WP-2024-007

# Multidimensional Index: A Note

Srijit Mishra

Indira Gandhi Institute of Development Research, Mumbai
April 2024

# Multidimensional Index: A Note

**Srijit Mishra**

Email(corresponding author): srijit@igidr.ac.in

**Abstract**

*This paper identifies the problem of information loss in computation of multidimensional index on account of cut-offs between and within dimensions. The former refers to the $k$ cutoff across dimensions on account of headcount fetish and the latter points to the restrictions on normalised deprivation score on account of deprivation line as benchmark. One also observes information loss when available data is converted to binary form. An empirical exercise substantiates the points. We suggest that the computation of multidimensional index in deprivation (poverty) and attainment (empowerment) should do away with these information loss by not censoring, by replacing deprivation line with maxima, where appropriate, as benchmark for computing normalised deprivation score, and by not converting available data to binary form.*

**Keywords: Attainment, Censoring, Deprivation, Headcount fetish, Information loss, Multidimensional index**

**JEL Code: C38, D63, H89, I32, O10, Z13, Z18**

**Acknowledgements:**

This paper has benefitted from discussions and the questions posed by students at IGIDR and the notes/recall from an in-class Talismanic exercise that was limited to computing multidimensional index with and without censoring. The concern for information loss in multidimensional index came up from weekly meetings with colleagues as part of the E²IMPA₹T study. Usual disclaimers apply.

Multidimensional Index: A Note

Srijit Mishra

Indira Gandhi Institute of Development Research, Mumbai, India
Email: srijit@igidr.ac.in

23 April 2024

*Abstract:*
*This paper identifies the problem of information loss in computation of multidimensional index on account of cut-offs between and within dimensions. The former refers to the $k$ cut-off across dimensions on account of headcount fetish and the latter points to the restrictions on normalised deprivation score on account of deprivation line as benchmark. One also observes information loss when available data is converted to binary form. An empirical exercise substantiates the points. We suggest that the computation of multidimensional index in deprivation (poverty) and attainment (empowerment) should do away with these information loss by not censoring, by replacing deprivation line with maxima, where appropriate, as benchmark for computing normalised deprivation score, and by not converting available data to binary form.*

Keywords: Attainment, Censoring, Deprivation, Headcount fetish, Information loss, Multidimensional index

JEL Codes: C38, D63, H89, I32, O10, Z13, Z18

Acknowledgements: This paper has benefitted from discussions and the questions posed by students at IGIDR and the notes/recall from an in-class Talismanic exercise that was limited to computing multidimensional index with and without censoring. The concern for information loss in multidimensional index came up from weekly meetings with colleagues as part of the E$^2$IMPA₹T study. Usual disclaimers apply.

## 1. Introduction

In recent times, the multidimensional index has attracted attention, particularly its application to poverty following Alkire and Foster (2011) and its extension to empowerment (Alkire and Meinzen-Dick et al., 2013), that is, deprivation and attainment respectively. The method in terms of poverty has had a detailed discussion in a book (Alkire and Foster et al., 2015) that has also been critically reviewed (Pattanaik and Xu, 2018). A positive aspect of bringing in a multidimensional index is that it has the potential to combine dimensions that are either cardinal (either continuous or discrete) or ordinal (either binary or more than two categories). This advantage is, however, lost because of the cut-off across dimensions. This surmounts to some information loss. Concerns on information loss can also be identified with the usage of deprivation line to compute normalised deprivation score. Besides, converting data to binary form will also lead to information loss.

The rest of this paper has notations in section 2, some alternative computations of multidimensional index in section 3. Information loss has been discussed in sections 4-6 referring to the headcount fetish leading to censoring, the use of a deprivation line to normalise deprivation scores, and the converting of data to binary form, respectively. An example based on data from an in-class Talismanic exercise is given in section 7. Some other observations and concluding remarks have been given in sections 8 and 9, respectively.

## 2. Notations

Let $x_{ij} > 0$ be a positive attainment of individual $i$ (also $i'$; $i \neq i'$) in dimension $j$ where $i = 1, \dots, n$ and $j = 1, \dots, d$. In each dimension there is a deprivation line, $z_j$. For unique identification (UID) of individuals and dimensions, we use $u_i$ and $v_j$, respectively.

The weight for individual $i$ is $w_i$ and the weight for dimension $j$ is $w_j$ such that the weight when both are taken together is $w_{ij} = w_i w_j$. If all individuals have an equal weight, then $w_i = 1/n$; similarly, if all dimensions have an equal weight, then $w_j = 1/d$. If $w_i = 1/n$ and $w_j = 1/d$, then $w_{ij} = 1/nd$.

Now, let normalised deprivation score of individual $i$ in dimension $j$ be ${g_{ij}}^{\alpha} \in [0,1]$ such that ${g_{ij}}^{\alpha} = \left[(z_j - x_{ij})/z_j\right]^{\alpha}$; $x_{ij} < z_j$, otherwise, ${g_{ij}}^{\alpha} = 0$; $x_{ij} \geq z_j$. Note that for $\alpha = 0,1,2$, the weighted sum over individuals in dimension $j$ provides $\alpha$-class deprivation measure, $P_{\alpha j} = \sum_i w_i {g_{ij}}^{\alpha}\ \forall\ j$, which can be identified with dimension- specific headcount ratio, $H_j$, poverty gap, $G_j$, and poverty-gap squared, $S_j$, respectively (Foster et al., 1984). Similarly, $P_{\alpha i} = \sum_j w_j {g_{ij}}^{\alpha}\ \forall\ i$. Instead of $z_j$ one can also use a dimension-specific maxima, $\hat{z}_j$, if appropriate.

