ELSEVIER

Journal of Econometrics 112 (2003) 207–223

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam

Adam Wagstaff[a,b,*], Eddy van Doorslaer[c], Naoko Watanabe[a]

[a] *The World Bank, 1818 H St. NW, Washington, DC, 20433, USA*
[b] *University of Sussex, Brighton, BN1 6HG, UK*
[c] *Erasmus University, 3000 DR Rotterdam, The Netherlands*

Abstract

Inequalities across the income distribution in a variable y can be decomposed into their causes, and changes in inequality in y can be decomposed into the effects of changes in the means and inequalities in the determinants of y, and changes in the effects of the determinants of y. Inequalities in height-for-age in Vietnam in 1993 and 1998 are largely accounted for by inequalities in consumption and in unobserved commune-level influences. Rising inequalities are largely accounted for by increases in average consumption and its protective effect, and rising inequality and general improvements at the commune level.

JEL classification: D30; I12

Keywords: Inequality; Decomposition; Malnutrition; Vietnam

1. Introduction

The large inequalities that exist in the health sector—between the poor and better-off—continue to be a cause for concern, in both the industrialized and the developing worlds. These inequalities are manifest in health outcomes (see, e.g. Van Doorslaer et al., 1997; Gwatkin et al., 2000; Wagstaff, 2000), the utilization of health services (see, e.g. Gwatkin et al., 2000), and in the benefits received from public expenditures on health services (see, e.g. Castro-Leal et al., 1999, 2000; Sahn and Younger, 2000).

In this paper, we present and apply some decomposition methods relevant to addressing three types of question. The first concerns the causes of health sector inequalities

* Corresponding author. Tel.: +1-202-473–0566; fax: +1-202-522-1153.
E-mail address: awagstaff@worldbank.org (A. Wagstaff).

0304-4076/02/$ - see front matter
PII: S0304-4076(02)00161-6

at a point in time. These inequalities stem from inequalities in the determinants of the variable of interest. For example, inequality in health sector subsidies presumably reflects inequalities in determinants of health service utilization (e.g., the quality of local health facilities, access to them, opportunity costs, etc.) and inequalities in the per unit subsidy (e.g. because of inequalities in liability for user fees). The issue arises: what is the relative contribution of each of these various inequalities in explaining subsidy inequalities? The second type of question concerns differences and changes in health sector inequalities. Countries vary substantially in the degree of inequality in different health sector outcomes (see, e.g. Gwatkin et al., 2000), and there is evidence that these inequalities have changed over time (see, e.g. Schalick et al., 2000; Victora et al., 2000). The obvious question is why these differences exist and why these changes have occurred. The third type of question in which we are interested concerns the impacts of policies and programs. The fact that inequalities appear to have widened over time in some countries does not mean necessarily that policies have been ineffective, let alone that they have *caused* the growth of inequality. The decomposition we present below can be useful in situations like this where one wants to separate out the effects on inequality of various changes, including the effects associated with programs that—inadvertently or otherwise—have effects on health sector inequalities.

In addition to presenting methods for unraveling the causes of health inequalities, we illustrate their use by analyzing the causes of levels of and changes in inequalities in child malnutrition in Vietnam over the period 1993–98. Whilst its child mortality figures are low by the standards of East Asia, Vietnam has a relatively high incidence of child malnutrition—albeit one that is falling (World Bank et al., 2001). By contrast, malnutrition inequalities were fairly small in Vietnam in 1993 by international standards (Wagstaff and Watanabe, 2000), but they have been rising (World Bank et al., 2001). The two empirical questions we seek to address, therefore, are: Why do inequalities in child malnutrition exist in Vietnam? And why did inequality in child malnutrition rise between 1993 and 1998?

The plan of the paper is as follows. In Section 2, we present the methods for decomposing the causes of health sector inequalities, focusing initially on levels and subsequently analyzing changes in inequality. In Section 3, we outline the empirical model and data we use to decompose the causes of levels of and changes in malnutrition inequalities in Vietnam. Section 4 presents and discusses our decomposition results, and Section 5 contains our conclusions.

2. Decomposing health sector inequalities: methods

2.1. Measuring health sector inequalities

We denote by y the variable in whose distribution by socioeconomic status (SES), we are interested.[1] The concentration curve, labeled L in Fig. 1, plots the cumulative

[1] The approach developed here could be used for the case where one wants to look at *pure* health inequality, in which case R would be the rank in the health distribution. The issue of which approach is more appropriate, and which measure of SES to use in the approach adopted here, are ethical ones and beyond the scope of this paper.

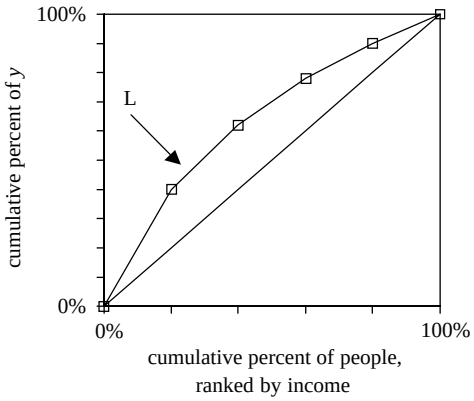

Fig. 1. Concentration curve.

proportion of y (on the vertical axis) against the cumulative proportion of the sample (on the horizontal axis), ranked by income (or some other measure of SES), beginning with the most disadvantaged person. If L lies above the diagonal, y is typically larger amongst the worse-off. The further L lies from the diagonal, the greater the degree of inequality in y across the income distribution. The concentration index, denoted below by C, is defined as twice the area between L and the diagonal (Wagstaff et al., 1991; Kakwani et al., 1997). C can be written in various ways, one (Kakwani et al., 1997) being

$$C = \frac{2}{n\mu}\sum_{i=1}^{n} y_i R_i - 1, \tag{1}$$

where μ is the mean of y, R_i is the fractional rank of the ith person in the income distribution. C, like the Gini coefficient, is a measure of *relative* inequality, so that a doubling of everyone's health leaves C unchanged. C takes a value of zero when L coincides with the diagonal, and is negative (positive) when L lies above (below) the diagonal.[2] In the case where y is a "bad"—like ill health or malnutrition—inequalities to the disadvantage of the poor (higher rates amongst the poor) push L above the diagonal and C below zero.

