Overall vs. Socioeconomic Health Inequality: A Measurement Framework and Two Empirical Illustrations

by

Adam Wagstaff[a] and Eddy van Doorslaer[b]

[a] Development Research Group and Human Development Network, The World Bank, Washington DC, USA
and
School of Social Sciences, The University of Sussex, Brighton, UK

[b] Erasmus University, 3000 DR Rotterdam, The Netherlands

January 2002

Summary

This paper outlines a framework for comparing empirically overall health inequality and socioeconomic health inequality. The framework, which is developed for both individual-level data and grouped data, is illustrated using data on malnutrition amongst Vietnamese children and on health utility amongst Canadian adults. In both cases, socioeconomic inequalities account for around 25% of overall inequality.

Corresponding author and contact details: Adam Wagstaff, The World Bank, 1818 H Street NW, Washington, D.C. 20433, USA. Tel. (202) 473-0566. Fax (202)-522 1153. Email: awagstaff@worldbank.org.

Keywords: Health inequality; socioeconomic inequality in health; socioeconomic health differentials.

Acknowledgements: The findings, interpretations and conclusions expressed in this paper are entirely those of the authors, and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

I. INTRODUCTION

In the literature on health inequalities two distinct strands are evident. The first examines overall inequalities in health and proceeds in much the same way as the literature on measuring income inequality [1, 2]. In this approach, which Wolfson and Rowe [3] refer to as the univariate approach, all inequalities in health are measured, irrespective of the other characteristics of the individuals involved. The second strand of the literature looks at a subset of health inequalities, namely those occurring across the distribution of some measure of socioeconomic status (SES) [4-10]—what Wolfson and Rowe term the bivariate approach.

This paper seeks to bring together these two approaches in a unified measurement methodology. We do not aim to resolve the much debated normative or ethical issue of which approach better captures policymakers' or societal concerns [11-16]. But by bringing the two approaches together, the paper may help show more clearly the normative issues involved. In addition to outlining a unified measurement framework, we present two empirical illustrations—one using data on child malnutrition in Vietnam, and the other using data on adult health in Canada.

II. SOME THEORY

Assume we have a scalar measure of health that is increasing in good health. If we rank individuals by their health, beginning with the least healthy, and graph on the x-axis the cumulative proportion of individuals ranked by health and on the y-axis the cumulative proportion of health, we obtain the Lorenz curve for health. Denote this by L_H. Twice the area between the diagonal (or line of equality) and the Lorenz curve equals the Gini coefficient, G, our measure of overall or "pure" health inequality.

Suppose we also have a scalar measure of SES that is increasing in SES. If we rank individuals by their SES, beginning with the most disadvantaged, and graph on the x-axis the cumulative proportion of individuals ranked by SES and on the y-axis the cumulative proportion of health, we obtain the concentration curve for health. Denote this by L_S. Twice the area between the diagonal (or line of equality) and the concentration curve equals the concentration index, C [17, 18], our measure of socioeconomic health inequality.

The concentration curve cannot lie below the Lorenz curve [19], since for any proportion of the population, p, the Lorenz curve reflects the health of the least healthy $100p$ percent of the population, while the concentration curve reflects the health of the most disadvantaged $100p$ percent of the population. If the rankings of individuals in the two distributions are the same, L_H and L_S will coincide, and G and C will be equal. But the rankings are unlikely to coincide. Some fairly disadvantaged people may enjoy better health than some much less disadvantaged people. Any difference in rankings between the two distributions will result in L_H lying below L_S, and G exceeding C.