$M_\alpha = \sum_i \sum_j w_{ij} {g_{ij}}^{\alpha}$ is a multidimensional index of deprivation, $q = (1/\sum_i w_i)\left(\sum_i w_i\left(\sum_j {g_{ij}}^{\alpha}\right)^0\right)$ is the number of deprived (note that $\left(\sum_j {g_{ij}}^{\alpha}\right)^0 = 0$ if ${g_{ij}}^{\alpha} = 0$), $H = q/n$ is headcount ratio and $A_\alpha = (1/q)\sum_i \sum_j w_j {g_{ij}}^{\alpha}$ is average $\alpha$-deprivation of the deprived. The contribution of individual $i$ and dimension $j$ to $M_\alpha$ is $M_{\alpha i}$ and $M_{\alpha j}$, respectively, such that, $M_\alpha = \sum_i M_{\alpha i}$ and $M_\alpha = \sum_j M_{\alpha j}$.

Further, $k \in (0,1]$ is the cut-off across dimensions for any individual to be considered as multidimensionally deprived. Now, we have censored normalised deprivation score, ${g_{ij|c}}^{\alpha}$; if $\sum_j w_j {g_{ij}}^0 \geq k$ then ${g_{ij|c}}^{\alpha} = {g_{ij}}^{\alpha}$, otherwise, if $\sum_j w_j {g_{ij}}^0 < k$ then ${g_{ij|c}}^{\alpha} = 0$. For discussing censoring, we will use $M_{\alpha|c}$, $M_{\alpha i|c}$, and $M_{\alpha j|c}$.

For a normalised deprivation score with maxima as benchmark we will use ${g_{ij|\hat{z}_j}}^{\alpha}$ and the multidimensional index of deprivation will be $M_{\alpha|\hat{z}_j}$ where contribution of individual $i$ and dimension $j$ will be denoted by $M_{\alpha i|\hat{z}_j}$ and $M_{\alpha j|\hat{z}_j}$, respectively. For binary, we will use ${i'}_b$, ${g_{ij|\,b}}^{\alpha}$, $M_{\alpha|b}$, $M_{\alpha i|b}$, and $M_{\alpha j|b}$. For maxima with binary, we will use $M_{\alpha|\hat{z}_j,b}$, $M_{\alpha i|\hat{z}_j,b}$, and $M_{\alpha j|\hat{z}_j,b}$.

Normalised attainment or empowerment score is ${e_{ij}}^{\alpha} = \left(1 - g_{ij}\right)^{\alpha}$. A multidimensional index of attainment or empowerment is $E_\alpha = 1 - M_\alpha$. This can also be written as $E_\alpha = \sum_j w_j I_j$ ; $I_j = 1 - P_{\alpha j}$, or, as $E_\alpha = \sum_i w_i I_i$ ; $I_i = 1 - P_{\alpha i}$.

To contrast, two attainment scenarios, with and without information loss, we will use, $E_{\alpha|c,b}$, $E_{\alpha i|c,b}$, $E_{\alpha j|c,b}$, $E_{\alpha|\hat{z}_j}$, $E_{\alpha i|\hat{z}_j}$, and $E_{\alpha j|\hat{z}_j}$. To refer to different situations (basic, censored, maxima, binary, maxima with binary) we use an add-on to the subscript like $M_{\alpha *}$, $M_{\alpha i *}$, and $M_{\alpha j *}$.

Before taking up the discussion on information loss, we would like to first introduce some alternative computation of $M_\alpha$. This is taken up in the next section.

## 3. Alternative computation of $M_\alpha$

To reiterate, multidimensional index of deprivation is a weighted average of normalised deprivation score of individual $i$ in dimension $j$ such that,

$$M_\alpha = \sum_i \sum_j w_{ij} g_{ij}{}^{\alpha}. \tag{1}$$

If all $i$ individuals have equal weight, $1/n$, then,

$$M_\alpha = \sum_i w_i \sum_j w_j g_{ij}{}^{\alpha} = \sum_i w_i P_{\alpha i} = (1/n) \sum_i P_{\alpha i}. \tag{2}$$

If all $j$ dimensions have equal weight, $1/d$, then,

$$M_\alpha = \sum_j w_j \sum_i w_i g_{ij}{}^{\alpha} = \sum_j w_j P_{\alpha j} = (1/d) \sum_j P_{\alpha j}. \tag{3}$$

If all $i$ individuals have equal weight, $1/n$, and all $j$ dimensions have equal weight, $1/d$, then,

$$M_\alpha = (1/nd) \sum_i \sum_j g_{ij}{}^{\alpha}. \tag{4}$$

Another way of capturing the multidimensional index of deprivation is headcount ratio times average $\alpha$-deprivation of the deprived such that,

$$M_\alpha = H A_\alpha. \tag{5}$$

The multiple ways of computing $M_\alpha$ is not exhaustive; it is indicative. This can be done for other multidimensional indices.

We now take up the discussion of the three types of information loss. To begin with, on headcount fetish leading to censoring.