2.2. Decomposing health sector inequalities

Suppose, we have a linear regression model linking our variable of interest, y, to a set of k determinants, x_k:

$$y_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i, \tag{2}$$

[2] C could be zero if L crosses the diagonal. This does not happen in our empirical illustration, but even if it did, C still provides a measure of the extent to which health is, on balance, concentrated amongst the poor (or better-off).

where the β_k are coefficients and ε_i is an error term. We assume that everyone in the selected sample or subsample—irrespective of their income—faces the same coefficient vector, β_k. Interpersonal variations in y are thus assumed to derive from systematic variations across income groups in the determinants of y, i.e., the x_k. We have the following result, which owes much to Rao's (1969) theorem in the income inequality literature (Podder, 1993), and which is proved in the appendix:

Result 1: Given the relationship between y_i and x_{ik} in Eq. (2), the concentration index for y, C, can be written as

$$C = \sum_k (\beta_k \bar{x}_k / \mu) C_k + GC_\varepsilon / \mu, \tag{3}$$

where μ is the mean of y, $\bar{x}_k$ is the mean of x_k, and C_k is the concentration index for x_k (defined analogously to C). In the last term (which can be computed as a residual), GC_ε is a generalized concentration index for ε_i, defined as

$$GC_\varepsilon = \frac{2}{n} \sum_{i=1}^{n} \varepsilon_i R_i, \tag{4}$$

which is analogous to the Gini coefficient corresponding to the generalized Lorenz curve (Shorrocks, 1983). Eq. (2) shows that C can be thought of as being made up of two components. The first is the deterministic component, equal to a weighted sum of the concentration indices of the k regressors, where the weight or "share" for x_k, is simply the elasticity of y with respect to x_k (evaluated at the sample mean). The second is a residual component, captured by the last term—this reflects the inequality in health that cannot be explained by systematic variation across income groups in the x_k.

2.3. Decomposing changes in health sector inequalities

The most general approach to unraveling the causes of changes in inequalities would be to allow for the possibility that all the components of the decomposition in Eq. (3) have changed and simply to take the difference of Eq. (3):

$$\Delta C = \sum_k (\beta_{kt} \bar{x}_{kt} / \mu_t) C_{kt} - \sum_k (\beta_{kt-1} \bar{x}_{kt-1} / \mu_{t-1}) C_{kt-1} + \Delta(GC_{\varepsilon t} / \mu_t). \tag{5}$$

This approach is somewhat uninformative. One might, for example, want to know how far changes in inequality in health were attributable to changes in *inequalities* in the determinants of health rather than to changes in the other influences on health inequality. Furthermore, some changes (for example, changes in the mean of x_k) might be offset by other changes (for example, changes in the extent of inequality in x_k).

A slightly more illuminating approach would be to apply an Oaxaca-type decomposition (Oaxaca, 1973) to Eq. (3). If we denote by η_{kt} the elasticity of y with respect to x_k at time t, and apply Oaxaca's method, we get

$$\Delta C = \sum_k \eta_{kt} (C_{kt} - C_{kt-1}) + \sum_k C_{kt-1} (\eta_{kt} - \eta_{kt-1}) + \Delta(GC_{\varepsilon t} / \mu_t), \tag{6}$$

with the alternative being:

$$\Delta C = \sum_k \eta_{kt-1}(C_{kt} - C_{kt-1}) + \sum_k C_k(\eta_{kt} - \eta_{kt-1}) + \Delta(GC_{\varepsilon t}/\mu_t). \tag{7}$$

This approach allows us to see—for each x_k in turn or for all x_k combined—the extent to which changes in health inequalities are due to changes in inequality in the determinants of health, rather than to changes in their elasticities.

Whilst more illuminating than Eq. (5), Eqs. (6) and (7) still conceal a lot. One cannot disentangle changes going on within the elasticity η_{kt}. It may be that the change in C owes far more to changes in β_k than to changes in the mean of x_k, or vice versa. Indeed, the components of η_{kt} may change in different directions, possibly having exactly offsetting effects. This would be especially worrisome if one's interest lay in the effects of a program thought to have influenced only one component of η_k (e.g. one of the β_k) at a time when one of the other components of η_k (e.g. the mean of x_k) was also changing.

A third possibility is to take the total differential of Eq. (3), allowing for changes *in turn* in each of the following: α, the β_k, the $\bar{x}_k$, and the C_k. We allow these changes to alter C directly and indirectly through μ. Doing this, we obtain the following result, which is also proved in the appendix:

Result 2: The change in C, ΔC, can be approximated by

$$\begin{aligned} \mathrm{d}C &= \frac{\mathrm{d}C}{\mathrm{d}\alpha}\mathrm{d}\alpha + \sum_k \frac{\mathrm{d}C}{\mathrm{d}\beta_k}\mathrm{d}\beta_k + \sum_k \frac{\mathrm{d}C}{\mathrm{d}\bar{x}_k}\mathrm{d}\bar{x}_k + \sum_k \frac{\mathrm{d}C}{\mathrm{d}C_k}\mathrm{d}C_k + \mathrm{d}\frac{GC_\varepsilon}{\mu} \\ &= -\frac{C}{\mu}\mathrm{d}\alpha + \sum_k \frac{\bar{x}_k}{\mu}(C_k - C)\,\mathrm{d}\beta_k + \sum_k \frac{\beta_k}{\mu}(C_k - C)\,\mathrm{d}\bar{x}_k \\ &\quad + \sum_k \frac{\beta_k \bar{x}_k}{\mu}\mathrm{d}C_k + \mathrm{d}\frac{GC_\varepsilon}{\mu}. \end{aligned} \tag{8}$$