We can derive expressions linking G and C that reinforce this point. The Gini coefficient can be written [20] as:

(1) $$G = \frac{2}{\mu}\operatorname{cov}(h, r_h),$$

where μ is mean health, h is health, and r_h is the person's fractional rank in the health distribution. A similar expression can be written down for C, where r_h is replaced by r_s (the person's fractional rank in the SES distribution). The ratio G/C can thus be written [21, 22]:

(2) $$\frac{G}{C} = \frac{\operatorname{cov}(h, r_h)}{\operatorname{cov}(h, r_s)},$$

which equals 1 if the rankings in the health and SES distributions coincide. We can also write down an expression for the difference between G and C. The Gini coefficient can be written [18]:

(3) $$G = \frac{2}{n \cdot \mu}\sum_{i=1}^{n} h_i r_{hi} - 1 ,$$

and a similar expression can be written down for C, replacing r_h by r_s. Thus we can write:

(4) $$\begin{aligned} G - C &= \frac{2}{n\mu}\sum_i h_i r_{hi} - 1 - \frac{2}{n\mu}\sum_i h_i R_{si} + 1 \\ &= \frac{2}{n\mu}\sum_i h_i \Delta R_i \\ &= \frac{2}{\mu}\operatorname{cov}(h, \Delta R), \end{aligned}$$

where $\Delta R_i = R_{hi} - R_{si}$ is the difference between the two fractional rank variables, which has a zero mean. G and C will be equal if the rankings in the two distributions coincide (i.e. $\Delta R_i = 0$ for all i), or if health and rank difference do not covary. Alternatively, we can think of G as being made up of two parts:

(5) $$G = C + R,$$

where $R = (2/\mu)\operatorname{cov}(h, \Delta R)$ is twice the area between the concentration curve and the Lorenz curve. R captures the change in the ranking in the "move" from the health distribution to the SES distribution, and is non-negative given the above.

The above assumes we have a continuous measure of SES. A similar decomposition [23] is available for the case where SES is measured by means of a categorical variable, such as the household's income class, or the grade of education completed by the household head. Suppose we have n SES groups, and we know the health of each individual. In this case, the decomposition becomes

(6) $$G = C_B + C_W + R_G .$$

In the first term, we attribute to everyone in the *k*th SES group the mean health of that group, μ_k, and line groups up in ascending order of the SES variable. Denote the resultant concentration curve by L_B, and the corresponding concentration index by C_B. The latter is known as the between-group concentration index. In the second term we keep the sample ranked by SES, but rerank individuals within each SES group by their health, beginning with the least healthy. Denote the resultant concentration curve by L_R and the resultant concentration index by C_R. C_W in eqn (6) is equal to twice the area between L_B and L_R, or equivalently the difference between C_R and C_B. This can be thought of as corresponding to within-group inequality. In the final term we rerank further, moving to the health ordering, thereby allowing, for example, for the possibility that the least healthy in SES group 2 may be less healthy than the most healthy in SES group 1. Doing this we obtain the Lorenz curve, L_H, and the Gini coefficient. The term R_G in eqn (6) is twice the area between L_H and the concentration curve L_R, or equivalently the difference between the Gini coefficient, G, and the concentration index C_R.

Three points are worth making. First, when health inequality at the grouped level is measured by the Gini coefficient, one cannot partition inequality into just within and between group inequality. There is a third term reflecting the fact that people's rankings are different in the health and SES distributions. Second, as the number of groups increases, C_W decreases, reaching zero in the limit, and C_B increases, becoming equal to C_R in the limit, where one is left with eqn (5). Third, in some grouped-data applications one will not know the health of individuals—only the SES group means, and the group sizes. In this case, the only measurable part of socioeconomic health inequality is C_B.

III. **EMPIRICAL ILLUSTRATIONS**

Inequality in adult health utility in Canada

Our data for Canada are taken from the 1994 National Population Health Survey (NPHS). We measure adults' health using the McMaster Health Utility Index (HUI) [24, 25]. This provides a description of an individual's overall functional health, based on eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. It assigns a single numerical value, between zero and one, for all possible combinations of levels of these eight self-reported health attributes, a score of one indicating perfect health. The HUI also embodies the views of society concerning health status, inasmuch as preferences about various health states are elicited from a representative sample of individuals. Income was defined in the NPHS as total annual household income before taxes and deductions, and respondents had to chose between one of nine income classes, ranging from less than Can$ 10,000 to over Can$ 80,000. Excluding cases with missing values and children under 12, our sample has 16 633 observations. Sampling weights were used throughout what follows.

Table 1 presents the results of the decomposition analysis using eqn (6). About 25% of adult health inequality (G) is generated by between-income group inequality (C_B), only 11% by within-group inequality (C_W). The remaining part is due to the overlapping term (R_G).

