Social Science & Medicine

Comparing indices of relative deprivation using behavioural evidence

Hilda Osafo Hounkpatin\textsuperscript{a*}, Alex M. Wood\textsuperscript{b} and Gordon D. A. Brown\textsuperscript{c}

\textsuperscript{a}School of Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK

\textsuperscript{b}Department of Psychological and Behavioural Science, London School of Economics and Political Science, London, WC2A 3LJ, UK.

\textsuperscript{c}Department of Psychology, University of Warwick, Coventry, CV4 7AL, UK

\* Corresponding author. School of Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK. E-mail addresses: H.O.Hounkpatin@soton.ac.uk (H. O. Hounkpatin), alex.wooduk@gmail.com (A. M. Wood), g.d.a.brown@warwick.ac.uk (G. D. A. Brown).
Abstract
What measure of relative deprivation best predicts health? While numerous indices of relative deprivation exist, few studies have compared how well different measures account for empirical data. Hounkpatin et al. (2016) demonstrated that the relative ranked position of an individual $i$'s income within a comparison group (their relative rank) was a better predictor of $i$'s health than $i$'s relative deprivation as assessed by the widely-used Yitzhaki index. In their commentary, Stark and Jakubek (2020) argue that both relative rank and relative deprivation may matter, and they develop a composite index. Here we identify some issues with their composite index, develop an alternative based on behavioral evidence, and test the various indices against data. Although almost all existing indices assume that the significance of an income $y_j$ to an individual with income $y_i$ ($y_j > y_i$) will be some increasing function of the difference between $y_i$ and $y_j$, we find that the influence of $j$'s income on $i$'s health is actually a reducing function of $(y_j - y_i)$. This finding — that less significance is assigned to distant higher incomes than to near higher incomes — is consistent with the well-established idea that we compare ourselves primarily to similar others.


**Introduction**

We welcome the rejoinder by Stark and Jakubek (2020) (SJ) on the issue of how best to model the psychosocial pathway linking relative deprivation to health and well-being. Their rejoinder, like our original paper and consistent with a variety of evidence (e.g., Boyce, Brown, & Moore, 2010; Daly, Boyce, & Wood, 2015; Mendelson, Kubbansky, & Thurston, 2008; Wetherall, Daly, Robb, Wood, & O'Connor, 2015), assumes that relative rank of income matters. Thus the main point at issue is whether there is only an effect of relative rank, as suggested in our original paper, or whether there is also an effect of relative deprivation as defined in terms of the summed weight of incomes above an individual’s own income (as in the Yitzhaki index). Although we believe that relative rank (RR) can legitimately be interpreted as one basis for measuring relative deprivation, for clarity here we retain SJ’s usage of the term "relative deprivation" (RD) to refer specifically to the measure given by the Yitzhaki index.

In this commentary we evaluate the ability of different indices of relative deprivation to fit empirical data. First, we take issue with some features of the composite index that SJ propose. We then offer an alternative formulation in which the RR and RD hypotheses are special cases of a more general equation. Third, we evaluate the RR and RD hypotheses against empirical data. We also note some psychologically implausible characteristics of the Yitzhaki index as a measure of RD. Finally, taking up SJ’s point about the detrimental effects of income inequality, we argue that there is no contradiction between the RR hypothesis and effects of societal income inequality.

**The Stark composite measure**

The Yitzhaki index of deprivation, $RD_i$, for individual $i$, captures the extent to which other incomes in the relevant social comparison group exceed $i$’s income. The measure is normalised by the total number $n$ of individuals in the comparison group. As SJ show (their equation 2), the index can be expressed as a product of two terms: a term that specifies the proportion of the population with incomes higher than $i$’s (i.e., $\hat{R}_i$, the approximate complement of relative rank as we use the term in the original paper), and a term that specifies the average amount by which these higher incomes exceed $i$’s income.

