Beyond the Health Concentration Index: An Atkinson Alternative for the Measurement of the Socioeconomic Inequality of Health

Guido Erreygers

Department of Economics, University of Antwerp

City Campus, Prinsstraat 13, 2000 Antwerpen, Belgium

guido.erreygers@ua.ac.be

November 10, 2006

Abstract

The Health Concentration Index is a frequently used indicator for the measurement of the socioeconomic inequality of health. This note starts from a discussion of some of the weaknesses of this index. It then presents two possible alternative measures. The first is an adaptation of the Concentration Index. The second and more important of the two is constructed by following an Atkinson approach.

Keywords: Health inequality, Socioeconomic inequality, Concentration Index, Atkinson Index

JEL Classification Number: D63, I10

1

1 Introduction

In his widely praised masterpiece *Poverty. A Study of Town Life* (1901) Benjamin Seebohm Rowntree devoted a whole chapter to the relation of poverty to health. The conclusion he reached was that his tests "point[ed] clearly to the low standard of health amongst those living in poverty" (p. 215). A telling example of such a test had to do with the general physical condition of children from working classes. A qualified investigator was asked to classify the physical condition of children under four headings: "Very good", "Good", "Fair", or "Bad". When the results were expressed in function of the income groups to which these children belonged, the following was obtained:

Table 1: Poverty and health in York (1901)

	Boys			
	Very good	Good	Fair	Bad
Section 1 (poorest)	2.8%	14.6%	31.0%	51.6%
Section 2 (middle)	7.4%	20.1%	53.7%	18.8%
Section 3 (highest)	27.4%	33.8%	27.4%	11.4%
	Girls			
Section 1 (poorest)	2.1%	14.6%	31.0%	52.3%
Section 2 (middle)	7.5%	21.2%	50.4%	20.9%
Section 3 (highest)	27.2%	38.0%	23.1%	11.7%

Source: Seebohm Rowntree (1901: 213-214).

It jumps from this table that the general health of the children was positively and strongly correlated to their income group. This is an early example of what epidemiologists have defined as the "social gradient in health and disease", the widely observed fact that rich people tend to be in better health than poor people.[1]

Not only epidemiologists, but also economists and other social scientists have studied the issue, and there is now a growing body of literature which deals with the measurement of the socioeconomic inequality of health. The main question asked by researchers in this field has been formulated as follows: "To what extent are there inequalities in health that are systematically related to socioeconomic status?"[2] Different indicators have been proposed for the measurement of this type of inequality. Among health economists the health Concentration Index, a measure closely related to the well-known Gini coefficient mainly used for the calculation of income inequality, seems to be perceived as the best available indicator. An impressive number of studies is

---

[1] See, for instance, Marmot (2003), and Daniels, Kennedy and Kawachi (2004).

[2] Wagstaff, Paci and van Doorslaer (1991: 546).

now available suggesting that the health Concentration Index provides useful insights into important aspects of the socioeconomic inequality of health.[3] There is also interesting work on the social welfare theoretic foundations of the health Concentration Index.[4]

In this paper I challenge the usefulness of the health Concentration Index as an indicator of the socioeconomic inequality of health in general. With "health in general" I refer to the overall health status of individuals, for instance such as recorded in surveys of self-reported health, where respondents are asked to describe their health status in terms like "excellent", "fair" or "bad". I argue that in those cases the value of the index must be treated with caution; in fact, the informational content of the indicator is rather poor. Moreover, the information which the indicator does provide can also (and easily) be obtained without having to calculate the health Concentration Index. The indicator might be useful for the assessment of some aspects of health inequality, but the number of these seems limited.

The main aim of this paper is to propose an alternative for the health Concentration Index. In fact, I will propose two alternatives. The first mirrors the procedure by which the health Concentration Index is constructed, but in a way which avoids its main shortcoming The second and more important measure builds upon the approach of Atkinson.

2   Assumptions

We consider a given population of n individuals represented by the set $N = \{1, 2, ..., n\}$. Our goal is to find out whether some kind of systematic relation exists between the socioeconomic and health statuses of these individuals. The socioeconomic status of individual i will be measured by her income y_i, a nonnegative real number. The health status of individual i will be measured by h_i, a nonnegative real number which represents her health situation. A social state is described by the two vectors $y = [y_1, y_2, \ldots, y_n]$ and $h = [h_1, h_2, \ldots, h_n]$. The average income of the population will be denoted as μ_y and the average health as μ_h:

$$\mu_y \equiv \frac{1}{n} \sum_i y_i, \quad \mu_h \equiv \frac{1}{n} \sum_i h_i \tag{1}$$

---

[3] One might for instance consult the list of working papers of the ECuity projects, http://www2.eur.nl/bmg/ecuity/.

[4] See, for instance, Bommier and Stecklov (2002), Fleurbaey (2006), and Bleichrodt and van Doorslaer (2006).

In addition, we denote the rank of individual i in the income distribution by λ_i, with the richest individual ranked first, and the rank of individual i in the health distribution by ρ_i, with the healthiest individual ranked first. In the case of ties, we assign to each member of the tied group the average rank of the group.

3 The Health Concentration Index

A huge variety of measures has been developed for the measurement of income inequality. If we are interested in the distribution of health as such, i.e. not in its relation to income, we could in principle take any measure used in connection with income inequality and apply it to the distribution of health. In this way we can define the Gini coefficient of health inequality G_h:

$$G_h \equiv 1 - \frac{\sum_i (2\rho_i - 1)h_i}{n^2 \mu_h} \tag{2}$$

In this formula the weight given to the health level of an individual is determined by the rank of that individual in the health distribution (ρ_i).

For the measurement of the socioeconomic inequality of health, i.e. the inequality of health in relation to the socioeconomic position of individuals, other indices are required. One such measure is the health Concentration Index C_h:

$$C_h \equiv 1 - \frac{\sum_i (2\lambda_i - 1)h_i}{n^2 \mu_h} \tag{3}$$

The main difference between G_h and C_h is that the latter uses weights determined by the rank of individuals in the income distribution (λ_i).

The properties of the Concentration Index and its relation to the Gini coefficient are well-known.[5] Whereas the Gini coefficient lies between 0 and 1 (or more exactly $(n-1)/n$), the Concentration Index can take any value between -1 and $+1$ (the exact limits are $-(n-1)/n$ and $+(n-1)/n$). $C_h = 0$ is seen as an indication that there is no systematic correlation between the distribution of health and the distribution of income, $C_h > 0$ that there is positive correlation, and $C_h < 0$ that there is negative correlation.

An interesting formula can be derived which explains this property. Suppose that the chance of your having a high or a low income is not related to your health status. This means that whatever your health status, the best guess we can make about your income is that it is equal to the average income

---

[5]See, for instance, Kakwani (1980: 173-181).

