

Correcting the Concentration Index

Guido Erreygers

Department of Economics, University of Antwerp
City Campus, Prinsstraat 13, 2000 Antwerpen, Belgium
guido.erreygers@ua.ac.be

November 10, 2006

Abstract

In recent years attention has been drawn to several shortcomings of the Concentration Index, a frequently used indicator of the socioeconomic inequality of health. Some modifications have been suggested, but these are only partial remedies. This paper proposes a corrected version of the Concentration Index which is superior to the original Concentration Index and its variants, in the sense that it is the only rank-dependent indicator which satisfies four key requirements (transfer, level independence, cardinal consistency, and mirror). The paper also shows how the corrected Concentration Index can be decomposed and generalized.

Keywords: Health inequality, Socioeconomic inequality, Concentration Index

JEL Classification Number: D63, I10

Acknowledgements: I am very grateful to Tadele Ferede, Peter Lambert, Ellen Van de Poel, Tom Van Ourti and Buhong Zheng for discussions on the subject matter of this note, and for comments on previous drafts. All errors are my responsibility.

1 Introduction

The Concentration Index enjoys an increasing popularity as an indicator of the inequality of health in relation to the socioeconomic position of individuals. Among the factors which have contributed to its success as a measure of the socioeconomic inequality of health we can mention its affinity with the Gini coefficient, its visual representation by means of the Concentration Curve, and the ease with which it can be decomposed. But there is also a growing awareness that it is far from being a perfect measure. In recent years attention has been drawn to the fact that the bounds of the Concentration Index may depend upon the mean of the health variable and hence make a comparison of populations with different mean health levels problematic (Wagstaff, 2005), that different rankings are obtained if inequalities in ill health rather than inequalities in health are considered (Clarke e.a., 2002), and that the value of the index is to a large extent arbitrary if the health variable is of a qualitative nature (Erreygers, 2006 and Zheng, 2006).

Various suggestions have been made to remedy some of these defects. What is lacking, though, is an attempt to come to grips with *all* of these problems. In this paper, I propose a corrected version of the Concentration Index which goes a long way in that direction. After a discussion of the three issues, I analyze the modifications which have been proposed in the literature. In the core of the paper I follow an axiomatic approach which allows me to derive a new correction formula. I show that the resulting indicator is the only one of its type which satisfies four key requirements, and indicate how the approach can be generalized. I end with a brief empirical illustration.

2 The Health Concentration Index

Let us consider a given population of n individuals represented by the set $N = \{1, 2, \dots, n\}$. To each individual we assign a rank λ_i based upon this person's socioeconomic position, with the best well-off individual ranked first and the least well-off ranked last. In the case of ties, we assign to each member of the tied group the average rank of the group. The variable in whose distribution we are interested is health. Suppose that we have information about the health situation of the population, e.g. about every person's general health condition, or about one particular aspect of it. This information is translated into the health variable h_i , a real number which measures the health status of individual i . The vector $h = (h_1, h_2, \dots, h_n)$ represents the health situation of the whole population.

A higher value of h_i indicates a better health situation of individual i .

In addition, I assume that the health status variable of each individual lies between certain bounds. Let a_h and b_h designate respectively the lower and the upper bound of the health variable. I assume that both the lower limit a_h and the upper limit b_h are finite, that $a_h < b_h$, and, at least for a start, that the lower limit is nonnegative. So for each person i we have $0 \leq a_h \leq h_i \leq b_h < +\infty$. The average health of the population will be denoted as μ_h :

$$\mu_h \equiv \frac{1}{n} \sum_{i=1}^n h_i \quad (1)$$

The health Concentration Index $C(h)$, introduced by Kakwani (1980: 173-178) and Wagstaff, Paci and Van Doorslaer (1991), is defined as follows:

$$C(h) \equiv 1 - \frac{\sum_{i=1}^n (2\lambda_i - 1)h_i}{n^2\mu_h} \quad (2)$$

It is a measure of the socioeconomic inequality of health based upon information on the socioeconomic ranks and the health levels of all individuals in the population. A positive value of $C(h)$ indicates that health is distributed in favour of the rich, and a negative one that it is distributed in favour of the poor. The higher the absolute value of the index, the more extreme the pro-rich or pro-poor character of the distribution is supposed to be.

3 The Bounds of the Concentration Index

Let us begin by examining the bounds issue. It is well-known that the value of $C(h)$ can never be lower than -1 and higher than $+1$. The exact bounds are $-(n-1)/n$ and $+(n-1)/n$, which would be obtained if either the least well-off or the best well-off monopolized health. Recently, however, Wagstaff (2005) has observed that if the health variable is binary, i.e. such that each h_i takes either the value of 0 or the value of 1, then the bounds of the Concentration Index depend upon the mean of the health variable. The bounds turn out to be much wider for populations with a low mean (i.e. close to 0) than for populations with a high mean (i.e. close to 1).

At first sight the result seems to be surprising, and one might suspect that it is caused by the binary nature of the health variable. But varying bounds are by no means limited to the case of a binary variable. In fact, it can be shown that any health variable with a finite upper value or a positive lower value gives rise to varying bounds of the Concentration Index. From what I will call further in the paper the “transfer” property, it follows that the

most extreme pro-rich distribution of health which is compatible with a given mean health, consists of a “bipolar” distribution in which the richest persons have maximum health and the rest minimum health. For a population with average health μ_h , let us define an auxiliary variable ψ_h :

$$\psi_h \equiv \frac{(\mu_h - a_h)}{(b_h - a_h)} \quad (3)$$

This variable indicates where the mean lies between the two extremes of a_h and b_h . (For instance, a value of $\psi_h = \frac{1}{2}$ indicates that μ_h lies exactly halfway between a_h and b_h .) The most extreme pro-rich distribution which is compatible with mean health μ_h , consists of the $\psi_h n$ richest persons having maximum health and the $(1 - \psi_h)n$ poorest persons having minimum health. Likewise, the most extreme pro-poor distribution of health which is compatible with this mean, consists of the $\psi_h n$ poorest persons having maximum health and the $(1 - \psi_h)n$ richest persons having minimum health.¹ We have the following:

Proposition 1 *If the health variable varies between a_h and b_h , with $0 \leq a_h < b_h < +\infty$, then the health Concentration Index $C(h)$ of a population with average health μ_h is such that:*

$$-\frac{(b_h - \mu_h)(\mu_h - a_h)}{\mu_h(b_h - a_h)} \leq C(h) \leq +\frac{(b_h - \mu_h)(\mu_h - a_h)}{\mu_h(b_h - a_h)} \quad (4)$$

Proof. The upper bound of $C(h)$ is reached when the $\psi_h n$ richest persons have the maximum health level b_h and the $(1 - \psi_h)n$ poorest persons the minimum health level a_h . It can then be written as:

$$C(h) = 1 - \frac{\sum_{i=1}^{\psi_h n} (2\lambda_i - 1)b_h + \sum_{i=\psi_h n+1}^n (2\lambda_i - 1)a_h}{n^2 \mu_h} \quad (5)$$

Since $\sum_{i=1}^{\psi_h n} (2\lambda_i - 1) = (\psi_h n)^2$ and $\sum_{i=\psi_h n+1}^n (2\lambda_i - 1) = n^2 - (\psi_h n)^2$, and making use of (3), we can simplify (5) as:

$$C(h) = \frac{(\mu_h - a_h)}{\mu_h} - \frac{(\mu_h - a_h)^2}{\mu_h(b_h - a_h)} \quad (6)$$

¹In both cases I assume for simplicity that $\psi_h n$ is a whole number.