We find SJ’s reformulation intuitive and attractive. However, we are less convinced of the psychological interpretability of the further step taken in SJ’s equation 3, in which it
proposed to add weightings to each of the two terms to enable the relative contributions of RR and RD to ill health to be estimated using empirical data. SJ define their composite measure of relative deprivation, \( CRD_i(\gamma) \), as:

\[
CRD_i(\gamma) = \tilde{R}_i^{\gamma} (\tilde{y}_i - y_i)^{1-\gamma}
\]

where \( y_i \) is the income of individual \( i \), \( \tilde{y}_i \) is the mean of incomes higher than \( y_i \), and \( \gamma \in [0,1] \) is a weighting parameter. SJ suggest that the parameter \( \gamma \) could be estimated, allowing the relative contributions of the magnitude of income differences and a purely rank-based measure to be estimated with appropriate empirical data.

One issue with this formulation is that there is no value of \( \gamma \) that allows the \( CRD_i(\gamma) \) index to represent \( RD_i \) as defined by Yitzhaki. When \( \gamma = 1 \), \( CRD_i(\gamma) \) does reduce to a pure rank-based measure. When \( \gamma = 0 \), however, \( CRD_i(\gamma) \) does not reduce to a pure measure of \( RD_i \). Instead, it becomes simply \( (\tilde{y}_i - y_i) \) which differs from the Yitzhaki index of RD because the sum of incomes higher than \( i \)'s are not now normalized by \( n \) (the total number of incomes in the reference group), as they are in Yitzhaki’s index, but instead are normalized only by the number of incomes higher than \( i \)'s. We can illustrate the problem using incomes 1,2,3,4,5, and the income of the third individual (i.e., 3). With \( \gamma = 0 \), we have \( CRD_3(0) = (4.5 - 3) = 1.5 \). However \( CRD_3(0) \) would take the same value if the set of comparison incomes was just \( y = [3 4 5] \), or \( y = [0.5 1 1.5 2 2.5 3 4 5] \), or if \( y = [1 2 3 3.5 4 5 5.5] \). We suggest that the fact that \( CRD_3(0) \) may be invariant to the number of incomes higher than \( y_i \), and will always be invariant to the number of incomes lower than \( y_i \), renders it both (a) very different from the Yitzhaki formulation, and (b) implausible as a candidate measure of relative deprivation.

As SJ note, for the special case where \( \gamma = .5 \), \( CRD_i(\gamma) \) becomes \( \sqrt{RD_i} \). But \( \sqrt{RD_i} \) is not the same as \( RD_i \), and moreover values of \( \gamma \) between 1 and .5 cannot be straightforwardly interpreted in terms of the relative weights given to a rank-only measure and the RD measure. For empirical estimation, we therefore need an alternative formulation — one in which rank-based and Yitzhaki-based measures can be interpreted as special cases. We develop this in the next section.

**An alternative measure**
Numerous measures of relative deprivation have been proposed, many of which build on Yitzhaki (1979) (see Adjaye-Gbewonyo & Kawachi, 2012). Almost all, following Runciman (1966), focus on the idea that only incomes higher than $i$’s will influence $i$’s sense of relative deprivation (although see D’Ambrosio & Frick, 2012). The number of incomes lower than $i$’s does influence relative deprivation according to most metrics, however, because population size enters into the denominator of the Yitzhaki index and most of its successors.

The indices differ primarily in the extent to which the influence of $j$’s income on $i$’s relative deprivation ($y_j > y_i$) depends on the distance between $y_j$ and $y_i$. Thus the Yitzhaki index assumes that the effect will depend on $(y_j - y_i)$, as does a recent dynamic index which takes account of the number of people who have overtaken $i$ between one time period and the next (Bossert & D’Ambrosio, 2007; Bossert & D’Ambrosio, 2020). Several approaches assume instead that the significance of $y_j$ for the relative deprivation of $i$ will be some increasing but concave function of $(y_j - y_i)$, such that $i$’s deprivation increases less than proportionately with the amount by which $y_j$ exceeds $y_i$ (Bossert & D’Ambrosio, 2014; Chakravarty & Chakraborty, 1984; Esposito, 2010; Paul, 1991; Podder, 1996; Stark, Bielawski, & Falniowski, 2017). These approaches allow the distribution of incomes $> y_i$ to influence $i$’s deprivation. The concavity is sometimes enforced by the functional form of the relevant deprivation index, and sometimes by assumptions about parameter bounds. In either case, however, there is no allowance for the possibility that incomes much higher than $y_i$ might have less significance for $i$’s deprivation than incomes also higher than but closer to $y_i$. This possibility is in contrast allowed for in the expression we now develop and then test against data.