μ_y. One might also say that whatever your rank in the health distribution, the best guess we can make about your rank in the income distribution is that it is equal to the average rank $(n+1)/2$. If this holds for every member of the population, the best estimate of the Concentration Index is:

$$\hat{C}_h = 1 - \frac{\sum_i (2\left[\frac{n+1}{2}\right] - 1)h_i}{n^2 \mu_h} = 0 \qquad (4)$$

More generally, assume that the best guess we can make about person i's income rank, $\hat{\lambda}_i$, is a simple linear function of that person's health rank ρ_i, i.e. assume that we have:

$$\hat{\lambda}_i = \hat{\alpha} + \hat{\beta}\rho_i \qquad (5)$$

Let the values of $\hat{\alpha}$ and $\hat{\beta}$ be estimated by means of an OLS regression of the equation $\lambda_i = \alpha + \beta\rho_i + u_i$. In that case $\hat{\beta}$ is equal to the coefficient of correlation between the ranks of income and the ranks of health:

$$\hat{\beta} = \frac{Cov(\lambda, \rho)}{\sigma_\lambda \sigma_\rho} \qquad (6)$$

and furthermore we have:

$$\hat{\alpha} = \frac{(1 - \hat{\beta})(n + 1)}{2} \qquad (7)$$

The last equation implies that $\sum_i \hat{\lambda}_i = \sum_i \lambda_i = \dfrac{n(n + 1)}{2}$. We have the following result:

Proposition 1 *If income ranks are estimated from health ranks by the relation $\hat{\lambda}_i = \hat{\alpha} + \hat{\beta}\rho_i$, where $\hat{\beta}$ is the coefficient of correlation between the ranks of income and the ranks of health, the estimated value of the health Concentration Index is $\hat{C}_h = \hat{\beta}G_h$.*

Proof. Replacing λ_i by $\hat{\lambda}_i$ in the definition of the health Concentration

Index, we obtain the following estimated value of the index:

$$
\begin{aligned}
\hat{C}_h &= 1 - \frac{\sum_i \left\{ 2 \left[\frac{(1-\hat{\beta})(n+1)}{2} + \hat{\beta}\rho_i \right] - 1 \right\} h_i}{n^2 \mu_h} \\
&= 1 - \frac{(1-\hat{\beta})(n+1)n\mu_h + \sum_i (2\hat{\beta}\rho_i - 1 + \hat{\beta} - \hat{\beta})h_i}{n^2 \mu_h} \\
&= 1 - \frac{(1-\hat{\beta})(n+1)n\mu_h + (\hat{\beta}-1)n\mu_h + \hat{\beta}\sum_i (2\rho_i - 1)h_i}{n^2 \mu_h} \\
&= 1 - (1-\hat{\beta}) - \hat{\beta}(1 - G_h) \\
&= \hat{\beta}G_h
\end{aligned}
$$

This establishes the result. ∎

Hence, if the ranks of income and health are uncorrelated (i.e. $\hat{\beta} = 0$ and $\hat{\alpha} = (n+1)/2$, the case we examined previously) the expected value of the health Concentration Index is zero. If, however, they are positively correlated ($\hat{\beta} > 0$), we expect the health Concentration Index to be positive; and of course if they are negatively correlated, we expect the index to be negative. What the formula $\hat{C}_h = \hat{\beta}G_h$ highlights is that the Concentration Index is essentially the coefficient of correlation between the ranks of income and health scaled by the Gini coefficient of health inequality.[6]

4   Critique of the Health Concentration Index

Let me begin by mentioning a well-known, but relatively minor shortcoming of the health Concentration Index. Although we would like to find out whether a systematic relation exists between income and health, with respect to income the index takes into account only the ranks and not the levels of income. But a given ranking of incomes may hide very different levels of income. Both a relatively equal and a relatively unequal distribution of income are compatible with any given ranking. Moreover, if changes occur in the distribution of income which do not affect the income ranks (e.g. a series of transfers which make the distribution more equal), no effect will be seen on the Concentration Index.

---

[6]This formula seems to be new. The closest I have been able to find is a result by Kakwani (1980: 174, Theorem 8.4). Applied to our case, he proves that C_h is equal to G_h multiplied by the ratio of the coefficient of correlation between h and ρ and the coefficient of correlation between h and λ.

The major problem with the health Concentration Index C_h - and with the Gini coefficient of health inequality G_h - is that the measurement of health is fundamentally different from the measurement of income. As a matter of fact, there is no "natural unit" for the measurement of health, and any particular unit seems to be as good (or as bad) as any other. Should health be measured on a scale between 0 and 1? Why not between 1 and 20, or between -2 and 50? Does it make sense to say that your health has "doubled" when your health status changes from h_i to $2h_i$? The position defended here is that this is highly doubtful. The health status indicator is an essentially qualitative variable which might be used to order people according to their health situation, but no conclusion can be drawn about the intensity corresponding to a specific value of the health indicator. In other words, the health indicator is intrinsically an ordinal variable, not a cardinal one.

This observation is by no means new, but its consequences should not be underestimated. No clear meaning can be given to the average of an ordinal variable. Therefore, all measures which refer to the average level of health in society - and the health Concentration Index is one of them - must be treated with extreme caution.[7] Those who advocate the use of this index appear to be aware of the problem and stress that the health indicator should be a cardinal variable. They have developed ingenious methods to transform ordinal health indicators (such as self-reported health) into cardinal ones.[8] Legitimate doubts may be raised about this alchemy which turns iron into gold. Moreover, it may not even be enough. In fact, we would like to have an indicator that enables us not only to identify the type of correlation (positive or negative), but also to say something about its magnitude (large or small). Yet if the health indicator is a cardinal variable but not a ratio-scale variable, then the value of the Concentration Index C_h (and of the Gini coefficient G_h as well) is arbitrary. In other words, the issue is whether we can measure health like we measure temperature (by a cardinal variable) or like we measure length (by a ratio-scale variable). We can go from any temperature indicator t to an alternative temperature indicator $\tilde{t}$ by means of a positive linear transformation $\tilde{t} = a + bt$, where b is a positive scalar. For the measurement of length, however, two alternative indicators l and $\tilde{l}$ must be such that $\tilde{l} = bl$.[9]

---

[7]For a clear presentation of the issues involved, see Allison and Foster (2004). With regard to qualitative variables like health they suggest to use median-based instead of mean-based inequality measures.

[8]A recent example is van Doorslaer and Jones (2003).

[9]An excellent overview of the different types of measurement is given by Sen (1973: 3-5).

It is easy to show that the value of the Concentration Index C_h is to a large extent arbitrary if the health indicator is a cardinal but not a ratio-scale variable. Suppose that we transform the original health indicator h into an alternative one $\tilde{h}$ by means of the function $\tilde{h} = a + bh$, where b is a positive scalar (if we are not prepared to contemplate negative health levels, we should add the requirements that h and $\tilde{h}$ must be non negative). Observe that $\mu_{\tilde{h}} = a + b\mu_h$. The Concentration Index of socioeconomic health inequality now becomes:

$$C_{\tilde{h}} = 1 - \frac{\sum_i (2\lambda_i - 1)(a + bh_i)}{n^2(a + b\mu_h)} \tag{8}$$

which after some manipulations can be reduced to[10]:

$$C_{\tilde{h}} = \left[\frac{b\mu_h}{a + b\mu_h} \right] C_h \tag{9}$$

Unless we always have $a = 0$, i.e. unless the health indicator is a ratio-scale variable, the value of the Concentration Index can be changed at will. The scale chosen for the measurement of the health status therefore determines the exact value of the index. As a result, it is simply impossible to say whether the measured inequality is large or small; a value close to zero does not mean that the socioeconomic inequality is necessarily small, and a value close to 1 does not mean that socioeconomic inequality is necessarily large.