After a rearrangement of terms this can be reduced to:

$$C(h) = \frac{(b_h - \mu_h)(\mu_h - a_h)}{\mu_h(b_h - a_h)} \quad (7)$$

The lower bound can be determined by an analogous procedure. ■

As an aside, I mention that this result can be extended to the Gini coefficient.² In the present case we have to distinguish two cases. First I assume that the lower bound of the health variable a_h is zero, and second that it is positive.

When $a_h = 0$, the bounds of the health Concentration Index can be rewritten as:

$$-\frac{(b_h - \mu_h)}{b_h} \leq C(h) \leq +\frac{(b_h - \mu_h)}{b_h} \quad (8)$$

For samples with a low average, i.e. such that $\mu_h \rightarrow a_h = 0$, the bounds of the health Concentration Index tend to -1 and $+1$. It is easy to check that if the best well-off person has a positive health level h_1 while all the others have the minimum health level $a_h = 0$, then we have $\mu_h = h_1/n$ and $C(h) = (n - 1)/n$. If the average increases, however, the bounds between which the health Concentration varies become narrower and narrower. In fact they decrease linearly, and if $\mu_h \rightarrow b_h$ both the upper and lower bounds tend to zero. This is the phenomenon described by Wagstaff (2005), in the case of a binary variable with $a_h = 0$ and $b_h = 1$.

When $a_h > 0$, things look a bit different. For samples with a low average, i.e. such that $\mu_h \rightarrow a_h$, the proposition shows that the bounds of the health Concentration Index tend to zero. If the average increases, the bounds widen and reach a maximum when $\mu_h = \sqrt{a_h b_h}$. At that point the bounds are $-\frac{(\sqrt{b_h} - \sqrt{a_h})^2}{(b_h - a_h)}$ and $+\frac{(\sqrt{b_h} - \sqrt{a_h})^2}{(b_h - a_h)}$. If the average further increases, the

²Let the Gini coefficient of health inequality, $G(h)$, be defined as:

$$G(h) \equiv 1 - \frac{\sum_{i=1}^n (2\rho_i - 1)h_i}{n^2\mu_h}$$

where ρ_i stands for the rank of individual i in the distribution of health. The most unequal distribution of health which is compatible with mean health μ_h consists of $\psi_h n$ persons having maximum health and $(1 - \psi_h)n$ persons having minimum health. It follows that the bounds of the Gini coefficient are:

$$0 \leq G(h) \leq \frac{(b_h - \mu_h)(\mu_h - a_h)}{\mu_h(b_h - a_h)}$$

This result has already been established by Gastwirth (1972: 308, Lemma 4).

bounds become narrower and if $\mu_h \rightarrow b_h$ they tend to zero again.³

Knowing that the bounds of the health Concentration Index in general depend upon the mean of the variable under examination, we should ask ourselves whether the unmodified Concentration Index provides a good measure of the degree of inequality if populations with different means are compared. Perhaps we should inflate the Concentration Index when the bounds are narrow, or deflate it when the bounds are wide. Whether or not a modification is required, depends upon whether or not the unmodified health Concentration Index is “distorted”. One cause of distortion concerns the way we look at health: we can indeed decide to concentrate either on attainments or on shortfalls. In the next section I deal with this issue.

4 Health and Ill Health

In empirical work one often looks at the distribution of a negative characteristic of health, such as sickness or malnutrition, rather than at positive aspects. To each health variable with a finite upper bound we can easily associate an ill health variable by calculating the shortfall with regard to the maximum. So, starting from the health variable h_i let us define the ill health variable s_i by the following transformation:

$$s_i \equiv b_h - h_i \tag{9}$$

For the population as a whole we have $s = (s_1, s_2, \dots, s_n)$. The ill health variable takes nonnegative values only, just like the original health variable. It has well-defined lower and upper limits: $a_s = 0$ and $b_s = (b_h - a_h)$. Observe moreover that the averages of the two variables are related to one another by the formula $\mu_s = b_h - \mu_h$. We can define the ill health Concentration Index $C(s)$ by analogy with the health Concentration Index $C(h)$:

$$C(s) \equiv 1 - \frac{\sum_{i=1}^n (2\lambda_i - 1)s_i}{n^2\mu_s} \tag{10}$$

Since the ill health variable is derived from the health variable, the two Concentration Indices must clearly be related to each other. It can be checked

³For completeness, we can also see what happens if the health variable has no finite upper bound. By calculating the limit of the bounds for $b_h \rightarrow +\infty$, we find that they are equal to $-(\mu_h - a_h)/\mu_h$ and $+(\mu_h - a_h)/\mu_h$. This means that in the case of a health variable with a positive lower value but an unbounded upper value, the bounds of the health Concentration Index widen hyperbolically as the mean of the health variable increases. If the health variable’s lower bound is equal to zero, the bounds of the Concentration Index are always $-(n - 1)/n$ and $+(n - 1)/n$.

easily that:

$$C(s) = -\frac{\mu_h}{\mu_s}C(h) \quad (11)$$

Hence, as could be expected, the two Concentration Indices always have opposite signs; a positive health Concentration Index $C(h)$, indicating a pro-rich distribution of health, translates into a negative ill health Concentration Index $C(s)$, indicating a pro-poor distribution of ill health. Given that $a_s = 0$ we can also infer the bounds of the ill health Concentration Index:

$$-\frac{(b_s - \mu_s)}{b_s} \leq C(s) \leq +\frac{(b_s - \mu_s)}{b_s} \quad (12)$$

or equivalently:

$$-\frac{(\mu_h - a_h)}{(b_h - a_h)} \leq C(s) \leq +\frac{(\mu_h - a_h)}{(b_h - a_h)} \quad (13)$$

Since health and ill health are mirrors of one another (e.g. either you are immunized against a certain disease or you are not), it seems reasonable to expect that the two Concentration Indices give mirror images of the degree of inequality. Suppose that we compare two distributions, h and h' (e.g. two countries at one moment of time, or one country at different moments of time). As long as the comparison is made between distributions with the same mean, the two indices yield the expected result: the indices move in the opposite direction (if $C(h) > C(h')$, then $C(s) < C(s')$), but in relative terms the change measured by the two indices is exactly the same ($C(h)/C(h') = C(s)/C(s')$). If, however, the distributions have different means, the comparison becomes blurred. The indices need not move into opposite directions, and certainly their ratios will be different.⁴ As shown by Clarke e.a. (2002), when comparing inequality between countries this can lead to different rankings according to whether you look at health or ill health.

So perhaps we should transform or normalize the indices, so as to eliminate the disturbing effect of the mean upon the comparison. The above discussion suggests at least one condition to be imposed upon the normalization procedure: after normalization, we want the health and ill health Concentration Indices to provide mirror images of one another. This is what I will call the “mirror” condition.

⁴See Sen (1992: 89-93) for a general discussion of the issues involved in measuring inequality in terms of attainments or in terms of shortfalls.