Our alternative measure, which we denote by $\text{CR}_i$, is based in part on a formulation developed by Brown et al. (2008), but differs from the more general treatment there in that (a) like SJ we focus on $\bar{R}_i$, (i.e., the complement of relative ranked position) rather than RR, (b) we ignore the possibility of loss aversion, and (c) we do not include a term allowing incomes below $y_i$ to be weighted according to how far below $y_i$ they are. The latter two points are discussed briefly below. More recent but related approaches can be found in Stark, Bielawski and Falniowski (2017) and in Bossert and D’Ambrosio (2020).

Let $y = (y_1, \ldots, y_n)$ be an ordered vector of $n$ increasing incomes. We first note that, for large $n$:
\[ R_i \approx \frac{(n - i)}{(n - 1)} = \frac{\sum_{k=i+1}^{n}(y_k - y_i)^0}{(n - 1)} \]

We first then add an additional parameter, \( \alpha \), and define our measure as:

\[ \bar{C}R_i = \frac{\sum_{k=i+1}^{n}(y_k - y_i)^\alpha}{(n - 1)} \]

Thus when \( \alpha = 0 \), \( \bar{C}R_i \approx \bar{R}_i \), and when \( \alpha = 1 \), in contrast, \( \bar{C}R_i \approx RD_i \) for large \( n \) (the approximation reflecting the fact that the denominator is \( (n-1) \) instead of \( n \)). An alternative formation is needed if the measure is to be bounded between 0 and 1; to preserve comparability with SJ we do not use that formulation here.

The \( \bar{C}R_i \) index therefore allows smooth transition between RR and RD models as \( \alpha \) varies and allows for meaningful interpretation of both negative and positive values of \( \alpha \). To adopt SJ’s example, consider \( \bar{C}R_3 \) in the context of incomes \( y = [1, 2, 3, 4, 5] \). When \( \alpha = 0 \), the contribution of the highest income (i.e., 5) to the index is \( (5 - 3)^0 = 1 \), and this contribution would not change if the income was 10 rather than 5. If \( \alpha = 1 \) the contribution of the highest income is \( (5 - 3)^1 = 2 \), and as in the Yitzhaki index this contribution would increase if the highest income was 10 rather than 5 (because \( (10 - 3)^1 = 7 \)). However, the contribution of higher incomes to deprivation need not be either independent of their distance above \( y_i \) (\( \alpha = 0 \)) or directly proportional to their distance above \( y_i \) (\( \alpha = 1 \)). Instead, the contribution of higher incomes might increase as a convex function of their distance above \( y_i \) (\( \alpha > 1 \)) or a concave increasing function (\( 0 < \alpha < 1 \)). Finally, and of particular interest in the present context, the \( \bar{C}R_i \) index allows for the possibility that the contribution of higher incomes might decrease with their distance above \( y_i \) (\( \alpha < 0 \)). For example, when \( \alpha = -0.5 \), the influence of the highest income on \( \bar{C}R_3 \) is smaller when it is 10: \( (10 - 3)^{-0.5} = .378 \) than when it is 5: \( (5 - 3)^{-0.5} = .707 \).

The idea that higher incomes might have less influence on feelings of deprivation as their distance increase is, we argue, psychologically plausible. It is well established in many domains of psychology that people tend to compare themselves mostly with similar others (e.g., Festinger, 1954), and such similarity may be defined in terms of income as well as in terms of characteristics such as age and education level. Perhaps, for example, the feeling of
deprivation felt by a professor of economics may be increased more by the presence of a departmental colleague earning just a little more than they do than by the much higher salary of the university president, and perhaps the dissimilarity of the president is defined psychologically at least partly in terms of the income difference.

We suggest that it may therefore be useful to use an index of i’s deprivation that not only includes a rank-only index and the Yitzhaki index as special cases but also allows for the influence of higher incomes to be either positively or negatively related to the amount by which they exceed i’s income. The index that we have proposed has these properties, and in the next section we therefore evaluate it against empirical data.