The same goes for the Gini coefficient of health inequality since we have:

$$G_{\tilde{h}} = \left[\frac{b\mu_h}{a + b\mu_h} \right] G_h \tag{10}$$

Hence a positive linear transformation of the measure of health affects the Concentration Index and the Gini coefficient in the same way. It follows that the ratio of the two will remain constant, and in particular we have, for the correlation tendency specified in (5):

$$\frac{\hat{C}_h}{G_h} = \frac{\hat{C}_{\tilde{h}}}{G_{\tilde{h}}} = \hat{\beta} \tag{11}$$

This relationship continues to hold even if our health indicator is a purely ordinal variable which is subjected to a positive monotonic transformation.

The upshot is that unless health is measured by a ratio-scale variable, the only solid information which is conveyed by the Concentration Index C_h and

---

[10]This result was already proved by Kakwani (1980: 176).

the Gini coefficient G_h together is that we expect their ratio to be equal to the coefficient of correlation between the ranks of income and health. Given how hard it is to give meaning to the notions of "zero health" and "doubling health", it seems extremely difficult to maintain that it is possible to measure health by a ratio-scale indicator. So perhaps we should limit ourselves to the calculation of the coefficient of correlation between the ranks of income and health, and abandon the idea that the health Concentration Index can be expressed more precisely. Or at least we should restrict the use of the health Concentration Index to those aspects of health which can be measured on a ratio-scale.

This difficulty with regard to the interpretation of the Concentration Index is reminiscent of the critique levied against Dalton's measure of inequality by Atkinson. Dalton (1920) proposed to measure income inequality by comparing the actual social welfare level to the maximum possible social welfare level. Atkinson (1970) pointed out that the value of this measure changes as a result of a positive linear transformation of the individual welfare function. He also showed that an alternative measure can be defined by comparing income levels instead of welfare levels. This is the approach which I follow in the remainder of the paper.

5   An Alternative Concentration Index

Unlike health, income *is* measured by a ratio-scale indicator. That is the reason why the Gini coefficient of income inequality G_y, defined as:

$$G_y \equiv 1 - \frac{\sum_i (2\lambda_i - 1)y_i}{n^2 \mu_y} \tag{12}$$

remains constant if we change the unit by which income is measured. This implies that the value of the Gini coefficient G_y can be used to assess the magnitude of income inequality.

For the measurement of the socioeconomic inequality of health we need to take into account both the distribution of income and the distribution of health. Health status is essentially an ordinal variable. We can safely rely on information concerning health status rank, but we should avoid to refer to health status levels. The health Concentration Index C_h is based upon health levels and income ranks. In view of the nature of income and health, it seems advisable to turn things around and to construct an indicator based upon income levels and health ranks. Instead of weighing health status levels by income ranks, let us weigh income levels by health status ranks. This gives

us the Concentration Index of health-ranked income inequality C_y:

$$C_y \equiv 1 - \frac{\sum\limits_{i}(2\rho_i - 1)y_i}{n^2\mu_y} \tag{13}$$

This alternative Concentration Index C_y has similar properties as the index C_h. Suppose that there is no systematic correlation between income and health: whether you are rich or poor has no influence on your rank in the health distribution. In other words, for every individual the expected rank in the health distribution is the same and equal to $(n+1)/2$. In these circumstances, the expected value of the Concentration Index is equal to:

$$\hat{C}_y = 1 - \frac{\sum\limits_{i}(2\left[\frac{n+1}{2}\right] - 1)y_i}{n^2\mu_y} = 0 \tag{14}$$

If, however, there is a systematic correlation between income and health, we can expect that the value of the Concentration Index will be different from zero. Assume that the health rank ρ_i and the income rank λ_i are correlated as follows[11]:

$$\hat{\rho}_i = \hat{\alpha} + \hat{\beta}\lambda_i \tag{15}$$

Following the same procedure as above, we can prove:

Proposition 2 *If health ranks are estimated from income ranks by the relation $\hat{\rho}_i = \hat{\alpha} + \hat{\beta}\lambda_i$, where $\hat{\beta}$ is the coefficient of correlation between the ranks of income and the ranks of health, the estimated value of the Concentration Index of health-ranked income inequality is $\hat{C}_y = \hat{\beta}G_y$.*

Several remarks should be made about the Concentration Index C_y. First, observe that when the health status variable h is subjected to a positive monotonic transformation, the value of index C_y remains the same, unlike that of index C_h. Moreover, a proportional change of all income levels has also no effect upon index C_y. Hence C_y is scale independent and some kind of absolute value may be attached to it.

Second, although the value of the Concentration Index theoretically lies between -1 and $+1$, Proposition 2 suggests that we can expect its value to lie between $-G_y$ and $+G_y$, which is a much narrower band.

Third, Proposition 2 clarifies that the Concentration Index is a combination of two other measures: the Gini coefficient G_y, which measures the 'pure' income inequality, and the rank correlation coefficient $\hat{\beta}$, which measures the

---

[11]The $\hat{\alpha}$ and $\hat{\beta}$ coefficients are the same as in (5).

correlation between the ranks of income and health. One might say that the second is the more important of the two, since it exclusively determines the sign of the Concentration Index, and it jointly determines its absolute level.

Can a case be made to replace the Concentration Index C_h by the Concentration Index C_y as an indicator of the socioeconomic inequality of health? There are at least two arguments in favour of this: first, the insensitivity of C_y to changes in the measurement of income and health, and second, the fact that the sign of the expected value of C_y is always the same as the sign of the expected value of C_h. An argument against such a change, however, is that the ranking of social states generated by index C_y is not necessarily the same as the one generated by index C_h. It might even be that a change which is recorded as inequality-increasing by one index, is recorded as inequality-decreasing or neutral by the other. Here is an example: suppose income is transferred from a relatively poor person to a relatively rich one, but without changing the income rank of any person, and without modifying the health status of any person. This transfer, which is clearly income inequality increasing, definitely increases the Gini coefficient G_y. Whether it also increases the Concentration Index C_y is not sure; if the rich person happens to be healthier than the poor person, C_y will increase; in the reverse case, C_y will decrease. So the exact effect upon C_y depends upon chance. We can, however, use the expected value formula of Proposition 2 to derive that we expect the index to increase if income and health are positively correlated ($\hat{\beta} > 0$) and to decrease if they are negatively correlated ($\hat{\beta} < 0$). Since health ranks and income ranks have not changed, the coefficient of correlation $\hat{\beta}$ remains the same as before the transfer. Moreover, given that nothing has been changed in the distribution of health and the income ranks remain the same, the transfer does not affect the Concentration Index C_h. So a change which has no effect at all upon the index C_h may well have an effect upon the index C_y (and vice versa).