5 Cardinality

Another issue concerns the nature of the health variable. Depending upon the quality of the information on individuals' health, the health variable can be ordinal, cardinal or of the ratio-scale variety. When the information is good enough only to rank individuals according to their health status, health is an ordinal variable. If we can also compare the differences between health states, the variable is cardinal. And if on top of that the situation of zero health is fixed unambiguously, we have a ratio-scale variable. In the case of a purely ordinal health variable, the indicator used should be such that it yields the same result if the variable is subjected to a positive monotone transformation. When health is measured cardinally, the indicator should remain constant when the variable is subjected to a positive linear transformation. And for a ratio-scale variable, the indicator must be invariant under any positive proportional transformation of the variable.

The value of the health Concentration Index remains unchanged only under a positive proportional transformation (see, for instance, Erreygers, 2006). Hence it requires that health is measured on a ratio-scale, like length and income. This can certainly be done for some aspects of health, but often it is simply impossible. Self-assessed health, for example, is typically expressed in qualitative categories such as "Very good", "Fair" and "Poor", which are hard to translate into a ratio-scale. It would already be a great step forward if we could adjust the Concentration Index so as to make it invariant to any positive linear transformation of the health variable. In that case it can be used in situations where health is measured on a cardinal scale, like temperature.

From now on I assume that the health variable is cardinal. I would like to know whether it is possible to modify the Concentration Index so that its value does not change when the health variable is subjected to a positive linear transformation. In formal terms, we have to look at the effect of transforming the health vector h into \tilde{h} , where $\tilde{h}_i \equiv \alpha + \beta h_i$, α and β are constants, and $\beta > 0$. This transformation evidently affects the value of the mean, since $\mu_{\tilde{h}} = \alpha + \beta\mu_h$, but it also affects the lower and upper limits of the health variable, which after the transformation are equal to $a_{\tilde{h}} = \alpha + \beta a_h$ and $b_{\tilde{h}} = \alpha + \beta b_h$. For the unmodified Concentration Index we have:

$$C(\tilde{h}) = \frac{\beta\mu_h}{(\alpha + \beta\mu_h)}C(h) \quad (14)$$

which clarifies that its value changes whenever $\alpha \neq 0$.

6 The Wagstaff Normalization

Wagstaff (2005) has suggested a normalization formula aimed at remedying the bounds issue. For a binary health variable with $a_h = 0$ and $b_h = 1$, he proposes to divide the health Concentration Index by its (upper) bound, which leads to a normalized health Concentration Index equal to $C(h)/(1 - \mu_h)$. Generalizing this procedure to all possible cases with lower bound a_h and finite upper bound b_h , we can define the Wagstaff-normalized health Concentration Index $W(h)$ as follows:

$$W(h) \equiv \frac{\mu_h(b_h - a_h)}{(b_h - \mu_h)(\mu_h - a_h)} C(h) \quad (15)$$

The minimum and maximum values of this index are always equal to -1 and $+1$, and so one might say that the Wagstaff procedure tackles the bounds issue by stretching the index in such a way that it always has a uniform range. I will examine in a moment whether this constitutes an appropriate answer, but first I would like to show that as a kind of positive externality the procedure also addresses the ill health and cardinality issues.

First, let us apply the Wagstaff procedure to the associated ill health variable with lower bound $a_s = 0$ and upper bound $b_s = b_h - a_h$. The Wagstaff-normalized ill health Concentration Index $W(s)$ is equal to:

$$W(s) \equiv \frac{\mu_s(b_s - a_s)}{(b_s - \mu_s)(\mu_s - a_s)} C(s) = \frac{(b_h - a_h)}{(\mu_h - a_h)} C(s) \quad (16)$$

Comparing the definitions of the normalized indices and taking into account the relation between $C(h)$ and $C(s)$, we see that the Wagstaff normalization procedure satisfies the mirror condition, since:

$$\frac{W(h)}{W(s)} = \frac{\mu_h}{(b_h - \mu_h)} \cdot \frac{C(h)}{C(s)} = -1 \quad (17)$$

Second, it is easy to show that the Wagstaff index is invariant to a positive linear transformation of the health variable. Using the notation of the previous section, we have:

$$W(\tilde{h}) = \frac{\mu_{\tilde{h}}(b_{\tilde{h}} - a_{\tilde{h}})}{(b_{\tilde{h}} - \mu_{\tilde{h}})(\mu_{\tilde{h}} - a_{\tilde{h}})} C(\tilde{h}) \quad (18)$$

Since $\frac{\mu_{\tilde{h}}(b_{\tilde{h}} - a_{\tilde{h}})}{(b_{\tilde{h}} - \mu_{\tilde{h}})(\mu_{\tilde{h}} - a_{\tilde{h}})} = \frac{(\alpha + \beta\mu_h)(b_h - a_h)}{\beta(b_h - \mu_h)(\mu_h - a_h)}$, and making use of (14), we easily derive that $W(\tilde{h}) = W(h)$. This means that for the Wagstaff index the health variable need not be a ratio-scale variable.

It seems, therefore, that the Wagstaff procedure solves three problems at one go. I have to point out, however, that by forcing the bounds of the index always to the same values, Wagstaff's normalization formula in some cases produces artificial or counterintuitive results. There may be changes in the distribution of health to which the Wagstaff index remains insensitive, or to which it reacts in a counterintuitive way. Consider the following bipolar distribution of health over a population of 100 persons: the 10 richest persons have maximum health and all the rest have minimal health. Suppose now that there is a change in health leading to a distribution in which the 20 richest persons have maximal health and the 80 poorest persons have minimal health. Clearly the situation has become more pro-rich: 10 rich persons are now in better health, while everything else remains the same. One would expect such a move in favour of the rich to be translated into an increase of the inequality indicator. But the Wagstaff index does not have this property: for both situations it is exactly equal to its maximum value of +1. Moreover, if the change in health were such that the 10 rich persons whose health improves all attain health levels below the maximum, then the Wagstaff index would *decrease*. In this case, an obvious pro-rich move is translated into a pro-poor change of the index.

7 The Generalized Concentration Index

The Wagstaff procedure apparently somehow exaggerates the correction it applies to the index and to its bounds. An alternative normalization procedure, introduced in order to deal with the health / ill health issue, has been used by Clarke e.a. (2002), although it has been formulated originally by Wagstaff, Paci and Van Doorslaer (1991). This is the so-called Generalized health Concentration Index $V(h)$, which is obtained by multiplying the health Concentration Index by the average health level:

$$V(h) \equiv \mu_h C(h) \tag{19}$$

Applying the same reasoning to the ill health variable, we obtain likewise:

$$V(s) \equiv \mu_s C(s) \tag{20}$$

Since $C(s) = -(\mu_h/\mu_s)C(h)$, it is immediately obvious that $V(h) = -V(s)$, and so the mirror condition is satisfied. It is also possible to show that this index always translates a pro-rich change of the health distribution into an increase, and a pro-poor change into a decrease. Hence, this index succeeds where the previous one failed.

It seems, however, that a price has to be paid for this. As a matter of fact, a positive linear transformation of the health variable can change the value of the Generalized Concentration Index. As observed by Clarke e.a. (2002: 1925), adding a constant α to the value of the health variable does not matter, but multiplying it by a positive scalar β different from 1 does, because:

$$V(\tilde{h}) = \beta V(h) \quad (21)$$

In other words, the Generalized Concentration Index $V(h)$ is sensitive to the scale or unit of the health variable. An equivalent way of saying is that $V(h)$ is an absolute measure of inequality, not a relative one.