Table 1: Health inequality decompositions on grouped data, Canada and Vietnam

	Canada	Vietnam	
	Income groups	Consumption deciles	Consumption percentiles
C_B	0.017	0.158	0.160
C_W	0.006	0.070	0.011
C_R	0.024	0.227	0.171
R_G	0.042	0.460	0.516
G	0.065	0.687	0.687

Inequalities in child malnutrition in Vietnam

Our data for Vietnam are from the 1998 Vietnam Living Standards Survey (VLSS). We measure malnutrition by the child's height-for-age percentile score (HAP) in a hypothetical population of well-nourished children assembled by the US National Center for Health Statistics (NCHS). Thus a score of 50 means the child in question is at the median height-for-age in the well-nourished reference population. This indicator conveys information on the depth of malnutrition rather than simply whether or not a child was malnourished. Furthermore, it is increasing in nutrition, and is non-negative. Only children under the age of ten are included in our sample (N=5214). We used sample weights throughout. The (weighted) sample mean of HAP is 14.59, indicating that the average Vietnamese child in 1998 was well below the median in the well-nourished reference population.

The value of G for these data is 0.687. To compute C and R, we rank children by per capita household consumption. This is a better measure of living standards than income or expenditure, since it captures what households consume whether or not they purchase it or produce it themselves, and whether they finance it through current, future or past income. We obtain a value of C equal to 0.164, and hence a value of R equal to 0.523. Thus, as in the Canadian data, one quarter of overall health inequality is accounted for by socioeconomic inequality. We can illustrate the effects of increasing the number of groups in eqn (6) by first dividing the sample into per capita consumption deciles, and then into percentiles. The results in Table 1 confirm that as the number of groups rises, C_B gets closer to C, C_W shrinks, and R_G rises. Interestingly, even with deciles, the contribution to overall health inequality of inequality within consumption groups (C_W) is relatively unimportant compared to between-group inequality and reranking.

IV. **CONCLUSIONS**

To summarize, overall health inequality (G) can be thought of as being made up of two parts: socioeconomic inequality (C) and a component capturing the closeness of the rankings in the health and SES distributions (R). In the grouped-data case, there is an additional term (C_W in eqn (6)), reflecting health inequalities within socioeconomic groups. In both empirical illustrations presented, socioeconomic inequality accounted for about a quarter of overall health inequality. It would, of course, be unwise to generalize from two sets of results. However, the two applications are quite different, in terms of the age of the individuals concerned, the health status indicator used, and level of economic development of the country studied. They must therefore cast some doubt on Smith's [26] claim that "even if the social class gradient was magically eliminated, dispersion in health outcomes in the population would remain very much the same" (p.164). Our results also suggest that socioeconomic inequalities may well comprise only a minority of overall health inequality. In the ethical debate over which type of inequality researchers ought to be measuring, there seems to be broad agreement that socioeconomic inequalities are indeed inequitable and unjust. The issue at stake is whether these are the *only* inequalities about which policymakers are—or ought to be—concerned. Insofar as they can be generalized, the results here suggest this debate is a nontrivial one—health inequalities reflect socioeconomic inequalities to a large degree, but other health inequalities clearly exist. The challenge is to be clearer about which of these—if any—might also be considered unjust.

References

1. Le Grand, J., Inequalities in health: some international comparisons. *European Economic Review*, 1987. **31**: p. 182-91.

2. Gakidou, E., C. Murray, and J. Frenk, Defining and measuring health inequality. *Bulletin of the World Health Organization*, 2000. **78**(1): p. 42-54.

3. Wolfson, M. and G. Rowe, On measuring inequalities in health. *Bull World Health Organ*, 2001. **79**(6): p. 553-60.

4. Kunst, A.E., J.J. Geurts, and J. van den Berg, International variation in socioeconomic inequalities in self reported health. *J Epidemiol Community Health*, 1995. **49**(2): p. 117-23.

5. Kunst, A.E. and J.P. Mackenbach, Size of mortality differences associated with educational level in nine industrialized countries. *American Journal of Public Health*, 1994. **84**: p. 932-937.

