

12

Explaining Differences between Groups: Oaxaca Decomposition

After inequalities in the health sector are measured, a natural next step is to seek to explain them. Why do inequalities in health exist between the poor and better-off in many countries despite health systems explicitly aimed at eliminating inequalities in access to health care? Why is inequality in the incidence of health sector subsidies greater in one country than in another? Why has the distribution of health or health care changed over time?

In this chapter and the next, we consider methods of decomposing inequality in health or health care into contributing factors. The core idea is to explain the distribution of the outcome variable in question by a set of factors that vary systematically with socioeconomic status. For example, variations in health may be explained by variations in education, income, insurance coverage, distance to health facilities, and quality of care at local facilities. Even if policy makers have managed to eliminate inequalities in some of these dimensions, inequalities between the poor and better-off may remain in others. The decomposition methods reveal how far inequalities in health can be explained by inequalities in, say, insurance coverage rather than inequalities in, say, distance to health facilities. The decompositions in this chapter and the next are based on regression analysis of the relationships between the health variable of interest and its correlates. Such analyses are usually purely descriptive, revealing the associations that characterize the health inequality, but if data are sufficient to allow the estimation of causal effects, then it is possible to identify the factors that generate inequality in the variable of interest. In cases in which causal effects have not been obtained, the decomposition provides an explanation in the statistical sense, and the results will not necessarily be a good guide to policy making. For example, the results will not help us predict how inequalities in Y would change if policy makers were to reduce inequalities in X , or reduce the effect of X and Y (e.g., by expanding facilities serving remote populations if X were distance to provider). By contrast, if causal effects have been obtained, the decomposition results ought to shed light on such issues.

The decomposition method outlined in this chapter, known as the Oaxaca decomposition (Oaxaca 1973), explains the gap in the means of an outcome variable between two groups (e.g., between the poor and the nonpoor). The gap is decomposed into that part that is due to group differences in the magnitudes of the determinants of the outcome in question, on the one hand, and group differences in the effects of these determinants, on the other. For example, poor children may be less healthy not only because they have less access to piped water but also because

their parents are less knowledgeable about how to obtain the maximum health benefits from piped water (Jalan and Ravallion 2003; Wagstaff and Nguyen 2003). The decomposition technique considered in the next chapter does not permit such a distinction between the contributions of differences in the magnitudes and the effects of determinants. In its favor, however, it does allow us to decompose inequalities in health or health care across the full distribution of say, income, rather than simply between the poor and the better-off.

Oaxaca-type decompositions

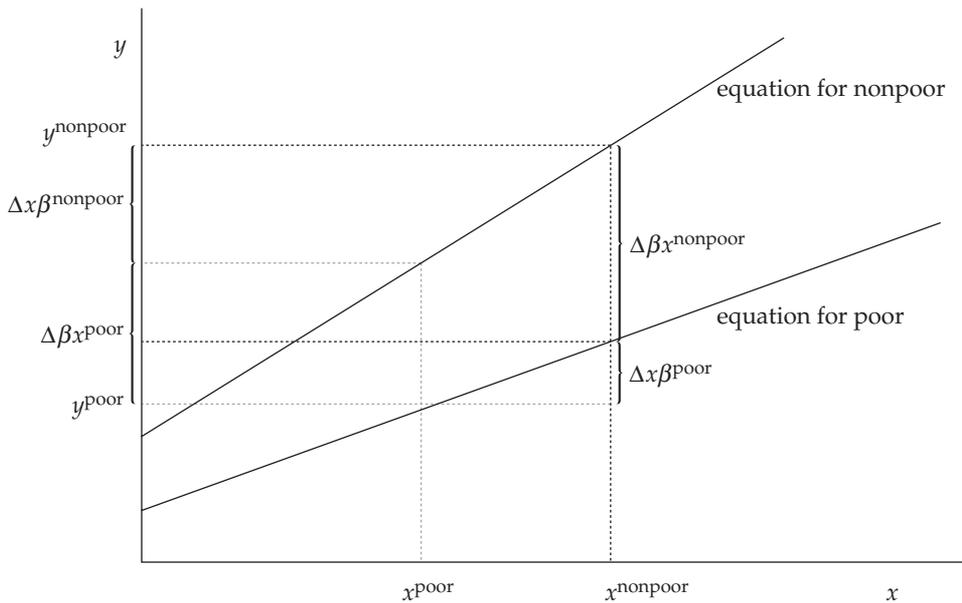
Some preliminaries

Suppose we have a variable, y , which is our outcome variable of interest. We have two groups, which we shall call the poor and the nonpoor. We assume y is explained by a vector of determinants, x , according to a regression model:

$$(12.1) \quad y_i = \begin{cases} \beta^{poor} x_i + \varepsilon_i^{poor} & \text{if } poor \\ \beta^{nonpoor} x_i + \varepsilon_i^{nonpoor} & \text{if } nonpoor \end{cases}$$

where the vectors of β parameters include intercepts. In the case of a single regressor, drawn in figure 12.1, the nonpoor are assumed to have a more advantageous regression line than the poor. At each value of x , the outcome, y , is better. In addition, the nonpoor are assumed to have a higher mean of x . The result is that the poor have a lower mean value of y than do the nonpoor.¹

Figure 12.1 Oaxaca Decomposition



Source: Authors.