8 An Axiomatic Approach

So far we have encountered three closely related measures of socioeconomic inequality: the Concentration Index $C(h)$, the Wagstaff index $W(h)$, and the Generalized Concentration Index $V(h)$. These can all be seen as members of a family of rank-dependent socioeconomic indicators. The general form of the indicators of this family is the following:

$$I(h) = f(H) \sum_{i=1}^n z_i h_i \quad (22)$$

where:

$$z_i = \frac{n+1}{2} - \lambda_i \quad (23)$$

The general form consists of two parts. The part which is common to all indices, $\sum_{i=1}^n z_i h_i$, expresses the rank-dependence character.⁵ It is a weighted sum of all individual health levels. The health h_i of individual i is weighted by a factor determined by the socioeconomic rank λ_i of this individual. The sign of the weight separates the “rich” from the “poor”: the first half of the population, composed of those with ranks smaller than $(n+1)/2$, have positive weights and constitute the group of the rich; the second half of the population, composed of those with ranks higher than $(n+1)/2$, have

⁵It turns out that there are at least three equivalent ways to express this sum. Observe that $\sum_{i=1}^n z_i h_i = \sum_{i=1}^n (\mu_\lambda - \lambda_i) h_i$, where $\mu_\lambda = (n+1)/2$. We have:

$$\sum_{i=1}^n (\mu_\lambda - \lambda_i) h_i = \sum_{i=1}^n (\mu_\lambda - \lambda_i) (h_i - \mu_h) = \sum_{i=1}^n \lambda_i (h_i - \mu_h)$$

negative weights and constitute the group of the poor. The absolute value of the weight increases linearly with the distance of the rank from the middle. This means, for example, that the richest and the poorest individual have exactly the same weight, albeit of the opposite sign.

The second part is a function $f(H)$ which determines the specific form of the indicator. I assume that the arguments of this function are health-related variables, captured by the vector H . More specifically I take $H = (a_h, b_h, \mu_h)$. (By analogy, in the case of ill health we work with a function $f(S)$ with arguments $S = (a_s, b_s, \mu_s)$.) Apart from requiring that $f(H) > 0$, I put no other constraints upon this function. For the three indicators I have discussed above, the specific form of the function can be deduced from the following expressions:

$$C(h) = \frac{2}{n^2 \mu_h} \sum_{i=1}^n z_i h_i \quad (24)$$

$$W(h) = \frac{2(b_h - a_h)}{n^2(b_h - \mu_h)(\mu_h - a_h)} \sum_{i=1}^n z_i h_i \quad (25)$$

$$V(h) = \frac{2}{n^2} \sum_{i=1}^n z_i h_i \quad (26)$$

The following proposition specifies some properties which are valid for the family of rank-dependent socioeconomic indicators as a whole.

Proposition 2 *All rank-dependent socioeconomic indicators of health have the same sign. For a distribution in which all persons have the same level of health, they are all equal to zero.*

Proof. Given that $f(H) > 0$, the sign of $I(h)$ is always equal to the sign of $\sum_{i=1}^n z_i h_i$. In particular we have $I(h) = 0 \Leftrightarrow \sum_{i=1}^n z_i h_i = 0$. When

$h_1 = h_2 = \dots = h_n$, we have $\sum_{i=1}^n z_i h_i = 0$, and so in this case the value of any rank-dependent socioeconomic indicator is equal to zero. ■

Two remarks are in order here. First, the value of the indices can be zero even if health is not distributed equally. In other words, we can have $\sum_{i=1}^n z_i h_i = 0$ in cases where health is distributed unequally. Second, although any two rank-dependent indices always have the same sign and are therefore always unanimous in their verdict on whether a distribution is pro-rich or pro-poor, this does not mean that they rank health distributions in the same order of socioeconomic inequality. According to one index distribution h may

be more pro-rich than distribution h' , but according to another the reverse might be true.

If we want to say more, we must go beyond the mere specification that $f(H) > 0$. It seems reasonable to let the choice for a specific form of $f(H)$ be guided by the properties which one wants the indicator to possess. Before I introduce these properties, let me draw attention to a subtle point of distinction. The statement “ $I(h)$ increases in value” is not identical to the statement “the socioeconomic inequality of health increases”. Whether a positive change in the index corresponds to an increase in inequality depends upon the starting value of the index. If one starts from a pro-rich distribution of health ($I(h) > 0$), a positive change increases the inequality, but if one starts from a pro-poor distribution ($I(h) < 0$), such a change decreases the inequality. Put differently, I speak of an increase in inequality when the index increases in absolute value (without changing sign), and of a decrease in inequality when the index decreases in absolute value (without changing sign).

The conditions which I impose upon the indicator are suggested by the three issues which I have discussed above. The first two relate to the bounds issue, and deal with the effects of changing the health position of two persons and of changing the health position of all individuals. The third condition concerns the cardinality issue and deals with the effects of changing the unit of measurement of health, and the fourth is about the effect of measuring ill health instead of health.

The first property refers to the case where a change in the health situation of one person is exactly offset by a change in the health situation of another person:

Transfer: Suppose the health of person j increases by a certain amount and the health of person k decreases by the same amount. If person j is richer than person k , the value of $I(h)$ increases; if person j is poorer than person k , the value of $I(h)$ decreases.

The transfer property states that a mean-preserving change of the distribution which favours the better-off is translated into a pro-rich change of the index, and one in favour of the worse-off into a pro-poor change. It is of course inspired by the well-known Pigou-Dalton principle of transfers with regard to the measurement of income inequality. Variants of this principle have been used in connection with the health Concentration Index, especially in the context of the social welfare interpretation of the index (e.g. Bleichrodt and Van Doorslaer, 2006).

The second property relates to the case where the health situation of all

individuals is changed in the same way:

Level independence: If the health levels of all persons change by the same absolute amount, the value of $I(h)$ remains constant.

This condition is similar, but not identical, to the translation invariance assumption used in the literature on poverty measurement (e.g. Zheng, 1994: 1453). In the present framework, translation invariance would mean that the index remains constant if all health levels increase by the same amount *and* if both the lower and the upper limit of the health variable shift by the same amount. When dealing with level independence, however, we do not change the lower and upper limits of the health variable. The motivation for this condition comes from the observation that if we start from a distribution in which all persons have the same level of health (which implies that the index is zero), and then increase all health levels by the same amount, then the index remains unchanged. The level independence condition extends this property to a broader set of distributions.

The third property deals with changes of the unit of measurement of health:

Cardinal consistency: A positive linear transformation of the health variable does not affect the value of $I(h)$.

In other words, I impose both scale and translation invariance. The final property concerns the health / ill health issue:

Mirror: The health index $I(h)$ and the associated ill health index $I(s)$ are equal in absolute value, but have opposite signs.

This follows directly from our discussion of the measurement of the inequality of ill health.

To check whether a given member of our family of rank-dependent measures satisfies any of these properties, it will be helpful to derive what conditions they impose upon the function $f(H)$.