**Empirical data**

As in our original paper, we used data from a nationally representative sample of 14,224 observations from 9,404 participants across three waves (2004, 2008, and 2012) of the English Longitudinal Study of Aging (ELSA). Multilevel regression models (based on lagged and contemporaneous values for each relative deprivation measure) were used to compare the fit (determined by goodness of fit statistics) of the RR model, RD model, SJ’s $CRD_i(y)$ and our $\tilde{CR}_i$ index (Table 1). The results indicated $\tilde{CR}_i$ was the best predictor of both self-rated health and allostatic load (Table 1). For self-rated health, the best fitting $\tilde{CR}_i$ was derived using $\alpha = -0.30$ for contemporary $\tilde{CR}_i$ and $\alpha = -0.40$ for lagged $\tilde{CR}_i$. For allostatic load, the best fitting $\tilde{CR}_i$ was derived using $\alpha = -0.30$ for contemporary $\tilde{CR}_i$ and $\alpha = -0.50$ for lagged $\tilde{CR}_i$.

The fact that the value of $\alpha$ that leads to the best fit is negative means that the relative deprivation of i will be less influenced by incomes much higher than i’s income than by higher incomes that are closer to i’s. Following SJ, we can illustrate this with incomes [1 2 3 4 5]. When $\alpha = 0$, the deprivation associated with income 3 is simply 0.5 (3’s relative rank). When $\alpha = -.5$ or 1, the deprivations are .43 and .75 respectively. If we change the highest income from 5 to 10, and $\alpha = 0$, the relative rank does not change. If $\alpha = -.5$, the deprivation reduces to 0.34, while if $\alpha = 1$ it increases to 2.0.

Note that the relationship between our index of relative deprivation (when $\alpha$ is negative) and income is qualitatively different to the relationship between income and either RR or RD. For example, if we compare $\bar{R}_i$ with $\log(y_i)$ we observe that $\bar{R}_i$ reduces rather slowly with income at both low and high levels of income, but reduces rather more sharply for intermediate levels of log income. The Yitzhaki index has the same qualitative property,
in that it falls most steeply for intermediate levels of income, but is qualitatively different in that it is less flat than $\tilde{R}_i$ for low and high incomes and correspondingly falls less steeply for intermediate incomes.

With our $\tilde{C}R_i$ index and $\alpha < 0$, however, the pattern is qualitatively different in that $\tilde{C}R_i$ does not necessarily decline monotonically with income. Instead, deprivation may be highest for intermediate incomes if those incomes are in the most dense region of the income distribution. This is because those incomes are surrounded by many other incomes that are just a little bit higher than they are themselves, and it is these higher-but-nearby incomes that have the greatest effect when $\alpha$ is negative. An alternative psychological interpretation, which we cannot exclude with present data, is that people’s sense of relative deprivation is based on the relative ranked position of their income within a reference group that is sampled from a narrower range of incomes than is typically considered in studies of relative deprivation.

The psychological implausibility of the Yitzhaki index

We note some potential psychological limitations with both the Yitzhaki index and some of the generalizations of it (including the ones discussed here). Calculation of the Yitzhaki index often uses untransformed incomes and differences in those incomes. Moreover, it only takes into account the magnitude of higher incomes and ignores the magnitudes of lower incomes. However one of the most prominent models in behavioural economics, Prospect Theory (Kahneman & Tversky, 1979), assumes that judgements are typically made relative to a reference point (which in the current context would be $y_i$) and are influenced by both losses and gains relative to that reference point. A key tenet of Prospect Theory is loss aversion - the idea that “losses loom larger than gains.” Consistent with the applicability of Prospect Theory to people's perceptions of their own incomes in the context of others' incomes, Boyce et al. (2013) show that the effects of income changes on life satisfaction are asymmetrical: The reductions in subjective well-being associated with reductions in income are in absolute terms greater than the increases in subjective well-being associated with increases in income. In the context of relative deprivation, it seems plausible that higher incomes would represent "losses" and lower incomes would represent “gains”, each relative to the reference point represented by income. We might therefore expect that the effect of lower incomes on the subjective deprivation might, like the effect of higher incomes, depend on how far below $y_i$ they fall. This possibility is not incorporated in any of
the formulations we have considered here, but can be included in more general formulations (Brown et al., 2008; D'Ambrosio & Frick, 2012).