¹In the case of the poor, we read off the equation for the poor above x^{poor} , giving a value of y equal to y^{poor} . In the case of the nonpoor, we read off the equation for the nonpoor above $x^{nonpoor}$, giving a value of y equal to $y^{nonpoor}$.

The gap between the mean outcomes, $y^{nonpoor}$ and y^{poor} , is equal to

$$(12.2) \quad y^{nonpoor} - y^{poor} = \beta^{nonpoor} x^{nonpoor} - \beta^{poor} x^{poor},$$

where $x^{nonpoor}$ and x^{poor} are vectors of explanatory variables evaluated at the means for the nonpoor and the poor, respectively.² For example, if we have just two x 's, x_1 and x_2 , we can write the following:

$$(12.3) \quad \begin{aligned} y^{nonpoor} - y^{poor} &= (\beta_0^{nonpoor} - \beta_0^{poor}) + (\beta_1^{nonpoor} x_1^{nonpoor} - \beta_1^{poor} x_1^{poor}) + (\beta_2^{nonpoor} x_2^{nonpoor} - \beta_2^{poor} x_2^{poor}) \\ &= G_0 + G_1 + G_2 \end{aligned}$$

so that the gap in y between the poor and the nonpoor can be thought of as being due in part to (i) differences in the intercepts (G_0), (ii) differences in x_1 and β_1 (G_1), and (iii) differences in x_2 and β_2 (G_2). For example, G_1 might measure the part of the gap in mean health status (y) due to differences in educational attainment (x_1) and the effects of educational attainment (β_1), and G_2 might measure the part of the gap due to the gap in accessibility to health facilities (x_2) and differences in the effects of accessibility (β_2).

Estimates of the difference in the gap in mean outcomes can be obtained by substituting sample means of the x 's and estimates of the parameters β 's into equation 12.2. In the rest, we make such a substitution but do not make it explicit in the notation.

Oaxaca's decomposition

We could stop here. But we might want to go further and ask how much of the overall gap or the gap specific to any one of the x 's (e.g., G_1 or G_2) is attributable to (i) differences in the x 's (sometimes called the explained component) rather than (ii) differences in the β 's (sometimes called the unexplained component). The Oaxaca and related decompositions seek to do just that.

From figure 12.1, it is clear that the gap between the two outcomes could be expressed in either of two ways:

$$(12.4) \quad y^{nonpoor} - y^{poor} = \Delta x \beta^{poor} + \Delta \beta x^{nonpoor}$$

where $\Delta x = x^{nonpoor} - x^{poor}$ and $\Delta \beta = \beta^{nonpoor} - \beta^{poor}$, or as

$$(12.5) \quad y^{nonpoor} - y^{poor} = \Delta x \beta^{nonpoor} + \Delta \beta x^{poor}$$

As the figure makes clear, these decompositions are equally valid. In the first, the differences in the x 's are weighted by the coefficients of the poor group and the differences in the coefficients are weighted by the x 's of the nonpoor group, whereas in the second, the differences in the x 's are weighted by the coefficients of the nonpoor group and the differences in the coefficients are weighted by the x 's of the poor group. Either way, we have a way of partitioning the gap in outcomes between

²Assuming exogeneity, the conditional expectations of the error terms in (12.1) are zero.

the poor and nonpoor into a part attributable to the fact that the poor have worse x 's than the nonpoor, and a part attributable to the fact that *ex hypothesi* they have worse β 's than the nonpoor.