But perhaps the problem is not as hard as it seems. At least the expected values of the two concentration indices - and therefore the ranking of social states according to the expected values of the two indices - are clearly related to one another. By combining Propositions 1 and 2 we obtain:

$$\hat{C}_h = \frac{G_h}{G_y}\hat{C}_y \qquad (16)$$

Hence any ordering of social states generated by the estimated indicator $\hat{C}_y$ can be transformed into an ordering of social states generated by the estimated indicator $\hat{C}_h$ by "correcting" the values of the first indicator by the corresponding ratio of the Gini coefficients, G_h/G_y.

6 A New Indicator

6.1 The Atkinson approach

Another way of measuring the socioeconomic inequality of health consists of following an Atkinson approach.[12] Let social welfare W be a simple sum of individual welfare levels u_i:

$$W = \sum_i u_i \tag{17}$$

Atkinson assumed that the welfare of individual i is determined exclusively by her income y_i, and that for all individuals an identical individual welfare function is used, which is concave in income. Hence the individual welfare of person i generated by income y_i is $u_i = u(y_i)$, and the social welfare generated by income distribution y is $W(y) = \sum_i u(y_i)$. To measure the inequality of this distribution, Atkinson looked for an equivalent income y^e which, if given to everyone, would generate exactly the same level of social welfare as the one generated by the existing unequal distribution of incomes. The Atkinson measure of inequality A is defined as:

$$A = 1 - \frac{y^e}{\mu_y} = \frac{\sum_i (y_i - y^e)}{\sum_i y_i} \tag{18}$$

The concavity of the individual welfare function ensures that the equivalent income cannot exceed the average income, and therefore A lies between 0 (perfect equality) and 1 (extreme inequality).

Suppose now that the welfare of individual i is determined both by her income y_i and by her health status h_i. As before we assume that income y_i is measured by a nonnegative real number, but with regard to health we adopt an approach which seems more appropriate to represent its qualitative nature. Let us assume that the health status of an individual can take a finite number of "values", say k different values $h(1), h(2), ..., h(k)$.[13] The individual welfare of a person with income y_i and health status h_i is described

---

[12]My main source of inspiration is Atkinson (1970). I use his framework to compare the inequality of income to the inequality of health in a *given* social state. The work of Atkinson and Bourguignon (1982) deals explicitly with multidimensional inequality, but tries to establish criteria to compare *different* social states.

[13]These values might correspond to purely qualitative designations of health, such as "very good", "good", "average", "bad" and "very bad". Our framework can be adapted easily to continuous health measures.

by a function $u_i = u(y_i, h_i)^{14}$, and the social welfare generated by incomes y and health h by:

$$W(y, h) = \sum_i u(y_i, h_i) \tag{19}$$

The purpose of a socioeconomic indicator of health inequality is to find out whether the distribution of health follows the same pattern as the distribution of income. In the Atkinson spirit we would like to test this by comparing the existing situation of income and health to another, "equivalent" situation of income and health, which somehow reflects the idea of income and health being distributed without correlation over the whole population. The difficulty is how to define this uncorrelated distribution of income and health. Could it be a situation in which we assign to every member of the population the same health status? Following Atkinson we could try and find the health status $h(j^*)$ such that

$$\sum_i u(y_i, h(j^*)) = \sum_i u(y_i, h_i) \tag{20}$$

Even if there exists such a health status level - which is doubtful - it is hard to see what use could be made of it. There is, in fact, no meaningful "average" health status to which it can be compared. As in the case of the health Concentration Index, little or nothing can be expected from a comparison of health levels. What we can do, however, is compare *probabilities* of attaining different health statuses.

6.2   Risk profiles

The actual health status of individual i, h_i, is drawn from the set of possible health statuses, $\{h(1), h(2), ..., h(k)\}$. If rich people tend to have better health than poor people, this means that for rich people the probabilities of drawing a good health status are higher than those for poor people. This can be formalized as follows. Let $\pi_i(j) \geq 0$ be the probability that person i has health status $h(j)$. Evidently we must have, for every person i:

$$\sum_j \pi_i(j) = 1 \tag{21}$$

where j goes from 1 to k. We would like to know whether the individual probability distribution vectors (or risk profiles) $\pi_i = [\pi_i(1), \pi_i(2), \dots, \pi_i(j)]$

---

[14]It would be more accurate to speak of a family of functions $u(y_i, h(j))$, one for each of the k possible health status levels $h(j)$. If $h_i = h(s)$, we have $u(y_i, h_i) = u(y_i, h(s))$. In other words, I am using here the concept of a state-dependent utility function (cf. Viscusi and Evans, 1990, and Evans and Viscusi, 1991).

differ systematically for different income groups, and whether a pattern can be found in the differences.[15]

Let us reflect for a moment on the notion of individual risk profile. The idea is that your chances of attaining a certain health status depend upon the group to which you belong. It might very well be that we have reliable information on systematic differences between the young and the old, between men and women, between urban and rural populations, etc. For our purposes we need a partition of the population into Q different subsets $N_1, N_2, \ldots, N_Q$, such that $N_1 \cup N_2 \cup \ldots \cup N_Q = N$ and no two subsets overlap. Each individual belongs to exactly one subset, or reference group as I will call it from now on. For each reference group R we then define a specific risk profile $\pi_R = [\pi_R(1), \pi_R(2), \ldots, \pi_R(j)]$ by calculating the frequency with which each health status occurs in this group. Let reference group N_R consists of n_R individuals, and suppose that $n_R(j)$ of them have health status $h(j)$. Then we take:

$$\pi_R(j) \equiv \frac{n_R(j)}{n_R} \tag{22}$$

If individual i belongs to reference group R, her individual risk profile π_i is equal to the risk profile of her reference group π_R.

One can think of various ways of dividing the population into reference groups, using criteria like age, sex, ethnicity, etc. or combinations of these. One possibility is to define reference groups on the basis of income; one could, for instance, divide the population into income deciles and derive the risk profile of each decile. Since what interests us here is the relationship between income and health, we assume in what follows that our reference groups are defined exclusively on the basis of income.[16]

Above we have shown how for each reference group a typical risk profile can be defined. Using the same procedure we can also define a risk profile for society as a whole, which I call the *average* risk profile of society. The average risk profile $\pi = [\pi(1), \pi(2), \ldots, \pi(j)]$ is found by calculating the frequency with which each health status occurs in society. If there are in total $n(j)$ individuals with health status $h(j)$, we have:

$$\pi(j) \equiv \frac{n(j)}{n} \tag{23}$$

Obviously, we have:

$$\pi = \sum_R \frac{n_R}{n} \pi_R = \sum_i \frac{1}{n} \pi_i \tag{24}$$

---

[15]Zheng (2006) has followed a similar approach.

[16]If reference groups are defined otherwise, the indicators which I introduce further in this section must be slightly modified.