## 4. The headcount fetish while proposing $M_\alpha$

In public policy, if a multidimensional index of deprivation is being proposed as an alternative to the conventional unidimensional income index of poverty, then a normal question is how many people are deprived? Now, under a multidimensional setting, this could give a higher number because an individual who is deprived in any dimension will get counted under the headcount, the union approach. An alternative is to consider an individual as deprived if the individual is deprived in all the dimensions, the intersection approach. This is restrictive and would give a lower number.

Hence, a way out was to suggest some intermediary that uses $k$ cut-off across dimensions such that an individual who is deprived in at least $k$ dimensions will be considered as deprived. For these individuals, one can compute the censored normalised deprivation score, ${g_{ij|c}}^{\alpha}$.

The dimensional cut-off helps in providing a headcount ratio that may not be too large or not too small. But then, this is like throwing the baby with the bath water. The purpose of $M_\alpha$ is not to arrive at a headcount ratio, but rather to provide a multidimensional index where the headcount ratio, $H$, is adjusted to the average $\alpha$-deprivation of the deprived, $A_\alpha$. In other words, if an individual is deprived in only one dimension, that is in dimension $j$ then the individual's contribution to the overall deprivation indicated in $M_\alpha$ is $w_j {g_{ij}}^{\alpha}$ and not $\sum_j w_j = 1$ that may get implied under $H$.

Proposition 1: There can be information loss in multidimensional index of deprivation on account of censoring, $M_{\alpha|c} \leq M_\alpha$.

Proof: This is so because ${g_{ij|c}}^{\alpha} \leq {g_{ij}}^{\alpha}\ \forall\ i, j$. It follows that there can be information loss for computations that involve $g_{ij|c}$ as against $g_{ij}$.

We suggest that the ingenuity of $M_\alpha$ be restored such that the possible information loss is addressed. A question that comes to mind is that if a restriction across dimensions can lead to information loss, then would a restriction within each dimension like the use of $z_j$ as benchmark in computing normalised deprivation score for all individuals in that dimension also lead to information loss. We take that up now.

**5. Is dimension-specific deprivation line also a problem?**

In dimension-specific deprivation line, an individual at the deprivation line or a wee-bit above that line will be considered non-deprived in that dimension. This, of course, is information loss. This is also a point of contention in poverty literature and there have been efforts to compute a measure of deprivation without a deprivation line or rather by capturing each and every individual's shortfall from a limiting maximum (Kumar et al., 2009). The human development index computed in the *Human Development Reports* since 1990 has used a dimension-specific normalisation using maxima and minima, which as indicated in Dutta et al. (1997) has since 1995 started using global maxima and global minima.

Proposition 2: There can be information loss in multidimensional index of deprivation on account of dimension-specific deprivation lines as against maxima as benchmark for computing normalised deprivation score, $M_\alpha \leq M_{\alpha|\hat{z}_j}$.

Proof: This is so because ${g_{ij}}^{\alpha} \leq {g_{ij|\hat{z}_j}}^{\alpha}\ \forall\ i, j$. It follows that there can be information loss for computations involving $g_{ij}$ as against $g_{ij|\hat{z}_j}$.

Hence, in the computation of multidimension index of deprivation one could consider $M_{\alpha,\hat{z}_j}$ over $M_\alpha$. However, caution needs to be used in the use of maxima. For instance, on norms like the nutritional requirement, which is also the deprivation line, one should continue using that as a benchmark for computing normalised deprivation score. Further, for nutritional requirement the norm is as per adult equivalent scale, and hence, the benchmark to obtain normalised deprivation score could be induvial-specific based on their age, gender, and nature of activity, as in Samal and Mishra (2023). While being cautious, we suggest doing away with deprivation line as benchmark for computing normalised deprivation score to address information loss in the computation of multidimensional index of deprivation.

In case of binary data, by default, maxima ought to be the deprivation line. Does this mean that there will be no information loss in case of binary data? The next section takes that up.

## 6. Does binary specification of dimensions help?

One of the problems with multidimensional index, as indicated in Pattanaik and Xu (2018), can be the following. Now, suppose two individuals $i$ and $i'$ have across five dimensions their $g._j{}^{\alpha}$ values for $\alpha = 1$ as $i = \{1,1,0,0,0\}$ and $i' = \{0,0,1,0.25,0.5\}$. With a $k$ cut-off of $3$ for both, one can say that $i$ is non-deprived but $i'$ is deprived. However, if one takes the average deprivation across dimensions for $\alpha = 1$ then $i$ and $i'$ have values of $0.4$ and $0.35$, but on account of $k$ cut-off $i$ is non-deprived while $i'$ is deprived. A way out of this could be if all the dimensions are represented in binary form.

In fact, the recent computations of multidimensional index of deprivation use dimensional information that is binary (including when the base information is cardinal), that is they take the value of either $1$ or $0$. This is fine if the data being used is binary to begin with. However, if available data are used through a threshold to convert them to binary then it surmounts to imposing a deprivation line that results in information loss. This information loss should be a matter of concern.

Getting back to the two individuals $i$ and $i'$, if data (which to being with could also be cardinal) were to be converted to binary, then $g._j{}^{\alpha}$ values for $\alpha = 1$ remains same for $i$ but changes for $i'$ such that $i'_b = \{0,0,1,1,1\}$ and average deprivation of $i'_b$ is now 0.6. But then this is on account of converting the normalised deprivation score to a binary form, which does surmount to information loss. Hence, once data have been collected, post-facto decision to convert them to binary should be seen as information loss.