6. Pamuk, E., Social class inequality in mortality from 1921-1972 in England and Wales. *Population Studies*, 1985. **39**: p. 17-31.

7. Pamuk, E., Social class inequality in infant mortality in England and Wales from 1921 to 1980. *European Journal of Population*, 1988. **4**: p. 1-21.

8. Schalick, L.M., W.C. Hadden, E. Pamuk, V. Navarro, and G. Pappas, The widening gap in death rates among income groups in the United States from 1967 to 1986. *International Journal of Health Services*, 2000. **30**(1): p. 13-26.

9. Vagero, D. and R. Erikson, Socioeconomic inequalities in morbidity and mortality in western Europe. *The Lancet*, 1997. **350**(August 16): p. 516-517.

10. Van Doorslaer, E., A. Wagstaff, H. Bleichrodt, S. Caonge, U.-G. Gerdtham, M. Gerfin, J. Geurts, et al., Income-related inequalities in health: Some international comparisons. *Journal of Health Economics*, 1997. **16**: p. 93-112.

11. Alleyne, G.A.O., J. Casas, and C. Castillo-Salgado, Equality, equity: why bother? *Bulletin of the World Health Organization*, 2000. **78**(1): p. 76-77.

12. Almeida, C., P. Braveman, M.R. Gold, C.L. Szwarcwald, J.M. Ribeiro, A. Miglionico, J.S. Millar, et al., Methodological concerns and recommendations on policy consequences of the World Health Report 2000. *The Lancet*, 2001. **357**(9268).

13. Braveman, P., B. Starfield, and H. Geiger, World Health Report 2000: how it removes equity from the agenda for public health monitoring and policy. *British Medical Journal*, 2001. **323**: p. 678-680.

14. Evans, T., M. Whitehead, F. Diderichsen, A. Bhuiya, and M. Wirth, Introduction, in *Challenging Inequities in Health: From Ethics to Action*, T. Evans, et al., Editors. 2001, Oxford University Press: Oxford.

15. Wagstaff, A., Economics, health and development: some ethical dilemmas facing the World Bank and the international community. *Journal of Medical Ethics*, 2001. **27**(4): p. 262-267.

16. Whitehead, M., The concepts and principles of equity and health. *International Journal of Health Services*, 1992. **22**(3): p. 429-445.

17. Wagstaff, A., P. Paci, and E. van Doorslaer, On the measurement of inequalities in health. *Social Science and Medicine*, 1991. **33**: p. 545-557.

18. Kakwani, N.C., A. Wagstaff, and E. Van Doorlsaer, Socioeconomic inequalities in health: Measurement, computation and statistical inference. *Journal of Econometrics*, 1997. **77**(1): p. 87-104.

19. Lambert, P., *The distribution and redistribution of income: A mathematical analysis*. 2nd ed. 1993, Manchester: Manchester University Press.

20. Lerman, R.I. and S. Yitzhaki, A note on the calculation and interpretation of the Gini index. *Economics Letters*, 1984. **15**: p. 363-368.

21. Kakwani, N.C., *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. 1980, New York: Oxford University Press.

22. Van Doorslaer, E. and A.M. Jones, *Inequalities in self-reported health: validation of a new approach to measurement*. 2001, Rotterdam: ECuity II Project Working Paper, Department of Health Policy and Management, Erasmus University, WP#2.

23. Lambert, P.J. and J.R. Aronson, Inequality Decomposition Analysis and the Gini Coefficient Revisited. *Economic Journal*, 1993. **103**(420): p. 1221-27.

24. Torrance, G.W., W. Furlong, D. Feeny, and M. Boyle, Multi-attribute preference functions: Health Utilities Index. *Pharmacoeconomics*, 1995. **7**(6): p. 503-520.

25. Feeny, D., W. Furlong, M. Boyle, and G.W. Torrance, Multiattribute Utility Function for a Comprehensive Health Status Classification System. *Medical Care*, 1996. **34**(7): p. 702-722.

26. Smith, J., Healthy bodies and thick wallets: the dual relation between health and socioeconomic status. *Journal of Economic Perspectives*, 1999. **13**: p. 145-66.