6.3   Equivalent incomes

If health is distributed unequally over the population and if this inequality is somehow related to income, then it will be reflected in differences between individual risk profiles. As a first step towards the measurement of socioeconomic health inequality, I translate the deviations between individual risk profiles and the average risk profile into income. Imagine for a moment that there were no systematic differences among individuals with regard to the probabilities of attaining a certain health status. This means that each individual's risk profile π_i would be equal to the average risk profile π. The expected individual welfare of person i with income y_i and risk profile π would be:

$$e(y_i, \pi) = \sum_j \pi(j) u(y_i, h(j)) \tag{25}$$

In reality, however, person i has health status h_i, and his actual individual welfare is $u(y_i, h_i)$. If $u(y_i, h_i) > e(y_i, \pi)$, he has been lucky in his draw of health status; if $u(y_i, h_i) < e(y_i, \pi)$, he has been unlucky. Now whether you have good or bad luck depends upon two circumstances: first upon the specific risk profile of the reference group to which you belong (this we can call social luck), and second upon your personal luck in drawing a specific health status. That is to say, even if you belong to a group with a high chance of drawing a bad health status, you might end up being in very fine health (and vice versa). What we are interested in is not the element of personal luck, but the systematic differences between income groups. One way of eliminating the personal luck factor, is to consider person i's expected individual welfare *according to the risk profile of his reference group* instead of his actual individual welfare. Let us define $e(y_i, \pi_i)$ as the group-specific expected individual welfare of person i with income y_i, i.e. the expected individual welfare of person i with income y_i and risk profile π_i:

$$e(y_i, \pi_i) = \sum_j \pi_i(j) u(y_i, h(j)) \tag{26}$$

If $e(y_i, \pi_i) > e(y_i, \pi)$, person i's reference group has a better than average risk profile and person i has good social luck; if $e(y_i, \pi_i) < e(y_i, \pi)$, person i's reference group has a worse than average risk profile and person i has bad social luck. In both cases we can ask ourselves the following question: which level of income x_i would ensure that person i's group-specific expected individual welfare is exactly equal to his expected individual welfare $e(y_i, \pi)$? Put more formally, we are looking for the income level x_i such that:

$$e(x_i, \pi_i) = e(y_i, \pi) \tag{27}$$

This income x_i is what I call the *equivalent income* of person i. It would ensure that person i with his specific risk profile can expect to be as well off as he could expect to be if his risk profile were that of society in general.

6.4   Properties of the expected welfare function

We started from an individual welfare function $u(y_i, h_i)$ defined over income and health status, but what we are ultimately interested in is the expected welfare function $e(y_i, \pi_i)$ defined over income and risk profile. It seems useful to impose some structure upon this function.[17]

The first and most obvious condition relates to the marginal expected welfare of income.

Condition 1 (Decreasing marginal expected welfare of income) *Given the risk profile π_i, the marginal expected welfare of income is positive and decreasing, i.e.*

$$\frac{\partial e(y_i, \pi_i)}{\partial y_i} > 0, \quad \frac{\partial^2 e(y_i, \pi_i)}{\partial y_i^2} < 0.$$

This condition implies that for a person with good social luck, i.e. for whom $e(y_i, \pi_i) > e(y_i, \pi)$, we have $x_i < y_i$, and for a person with bad social luck, i.e. $e(y_i, \pi_i) < e(y_i, \pi)$, we have $x_i > y_i$. If positive, the difference $y_i - x_i$ measures the income which can be taken from person i as a compensation for her good social luck; if negative, the difference $y_i - x_i$ measures the income which should be given to person i as a compensation for her bad social luck.

Our second condition makes a statement about the marginal expected welfare of income at different risk profiles:

Condition 2 (Social luck persistence) *Let π' and π'' be two different risk profiles. If at income level y_i we have $e(y_i, \pi') > e(y_i, \pi'')$, then*

$$\frac{\partial e(y_i, \pi')}{\partial y_i} \geq \frac{\partial e(y_i, \pi'')}{\partial y_i};$$

and if we have $e(y_i, \pi') = e(y_i, \pi'')$, then

$$\frac{\partial e(y_i, \pi')}{\partial y_i} = \frac{\partial e(y_i, \pi'')}{\partial y_i}.$$

---

[17]Clearly the properties of $e(y_i, \pi_i)$ are intimately connected to those of $u(y_i, h_i)$, i.e. those of $u(y_i, h(j))$. It would be interesting to know whether the conditions imposed upon $e(y_i, \pi_i)$ can be translated into an equivalent set of conditions upon the functions $u(y_i, h(j))$.

The condition of social luck persistence says that if you compare two persons having the same income but different risk profiles, an increase of income causes an increase in expected welfare which for the person with the better risk profile π' is at least as great as it is for the person with the worse risk profile π''. In other words, the person with the better risk profile must be considered as more capable of generating individual welfare from a given income.[18]

This condition has two important consequences. First of all, it implies that risk profiles can be ranked in order of preference independent of income, i.e. regardless of their income, all people rank risk profiles in the same order of preference.

Corollary 1 (Constant risk profile preferences) *The preference order of risk profiles is independent of income, i.e.*

1. *if for some y_i, π' is preferred to π'', i.e. if $e(y_i, \pi') > e(y_i, \pi'')$, then this holds for all y_i;*

2. *if for some y_i, π' and π'' are equally preferred, i.e. if $e(y_i, \pi') = e(y_i, \pi'')$, then this holds for all y_i.*

Second, it allows us to say something about the social luck effect, by which I mean the effect on expected individual welfare of a change in risk profile. Consider two different risk profiles, and let risk profile π' be preferred to risk profile π''. By Condition 1, if a person's income rises from y'' to y', his expected welfare will increase whether his risk profile is π' or π''. The condition of social luck persistence implies that when his risk profile is π' rather than π'', the increase in expected welfare will be at least as great. In formal terms, we have $e(y', \pi') - e(y'', \pi') \geq e(y', \pi'') - e(y'', \pi'')$. But this implies $e(y', \pi') - e(y', \pi'') \geq e(y'', \pi') - e(y'', \pi'')$, and so we can conclude that the effect of improving your risk profile does not decrease with income.

Corollary 2 (Non decreasing social luck effect) *Suppose that we have incomes y' and y'' and risk profiles π' and π'' such that $y' > y''$ and risk profile π' is preferred to π''. Then it follows that $e(y', \pi') - e(y', \pi'') \geq e(y'', \pi') - e(y'', \pi'')$.*

---

[18]In his utilitarian calculus, Edgeworth (1881: 77-78) drew attention to differences in the capacity for pleasure or happiness, e.g. between men and women. What I suggest here is that people with a relatively good risk profile tend to have higher individual welfare than people with a relatively bad one, and that the difference does not become smaller for higher incomes.