**Rank-based indices and inequality**

Finally, we note SJ’s point regarding the negative effect of income inequality on various indices of societal and individual well-being (Wilkinson & Pickett, 2009, 2018). Rank-only measures will, by definition, be insensitive to the distribution of incomes in a population. Some data are consistent with this observation. For example, self-rated life satisfaction has been argued to reflect the relative ranked position of individuals within a comparison group (Boyce et al., 2010) and, consistent with this, there is little or no effect of country-level inequality on average subjective life-satisfaction within countries (Kelley & Evans, 2017a, 2017b). However there is no inconsistency between claims that relative rank of income determines relative deprivation and the claims that country-or state-level income inequality is negative for either subjective or objective health. This is because income inequality appears to have effects through making people more materialistic and more status-conscious (e.g., Wilkinson & Pickett, 2018). For example, people perform more Internet searches related to status-related items such as designer clothes when income inequality is high (Walasek & Brown, 2015, 2016) and a similar association between income inequality and concerned with social status is seen in people's tweets (Walasek, Bhatia, & Brown, 2018). It is well established that increased concern with status and social comparison is negative for important aspects of health (Kasser, 2002); impaired health and well-being in more unequal societies may therefore reflect increased attention to well-being-negative aspects of life rather than individual-level sensitivity to cardinal properties of income distributions.

**Summary and conclusion**

We have presented an alternative formulation of a relative deprivation index that, we claim, offers a more natural psychological interpretation than SJ’s composite measure. Moreover, our formulation, which allows for the influence of higher incomes to be negatively related to the amount by which they exceed i’s income, better explained differences in self-rated health and allostatic load. Our results do however come with a number of caveats. The different indices of relative deprivation can often mimic each other rather closely, and further research will be needed to see how robust the selection of best-fitting function is to choice of dataset and decision made during analysis. Moreover, rank-based transformations can reduce...
outliers that may reduce the fit of non-rank models. However, together with the need to engage with policy makers to reduce impact of relative deprivation, it will be important to develop and test measures of relative deprivation that respect and reflect the relevant underlying psychological processes.

Acknowledgements

This study was supported by the Economic and Social Research Council (U.K.) (grant number ES/P008976/1) and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 788826).
References


Hounkpatin, H. O., Wood, A. M., & Dunn, G. (2016). Does income relate to health due to psychosocial or material factors? Consistent support for the psychosocial hypothesis requires operationalization with income rank not the Yitzhaki Index. *Social Science & Medicine, 150*, 76-84.


Kelley, J., & Evans, M. D. R. (2017a). The new income inequality and well-being paradigm: Inequality has no effect on happiness in rich nations and normal times, varied effects in extraordinary circumstances, increases happiness in poor nations, and interacts with individuals' perceptions, attitudes, politics, and expectations for the future. *Social Science Research, 62*, 39-74.


Table 1 Goodness of fit statistics for multilevel regression models

<table>
<thead>
<tr>
<th>Rank</th>
<th>Yitzhaki Index</th>
<th>CRD</th>
<th>( \overline{CR}_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>BIC</td>
<td>AIC</td>
</tr>
<tr>
<td>Region</td>
<td>36,080.94</td>
<td>36,322.94</td>
<td>36119.35</td>
</tr>
<tr>
<td>ELSA, self-rated health (N=14224)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>16,802.07</td>
<td>17,022.77</td>
<td>16811.04</td>
</tr>
<tr>
<td>ELSA, allostatic load (N=7310)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Best fit model in bold. Best fit for \( CRD_t(\gamma) \) measure for self-rated health used \( \gamma = 0.70 \) for both contemporary \( CRD_t(\gamma) \) and lagged \( CRD_t(\gamma) \). Best fit for \( CRD_t(\gamma) \) measure for allostatic load used \( \gamma = 0.80 \) for contemporary \( CRD_t(\gamma) \) and \( \gamma = 0.30 \) for lagged \( CRD_t(\gamma) \).