.000 (00.47)	0.001 (02.42)	0.008 (07.89)	0.003 (03.57)	0.001 (01.14)
Rajasthan	-0.002 (-02.02)	0.002 (02.64)	0.001 (01.38)	0.002 (04.88)	0.006 (07.43)	0.009 (11.22)	0.003 (04.44)
Tamil Nadu	0.001 (01.17)	0.002 (02.26)	0.002 (02.50)	0.011 (08.94)	0.001 (01.16)	0.004 (04.42)	0.001 (01.56)
Uttar Pradesh	0.001 (00.69)	0.002 (02.44)	0.002 (02.46)	0.003 (04.58)	0.003 (03.80)	0.004 (05.65)	0.011 (10.88)
West Bengal	0.003 (03.20)	0.006 (06.76)	0.004 (04.95)	0.009 (09.83)	0.007 (06.79)	0.011 (11.22)	0.016 (12.95)
All India	0.001 (01.33)	0.003 (03.51)	0.003 (03.87)	0.005 (06.21)	0.005 (04.41)	0.007 (06.65)	0.006 (05.43)

Source: Author's calculation using NSSO data.

(N.B. Figures in the parentheses indicate percentage of interaction in combined inequality.)

It is noteworthy that, in some states, the 'interaction' exceeds 10 percent, which might be a severe issue for a subgroup decomposable inequality measure. For instance, the interaction exceeded 10 percent for West Bengal and Uttar Pradesh in 2011-12, West Bengal, Maharashtra, Rajasthan in 2009-10, and Madhya Pradesh in 2004-05.

DOI: http://dx.doi.org/10.62424/VUJE.2024.18.00.11

7. Conclusion:

Cowell (2005) suggested that the Theil entropy measure has the special advantage of additive decomposability, which automatically came through the structure of the Theil entropy measure. The Theil's measure is likely effective for subgroup decomposition. We observe that Theil's measure is not strongly decomposable into within-group inequality and between-group inequality; a residual part is also present. The residual part is the weighted average of the product of within-group inequality and group mean difference or the between-group inequality. In other words, it is an interaction of within-group inequality and between-group inequality. This 'interaction' is neither an exclusive part of within-group inequality nor of between-group inequality; instead, it is the part where the two types of inequalities overlap. Our empirical results show that within inequality is as high as 99.27% in Kerala in 2011-12; between inequality is maximum at 32.94% (in Karnataka in 2009-10); whereas the interaction is highest at 12.95% in West Bengal in 2011-12.

References:

1. Anand, S. (1983). Inequality and poverty in Malaysia: Measurement and decomposition. *Oxford University Press.*
2. Bourguignon, F. (1979). Decomposable Income Inequality Measures. *Econometrica,* 47, 901-920.
3. Cowell, F. (2003). Theil, inequality and the structure of income distribution. *LSE STICERD Research Paper,* (67).
4. Cowell, F. A. (2006). Theil, inequality indices and decomposition. In Dynamics of inequality and poverty (pp. 341-356). *Emerald Group Publishing Limited.*
5. Ebert, U. (1988). On the decomposition of inequality: Partitions into nonoverlapping sub-groups. In *Measurement in Economics: Theory and Applications of Economics Indices* (pp. 399-412). Heidelberg: Physica-Verlag HD.
6. Foster, J. E. (1983). An axiomatic characterisation of the Theil measure of income inequality. *Journal of Economic Theory*, 31(1), 105-121.
7. Foster, J. E., & Shneyerov, A. A. (1999). A general class of additively decomposable inequality measures. *Economic Theory*, 14, 89-111.
8. Mishra, P., & Parikh, A. (1992). Household consumer expenditure inequalities in India: a decomposition analysis. *Review of Income and Wealth,* 38(2), 225-236.
9. Mondal, D., & Kayet, A. (2018). Trends and Patterns of Combined Inequality in India: An analysis across Major States from 1983 to 2011-12. *Sarvekshana* Vol. PDOS, 57, 69-95.
10. Shorrocks, A.F. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica,* 48, 613-625.
11. Shorrocks, A.F. (1984). Inequality Decomposition by Population Subgroups. *Econometrica,* 52, 1369-1385.
12. Theil, H. (1967). Economics and Information Theory. *North-Holland Publishing Company*, Amsterdam,