Proposition 3 *For any rank-dependent socioeconomic inequality indicator of the form $f(H) \sum_{i=1}^n z_i h_i$, we have: (i) the transfer property holds if and only if $\frac{\partial f(H)}{\partial h_j} = \frac{\partial f(H)}{\partial h_k}, \forall j, k$; (ii) the level independence property holds if and only if $\sum_{j=1}^n \frac{\partial f(H)}{\partial h_j} = 0$; (iii) the cardinal consistency property holds if and*

only if $f(H) = \beta f(\tilde{H})$, where $\tilde{h}_i = \alpha + \beta h_i$; (iv) the mirror property holds if and only if $f(H) = f(S)$, where $s_i = b_h - h_i$.

Proof. (i) A small change in the health of person j leads to the following change in the health index:

$$\frac{\partial I(h)}{\partial h_j} = \frac{\partial f(H)}{\partial h_j} \sum_{i=1}^n z_i h_i + f(H) z_j \quad (27)$$

For person k we have likewise:

$$\frac{\partial I(h)}{\partial h_k} = \frac{\partial f(H)}{\partial h_k} \sum_{i=1}^n z_i h_i + f(H) z_k \quad (28)$$

Changing the health of person j by the amount Δh_j and that of person k by the amount Δh_k , has the following effect upon the health index:

$$\Delta I(h) = \frac{\partial I(h)}{\partial h_j} \Delta h_j + \frac{\partial I(h)}{\partial h_k} \Delta h_k \quad (29)$$

Let $\Delta h_j = -\Delta h_k = \Delta > 0$. Rearranging terms and using the derivations of $\frac{\partial I(h)}{\partial h_j}$ and $\frac{\partial I(h)}{\partial h_k}$ we obtain:

$$\Delta I(h) = \left\{ \left[\frac{\partial f(H)}{\partial h_j} - \frac{\partial f(H)}{\partial h_k} \right] \sum_{i=1}^n z_i h_i + f(H) [z_j - z_k] \right\} \Delta \quad (30)$$

If j is richer than k , we have $z_j - z_k > 0$; and if j is poorer than k , we have $z_j - z_k < 0$. This means that the transfer property holds if and only if the sign of $\Delta I(h)$ is always equal to the sign of $z_j - z_k$. Since the sign of $\sum_{i=1}^n z_i h_i$ can be positive, negative or zero, it follows that the transfer property holds if and only if $\frac{\partial f(H)}{\partial h_j} = \frac{\partial f(H)}{\partial h_k}$.

(ii) If all the health levels change by the same amount Δ , the effect upon the health index is:

$$\Delta I(h) = \sum_{j=1}^n \frac{\partial I(h)}{\partial h_j} \Delta = \sum_{j=1}^n \left[\frac{\partial f(H)}{\partial h_j} \sum_{i=1}^n z_i h_i + f(H) z_j \right] \Delta \quad (31)$$

This can be reduced to:

$$\Delta I(h) = \sum_{j=1}^n \frac{\partial f(H)}{\partial h_j} \sum_{i=1}^n z_i h_i \Delta \quad (32)$$

Since the sign of $\sum_{i=1}^n z_i h_i$ can be positive, negative or zero, it follows that level independence holds if and only if $\sum_{j=1}^n \frac{\partial f(H)}{\partial h_j} = 0$.

(iii) By definition we have $I(h) = f(H) \sum_{i=1}^n z_i h_i$ and $I(\tilde{h}) = f(\tilde{H}) \sum_{i=1}^n z_i \tilde{h}_i$.

Since $\tilde{h}_i = \alpha + \beta h_i$, it follows that $I(\tilde{h}) = f(\tilde{H}) \sum_{i=1}^n z_i (\alpha + \beta h_i) = \beta f(\tilde{H}) \sum_{i=1}^n z_i h_i$.

Therefore, $I(h) = I(\tilde{h})$ if and only if $f(H) = \beta f(\tilde{H})$.

(iv) By definition we have $I(h) = f(H) \sum_{i=1}^n z_i h_i$ and $I(s) = f(S) \sum_{i=1}^n z_i s_i$.

Since $s_i = b_h - h_i$, it follows that $I(s) = f(S) \sum_{i=1}^n z_i (b_h - h_i) = -f(S) \sum_{i=1}^n z_i h_i$.

Therefore, $I(h) = -I(s)$ if and only if $f(H) = f(S)$. ■

Before I examine whether any of the rank-dependent socioeconomic indicators possesses these four properties, I introduce a fifth property:

Monotonicity: If the health of a “rich” person (i.e. someone for whom $z_j > 0$) increases, the value of $I(h)$ increases; if the health of a “poor” person (i.e. someone for whom $z_j < 0$) increases, the value of $I(h)$ decreases.

It is possible to demonstrate that this property is closely related to the first and the second:

Proposition 4 *For rank-dependent socioeconomic indicators of health, transfer and level independence together imply monotonicity, and vice versa.*

Proof. A small change in the health of person j leads to the following change in the health index:

$$\frac{\partial I(h)}{\partial h_j} = \frac{\partial f(H)}{\partial h_j} \sum_{i=1}^n z_i h_i + f(H) z_j \quad (33)$$

If j is rich, we have $f(H) z_j > 0$; if j is poor, we have $f(H) z_j < 0$. This means that the monotonicity property holds if and only if the sign of $\frac{\partial I(h)}{\partial h_j}$

always coincides with the sign of $f(H) z_j$. Since the sign of $\sum_{i=1}^n z_i h_i$ can be positive, negative or zero, it follows that the monotonicity property holds if and only if $\frac{\partial f(H)}{\partial h_j} = 0$ for all j .

The transfer property implies that $\frac{\partial f(H)}{\partial h_j} = \frac{\partial f(H)}{\partial h_k}$ for all j and k , and the level independence property that $\sum_{j=1}^n \frac{\partial f(H)}{\partial h_j} = 0$. Hence, if both properties hold we must have $\frac{\partial f(H)}{\partial h_j} = 0$ for all j , which implies that the monotonicity property holds. Inversely, if monotonicity holds, both transfer and level independence are implied. ■

I alluded to the property of monotonicity when I pointed out that the Wagstaff normalization procedure exaggerates the bounds correction. The property states that a pro-rich change in the distribution must be translated into a pro-rich change of the index (i.e. $\Delta I(h) > 0$), and a pro-poor change in the distribution into a pro-poor change of the index (i.e. $\Delta I(h) < 0$). I consider the monotonicity, cardinal consistency and mirror conditions to be the formal expressions of the kind of modifications suggested by the discussion of the bounds, cardinality and health / ill health issues.

9 A New Indicator

It is not very difficult to check that none of the three indicators which I have examined above satisfies all of these properties. The Concentration Index $C(h)$ satisfies only the transfer property⁶. The Wagstaff index $W(h)$ satisfies the transfer, cardinal consistency and mirror properties, but not that of level independence. The Generalized Concentration Index $V(h)$, finally, satisfies the transfer, mirror and level independence properties, but not that of cardinal consistency.

Is it possible to find a rank-dependent socioeconomic inequality index which satisfies *all* of these properties? Consider the following indicator:

$$E(h) = \frac{8}{n^2(b_h - a_h)} \sum_{i=1}^n z_i h_i \quad (34)$$

I now demonstrate both an existence and a uniqueness result:

Proposition 5 *Indicator $E(h)$ satisfies the properties of transfer, level independence, cardinal consistency and mirror.*

⁶This property provides the basis for the claim that the index reaches its maximum value when the $\psi_h n$ richest persons have maximum health and the $(1 - \psi_h)n$ poorest persons have minimum health (cf. Proposition 1).