The decompositions in equations 12.4 and 12.5 can be seen as special cases of a more general decomposition:³

$$(12.6) \quad y^{\text{nonpoor}} - y^{\text{poor}} = \Delta x \beta^{\text{poor}} + \Delta \beta x^{\text{poor}} + \Delta x \Delta \beta \\ = E + C + CE$$

so that the gap in mean outcomes can be thought of as deriving from a gap in endowments (E), a gap in coefficients (C), and a gap arising from the interaction of endowments and coefficients (CE). Equations 12.4 and 12.5 are special cases in which

$$(12.4) \quad y^{\text{nonpoor}} - y^{\text{poor}} = \Delta x \beta^{\text{poor}} + \Delta \beta x^{\text{nonpoor}} = E + (CE + C)$$

and

$$(12.5) \quad y^{\text{nonpoor}} - y^{\text{poor}} = \Delta x \beta^{\text{nonpoor}} + \Delta \beta x^{\text{poor}} = (E + CE) + C.$$

So, in effect, the first decomposition places the interaction in the unexplained part, whereas the second places it in the explained part.⁴

Related decompositions

We can also write Oaxaca's decomposition as a special case of another decomposition:

$$(12.7) \quad y^{\text{nonpoor}} - y^{\text{poor}} = \Delta x [D \beta^{\text{nonpoor}} + (I - D) \beta^{\text{poor}}] + \Delta \beta [x^{\text{nonpoor}} (I - D) + x^{\text{poor}} D],$$

where I is the identity matrix and D a matrix of weights. In the simple case, where x is a scalar rather than a vector, I is equal to one, and D is a weight. In this case, $D = 0$ in the first decomposition, equation 12.4, and $D = 1$ in the second, equation 12.5. In the case in which x is a vector, we have

$$(12.8) \quad D = 0 \text{ (Oaxaca) (equation 12.4')}$$

$$(12.9) \quad D = I \text{ (Oaxaca) (equation 12.5')}$$

Other formulations have been suggested. Cotton (1988) suggested weighting the differences in the x 's by the mean of the coefficient vectors, giving us

$$(12.10) \quad \text{diag}(D) = 0.5 \text{ (Cotton),}$$

³This notation is from Ben Jann's help file for his Stata `decompose` routine used later in the chapter.

⁴The rationale for this is that the decompositions were devised to look at discrimination in the labor market. The analog of the nonpoor would be whites or males, and the analog of the poor would be blacks or women. In the first decomposition the presumption is that it is blacks and women who are paid according to their characteristics, whereas whites and men receive unduly generous remuneration. In the second decomposition, the presumption is that whites and men are paid according to their characteristics, and it is blacks and women who are discriminated against.

where $\text{diag}(D)$ is the diagonal of D . Reimers (1983) suggested weighting the coefficient vectors by the proportions in the two groups, so that if f_{NP} is the sample fraction in the nonpoor group, we have

$$(12.11) \quad \text{diag}(D) = f_{NP} \text{ (Reimers)}.$$

In addition to Oaxaca's two decompositions and the additional two proposed by Cotton and Reimers, there is a fifth proposed by Neumark (1988), which makes use of the coefficients obtained from the pooled data regression, β^P :

$$(12.12) \quad y^{\text{nonpoor}} - y^{\text{poor}} = \Delta x \beta^P + \left[x^{\text{nonpoor}} (\beta^{\text{nonpoor}} - \beta^P) + x^{\text{poor}} (\beta^P - \beta^{\text{poor}}) \right] \text{ (Neumark)}.$$

Illustration: decomposing poor–nonpoor differences in child malnutrition in Vietnam

We illustrate the decompositions by means of an example. The setting is Vietnam. The aim of the exercise is to explain the difference between the poor and the nonpoor in child malnutrition, measured anthropometrically through height-for-age z-scores (see chapter 4).