Proposition 3: There can be information loss in multidimensional index of deprivation on account of converting dimension-specific data to binary form, $M_{\alpha|b} \geq M_{\alpha}$.

Proof: This is so because ${g_{ij|b}}^{\alpha} = 0{,}1$ while ${g_{ij}}^{\alpha} \in [0{,}1]$. It follows that there can be information loss for computations that involve converting dimension-specific data to binary form. Unlike in propositions 1 and 2 where information loss is identified with a lower value of multidimensional index of deprivation, why is it that $M_{\alpha|b} \geq M_{\alpha}$ is also considered as information loss. This is so because with the individuals identified as deprived in both the scenarios remaining the same, the individual $i$ who is deprived will in the binary scenario have ${g_{ij|b}}^{\alpha} = 1$ and with available data will have ${g_{ij}}^{\alpha} \leq 1$. In other words, in the case of deprived ${g_{ij|b}}^{\alpha}$ will take only one value while ${g_{ij}}^{\alpha}$ can take a wider range of values. It is in this sense that there is information loss in the binary form. For the same reason, $M_{\alpha|\hat{z}_j,b} \geq M_{\alpha|\hat{z}_j}$.

In the use of $M_{\alpha}$, we also suggest to not convert data to binary form to avoid information loss. Such conversion might do away with one problem, but it is like running away from the frying pan to the fire. In any case, no censoring and use of maxima, where appropriate, as benchmark in the computation of normalised deprivation score, would also do away with the problem.

The concerns of the three types of information loss in the computation of multidimensional index are important. We now take that up further through an empirical exercise.

## 7. An empirical exercise for information loss

To take advantage of the ingenuity and richness of $M_{\alpha}$ one should do away with information loss. We provide an empirical exercise that addresses the three concerns of information loss discussed earlier.

### *7.1 Data*

In an in-class exercise, students were first asked to spell out different dimensions that they consider to be important from the perspective of students. These dimensions are income (this could be in the form of fellowship or through other sources, a continuous variable in money units of Indian rupees, ₹), progress in education (a binary variable: satisfactory, 1; not satisfactory, 0), health status (categorical variable: reasonably healthy, 3; somewhat healthy, 2; somewhat unhealthy with mild issues, 1; and not being healthy on account of regular bouts of illness, 0), and access to resources (categorical variable: exceeds expectation, 4; meets expectation, 3; near expectation, 2; below expectation, 1; no resources, 0). It needs to be mentioned that the student cohorts are quantitatively trained and comfortable with numbers and after deliberation and mutual agreement they arrived at the numerical-cum-textual equivalence of the binary and categorical variables as part of the in-class exercise. In fact, it would not be incorrect to state that the numbers came out first and the equivalent text required additional deliberation.

Table 1: Attainments across four dimensions through a Talismanic exercise

| UID | Income, $v_1$; ('000 ₹) | Education Progress, $v_2$; (Satisfactory, 1; Not satisfactory, 0) | Health Status, $v_3$; (Reasonably healthy, 3; somewhat healthy, 2; somewhat unhealthy, 1; not healthy, 0). | Access to resources, $v_4$; (Exceeds expectation, 4; meets expectation, 3; near expectation, 2; below expectation, 1; no resources, 0) |
|---|---|---|---|---|
| $u_1$ | 5 | 1 | 1 | 3 |
| $u_2$ | 10 | 0 | 2 | 3 |
| $u_3$ | 8 | 0 | 2 | 2 |
| $u_4$ | 15 | 1 | 2 | 4 |
| $u_5$ | 9 | 1 | 2 | 1 |
| $u_6$ | 20 | 0 | 1 | 2 |
| $u_7$ | 10 | 0 | 1 | 2 |
| $u_8$ | 14 | 1 | 3 | 2 |
| $u_9$ | 12 | 1 | 1 | 3 |
| $u_{10}$ | 25 | 0 | 1 | 2 |

Source: Author's notes and recall from in-class Talismanic exercise

In the second step, following Mahatma Gandhi's Talisman, students were asked to give scores about another student that they can visualise (but not from their institute) in the four dimensions. The data for 10 individuals are given in Table 1.

*7.2 Multidimensional deprivation and information loss*

For the four dimensions, we propose deprivation lines of ₹15K for income, 1 for education progress, 2 for health status, and 3 for access to resources. Using these deprivation lines, we provide the normalised deprivation scores, $g_{ij}$, of the four dimensions and whether an individual is deprived in union, intersection, and intermediary ($k = 3$) methods in Table 2. From these, the headcount ratio for the union method is 0.9, for the intersection method is 0.1 and for the intermediary method is 0.4. It is true that the intermediary method lies somewhere in between the union and the intersection methods, but our purpose is to compute multidimensional index of deprivation.