From Eq. (8), it emerges that although α does not enter the decomposition for *levels*, i.e. Eq. (3), *changes* in α do produce *changes* in C. Take the case where y is a measure of good health, and has a positive mean and a positive C (good health is concentrated amongst the better off). In this case, $\mathrm{d}C/\mathrm{d}\alpha < 0$. A rise in α ($\mathrm{d}\alpha > 0$) amounts to an equal increase in everyone's health, and (relative) inequality in health falls, in just the same way as an equal increase in income for everyone reduces relative income inequality (Podder, 1993). The reduction in inequality is larger, the larger is C and the smaller is μ. The case we consider in the empirical analysis is somewhat different–we look at inequality in *ill health*, our y-variable being an increasing function of child malnutrition. We have a positive mean (average malnutrition is positive) and a negative value of C (levels of malnutrition are higher amongst the poor). In this case, $\mathrm{d}C/\mathrm{d}\alpha > 0$. Suppose there is a *reduction* in α ($\mathrm{d}\alpha < 0$). This amounts to an equal reduction in everyone's level of malnutrition, and the first term on the RHS of (8) is negative—i.e. C becomes more negative and inequality worsens. This is the mirror image of the case where y is a measure of good health—there a given increase in health represents a bigger proportional increase for poor people, while in the case

where y is a measure of ill health a given decrease in ill health represents a bigger proportional reduction for better-off people.

The second and third terms on the RHS of Eq. (8) show that the sign of the effect on C of a change in β_k, or of a change in $\bar{x}_k$, depends on whether x_k is more or less unequally distributed than y. These results reflect two channels of influence—the direct effect of the change in β_k (or $\bar{x}_k$) on C, and the indirect effect operating through μ. If the variable in question is equally distributed ($C_k = 0$), the direct effect is zero. Take the case where y is a measure of ill health, with positive mean and negative concentration index. Assume x_k also has a positive mean, and has a dampening effect on ill health ($\beta_k < 0$). Consider the effect of a rise in $\bar{x}_k$, holding constant the degree of inequality in x_k—i.e. an equiproportionate rise in x_k. The direct effect of this change is a reduction (in numerical value) in the size of C, the reason being that the existing inequality in x_k generates more inequality in y to the disadvantage of the poor. But there is an additional effect, operating through the mean. The rise in $\bar{x}_k$ lowers average ill health, which, holding all else constant, makes for more *relative* inequality in y (i.e. makes for a more negative C). In this case, the two effects reinforce one another. This will not always be so. Take the case where y is a measure of good health with positive mean and concentration index. Assume x_k contributes to good health and is unequally distributed to the advantage of the better-off ($C_k > 0$). The direct effect of an increase in $\bar{x}_k$ is to raise inequality (C becomes more positive), since the existing inequality in x_k generates more inequality in y. But the rise in $\bar{x}_k$ raises the mean of y which, all else constant, lowers inequality in y. Whether the net effect of the rise in $\bar{x}_k$ is to raise or lower inequality in y depends on whether x_k is more unequally distributed than y itself (i.e. whether C_k-C is positive or negative). Similar remarks apply to the case of a change in β_k.

Finally, and more straightforwardly, an increase (decrease) in inequality in x_k (i.e. C_k) will increase (reduce), the degree of inequality in y. The impact is an increasing function of β_k and $\bar{x}_k$, and a decreasing function of μ. So, for example, if y is increasing in ill health, $C < 0$, and x_k reduces ill health, a rise in inequality in x_k will make for a reduction (in numerical size) in C (i.e. C becomes more negative).

3. Model, data and variable definitions

Our data are from the 1993 and 1998 Vietnam Living Standards Surveys (VLSS). We focus on inequalities in stunting (low height-for-age), which we measure using the negative of the child's height-for-age z-score, with the US National Center for Health Statistics (NCHS) data providing the reference.[3] We have two reasons for

[3] Bhargava (2000) has shown that distributions of anthropometric scores can be sensitive to the reference standards chosen. We therefore recomputed our 1998 results using a UK reference scale (Freeman et al., 1995). Although the mean z-score changed slightly, the level of inequality was virtually identical to that obtained using the US reference data—C was −0.110 rather than −0.099. The regression results were very similar too, as inevitably were the decompositions. For example, the UK-based figures for the contributions of household consumption and commune fixed effects for 1998 in Table 2 were −0.055 and −0.047, respectively. Tables showing these results are available upon request from the authors.

favoring the z-score over a binary variable indicating whether or not the child in question was stunted (i.e. two standard deviations or more below the NCHS mean). First, it conveys information on the depth of malnutrition rather than simply whether or not a child was malnourished. Second, it is amenable to linear regression analysis (Lavy et al., 1996; Thomas et al., 1996; Ponce et al., 1998). This is essential to our decomposition method.[4] The use of the z-score in the analysis of inequality does require that we accept the value judgment that "taller is always better", but this seems relatively innocuous.[5] We use the *negative* of the z-score to make our malnutrition variable easier to interpret—it is increasing in malnutrition, and in both years has a positive mean.[6] Like Ponce et al. (1998), we confine our attention to children under the age of ten, there being evidence that over the age of nine genetic factors start to seriously constrain growth (Martorell and Habicht, 1986; Kostermans, 1994). We have 5067 children under the age of ten in 1993, and 4796 in 1998.[7]

Mean values of (the negative of) the height-for-age variable in the 1993 and 1998 samples were 2.036 and 1.608, respectively, indicating an appreciable improvement in average nutritional status between 1993 and 1998.[8] To compute the concentration indices, we ranked children by per capita household consumption in 1998 prices.[9] Our concentration indices for 1993 and 1998 were −0.077 and −0.099, respectively, indicating a concentration of malnutrition amongst the poor in each year, and an appreciable worsening in inequality between 1993 and 1998.