Proof. This index is characterized by the function $f(H) = \frac{8}{n^2(b_h - a_h)}$, which implies $\frac{\partial f(H)}{\partial h_j} = \frac{\partial f(H)}{\partial h_k} = 0$, $f(S) = \frac{8}{n^2(b_s - a_s)} = \frac{8}{n^2(b_h - a_h)} = f(H)$, and $f(\tilde{H}) = \frac{8}{n^2[(\alpha + \beta b_h) - (\alpha + \beta a_h)]} = \frac{f(H)}{\beta}$. Hence the four properties are satisfied. ■

Proposition 6 *All rank-dependent socio-economic indicators which satisfy the properties of transfer, level independence and cardinal consistency are positive scalar multiples of indicator $E(h)$, and such that the factor of proportionality is independent of the unit of measurement of health.*

Proof. Suppose there exists another rank-dependent socioeconomic indicator $K(h)$ which satisfies the properties of transfer and level independence. Since both $E(h)$ and $K(h)$ are rank-dependent indicators we can write $E(h) = f_E(H) \sum_{i=1}^n z_i h_i$ and $K(h) = f_K(H) \sum_{i=1}^n z_i h_i$, and so we have:

$$K(h) = \frac{f_K(H)}{f_E(H)} E(h) \quad (35)$$

Now suppose there is a change Δh_j in the health position of individual j , which leads to the new health distribution h' . The values of the indicators change to $E(h')$ and $K(h')$, which are such that

$$E(h') = E(h) + \frac{\partial E(h)}{\partial h_j} \Delta h_j \quad (36)$$

$$K(h') = K(h) + \frac{\partial K(h)}{\partial h_j} \Delta h_j \quad (37)$$

Since both indicators satisfy the transfer and level independence properties we have:

$$\frac{\partial E(h)}{\partial h_j} = f_E(H) z_j \quad (38)$$

$$\frac{\partial K(h)}{\partial h_j} = f_K(H) z_j \quad (39)$$

Therefore we obtain:

$$E(h') = E(h) + f_E(H) z_j \Delta h_j \quad (40)$$

$$K(h') = K(h) + f_K(H) z_j \Delta h_j \quad (41)$$

Making use of (35) the last equation can be transformed into:

$$K(h') = \frac{f_K(H)}{f_E(H)} E(h) + f_K(H) z_j \Delta h_j \quad (42)$$

which leads to:

$$K(h') = \frac{f_K(H)}{f_E(H)} E(h') \quad (43)$$

Therefore, for the given unit of measurement of health, the values of $K(h)$ and $E(h)$ always stand in the same positive proportion to each other.

Now suppose that we change the unit of measurement and transform h into \tilde{h} , where $\tilde{h}_i = \alpha + \beta h_i$. Since both indicators satisfy the property of cardinal consistency, we have $f_E(\tilde{H}) = \frac{f_E(H)}{\beta}$ and $f_K(\tilde{H}) = \frac{f_K(H)}{\beta}$, and therefore:

$$K(\tilde{h}) = \frac{f_K(\tilde{H})}{f_E(\tilde{H})} E(\tilde{h}) = \frac{f_K(H)}{f_E(H)} E(h) = K(h) \quad (44)$$

Hence, whatever the unit of measurement of health, the values of the two indicators always stand in the same proportion to one another. ■

The upshot is that, up to a scalar multiple, $E(h)$ is the only rank-dependent socioeconomic indicator which satisfies the four properties. The reason why I have chosen $E(h)$ is that its maximum bounds are equal to -1 and $+1$. It can be shown quite easily that:

$$-\frac{4(b_h - \mu_h)(\mu_h - a_h)}{(b_h - a_h)^2} \leq E(h) \leq +\frac{4(b_h - \mu_h)(\mu_h - a_h)}{(b_h - a_h)^2} \quad (45)$$

The bounds tend to zero if $\mu_h \rightarrow a_h$. They increase if the mean rises, and reach a maximum when μ_h is exactly halfway between a_h and b_h , at which point the bounds are equal to -1 and $+1$ (when n is even), and to $-(n^2 - 1)/n^2$ and $+(n^2 - 1)/n^2$ (when n is odd). Then they decrease, and tend to zero again if $\mu_h \rightarrow b_h$.

10 A Decomposition Formula

The popularity of the Concentration Index can to some extent be explained by the easiness with which it can be decomposed into contributing factors. For instance, Wagstaff, Van Doorslaer and Watanabe (2003) have shown that if we start from an explanation of health expressed by the following linear regression equation:

$$h_i = \gamma + \sum_{j=1}^q \lambda_j x_{ji} + \varepsilon_i \quad (46)$$

where x_1, x_2, \dots, x_q stand for the q explanatory variables, ε represents the error term, and γ and $\lambda_1, \lambda_2, \dots, \lambda_q$ are the coefficients, then the health Concentration Index can be written as:

$$C(h) = \sum_{j=1}^q \lambda_j \frac{\mu_j}{\mu_h} C(x_j) + \frac{1}{\mu_h} V(\varepsilon) \quad (47)$$

where μ_j is the average of x_j , $C(x_j)$ the Concentration Index of x_j , and $V(\varepsilon)$ the Generalized Concentration Index of ε . This formula can then be used to explain the changes in the Concentration Index.

It should be noted, however, that if health is a cardinal but not a ratio-scale variable, the coefficients of (46) depend upon the particular unit chosen for the measurement of health, and this may affect the decomposition results. Assuming that the explanatory variables are all ratio-scale variables, we can obtain an analogous decomposition formula for the corrected Concentration Index $E(h)$. Instead of starting from (46), let us start from the regression:

$$\frac{h_i - a_h}{b_h - a_h} = \gamma^* + \sum_{j=1}^q \lambda_j^* x_{ji} + \varepsilon_i^* \quad (48)$$

Since the coefficients and the error term of this regression are independent of the particular unit chosen for the health variable, I have added asterisks to them. I now show:

Proposition 7 *Given the linear regression equation (48), the corrected Concentration Index $E(h)$ can be expressed as:*

$$E(h) = 4 \left[\sum_{j=1}^q \lambda_j^* V(x_j) + V(\varepsilon^*) \right] \quad (49)$$

where $V(x_j)$ is the Generalized Concentration Index of x_j and $V(\varepsilon^*)$ that of ε^* .

Proof. From (48) I deduce:

$$h_i = (b_h - a_h) \left[\gamma^* + \sum_{j=1}^q \lambda_j^* x_{ji} + \varepsilon_i^* \right] + a_h \quad (50)$$

In the definition of $E(h)$ I then replace h_i by the right-hand side of (50).