Table 2: The normalised deprivation scores in the four dimensions and individual's deprivation in union, intersection, and intermediary methods

| UID | Income, $v_1$ | Education Progress, $v_2$ | Health Status, $v_3$ | Access to Resources, $v_4$ | Deprivation | | |
|---|---|---|---|---|---|---|---|
| | | | | | Union, ∪ | Intersection, ∩ | Intermediary, $k = 3$ |
| $u_1$ | 0.67 | 0 | 0.5 | 0 | 1 | 0 | 0 |
| $u_2$ | 0.33 | 1 | 0 | 0 | 1 | 0 | 0 |
| $u_3$ | 0.47 | 1 | 0 | 0.33 | 1 | 0 | 1 |
| $u_4$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $u_5$ | 0.40 | 0 | 0 | 0.67 | 1 | 0 | 0 |
| $u_6$ | 0 | 1 | 0.5 | 0.33 | 1 | 0 | 1 |
| $u_7$ | 0.33 | 1 | 0.5 | 0.33 | 1 | 1 | 1 |
| $u_8$ | 0.07 | 0 | 0 | 0.33 | 1 | 0 | 0 |
| $u_9$ | 0.20 | 0 | 0.5 | 0 | 1 | 0 | 0 |
| $u_{10}$ | 0 | 1 | 0.5 | 0.33 | 1 | 0 | 1 |

Source: Author's calculation based on data from in-class Talismanic exercise.

In Table 3, the comparison of contributions to multidimensional index of deprivation for $\alpha = 1$ indicates that $M_1 = 0.308$ and $M_{1|c} = 0.191$. For $\alpha = 0$, if $g_{ij} > 0$ then we will have ${g_{ij}}^0 = 1$ such that $w_{ij}{g_{ij}}^0 = 0.025$ (assuming equal weight across individuals, $w_i = 0.1$, and equal weight across dimensions, $w_j = 0.25$). Sum of these, without and with censoring, gives us $M_0 = 0.575$ and $M_{0|c} = 0.325$. Similarly, one can compute for other values of $\alpha$. In short, we have $M_\alpha \geq M_{\alpha|c}$, as indicated in Proposition 1.

Table 3: Comparing contributions to multidimensional index of deprivation with censored multidimensional index of deprivation when $\alpha = 1$

| UID | Contributions to Multidimensional Index of Deprivation, $M_1$ | | | | | Contributions to Censored Multidimensional Index of Deprivation, $M_{1\|c}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{0i}$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{0i\|c}$ |
| $u_1$ | 0.017 | 0 | 0.013 | 0 | 0.029 | 0 | 0 | 0 | 0 | 0 |
| $u_2$ | 0.008 | 0.025 | 0 | 0 | 0.033 | 0 | 0 | 0 | 0 | 0 |
| $u_3$ | 0.012 | 0.025 | 0 | 0.008 | 0.045 | 0.012 | 0.025 | 0 | 0.008 | 0.045 |
| $u_4$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $u_5$ | 0.010 | 0 | 0 | 0.017 | 0.027 | 0 | 0 | 0 | 0 | 0 |
| $u_6$ | 0 | 0.025 | 0.013 | 0.008 | 0.046 | 0 | 0.025 | 0.013 | 0.008 | 0.046 |
| $u_7$ | 0.008 | 0.025 | 0.013 | 0.008 | 0.054 | 0.008 | 0.025 | 0.013 | 0.008 | 0.054 |
| $u_8$ | 0.002 | 0 | 0 | 0.008 | 0.010 | 0 | 0 | 0 | 0 | 0 |
| $u_9$ | 0.005 | 0 | 0.013 | 0 | 0.018 | 0 | 0 | 0 | 0 | 0 |
| $u_{10}$ | 0 | 0.025 | 0.013 | 0.008 | 0.046 | 0 | 0.025 | 0.013 | 0.008 | 0.046 |
| $M_{0j}$ | 0.062 | 0.125 | 0.063 | 0.058 | 0.308 | | | | | |
| $M_{0j\|c}$ | | | | | | 0.020 | 0.100 | 0.038 | 0.033 | 0.191 |

Source: Author's calculation based on data from in-class Talismanic exercise.

Table 4: Contributions to multidimensional index of deprivation with maxima as benchmark for computing normalised deprivation scores

| UID | Contributions to Multidimensional Index of Deprivation with Maxima as Benchmark for Computing Normalised Deprivation Scores, $M_{0\|\hat{z}_j}$ | | | | | Contributions to Multidimensional Index of Deprivation with Maxima as Benchmark for Computing Normalised Deprivation Scores, $M_{1\|\hat{z}_j}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{0i\|\hat{z}_j}$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{1i\|\hat{z}_j}$ |
| $u_1$ | 0.025 | 0 | 0.025 | 0.025 | 0.075 | 0.020 | 0 | 0.017 | 0.006 | 0.043 |
| $u_2$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.015 | 0.025 | 0.008 | 0.006 | 0.055 |
| $u_3$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.017 | 0.025 | 0.008 | 0.013 | 0.063 |
| $u_4$ | 0.025 | 0 | 0.025 | 0 | 0.050 | 0.010 | 0 | 0.008 | 0 | 0.018 |
| $u_5$ | 0.025 | 0 | 0.025 | 0.025 | 0.075 | 0.016 | 0 | 0.008 | 0.019 | 0.043 |
| $u_6$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.005 | 0.025 | 0.017 | 0.013 | 0.059 |
| $u_7$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.015 | 0.025 | 0.017 | 0.013 | 0.069 |
| $u_8$ | 0.025 | 0 | 0 | 0.025 | 0.050 | 0.011 | 0 | 0 | 0.013 | 0.024 |
| $u_9$ | 0.025 | 0 | 0.025 | 0.025 | 0.075 | 0.013 | 0 | 0.017 | 0.006 | 0.036 |
| $u_{10}$ | 0 | 0.025 | 0.025 | 0.025 | 0.075 | 0 | 0.025 | 0.017 | 0.013 | 0.054 |
| $M_{0j\|\hat{z}_j}$ | 0.225 | 0.125 | 0.225 | 0.225 | 0.800 | | | | | |
| $M_{1j\|\hat{z}_j}$ | | | | | | 0.122 | 0.125 | 0.117 | 0.100 | 0.464 |

Source: Author's calculation based on data from in-class Talismanic exercise.