In what follows I assume that Conditions 1 and 2 are satisfied. An additional condition which might be imposed relates to the ratio of the marginal expected individual welfare of income at different risk levels. If this ratio is the same whatever the income may be, then we speak of constant risk impact:

Condition 3 (Constant risk impact) *Let π' and π'' be two given risk profiles. Then for any income level y_i the ratio $\dfrac{\partial e(y_i, \pi')}{\partial y_i} / \dfrac{\partial e(y_i, \pi'')}{\partial y_i}$ is constant.*

An example of a family of individual welfare functions which generates an expected welfare function with all the above mentioned properties is the following. For each possible health level $h(j)$ we start from the function:

$$u(y_i, h(j)) = \beta + \alpha(j)v(y_i) \tag{28}$$

where β and $\alpha(j)$ are scalars, $\partial v(y_i)/\partial y_i > 0$, and $\partial^2 v(y_i)/\partial y_i^2 < 0$. The scalars $\alpha(j)$ measure the impact of the different health statuses upon individual welfare. If your income remains the same but your health status changes from $h(s)$ to $h(t)$, your individual welfare changes by the amount $[\alpha(s) - \alpha(t)]\,v(y_i)$. If two people with the same income but different health statuses $h(s)$ and $h(t)$ see their income increase by the same amount, the ratio of their welfare increases is equal to $\alpha(s)/\alpha(t)$. For simplicity, I choose the following simple specification for function $v(y_i)$:

$$v(y_i) = [y_i]^\epsilon \tag{29}$$

where $0 < \epsilon < 1$. Since for positive incomes the values of $v(y_i)$ are positive, the health status $h(j)$ with the highest $\alpha(j)$ coefficient is obviously the one which is the most desired; without loss of generality we can assume that $\alpha(1) > \alpha(2) > \ldots > \alpha(k) > 0$.

To check whether the conditions specified above hold, let us define scalars θ and θ_i as follows:

$$\theta = \sum_j \pi(j)\alpha(j) > 0 \tag{30}$$

$$\theta_i = \sum_j \pi_i(j)\alpha(j) > 0 \tag{31}$$

It is easy to see that we have $e(y_i, \pi) = k\beta + \theta\,[y_i]^\epsilon$ and $e(y_i, \pi_i) = k\beta + \theta_i\,[y_i]^\epsilon$, which implies that he ratio θ_i/θ can be seen as a social luck indicator: if $\theta_i/\theta > 1$, risk profile π_i is better than the average risk profile π, and if $\theta_i/\theta <$

1, risk profile π_i is worse than the average risk profile π. Condition 1 follows from $e(y_i, \pi_i) = k\beta + \theta_i [y_i]^\epsilon$, $\theta_i > 0$, and the fact that $0 < \epsilon < 1$. Condition 2 follows from the fact that $\dfrac{\partial e(y_i, \pi')}{\partial y_i} = \theta'\epsilon [y_i]^{\epsilon-1}$, $\dfrac{\partial e(y_i, \pi'')}{\partial y_i} = \theta''\epsilon [y_i]^{\epsilon-1}$, and $e(y_i, \pi') > e(y_i, \pi'')$ if and only if $\theta' > \theta''$. Finally, Condition 3 is also satisfied since $\dfrac{\partial e(y_i, \pi')}{\partial y_i} \Big/ \dfrac{\partial e(y_i, \pi'')}{\partial y_i} = \theta'/\theta''$, which is constant.

For this specification the equivalent incomes can be calculated very easily. It turns out that the equivalent income x_i is always a fixed proportion of the actual income y_i, with the factor of proportionality inversely related to the degree of social luck θ_i/θ:

$$x_i = \left[\frac{\theta}{\theta_i}\right]^{1/\epsilon} y_i \tag{32}$$

6.5 An Atkinson measure

I propose to measure the socioeconomic inequality of health by comparing the actual income distribution y to the equivalent income distribution x. More specifically, I suggest to construct an indicator taking the individual income differences $y_i - x_i$ as the basic units. Following Atkinson, let us see which of the two distributions has the lowest total income (and since both distributions involve the same number of people, this is equivalent to checking which of the two has the lowest average income). If distribution y has the lowest total income, we have $\sum_i (y_i - x_i) < 0$, and if distribution x has the lowest total income, we have $\sum_i (y_i - x_i) > 0$. The case $\sum_i (y_i - x_i) < 0$ means that the aggregate amount of money which can be taken from those with good social luck is lower than the amount of money which has to be given to those with bad social luck, and the case $\sum_i (y_i - x_i) > 0$ that it is higher. That in itself may be an interesting fact, but what does it imply for the socioeconomic inequality of health?

The answer is found by looking at the way in which the equivalent income is defined. As explained above, the equivalent income x_i is such that $e(x_i, \pi_i) = e(y_i, \pi)$, i.e. it is derived from a comparison of a situation with income x_i and risk profile π_i and a situation with income y_i and risk profile π. From the definition it follows trivially that

$$e(x_i, \pi_i) - e(y_i, \pi_i) = e(y_i, \pi) - e(y_i, \pi_i) \tag{33}$$

Imagine that we move from the situation with income y_i and risk profile π to the situation with income x_i and risk profile π_i. This change can be

decomposed into two movements: a first in which your risk profile changes from π to π_i, and a second in which your income changes from y_i to x_i. The right-hand side of (33) captures the first or "social luck effect", while the left-hand side captures the second or "income effect". By construction the change in income must be such that the income effect exactly offsets the social luck effect. Now let us compare two persons with the same risk profile but with different incomes. Assume first that they have bad social luck. Corollary 2 with regard to the expected welfare function implies that the social luck effect will be at least as strong for the richer as it is for the poorer person, since the difference $e(y_i, \pi) - e(y_i, \pi_i)$ does not decrease with income. This implies that for the richer person, the income change must cause an income effect which is higher than or equal to that of the poorer person. Condition 1 stipulates that the richer person has a lower marginal expected welfare of income than the poorer one. Hence, the combination of an equal or higher social luck effect and a lower marginal expected welfare of income entails that the richer person should be given a higher amount of money in compensation for her bad social luck. Alternatively, assume that our two individuals have good social luck. Again, the social luck effect will be at least as strong for the richer as it is for the poorer person. And because of the richer person's lower marginal expected welfare of income, a larger amount of money can be taken away from him to compensate him for his good social luck. Therefore, both in case of good and of bad social luck, the compensatory amounts $\mid y_i - x_i \mid$ are higher for rich than for poor people, and increase with income.

This is precisely the kind of metric we need to be able to use an Atkinson type of indicator for the measurement of socioeconomic health inequality. We have to distinguish three possibilities. First, suppose that good and bad risk profiles are distributed at random over the whole population. Both among the poor and among the rich there will be income groups with good and with bad risk profiles. On average we can expect that for both poor and rich the compensatory payments will be roughly equal to the compensatory receipts. Hence for society as a whole we expect to find $\sum_i (y_i - x_i) = 0$, or $\mu_y = \mu_x$. Second, suppose that the distribution of risk profiles is such that the bad profiles occur more frequently among the poor, and the good more frequently among the rich. In that case the aggregate sum which has to be paid to compensate those with bad risk profiles will be smaller than the aggregate sum which can be taken from those with good risk profiles. Hence for society as a whole we will have $\sum_i (y_i - x_i) > 0$, or $\mu_y > \mu_x$. Third, suppose that the distribution is such that the bad profiles occur more frequently among the rich, and the good more frequently among the poor. Then the aggregate sum which has to be paid to compensate those with

bad risk profiles will be higher than the aggregate sum which can be taken from those with good risk profiles. Hence for society as a whole we will have $\sum_i (y_i - x_i) < 0$, or $\mu_y < \mu_x$. Summarizing: $\mu_y = \mu_x$ is a sign that risk profiles are distributed at random, $\mu_y > \mu_x$ a sign that there is some kind of positive correlation between the distribution of good risk profiles and the distribution of income, and $\mu_y < \mu_x$ a sign that there is some kind of negative correlation between the distribution of good risk profiles and the distribution of income. It should also be clear that the more pronounced the (positive or negative) correlation will be, the greater the deviation between μ_y and μ_x will be. Therefore we can use the following Atkinson-like measure as an indicator of socioeconomic inequality of health:

$$A_h = 1 - \frac{\mu_x}{\mu_y} = \frac{\sum_i (y_i - x_i)}{\sum_i y_i} \qquad (34)$$

The interpretation of the index must be adapted slightly. In contrast to the original Atkinson index, the index A_h can take positive as well as negative values. A positive value indicates that, on average, income and health are positively correlated; and a negative value, that they are negatively correlated.