Since $\sum_{i=1}^n z_i a_h = \sum_{i=1}^n z_i \gamma^* = 0$, I can simplify the resulting expression to:

$$E(h) = \frac{8}{n^2} \left[\sum_{j=1}^q \lambda_j^* \sum_{i=1}^n z_i x_{ji} + \sum_{i=1}^n z_i \varepsilon_i^* \right] \quad (51)$$

Applying the definition of the Generalized Concentration Index to x_j and ε^* , I obtain:

$$E(h) = 4 \left[\sum_{j=1}^q \lambda_j^* V(x_j) + V(\varepsilon^*) \right] \quad (52)$$

■

11 A Generalization

It is not too difficult to push the analysis one step further. What characterizes the rank-dependent measures is that the weights which are given to the individual health levels are determined by the socioeconomic ranks of these individuals. But instead of using the weights $z_i = \frac{n+1}{2} - \lambda_i$ other weights might be used. One might think of a generalization along the lines of the “extended” Concentration Index proposed by Wagstaff (2002), which makes it possible to incorporate varying degrees of aversion to inequality by an appropriate choice of the value of an aversion parameter.⁷ Here I suggest a more general framework. I start from the following generalization of (34):

$$F(h) = \frac{8}{n^2(b_h - a_h)} \sum_{i=1}^n w_i h_i \quad (53)$$

and impose the following restrictions upon the weights w_i :

$$w_j \geq w_k \text{ if } \lambda_j < \lambda_k \quad \text{and} \quad w_j = w_k \text{ if } \lambda_j = \lambda_k \quad (54)$$

$$\sum_{i=1}^n w_i = 0 \quad (55)$$

$$\sum_{i=1}^n |w_i| = \frac{n^2}{4} \quad (56)$$

The first condition ensures that the index translates a pro-rich distribution into a positive value and a pro-poor distribution into a negative one.⁸ The

⁷I take the opportunity to point out that in my opinion the following description of the weights used in the calculation of the health Concentration Index is slightly misleading: “Thus, the poorest person gets their health share weighted by a number close to two. The weights decline in a stepwise fashion, reaching a number close to zero for the richest person.” (Wagstaff, 2002: 630). The weights given to both are, in fact, equal, but of the opposite sign.

⁸The rich are those for which $w_i > 0$, and the poor those for which $w_i < 0$. The number of rich persons need not be the same as the number of poor persons.

second is necessary to guarantee that an equal distribution of health always leads to a zero value of the index. The third condition determines the maximum bounds of the index: it ensures that its value never falls outside the range $[-1, +1]$. The following result holds:

Proposition 8 *Under conditions (54) and (55) the indicator of socioeconomic inequality of health $F(h)$ exhibits the properties of transfer, level independence, cardinal consistency and mirror.*

(The proof is analogous to the proofs of propositions 3 and 5 and is omitted here.)

The conditions put upon the weights w_i are relatively weak, and allow a great variety of concerns to be introduced. They make it possible, for instance, to determine the weight of each individual in function of this person's income (or of whatever variable is used to classify people in the socioeconomic sphere). In this way the index can be enriched and refined.

12 An Example

To illustrate the importance of correcting the Concentration Index, let us consider a sample of countries for which data on stunting have been collected by Van de Poel and Hosseinpoor (2006) (see Table 1). For each country they have estimated the number of children who suffer from stunting (chronic malnutrition). Hence, for each child a binary ill health variable s can be defined which indicates whether the child must be considered as stunted ($s_i = 1$) or not ($s_i = 0$). The corresponding health variable h has values $h_i = 1$ for non-stunting and $h_i = 0$ for stunting. The average of the ill health variable (second column of Table 1) indicates the percentage of children suffering from stunting; the average of the health variable (third column) indicates the percentage of the population that do not suffer from stunting. The fourth column gives the negative of the ill health Concentration Index (as calculated by Van de Poel and Hosseinpoor), and the fifth the health Concentration Index. Both make it clear that stunting occurs more frequently among the poor. The sixth and seventh columns give the Wagstaff normalized Concentration Index $W(h)$ and the corrected Concentration Index $E(h)$.⁹

⁹In this example, where $b_h - a_h = 1$, we have the following relation between index $E(h)$ and the Generalized Concentration Index $V(h)$: $E(h) = 4V(h)$. Since $V(h)$ is simply a quarter of $E(h)$ I do not represent it separately in Table 1.

Table 1: Stunting in developing countries

Country	μ_s	μ_h	- C(s)	C(h)	W(h)	E(h)
Malawi 2000	0,4902	0,5098	0,0756	0,0727	0,1483	0,1482
Madagascar 2003/04	0,4834	0,5166	0,0041	0,0038	0,0079	0,0079
Ethiopia 2000	0,4730	0,5270	0,0390	0,0350	0,0740	0,0738
Mozambique 2003	0,4616	0,5384	0,0760	0,0652	0,1412	0,1403
Zambia 2001/02	0,4606	0,5394	0,0990	0,0845	0,1835	0,1824
Rwanda 2000	0,4237	0,5763	0,0878	0,0646	0,1524	0,1488
Niger 1998	0,4107	0,5893	0,0446	0,0311	0,0757	0,0733
Chad 2004	0,4094	0,5906	0,0537	0,0372	0,0909	0,0879
Uganda 2000/01	0,3860	0,6140	0,0541	0,0340	0,0881	0,0835
Burkina Faso 2003	0,3850	0,6150	0,0977	0,0612	0,1589	0,1505
Nigeria 2003	0,3845	0,6155	0,1612	0,1007	0,2619	0,2479
Mali 2001	0,3763	0,6237	0,0990	0,0597	0,1587	0,1490
Tanzania 2004	0,3705	0,6295	0,1187	0,0699	0,1886	0,1759
Nepal 2001	0,3626	0,6374	0,0935	0,0532	0,1467	0,1356
India 1998/99	0,3579	0,6421	0,0706	0,0394	0,1100	0,1011
Mauritania 2000/01	0,3450	0,6550	0,1108	0,0584	0,1692	0,1529
Comoros 1996	0,3377	0,6623	0,1317	0,0672	0,1989	0,1779
CAR 1994/95	0,3365	0,6635	0,0753	0,0382	0,1135	0,1014
Cameroon 2004	0,3165	0,6835	0,1698	0,0786	0,2484	0,2150
Kenya 2003	0,3056	0,6944	0,1265	0,0557	0,1822	0,1546
Benin 2001	0,3036	0,6964	0,1050	0,0458	0,1508	0,1275
Pakistan 1990/91	0,2997	0,7003	0,1182	0,0506	0,1688	0,1417
Ghana 2003	0,2943	0,7057	0,1743	0,0727	0,2470	0,2052
Cambodia 2000	0,2943	0,7057	0,0887	0,0370	0,1257	0,1044
Uzbekistan 1996	0,2823	0,7177	0,0719	0,0283	0,1002	0,0812
Zimbabwe 1999	0,2645	0,7355	0,1126	0,0405	0,1531	0,1191
Guinea 1999	0,2607	0,7393	0,1067	0,0376	0,1443	0,1113
Bangladesh 2004	0,2542	0,7458	0,1108	0,0378	0,1486	0,1127
Cote d'Ivoire 1998/99	0,2517	0,7483	0,1728	0,0581	0,2309	0,1740
Namibia 2000	0,2264	0,7736	0,1014	0,0297	0,1311	0,0918
Togo 1998	0,2172	0,7828	0,1600	0,0444	0,2044	0,1390
Gabon 2000	0,2065	0,7935	0,2397	0,0624	0,3021	0,1980
Kyrgyz Republic 1997	0,1328	0,8672	0,1627	0,0249	0,1876	0,0864
Egypt 2000	0,1288	0,8712	0,1531	0,0226	0,1757	0,0789
Guatemala 1998/99	0,1218	0,8782	0,2247	0,0312	0,2559	0,1095
Morocco 2003-2004	0,0930	0,9070	0,2475	0,0254	0,2729	0,0921
Haiti 2000	0,0797	0,9203	0,2154	0,0187	0,2341	0,0687
Armenia 2000	0,0766	0,9234	0,1770	0,0147	0,1917	0,0542
Bolivia 2003	0,0680	0,9320	0,2739	0,0200	0,2939	0,0745
Peru 2000	0,0501	0,9499	0,3676	0,0194	0,3870	0,0737
Nicaragua 2001	0,0422	0,9578	0,3304	0,0146	0,3450	0,0558
Paraguay 1990	0,0394	0,9606	0,3216	0,0132	0,3348	0,0507
Kazakhstan 1999	0,0359	0,9641	0,2624	0,0098	0,2722	0,0377
Turkey 1998	0,0358	0,9642	0,3388	0,0126	0,3514	0,0485
Dominican Republic 2002	0,0344	0,9656	0,2679	0,0095	0,2774	0,0369
Brazil 1996	0,0275	0,9725	0,4444	0,0126	0,4570	0,0489
Colombia 2005	0,0215	0,9785	0,2699	0,0059	0,2758	0,0232