To explain variations in height-for-age, we adopt a standard household production-type anthropometric regression framework (Lavy et al., 1996; Thomas et al., 1996; Ponce et al., 1998; Alderman, 2000), in which the (negative of the) child's height-for-age z-score is specified to be a linear function of a vector of child-level variables, X_1, a vector of household-level variables, X_2, and a commune fixed effect at the level of the

[4] We are, in fact, able to introduce some nonlinearities—we have a squared term for one variable, and enter another in logarithmic form.

[5] The same cannot be said, of course, of the analogous value judgement for weight-for-age. For both stunting and underweight, one could apply the methods outlined above with a malnutrition deficit variable, defined as the gap between the actual z-score and the threshold used to define malnourished children (minus two standard deviations below the mean of the original z-score). We did this in the case of stunting, and obtained broadly similar results to those presented here.

[6] Our method would break down if the mean were zero, which it would be if one used anthropometric z-scores on the NCHS children. This group of children is not, however, a representative sample of the US children, the NCHS data on children age 0–36 months of age having been collected over a long period of time from a population of middle-class, white, bottle-fed Americans. The problem of a zero mean even on these children would not, of course, arise if one used a malnutrition deficit variable rather than the z-score.

[7] This excludes children with missing information on any of the variables included in the regression model.

[8] We employed sample weights in computing the means for the 1998 survey, since the 1998 LSMS—unlike the 1993 LSMS—is not representative without weights.

[9] We also employed sample weights in generating the ranks in the consumption distribution for the 1998 data (they were unnecessary for the 1993 data). For both years, concentration indices were computed using the convenient covariance method (Jenkins, 1988), using sample weights in the case of 1998.

child's commune.[10] We interpret our estimating equation as a reduced-form demand equation—rather than a production function—and estimate it using OLS.[11,12]

In the X_1 vector we include the child's age, age squared, and gender. Age entered nonlinearly allows for the fact that Vietnamese children start off in the early months being close to the US reference in terms of height-for-age, but gradually fall behind (Ponce et al., 1998). Including gender allows for the possibility of systematic differences in stunting by gender. The vector X_2 includes household living standards, measured by the log of per capita household consumption,[13] dummy variables indicating whether or not the child's household had satisfactory drinking water and sanitation,[14] and the years of schooling of the household head and the mother. These latter variables are fairly standard covariates in anthropometric regressions (Alderman, 2000), and we expect all to reduce malnutrition.

4. Results

4.1. Regression results

Table 1 shows the regression results for 1993 and 1998. The (joint) hypothesis of time-invariant slope coefficients is rejected at just over the 5% level, and the hypothesis

[10] We had hoped to be able to isolate the contribution of inequalities in different community-level variables, but this proved impossible. Commune data were collected in both years, but only for communes in rural areas (and, in 1998, small towns). Furthermore, the 1993 community survey was rather limited in scope (Ponce et al., 1998). On top of this, health facility data were collected only for 1998, and even then were collected only for a limited set of facilities, and only in rural areas.

[11] We corrected standard errors both for heteroscedasticity and the effects of geographic clustering at the commune level (Deaton, 1997). Sample weights were used in the model estimation for 1998 (they were not required for the 1993 data).

[12] The issue arises of whether household consumption—one of our X_2 variables—is endogenous and therefore whether OLS is appropriate. There are various possible reasons why consumption might be endogenous, the most obvious being that mothers may base their work decisions in part on the health and nutritional status of their children, and that well-nourished children may be put to work (Ponce et al., 1998; Alderman, 2000). Consumption may also be subject to measurement error that is correlated with the error term ε. Suppose one were to use instrumental variables (IV) instead of OLS. Then one would need to replace the concentration index of consumption in Eq. (3) by the concentration index of *predicted* consumption from the first stage of the IV procedure. Furthermore, in order for the decomposition to hold, one would need to re-rank children by their predicted consumption in the computation of *all* concentration indices, including that of the malnutrition z-score. This would change the interpretation of the concentration index to be explained. In effect, one would be explaining inequalities purged of any simultaneity and measurement error. Arguably this might be a more interesting quantity to explain, but it is not *actual* inequality. Our results below are therefore based on OLS, but we acknowledge that this is not an open and shut case.

[13] The household consumption aggregates, taken from the 2000 version of the VLSS data files, were produced by VLSS staff using standard procedures, and in the case of the 1993 data were converted by VLSS staff in June 2000 to 1998 prices using region-specific price indices.

[14] We have defined the drinking water and sanitation variables along the lines proposed by UNICEF (Government of Vietnam, 2000). Safe drinking water was defined as: tap or standpipe; deep dug well with pump; hand-dug well; or rain water. Satisfactory sanitation was defined as: flush toilet or latrine. Both differ slightly from the definitions used by UNICEF, because the categories in the VLSS data are somewhat different from those used by UNICEF.