6.6   Properties of A_h

The indicator A_h has properties which seem to make it suitable as a measure of socioeconomic health inequality.

First, it is insensitive to a positive linear transformation of the individual welfare functions $u(y_i, h(j))$. Assume that instead of functions $u(y_i, h(j))$ we use functions $\tilde{u}(y_i, h(j))$ defined as

$$\tilde{u}(y_i, h(j)) = a + bu(y_i, h(j)) \qquad (35)$$

where b is a positive scalar. Clearly we have:

$$\tilde{e}(y_i, \pi) = a \sum_j \pi(j) + b \sum_j \pi(j) u(y_i, h(j)) = a + be(y_i, \pi) \qquad (36)$$

$$\tilde{e}(x_i, \pi_i) = a \sum_j \pi_i(j) + b \sum_j \pi_i(j) u(x_i, h(j)) = a + be(x_i, \pi_i) \qquad (37)$$

From this it follows easily that $\tilde{e}(x_i, \pi_i) = \tilde{e}(y_i, \pi)$ if and only if $e(x_i, \pi_i) = e(y_i, \pi)$. Hence the equivalent income x_i is the same if we replace the individual welfare function by a positive linear transformation of it.

Second, if all income groups have the same risk profile (i.e. if for all groups R and for all individuals i we have $\pi_R = \pi_i = \pi$), the value of the index A_h is equal to zero. This follows from the fact that in that case the equality $e(x_i, \pi_i) = e(y_i, \pi)$ can be satisfied only if $x_i = y_i$.

Next, let us examine the effect of a change in risk profiles. To keep things simple, let us suppose that the change does not modify the average risk profile π. More specifically, let us imagine that two income groups with the same number of people switch their risk profiles. If the poorest group is better off as a result, then the measure A_h decreases; if the richest group is better off, then the measure A_h increases. This can be shown as follows. Suppose that initially the richest group has risk profile π' and the poorest π'', and let π' be preferred to π''. Consider a member of the richest group with income y' and a member of the poorest group with income y''. Since $y' > y''$, it follows from Corollary 2 that $e(y', \pi') - e(y', \pi'') \geq e(y'', \pi') - e(y'', \pi'')$. The term on the left-hand side represents the expected welfare loss suffered by the member of the rich group in case of a switch from profile π' to profile π'', and the term on the right-hand side the simultaneous expected welfare gain of the member of the poor group. Since the rich person's loss is at least as great as the poor person's gain, and the rich person's marginal expected welfare of income is lower than that of the poor, the amount by which the rich person's equivalent income will increase is higher than the amount by which that of the poor person will decrease. Since the two income groups have the same number of people, the total equivalent income increases, and therefore also the average equivalent income. Hence A_h decreases. In case risk profile π'' is preferred to π', the opposite conclusion is reached.

To examine the effect of a change in incomes, I assume for simplicity that (28) and (29) hold. Moreover, let the changes be small, i.e. such that no person changes income group and such that average income remains the same. Suppose that an amount of money $\Delta > 0$ is transferred from a person with income y'' to a person with income y'. Since $x' = [\theta/\theta']^{1/\epsilon} y'$ and $x'' = [\theta/\theta'']^{1/\epsilon} y''$, the equivalent income of the person who receives Δ increases by the amount $[\theta/\theta']^{1/\epsilon} \Delta$, whereas the equivalent income of the person who loses Δ decreases by the amount $[\theta/\theta'']^{1/\epsilon} \Delta$. Hence, if $\theta' < \theta''$ (i.e. the receiving person had the worst risk profile of the two), the average equivalent income increases, which means that the index A_h diminishes in value. If, however, $\theta' > \theta''$ (i.e. the receiving person had the best risk profile of the two), the average equivalent income decreases, which means that the index A_h augments in value. Observe also that if $\theta' = \theta''$ the index does not change. This implies that income changes within the same income group have no

effect upon the index.[19]

7 An Example

To illustrate the calculation of the different indicators I have taken some data of the 2005 database of the *Behavioral Risk Factor Surveillance System* (BRFSS), a yearly telephone survey of the United States population.[20] The following table lists the number of people in each of eight income groups describing their overal health situation as "excellent", "very good", "good", "fair" or "poor":[21]

Table 2: General health in the USA (2005)

Income	Health situation					All
(× $1,000)	Excellent	Very good	Good	Fair	Poor	
0-10	1,494	2,744	5,010	5,128	3,856	18,232
10-15	1,573	3,525	6,337	5,038	3,087	19,560
15-20	2,508	5,481	8,856	5,572	2,655	25,072
20-25	3,902	8,299	11,314	5,532	2,181	31,228
25-35	6,466	13,296	14,607	5,795	1,832	41,996
35-50	10,043	18,975	16,462	5,000	1,411	51,891
50-75	12,156	21,069	14,301	3,429	855	51,810
75-	21,612	27,267	14,278	2,869	665	66,691
All	59,754	100,656	91,165	38,363	16,542	306,480

Source: BRFSS.

Table 2 allows us to calculate the risk profiles of each income group:

---

[19]This might be seen as an unwanted insensitivity of the index. The finer the income groups are defined, the smaller the problem will be. A practical advantage of it is that for each income group R we only need to calculate the equivalent income x_R which corresponds to the average income μ_R of this group, i.e. x_R is such that $e(x_R, \pi_R) = e(\mu_R, \pi)$. No matter how the incomes y_i are distributed in this group, we always have $\sum_{i \in R} x_i = n_R x_R$.

[20]For more details on the survey, see http://www.cdc.gov/brfss/ .

[21]More specifically I have used the answers to Question 1.1 "Would you say that in general your health is – ?" and to Question 13.9 "Is your annual household income from all sources – ?". I have eliminated all those who refused to answer any of these questions, or who said they did not know.

Table 3: Risk profiles in the USA (2005)

Income	Health situation				
(× $1,000)	Excellent	Very good	Good	Fair	Poor
0-10	8.19%	15.05%	27.48%	28.13%	21.15%
10-15	8.04%	18.02%	32.40%	25.76%	15.78%
15-20	10.00%	21.86%	35.32%	22.22%	10.59%
20-25	12.50%	26.58%	36.23%	17.71%	6.98%
25-35	15.40%	31.66%	34.78%	13.80%	4.36%
35-50	19.35%	36.57%	31.72%	9.64%	2.72%
50-75	23.46%	40.67%	27.60%	6.62%	1.65%
75-	32.41%	40.89%	21.41%	4.30%	1.00%
All	19.50%	32.84%	29.75%	12.52%	5.40%

Source: Own calculations.