Table 4 shows contribution to multidimensional index of deprivation with maxima as benchmark for computing normalised deprivation scores such that $M_{0|\hat{z}_j} = 0.8$ and $M_{1|\hat{z}_j} = 0.464$. Compared with values in Table 3 and subsequent computations indicated earlier, one observes that $M_{\alpha|\hat{z}_j} \geq M_\alpha$, as indicated in Proposition 2. Combining with Table 3 and drawing from propositions 1 and 2 one can state that $M_{\alpha|\hat{z}_j} \geq M_\alpha \geq M_{\alpha|c}$.

Table 5: Contributions to multidimensional index of deprivation with binary health status

| UID | Contributions to Multidimensional Index of Deprivation with Binary Health Status and Number Deprived as in Base, $M_{1\|b}$ | | | | | Contributions to Multidimensional Index of Deprivation with Binary Health Status and Number Deprived as in Maxima, $M_{1\|\hat{z}_j,b}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{1i\|b}$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $M_{1i\|\hat{z}_j,b}$ |
| $u_1$ | 0.020 | 0 | 0.025 | 0.006 | 0.051 | 0.020 | 0 | 0.025 | 0.006 | 0.051 |
| $u_2$ | 0.015 | 0.025 | 0 | 0.006 | 0.046 | 0.015 | 0.025 | 0.025 | 0.006 | 0.071 |
| $u_3$ | 0.017 | 0.025 | 0 | 0.013 | 0.055 | 0.017 | 0.025 | 0.025 | 0.013 | 0.080 |
| $u_4$ | 0.010 | 0 | 0 | 0 | 0.010 | 0.010 | 0 | 0.025 | 0 | 0.035 |
| $u_5$ | 0.016 | 0 | 0 | 0.019 | 0.035 | 0.016 | 0 | 0.025 | 0.019 | 0.060 |
| $u_6$ | 0.005 | 0.025 | 0.025 | 0.013 | 0.068 | 0.005 | 0.025 | 0.025 | 0.013 | 0.068 |
| $u_7$ | 0.015 | 0.025 | 0.025 | 0.013 | 0.078 | 0.015 | 0.025 | 0.025 | 0.013 | 0.078 |
| $u_8$ | 0.011 | 0 | 0 | 0.013 | 0.024 | 0.011 | 0 | 0 | 0.013 | 0.024 |
| $u_9$ | 0.013 | 0 | 0.025 | 0.006 | 0.044 | 0.013 | 0 | 0.025 | 0.006 | 0.044 |
| $u_{10}$ | 0 | 0.025 | 0.025 | 0.013 | 0.063 | 0 | 0.025 | 0.025 | 0.013 | 0.063 |
| $M_{1j\|b}$ | 0.122 | 0.125 | 0.125 | 0.100 | 0.472 | | | | | |
| $M_{1j\|\hat{z}_j,b}$ | | | | | | 0.122 | 0.125 | 0.225 | 0.100 | 0.572 |

Source: Author's calculation based on data from in-class Talismanic exercise.

So far, in health status (Tables 3 and 4), the contribution to $M_{\alpha*}$ was obtained from Talismanic health status with different deprivation lines. We now use health status with two different binary forms (one where number deprived is the same as that for uncensored in Table 3, reasonably healthy and somewhat healthy get 1 while somewhat unhealthy and not healthy get 0; and the other where the number deprived is the same as in maxima in Table 4, only reasonably healthy get 1 and the rest get 0).

The results for both the scenarios for $\alpha = 0$ will remain the same, $M_{0|b} = M_0 = 0.575$ and $M_{0|\hat{z}_j,b} = M_{0|\hat{z}_j} = 0.8$. The results for both the scenarios for $\alpha = 1$ are given in Table 5.

Comparing these with Tables 3 and 4, we observe that $M_{1|b} = 0.472 > M_1 = 0.308$ and $M_{1|\hat{z}_j,b} = 0.572 > M_{1|\hat{z}_j} = 0.464$. Thus, $M_{\alpha|b} \geq M_\alpha$ or $M_{\alpha|\hat{z}_j,b} \geq M_{\alpha|\hat{z}_j}$, as in Proposition 3.

### *7.3 Multidimensional index of attainment or empowerment and information loss*

Having shown the three scenarios of information loss (censoring through $k$ on account of headcount fetish, deprivation line as benchmark to compute normalised deprivation score such that information on or above the deprivation line are treated in the same manner, and limiting to binary when the richness of data could allow for using additional information), we now extend the exercise to multidimensional index of attainment or empowerment using the same data. In Table 6, we provide two scenarios, the possible contribution to attainment, with or without the three information losses, $E_{1|c,b}$, and $E_{1|\hat{z}_j}$. The former has censoring, deprivation line and uses health status in binary form while the latter does away with all these.