Table 1
Stunting regressions for 1993 and 1998

	1993		1998	
	Coefficient	*t*-Stat	Coefficient	*t*-Stat
Constant (mean commune fixed effect)	3.008	6.401	2.468	5.396
Child's age (in months)	0.038	14.163	0.039	11.630
Child's age squared	0.000	−13.514	0.000	−10.938
Child = male	0.086	2.567	0.143	3.743
Household consumption	−0.261	−4.042	−0.272	−4.652
Safe drinking water	−0.023	−0.284	−0.081	−1.420
Satisfactory sanitation	−0.128	−1.861	−0.050	−0.706
Years schooling household head	−0.005	−0.814	−0.003	−0.513
Years schooling mother	−0.012	−1.645	−0.001	−0.094
N	5067		4796	
R^2	0.188		0.247	
F for regression	38.76		35.21	
DF for F-test for regression	7, 4909		7, 4594	
F commune fixed effects = 0	2.91		2.98	
DF for F-test for commune fixed effects = 0	149, 4909		193, 4594	
p for F-test for commune fixed effects = 0	0.000		0.000	

t-test with 9701 degrees of freedom to test hypothesis of no change in mean commune fixed effect between 1993 and 1998: 45.61 ($p = 0.000$). Test is based on commune fixed effects obtained from regression on pooled sample. F-test with $(8, 9503)$ degrees of freedom to test joint hypothesis of no change in slope coefficients is $1.98(p = 0.051)$. Test based on interactions between X-variables and time dummy in pooled sample. Regressions undertaken using AREG routine in STATA with cluster option. Results for 1998 based on weighted data. F-test for commune fixed effects = 0 in 1998 computed with weights treated as analytic weights and without clustering on communes.

of time invariance in the commune fixed effects is decisively rejected. The mean of the commune fixed effects falls considerably between 1993 and 1998, and in each year the hypothesis of zero commune fixed effects is decisively rejected. Child's age has a significant inverted u-shaped relationship in both years (reaching its peak at around $6\frac{1}{2}$ years in 1998), with a slight strengthening of the relationship between 1993 and 1998. Boys are more prone to stunting than girls, and the gender gap—holding all else constant—apparently widened over the period 1993–98. Household consumption has a statistically significant negative effect on malnutrition in both years, but the effect was somewhat stronger in 1998 than 1993. Safe drinking water reduces malnutrition in both years, but the effect is stronger and closer to achieving statistical significance in 1998. Satisfactory sanitation also reduces malnutrition in both years, but the effect is smaller in 1998 and is insignificant in that year. Parents' education reduces malnutrition in both years, but the effect has fallen—dramatically so in the case of mother's education—and the larger impact of mother's education that is evident in 1993 is no longer evident in 1998.

4.2. Decomposition results

Table 2 shows the decompositions for the 2 years. The first two columns under the heading "contributions" make it clear that the bulk of inequality in malnutrition

Table 2
Inequality decompositions for 1993 and 1998, and change 1993–98

	Coefficients		Means		Elasticities		Concentration indices		Contributions to C		
	1993	1998	1993	1998	1993	1998	1993	1998	1993	1998	Change
Child's age (in months)	0.038	0.039	60.982	66.962	1.137	1.630	0.020	0.018	0.023	0.030	0.007
Child's age squared	0.000	0.000	4883.834	5616.139	−0.634	−0.880	0.030	0.028	−0.019	−0.025	−0.006
Child = male	0.086	0.143	0.514	0.506	0.022	0.045	0.003	0.014	0.000	0.001	0.001
Household consumption	−0.261	−0.272	7.300	7.611	−0.936	−1.288	0.038	0.040	−0.035	−0.052	−0.016
Safe drinking water	−0.023	−0.081	0.221	0.331	−0.003	−0.017	0.312	0.256	−0.001	−0.004	−0.003
Satisfactory sanitation	−0.128	−0.050	0.146	0.202	−0.009	−0.006	0.468	0.508	−0.004	−0.003	0.001
Years schooling household head	−0.005	−0.003	6.812	7.108	−0.017	−0.015	0.065	0.094	−0.001	−0.001	0.000
Years schooling mother	−0.012	−0.001	6.321	6.722	−0.037	−0.003	0.075	0.108	−0.003	0.000	0.003
Commune fixed effects			3.008	2.468	1.477	1.534	−0.024	−0.031	−0.035	−0.047	−0.012
Total									−0.075	−0.102	−0.027

Table 3
Oaxaca-type decomposition for change in inequality, 1993–98

	Eq. (6)		Eq. (7)		Total	
	$\Delta C\eta$	$\Delta \eta C$	$\Delta C\eta$	$\Delta \eta C$	Total	%
Child's age (in months)	−0.003	0.010	−0.002	0.009	0.007	−30
Child's age squared	0.001	−0.007	0.001	−0.007	−0.006	26
Child = male	0.000	0.000	0.000	0.000	0.001	−3
Household consumption	−0.003	−0.013	−0.002	−0.014	−0.016	74
Safe drinking water	0.001	−0.004	0.000	−0.004	−0.003	16
Satisfactory sanitation	0.000	0.001	0.000	0.002	0.001	−5
Years schooling household head	0.000	0.000	−0.001	0.000	0.000	1
Years schooling mother	0.000	0.003	−0.001	0.004	0.003	−11
Commune fixed effects	−0.011	−0.001	−0.010	−0.002	−0.012	55
"Residual"					0.005	−24
Total	−0.015	−0.012	−0.016	−0.012	−0.022	

in both 1993 and 1998 was caused by inequalities in household consumption and inequalities in the commune fixed effects, both disfavoring the poor. In the case of the commune fixed effects, the inference is that in both years poor children lived in communes that were likely to have characteristics that increased the likelihood of them being malnourished. The contributions from inequalities in age (higher income groups tend to have slightly older children) and age squared are evident, but the former is almost totally offset by the latter—the net effect of age in 1993 is equal to only 0.004 (i.e. 0.023–0.019). Inequalities in drinking water, sanitation, and parental schooling all disfavored the poor in both years, but their contributions to malnutrition inequalities were fairly small, accounting in total for only −0.009 points of a total of −0.075 in 1993, and only −0.008 of a total of −0.102 in 1998.