Even a superficial look at the data suffices to ascertain the existence of a clear social gradient. Let us check whether the indicators discussed above also bring this out. In order to calculate the value of indicator C_h I assume that "Excellent" is translated into a score of 5, "Very good" into a score of 4, "Good" into a score of 3, "Fair" into a score of 2, and "Poor" into a score of 1. These values imply that the average health μ_h is equal to 3.4852. Likewise, in order to calculate the value of indicator A_h, the individual welfare functions $u(y_i, h(j))$ must be specified. I take the welfare function introduced above, i.e. $u(y_i, h(j)) = \beta + \alpha(j) [y_i]^\epsilon$, and assume that $\beta = 0$, $\alpha(1) = 5$, $\alpha(2) = 4$, $\alpha(3) = 3$, $\alpha(4) = 2$, $\alpha(5) = 1$, and $\epsilon = \frac{1}{2}$. Moreover, I assume that the average per capita incomes in the eight income classes are $5,000, $12,500, $17,500, $22,500, $30,000, $42,500, $62,500, and $100,000. For these values it turns out that $\theta = 3.4852$. For each of the eight income classes Table 4 gives the θ_R values, the social luck indicators θ_R/θ, the average per capita income y_R and the equivalent per capita income x_R:

Table 4: Social luck indicators and equivalent incomes

Income (× $1,000)	θ_R	θ_R/θ	y_R	x_R	$y_R - x_R$
0-10	2.6101	0.7489	5,000	8,915	-3,915
10-15	2.7678	0.7942	12,500	19,820	-7,320
15-20	2.9846	0.8564	17,500	23,863	-6,363
20-25	3.1988	0.9178	22,500	26,710	-4,210
25-35	3.3993	0.9753	30,000	31,536	-1,536
35-50	3.6020	1.0335	42,500	39,789	+2,711
50-75	3.7767	1.0836	62,500	53,225	+9,275
75-	3.9940	1.1460	100,000	76,146	+23,854

Source: Own calculations.

The social luck indicators, and therefore also the equivalent incomes, reveal the existence of a social gradient. Only the three highest income groups have good social luck. The average per capita income μ_y is \$48,452, and the average per capita equivalent income μ_x is \$43,095.

The specific values of the three indicators discussed in this paper are:

$$C_h = 0.0684, \quad C_y = 0.1279, \quad A_h = 0.1106 \tag{38}$$

All values are positive, which confirms the existence of a social gradient in health. The values in themselves say very little; it is only when we compare different situations that they enable us to say whether inequality has decreased or increased, and eventually by how much. It is important to note that the indicators need not change in the same direction. As an illustration, suppose that there is transfer of income of \$7,500 from 400 persons with an income of \$30,000 and in fair health to 400 persons with an income of \$5,000 and in fair health. As a result, the number of people in fair health will be 4,728 in the first (i.e. lowest) income class, 5,438 in the second, 5,932 in the fourth, and 5,395 in the fifth (everything else remains the same). The three indicators would then be:

$$C_h = 0.0685, \quad C_y = 0.1279, \quad A_h = 0.1104 \tag{39}$$

The interesting fact is that C_h increases, C_y remains unchanged, and A_h decreases.

8   Conclusion

I have tried to show that Atkinson's approach to the measurement of the inequality of income can be adapted to the measurement of the socioeconomic inequality of health. There are two crucial steps in the procedure. First, one has to move from a comparison of health status levels to a comparison of probabilities of attaining different health status levels. Second, for each person an equivalent income must be defined by comparing her (expected) welfare under the average conditions of society with her (expected) welfare under the specific conditions characteristic for her economic position.

The construction avoids direct comparisons of health status levels, and in this sense improves upon the health Concentration Index. Health status levels are compared only indirectly, through their effect upon individual welfare. The advantage of the approach followed here is that it makes the comparison completely transparent. It is also explicitly founded on familiar concepts of social welfare theory.

References

[1] ALLISON, R. Andrew and James E. FOSTER (2004), "Measuring health inequality using qualitative data", *Journal of Health Economics*, 23(3): 505-524.

[2] ATKINSON, Anthony B. (1970), "On the measurement of inequality", *Journal of Economic Theory*, 2(3): 244-263.

[3] ATKINSON, Anthony B. and François BOURGUIGNON (1982), "The comparison of multi-dimensioned distributions of economic status", *Review of Economic Studies*, 49(2): 183-201.

[4] BLEICHRODT, Han and Eddy VAN DOORSLAER (2006), "A welfare economics foundation for health inequality measurement", *Journal of Health Economics*, 25(5): 945-957.

[5] BOMMIER, Antoine and Guy STECKLOV (2002), "Defining health inequality: why Rawls succeeds where social welfare theory fails", *Journal of Health Economics*, 21(3): 497-513.

[6] DALTON, Hugh (1920), "The measurement of the inequality of incomes", *Economic Journal*, 30(119): 348-361.

[7] DANIELS, Norman, Bruce KENNEDY and Ichiro KAWACHI (2004), "Health and inequality, or, why justice is good for our health", in: Sudhir Anand, Fabienne Peter and Amartya Sen (Eds.), *Public Health, Ethics, and Equity*, Oxford, Oxford University Press: 63-91.

[8] EDGEWORTH, Francis Ysidro (1881), *Mathematical Psychics. An Essay on the Application of Mathematics to the Social Sciences*, London, C. Kegan Paul & Co.

[9] EVANS, William N. and W. Kip VISCUSI (1991), "Estimation of state-dependent utility functions using survey data", *Review of Economics and Statistics*, 73(1): 94-104.

[10] FLEURBAEY, Marc (2006), "Health, equity and social welfare", *Annales d'Economie et de Statistique*, forthcoming.

[11] KAKWANI, Nanak. C. (1980), *Income Inequality and Poverty. Methods of Estimation and Policy Applications*, Oxford, Oxford University Press.

[12] MARMOT, Michael G. (2003), "Understanding social inequalities in health", *Perspectives in Biology and Medicine*, 46(3 supplement): S9-S23.

[13] SEEBOHM ROWNTREE, Benjamin (1901), *Poverty: A Study of Town Life*, London and New York, Macmillan.

[14] SEN, Amartya (1973), *On Economic Inequality*, Oxford, Clarendon Press.

[15] VAN DOORSLAER, Eddy and Andrew JONES (2003), "Inequalities in self-reported health: validation of a new approach to measurement", *Journal of Health Economics*, 22(1): 61-87.

[16] VISCUSI, W. Kip and William N. EVANS (1990), "Utility functions that depend upon health status: estimates and economic implications", *American Economic Review*, 80(3): 353-374.

[17] WAGSTAFF, Adam, Pierella PACI and Eddy VAN DOORSLAER (1991), "On the measurement of inequalities in health", *Social Science & Medicine*, 33(5): 545-557.

[18] ZHENG, Buhong (2006), "Measuring health inequality and health opportunity", Paper presented at the conference *Advanving Health Equity*, Helsinki, WIDER-UNU, September.