Table 6: Contributions to multidimensional index of attainment with and without information loss

| UID | Contributions to Multidimensional Index of Attainment with the Three Types of Information Loss, $E_{1\|c,b}$ | | | | | Contributions to Multidimensional Index of Attainment without the Three Types of Information Loss, $E_{1\|\hat{z}_j}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $E_{1i\|c,b}$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $E_{1i\|\hat{z}_j}$ |
| $u_1$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.005 | 0.025 | 0.008 | 0.019 | 0.0571 |
| $u_2$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.010 | 0 | 0.017 | 0.019 | 0.0454 |
| $u_3$ | 0.013 | 0 | 0.025 | 0.017 | 0.055 | 0.008 | 0 | 0.017 | 0.013 | 0.0372 |
| $u_4$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.015 | 0.025 | 0.017 | 0.025 | 0.0817 |
| $u_5$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.009 | 0.025 | 0.017 | 0.006 | 0.0569 |
| $u_6$ | 0.025 | 0 | 0 | 0.017 | 0.042 | 0.020 | 0 | 0.008 | 0.013 | 0.0408 |
| $u_7$ | 0.017 | 0 | 0 | 0.017 | 0.033 | 0.010 | 0 | 0.008 | 0.013 | 0.0308 |
| $u_8$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.014 | 0.025 | 0.025 | 0.013 | 0.0765 |
| $u_9$ | 0.025 | 0.025 | 0.025 | 0.025 | 0.100 | 0.012 | 0.025 | 0.008 | 0.019 | 0.0641 |
| $u_{10}$ | 0.025 | 0 | 0 | 0.017 | 0.042 | 0.025 | 0 | 0.008 | 0.013 | 0.0458 |
| $E_{1j\|c,b}$ | 0.230 | 0.150 | 0.175 | 0.217 | 0.772 | | | | | |
| $E_{1j\|\hat{z}_j}$ | | | | | | 0.128 | 0.125 | 0.133 | 0.150 | 0.5363 |

Source: Author's calculation based on data from in-class Talismanic exercise.

Given that $E_\alpha = 1 - M_\alpha$, one would expect that the effect of information loss for multidimensional index of attainment to be opposite of what one observed for deprivation.

This is also observed in Table 6 as $E_{1|c,b} = 0.772 > E_{1|\hat{z}_j} = 0.536$. In other words, information loss has the potential danger of showing greater attainment. It is also important to note that the contribution of individual $i$ to $E_{1|c,b}$ indicated under $E_{1i|c.b}$ shows that there are four distinct values (six individuals with 0.1, two individuals with 0.042 and one individual each with 0.033 and 0.055) whereas contribution of individual $i$ to $E_{1|\hat{z}_j}$ indicated under $E_{1i|\hat{z}_j}$ shows that all the ten individual have distinct values. In other words, information loss could lead to clustering that does away with the advantages that a multidimensional index like $E_\alpha$ or $M_\alpha$ can bring.

The Tables 3-6 reiterates the information loss in the computation of multidimensional index of deprivation and attainment. We now take up some other observations.

## 8. Some other observations

The computation provided in Tables 3-6 suggest that $M_\alpha$ (or $E_\alpha$), with or without different kinds of information loss, are decomposable. In other words, each and every individual's contribution across all the dimensions adds up to the overall measure, $M_\alpha = \sum_i \sum_j w_{ij}\, g_{ij}{}^{\alpha}$. Besides, each individual's contribution across all the dimensions adds up to $M_{\alpha i}$, and each dimension's contribution over all the individuals adds up to $M_{\alpha i}$. Further, $M_\alpha = \sum_i M_{\alpha i} = \sum_j M_{\alpha j}$ (see last row for column's under $M_{\alpha i*}$ or last column for row's under $M_{\alpha j*}$). Similarly, for $E_\alpha$.

One also observes in Tables 3-6, with or without information loss, that there is no reference to headcount ratio. And, when we do away with censoring there is no reference to the $k$ cut-off across dimensions. Given the purpose of $M_\alpha$, as indicated earlier, headcount fetish leading to censoring is akin to throwing the baby with the bathwater.

Second, censoring is counter intuitive. For instance, if a differently abled individual has her or his capabilities compromised but that individual's deprivation turns out to be below a $k$ cut-off, and hence, not considered deprived. In a similar vein, an individual who was considered deprived, but through public provisioning became a participant of a housing scheme and because of that the individual is now below the $k$ cut-off even if there is not much of a perceptible difference in the other dimensions. Nevertheless, the individual is now non-deprived.

One may argue that, technically speaking, $k$ cut-off across dimensions for any individual is like a deprivation line in any dimension, $z_j$. Well, not exactly. The semblance of similarity is in the sense that one is across dimensions (say, for each row) and the other is within each dimension (say, for each column). However, they are both not the same. It would be same if we bring in another cut-off indicating that if at least this many individuals are deprived in a dimension, then only the total number of deprived in that dimension would be considered, otherwise all individuals in that dimension would be considered as non-deprived. If this seems inappropriate then, so should be $k$ cut-off across dimensions.

To wit, usage of $H$ and $A_\alpha$ in the computation of $M_\alpha$ is an unnecessary detour. It aids in our understanding and that is another matter, but then reference to a high $H$ cannot be the basis to argue for $k$ cut-off. Hence, I rest my case against $k$ cut-off in the computation of $M_\alpha$.