The column headed "change" in the last column—the empirical analog of Eq. (5)—indicates that the bulk of the deterioration in malnutrition inequality between 1993 and 1998 was due to changes in respect of household consumption and changes at the commune level. The net change in respect of inequalities in child's age was slightly pro-poor—a change of +0.001. Changes in water and sanitation were in opposite directions, with changes in respect of water actually making for more inequality in malnutrition. Changes in respect of education of the household head were negligible, while changes in respect of the mother's education tended to narrow malnutrition inequalities slightly. Even combined, however, these changes were small relative to the changes in respect of household consumption and the commune fixed effects.

What Eq. (5)—and its empirical counterpart in the final column of Table 2—does not enable us to see is how far these changes were due to changes in elasticities rather than changes in inequality. The Oaxaca decomposition results in Table 3 allows us to answer this question. For both household consumption and the commune fixed effects, changes in the elasticities and in inequality reinforce one another. In the former case, Table 3 suggests it is the changing elasticity—rather than rising consumption inequality—that accounts for the bulk of the rise in inequality associated with changes

Table 4
Total differential decomposition of change in inequality, 1993–98

	β's	Means of x's	CIs	GC_ε	Total
Child's age (in months)	0.003	0.011	−0.002		0.012
Child's age squared	0.003	−0.010	0.001		−0.006
Child = male	0.001	0.000	0.000		0.001
Household consumption	−0.005	−0.005	−0.002		−0.011
Safe drinking water	−0.002	0.000	0.000		−0.003
Satisfactory sanitation	0.003	−0.002	0.000		0.001
Years schooling household head	0.001	0.000	−0.001		0.000
Years schooling mother	0.005	0.000	−0.001		0.004
Commune fixed effects	0.000	−0.014	−0.010		−0.025
"Residual"				0.005	0.005
Total	0.010	−0.021	−0.016	0.005	−0.021
Column as % total	−47%	98%	74%	−25%	

in respect of consumption. In the case of the commune fixed effects, by contrast, changing inequalities appear in Table 3 to be more important than changing elasticities. Overall—taking the changes of all the determinants of malnutrition into account—the rise in inequality in malnutrition is roughly equally attributable to changing elasticities and changing inequalities in the determinants of malnutrition.

The total differential decomposition allows us to "drill down" still further and pinpoint the changes within the changes in the elasticities, and to explore the possibility that the change in the elasticity might be small or even zero because of offsetting changes within it—the coefficient might rise, for example, while the mean of the determinant could fall. It should be recalled that this decomposition is based on an approximation and is accurate only for small changes. The fact that there are discrepancies between the "Total" column in Table 4 and the "Change" column in Table 2 reminds us of this. These discrepancies are especially pronounced for household consumption and the commune fixed effects—unsurprising, given these are the major drivers of rising malnutrition inequalities.

The total differential decomposition in Table 4 provides important additional information to Table 2, by revealing that the relative contributions of the increased regression coefficient (in absolute size) and the increased mean consumption on the rise in the consumption elasticity of malnutrition were similar in magnitude. The total differential decomposition also uncovers the fairly large contribution to rising malnutrition inequality coming from the changes in the means of the commune fixed effects, which was actually larger in absolute size than the change attributable to widening inequalities in the commune fixed effects.

The total differential approach also helps us establish the magnitudes involved in cases of largely offsetting changes in the various components of the elasticity. The changes in respect of sanitation provide an example—albeit one that is not quantitatively very important. From Table 2, it is evident that a higher proportion of under-ten children lived in households with satisfactory sanitation in 1998 than in 1993. This

tended to worsen malnutrition inequalities (i.e. make C even more negative). But over the same period, the protective effect of sanitation seems to have fallen. This effect tended to *narrow* malnutrition inequalities. We know from the Oaxaca decomposition that these effects roughly canceled each other out: $\Delta\eta C$ is around 0.001 or 0.002, depending on whether we use Eq. (6) or Eq. (7). What we do not know is the actual magnitudes of their impacts on C. Table 4 tells us what these are, namely that the change in the regression coefficient made C rise by 0.003, while the increased coverage made C fall by 0.002.

Table 4 also gives us an estimate of the overall impacts on malnutrition inequalities of: (a) changes in regression coefficients, (b) changes in the means of the determinants of malnutrition, and (c) changes in the degree of inequality in the determinants of malnutrition. Whilst changes in the means and inequalities of the determinants of malnutrition have, on balance, tended to worsen inequalities in malnutrition, the opposite is true, on balance, of changes in the slope coefficients. There are, course, exceptions to these patterns—changes in the regression coefficients of consumption and drinking water, for example, have tended to make malnutrition inequalities worse rather than better. One take-home message from the bottom row of Table 4 is that without the inequality-reducing effects of the changes in the regression coefficients, inequality in height-for-age would have changed by -0.036, rather than by -0.021. Another is that changes in the degree of inequality in the determinants of malnutrition made C change by -0.016, whereas the actual value of C changed by -0.021. There is, in other words, more to rising inequalities in malnutrition than rising inequalities in its determinants.