Source: Columns 2 and 4: Van de Poel and Hosseinpoor (2006); other columns: own calculations.

One of the issues explored in the literature on stunting is whether there exists a relationship between the prevalence of stunting (as measured by the average μ_s) and the socioeconomic inequality of stunting (as measured by an inequality indicator). There are four ways in which this issue can be tested here, since we have four different inequality measures. Table 2 gives the coefficients of correlation between the average degree of stunting (μ_s) and our four inequality measures:

Table 2: Coefficients of correlation

	$-C(s)$	$C(h)$	$W(h)$	$E(h)$
μ_s	-0,8967	+0,6773	-0,7666	+0,5190

The results are strikingly different. Two measures, $-C(s)$ and $W(h)$, indicate a strong negative relationship, whereas the other two, $C(h)$ and $E(h)$, point to a positive correlation. The negative correlation - a low prevalence of stunting tends to be associated with a high degree of socioeconomic inequality - was already apparent in one of the figures of Wagstaff and Watanabe (2000: 33), where they plotted the non normalized stunting Concentration Index $C(s)$ against the average stunting rate μ_s . The message which I want to convey here is that one should be extremely careful with this kind of comparisons.

13 Concluding Remarks

I started from a rather negative position, in the sense that I drew up a list of a number of shortcomings of the “traditional” health Concentration Index $C(h)$. I also showed that modified versions of it, such as Wagstaff’s index $W(h)$ and the Generalized Concentration Index $V(h)$, manage to remedy some of these shortcomings, but not of all of them. But gradually I adopted a more constructive attitude. Using an axiomatic approach I formulated four key requirements - transfer, level independence, cardinal consistency and mirror - which a corrected Concentration Index should satisfy. Then I introduced a new measure, $E(h)$, and demonstrated not only that it satisfies all of these requirements, but also that it is basically the only one that does so. Moreover, it can be decomposed easily.

I must stress that the applicability of the corrected Concentration Index is limited to the case where the health variable is cardinal and has finite lower and upper bounds. In another paper (Erreygers, 2006) I argue that

if the health variable is ordinal, the Concentration Index approach must be abandoned and replaced by a radically different perspective. What emerges from the present paper is that the Concentration Index approach need not be confined to situations where the health variable is of the ratio-scale type. The corrected Concentration Index proposed here substantially broadens the scope of the Concentration Index approach.

Whether or not the corrected Concentration Index $E(h)$ is truly superior to other variants of the index, depends to a large extent upon whether or not my set of four key requirements is a reasonable one. If we are working with a cardinal variable, it seems impossible to discard the condition of cardinal consistency. In order to deal with the health / ill health issue I see little or no alternatives for the mirror condition. What remains then is the monotonicity condition, which combines the transfer and level independence conditions. Here at least I think there might be room for something else, although I do not see clearly what it could be. Finding a substitute for the monotonicity condition is one possible avenue for further research.

The corrected Concentration Index $E(h)$ can be calculated very easily. Among the members of the family of rank-dependent indicators of socioeconomic inequality its closest cousin appears to be the Generalized Concentration Index $V(h)$. A simple correction, consisting of dividing index $V(h)$ by the range of the health variable $b_h - a_h$ and multiplying the result by 4, suffices to turn this absolute indicator of inequality into a relative one. Surprisingly, however, the index $E(h)$ is both a relative and an absolute indicator of inequality. This means that when measuring the socioeconomic inequality of health, we are not stuck in a dilemma which forces us to choose between either a relative or an absolute measure: our indicator can be both at the same time. By contrast, when it comes to the measurement of income inequality or poverty, the situation appears to be fundamentally different (see, e.g., Zheng, 1994).

References

- [1] BLEICHRODT, Han and Eddy VAN DOORSLAER (2006), “A welfare economics foundation for health inequality measurement”, *Journal of Health Economics*, 25(5): 945-957.
- [2] CLARKE, Philip M., Ulf-G. GERDTHAM, Magnus JOHANNESSON, Kerstin BINGEFORS and Len SMITH (2002), “On the measurement of relative and absolute income-related health inequality”, *Social Science & Medicine*, 55(11): 1923-1928.
- [3] ERREYGERS, Guido (2006), “Beyond the health Concentration Index: an Atkinson alternative for the measurement of the socioeconomic inequality of health”, Paper presented at the conference *Advancing Health Equity*, Helsinki, WIDER-UNU, September.
- [4] GASTWIRTH, Joseph L. (1972), “The estimation of the Lorenz curve and Gini index”, *Review of Economics and Statistics*, 54(3): 30-316.
- [5] KAKWANI, Nanak C. (1980), *Income Inequality and Poverty. Methods of Estimation and Policy Applications*, Oxford, Oxford University Press.
- [6] SEN, Amartya (1992), *Inequality Reexamined*, Oxford, Clarendon Press.
- [7] VAN DE POEL, Ellen and Ahmad Reza HOSSEINPOOR (2006), “Socioeconomic inequality in malnutrition in developing countries”, World Health Organisation, in progress.
- [8] WAGSTAFF, Adam (2002), “Inequality aversion, health inequalities, and health achievement”, *Journal of Health Economics*, 21(4): 627-641.
- [9] WAGSTAFF, Adam (2005) “The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality”, *Health Economics*, 14(4): 429-432.
- [10] WAGSTAFF, Adam and Naoko WATANABE (2000), “Socioeconomic inequalities in child malnutrition in the developing world”, World Bank, Policy Research Working Paper, WPS 2434.
- [11] WAGSTAFF, Adam, Pierella PACI and Eddy VAN DOORSLAER (1991), “On the measurement of inequalities in health”, *Social Science & Medicine*, 33(5): 545-557.

- [12] WAGSTAFF, Adam, Eddy VAN DOORSLAER and Naoko WATANABE (2003), “On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam”, *Journal of Econometrics*, 112(1): 207-223.
- [13] ZHENG, Buhong (1994), “Can a poverty index be both relative and absolute?”, *Econometrica*, 62(6): 1453-1458.
- [14] ZHENG, Buhong (2006), “Measuring health inequality and health opportunity”, Paper presented at the conference *Advancing Health Equity*, Helsinki, WIDER-UNU, September.