If deprivation line is technically not similar to $k$ cut-off, then what are the concerns regarding information loss with it. The use of deprivation line as a benchmark to compute normalised deprivation score creates a different problem. It does not distinguish between all the individuals who are on or above the deprivation line. Such a restriction might be relevant when one is discussing about the deprived alone, but if one is to discuss about the deprived in relation to the non-deprived then it leads to information loss, which is to be avoided. It is in this context that we propose the usage of maxima as benchmark, where appropriate, to derive dimension-specific normalised deprivation score.

The use of maxima will be both local and global in a specific exercise. However, if the dimension-specific normalised deprivation score or multidimensional index values are to be comparable then one should use appropriate global maxima.

The information loss while converting data to a binary form is of another type. It leads to clustering of information to two categories: those on or above the deprivation line to one category and those below the deprivation line to another category. The former is akin to the deprivation line in the sense that all non-deprived are treated in a similar manner while the latter implies that variation among the deprived is also not to be taken into consideration. In our empirical exercise, the Talismanic health status ordinal data with four categories was converted to binary data. It is true that quantification of ordinal data can raise concerns. Nevertheless, we used the Talismanic health status and other categorical data from an in-class exercise where the numerical-cum-textual equivalence were arrived at through mutual agreement that provides us with an indicative quantification. Keeping that aside, the point we want to make is post-facto binary categorisation at the analysis stage does lead to information loss. It is also possible for cardinal data to be used in a binary form. Such an imposition also leads to information loss.

Increasingly, $M_\alpha$ or $E_\alpha$ are being used as a variable in subsequent statistical analysis, either as an independent variable or as a variable that is of interest from the outcome perspective, say in a supposedly public policy intervention. In such situations, it would help if all the three types of information loss pointed out here are also avoided. We now provide our concluding remarks.

## 9. Conclusion

In any aggregation if multiple dimensions are being reduced to a single index, then that does surmount to information loss, which is inevitable, and this paper is not contesting that. Rather, the point is that all the available information in the multiple dimensions is not being utilised. In the multidimensional index of $M_\alpha$ or $E_\alpha$ the information loss of this latter type is on account of censoring, normalising deprivation score by using deprivation line as a benchmark (excluding situations where the deprivation line is a norm of requirement like for nutrition or situations where beyond a point greater numerical value is not to be identified with greater attainment), and converting of data to a binary form. By censoring, an individual's deprivation in some dimensions will not be used because the individual may not be considered as deprived. Normalising deprivation score by using a deprivation line as benchmark would mean that all attainments on or above that line will be considered in the same manner. Converting data to binary form would mean that the entire information set is just divided into two categories. The three types of information loss are also shown through an empirical exercise, which is based on data that came out of an in-class Talismanic exercise.

While raising the concerns of information loss on account of headcount fetish leading to censoring, normalising deprivation scores through deprivation line as benchmark, and converting data to binary form, we are not trying to contest the relevance of headcount ratio, deprivation line or converting of data from one form to another. Our contention in terms of information loss is restricted to their usage in computation of $M_\alpha$ or $E_\alpha$. It is in this context and to restore the ingenuity of the multidimensional index of $M_\alpha$ or $E_\alpha$ that we propose no censoring, we suggest use of maxima where appropriate, and we encourage taking advantage of the richness in available data. This is important for the computation of $M_\alpha$ or $E_\alpha$ as an independent exercise, but also when such multidimensional indices are used as a variable in subsequent statistical or econometric analysis.

**References**

Alkire, S., Foster, J. (2011). "Counting and Multidimensional Poverty Measurement," *Journal of Public Economics*, 95 (8-9): 476-487. DOI link accessed 18 Apr 2024.

Alkire, S., Foster, J., Seth, S., Santos, M.E., Roche, J.M., Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*, Oxford; online edition, Oxford Academic. DOI link accessed 18 Apr 2024.

Alkire, S., Meinzen-Dick, R., Peterman, A., Quisumbing, A., Seymour, G., Vaz, A. (2013). "The Women's Empowerment in Agriculture Index," *World Development*, 52: 71-91. DOI link accessed 18 Apr 2024.

Dutta, B., Panda, M., Wadhwa, W. (1997). "Human Development in India." In *Measurement of Inequality and Poverty*, edited by S. Subramanian, 329–357. New Delhi: Oxford University Press.

Foster, J., Greer, J., Thorbecke, E. (1984). ‘A Class of Decomposable Poverty Measures,” *Econometrica*, 52 (3): 761-766. JSTOR DOI link accessed 18 Apr 2024.

Kumar, T.K., Mallick, S., Holla, J. (2009). “Estimating Consumption Deprivation in India Using Survey Data: A State-Level Rural–Urban Analysis Before and During Reform Period,” *The Journal of Development Studies*, 45 (4): 441-470. DOI link accessed 19 Apr 2024,

Pattanaik, P.K., Xu, Y. (2018). “On Measuring Multidimensional Deprivation,” *Journal of Economic Literature*, 56 (2): 657-72. DOI link accessed 18 Apr 2024.

Samal, R.R., Mishra, S. (2023). “Food Security among Kandhas of Kandhamal, Odisha, India: A Mixed Method Study,” *Journal of Asian and African Studies*, 58(8): 1443-1464. DOI link accessed 18 Apr 2024.