5. Conclusions and discussion

Our main aim in this paper has been to present some decomposition methods to enable researchers to unravel the causes of health sector inequalities, and their change over time, or variations across countries. Inequalities are caused by inequalities in the determinants of the variable of interest, and our decomposition in Eq. (3) allows one to assess the relative importance of these different inequalities in generating inequalities in the variable of interest. Changes over time in inequality in the variable of interest can be due to changes in the degree of inequality in its determinants, or to changes in the means of the various determinants, or to changes in their impact on the variable of interest. The total differential decomposition in Eq. (8) allows one to disentangle these three possible causes of changing inequality. The decomposition also alerts us to a potential tradeoff between reducing relative inequality and improving the mean of the variable of interest—a tradeoff discussed by Contoyannis and Forster (1999a,b) and apparent in our empirical illustration. In the case we examined (inequalities in malnutrition), rising incomes were found to reduce malnutrition and hence reduce average malnutrition. But rising incomes—holding income inequality constant—increase relative inequality in malnutrition, directly by magnifying the inequality in malnutrition attributable to income inequality, and indirectly by reducing mean malnutrition. The decomposition is helpful in this regard, in that it allows one to see how much of the

rise in inequality in the variable of interest was associated with rising inequality in its determinants (changes which will not have improved the mean) and how much of the rise was due to changes in the means of the determinants or their impacts (changes which may have improved the mean).

Our empirical results, whilst intended primarily as an empirical illustration, and subject to the usual caveats regarding the causal interpretation of cross-sectional results, are of some interest in their own right. They suggest that inequalities in stunting amongst young children in Vietnam in both 1993 and 1998 were due largely to inequalities in household consumption and to inequalities in unobserved determinants at the commune level (poor children living in areas that are not conducive to good health). They also suggest that it was changes in these two factors that were largely responsible for the rise in inequality in malnutrition over the period 1993–98.

In the case of household consumption, rising inequality accounted for only a small part of the rise in malnutrition inequality. More important were the increases in average consumption and the protective effect of consumption on malnutrition. As far as commune-level factors are concerned, the picture appears to have been one of an improvement overall in the commune-level determinants of malnutrition and an increase in the inequality in their distribution. Both factors have made for more inequality in malnutrition, with our estimates suggesting that the rise in inequality in the commune-level determinants of malnutrition was slightly less important in terms of its impact on inequality in malnutrition. Overall, our results suggest that rising inequalities in stunting owed most to changes in the means of the determinants of malnutrition, with rising inequalities in the determinants of malnutrition being the next important factor. Comparatively little of the rise in inequality was attributable to changes in the impacts of the various determinants.

How plausible are our empirical results? Rising consumption, increased coverage of water and sanitation, and rising education levels have indeed all occurred in Vietnam during the 1990s, and it is well known that these factors reduce malnutrition. It is also known that inequality in consumption has risen somewhat (Glewwe et al., 2000) and that inequality in educational attainment has risen (Wagstaff and Nguyen, 2002). What is less clear is whether the changes attributable to changes in the regression coefficients and changes at the commune level are plausible. The former are in any case fairly small, and the largest single factor is the change brought about by the increased protective effect of household consumption. This seems fairly plausible. It is true that the price of drugs has fallen considerably in real terms over the period in question (World Bank et al., 2001), but several other changes seem likely to have increased (in absolute terms) the marginal impact of household consumption levels on nutritional outcomes. Examples in the health sector include rising user fees at public health facilities and the growth of the private health sector (World Bank et al., 2001). Changes in the market for foodstuffs have also resulted in a greater availability and variety of food, so that households with sufficient resources now have the opportunity to purchase quality foodstuffs throughout the year. It also seems likely that whilst, on balance, the commune-level determinants of malnutrition have improved, these improvements have not been spread equally across poor and better-off communes. Elsewhere (Wagstaff and Nguyen, 2002) one of us has noted that in some

respects the poor appear to be slipping backwards in respect of access to and utilization of public health facilities. Vaccination and antenatal visit coverage grew more slowly amongst the bottom three quintiles between 1993 and 1998, while the proportion of newborns delivered by skilled birth attendants actually *fell* between 1993 and 1998 in the bottom three quintiles. In short, whilst intended to be primarily illustrative, our results shed some light on the possible causes of rising inequalities in malnutrition in Vietnam.

Acknowledgements

Without wishing to implicate them in any way, we are grateful to the following for the helpful comments on an earlier version of the paper or research leading up to it: three anonymous referees; Anne Case, Angus Deaton, Christina Paxson and other participants at a seminar at Princeton; participants at the 2001 International Health Economics Association Meeting in York; Harold Alderman, Alok Bhargava, Deon Filmer, Berk Ozler, Martin Ravallion and Tom Van Ourti. The findings, interpretations and conclusions expressed in this paper are those of the authors, and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

Appendix A.

A.1. Proof of Result 1

Substitute Eq. (2) into Eq. (1) to obtain:

$$C = \frac{2}{n\mu}\left[\frac{\alpha}{2}n + \beta_1\sum_{i=1}^{n}x_{1i}R_i + \cdots + \beta_k\sum_{i=1}^{n}x_{ki}R_i + \sum_{i=1}^{n}\varepsilon_i R_i\right] - 1, \tag{A.1}$$

since the mean of R_i is one half. Using Eq. (1) to obtain an equation for the concentration index of x_k yields

$$\sum_i x_{ki}R_i = \frac{(C_k+1)}{2}n\bar{x}_k. \tag{A.2}$$

Substituting Eqs. (A.2) and (4) into Eq. (A.1) yields

$$C = \frac{2}{n\mu}\left[\frac{\alpha}{2}n + \beta_1\frac{(C_1+1)}{2}n\bar{x}_1 + \cdots + \beta_k\frac{(C_k+1)}{2}n\bar{x}_k + GC_\varepsilon\frac{n}{2}\right] - 1. \tag{A.3}$$

We obtain Eq. (3) by expanding terms and bearing in mind that

$$\mu = \alpha + \sum_k \beta_k\bar{x}_k.$$

A.2. Proof of Result 2

The derivatives in (8) can be obtained as (Podder, 1993)

$$\frac{dC}{d\alpha} = \frac{dC}{d\mu}\frac{d\mu}{d\alpha} = -\frac{1}{\mu}C,$$

$$\frac{dC}{d\beta_k} = \frac{\partial C}{\partial \beta_k} + \frac{dC}{d\mu}\frac{d\mu}{d\beta_k} = \frac{\bar{x}_k C_k}{\mu} - \frac{\bar{x}_k}{\mu}C = \frac{\bar{x}_k}{\mu}(C_k - C),$$

$$\frac{dC}{d\bar{x}_k} = \frac{\beta_k}{\mu}(C_k - C),$$

$$\frac{dC}{dC_k} = \frac{\beta_k \bar{x}_k}{\mu}$$

Substituting these derivatives into Eq. (8) yields the desired result.

References

Alderman, H., 2000. Anthropometry. In: Grosh, M., Glewwe, P. (Eds.), Designing Household Survey Questionnaires for Developing Countries. The World Bank, Washington, DC, pp. 251–272.

Bhargava, A., 2000. Modeling the effects of maternal nutritional status and socioeconomic variables on the anthropometric and psychological indicators of Kenyan infants from age 0–6 months. American Journal of Physical Anthropology 111 (1), 89–104.

Castro-Leal, F., Dayton, J., Demery, L., Mehra, K., 1999. Public social spending in Africa: do the poor benefit? World Bank Research Observer 14 (1), 49–72.

Castro-Leal, F., Dayton, J., Demery, L., Mehra, K., 2000. Public spending on health care in Africa: do the poor benefit? Bulletin of the World Health Organization 78 (1), 66–74.

Contoyannis, P., Forster, M., 1999a. Our healthier nation? Health Economics 8, 289–296.

Contoyannis, P., Forster, P., 1999b. The distribution of health: a theoretical framework. Journal of Health Economics 18 (5), 605–622.

Deaton, A., 1997. The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Johns Hopkins University Press, Baltimore, MD.

Freeman, J., et al., 1995. Cross-sectional stature and weight reference curves for the UK, 1990. Archives of Disease in Childhood 73, 17–24.

Glewwe, P., Gragnolati, M., Zaman, H., 2000. Who gained from Vietnam's boom in the 1990s? Policy Research Working Paper No. 2275. The World Bank, Washington, DC.

Government of Vietnam, 2000. Analysis of the Multiple Indicator Cluster Survey II (MICSII). Statistical Publishing House, Hanoi.

Gwatkin, D.R., Rutstein, S., Johnson, K., Pande, R., Wagstaff, A., 2000. Socioeconomic Differences in Health, Nutrition and Population. Health, Nutrition & Population Discussion Paper. The World Bank, Washington, DC.

Kakwani, N., Wagstaff, A., Van Doorslaer, E., 1997. Socioeconomic inequalities in health: measurement, computation and statistical inference. Journal of Econometrics 77 (1), 87–104.

Kostermans, K., 1994. Assessing the quality of anthropometric data. LSMS Working Paper No. 101. The World Bank, Washington, DC.

Lavy, V., Strauss, J., Thomas, D., de Vreyer, P., 1996. Quality of care, survival and health outcomes in Ghana. Journal of Health Economics 15, 333–357.

Martorell, R., Habicht, J.P., 1986. Growth in early childhood in developing countries. In: Falkner, F., Tanner, J.M. (Eds.), Human Growth: A Comprehensive Treatise. Plenum Press, New York.

Oaxaca, R.L., 1973. Male–female wage differentials in urban labor markets. International Economic Review 14, 693–709.

Podder, N., 1993. The disaggregation of the Gini coefficient by factor components and its applications to Australia. Review of Income and Wealth 39 (1), 51–61.

Ponce, N., Gertler, P., Glewwe, P., 1998. Will Vietnam grow out of malnutrition? In: Dollar, D., Glewwe, P., Litvack, J. (Eds.), Household Welfare and Vietnam's Transition. World Bank, Washington, DC, pp. 257–275.

Rao, V.M., 1969. Two decompositions of concentration ratio. Journal of Royal Statistical Society Series A 132 (3), 418–425.

Sahn, D.E., Younger, S.D., 2000. Expenditure incidence in Africa: microeconomic evidence. Fiscal Studies 21 (3), 329–348.

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

Shorrocks, A., 1983. Ranking income distributions. Economica 50, 1–17.

Thomas, D., Lavy, V., Strauss, D., 1996. Public policy and anthropometric outcomes in the Cote d'Ivoire. Journal of Public Economics 61, 155–192.

Van Doorslaer, E., Wagstaff, A., Bleichrodt, H., Caonge, S., Gerdtham, U.-G., et al., 1997. Income-related inequalities in health: some international comparisons. Journal of Health Economics 16, 93–112.

Victora, C., Barros, F.C., Vaughan, J.P., Silva, A.C., Tomasi, E., 2000. Explaining trends in inequities: evidence from Brazilian child health studies. Lancet 356 (9235), 1093–1098.

Wagstaff, A., 2000. Socioeconomic inequalities in child mortality: comparisons across nine developing countries. Bulletin of the World Health Organization 78 (1), 19–29.

Wagstaff, A., Nguyen, N., 2002. Poverty and survival prospects of Vietnamese children under Doi Moi. Policy Research Working Paper 2832. World Bank, Washington, DC.

Wagstaff, A., Watanabe, N., 2000. Socioeconomic inequalities in child malnutrition in the developing world. Policy Research Working Paper 2434. World Bank, Washington, DC.

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

World Bank, SIDA, AusAID, Royal Netherlands Embassy and Ministry of Health of Vietnam, 2001. Vietnam. Growing Healthy: A Review of Vietnam's Health Sector. The World Bank, Hanoi.