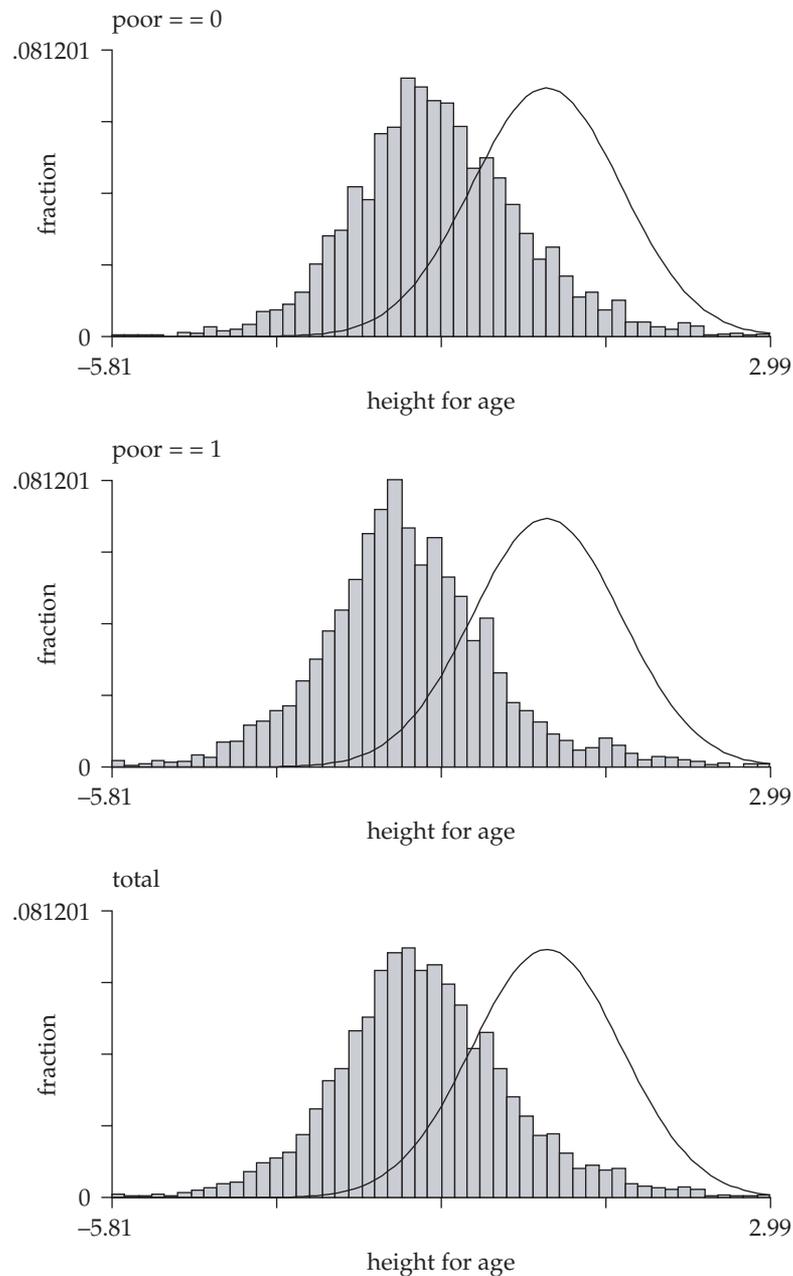
We classify (under-10) children as poor if they are below the poverty line of D 1,790,000 (D = Vietnamese dong), which is the classification developed by the World Bank (Glewwe, Gragnolati, and Zaman 2000) and used by the government of Vietnam. On this basis, using sample weights, we have 46 percent of under-10 children being classified as poor. Figure 12.2 shows that poor children (poor = 1) tend to have a height-for-age z-score (HAZ) lower than that of nonpoor children (poor = 0). The mean HAZ values among the nonpoor and poor are -1.44 and -1.86 , respectively. A mean of 0.00 would place the group in question at the 50th centile in the U.S. reference sample of well-nourished children (distribution sketched in figures), so even the average nonpoor child in Vietnam is substantially undernourished by U.S. standards. Our focus here is on explaining the gap of 0.42 between the mean HAZ of nonpoor and poor children.

Regression model and its estimation

In our setting, y is the HAZ malnutrition score. As in chapter 13, we use basically the same regression model as Wagstaff, van Doorslaer, and Watanabe (2003) and include the log of the child's age in months ($\ln\text{age}$), a dummy indicating whether the child in question is male (sex), dummies indicating whether the child's household has safe drinking water (safewtr) and satisfactory sanitation (oksan), the years of schooling of the child's mother (schmom), and the natural logarithm of household per capita consumption ($\ln\text{pcexp}$). Our poverty grouping variable is poor , which takes a value of 1 if the child's household is poor. The first step is to see whether the regression coefficient vector, β , differs systematically between the poor and nonpoor. The relevant Stata commands are as follows:

```
xi: regr haz poor i.poor|lnage i.poor|sex i.poor|safwtr
    i.poor|oksan i.poor|schmom i.poor|lnpcexp [pw=wt]
testparm poor _I*
```

Figure 12.2 Malnutrition Gaps between Poor and Nonpoor Children, Vietnam, 1998



Source: Authors.

The first command runs a regression with the `poor` dummy included alone and interacted with all the x 's. The second command tests the hypothesis that the coefficients on the `poor` dummy and its interactions are simultaneously equal to zero. The F-statistic, with 7 and 5154 degrees of freedom, is 2.03 and has a p -value of 0.0472. Thus the Oaxaca-type approach, which allows for different regression coefficients, makes some sense in this context, although rejection of the null of parameter homogeneity is somewhat marginal.

Decomposition

Ben Jann's Stata routine `decompose`, which is downloadable from the Stata Web site,⁵ allows all the decompositions outlined above to be computed in just one command:

```
decompose haz lnage sex safwtr oksan schmom lnpcexp [pw=wt],
  by(poor) detail estimates
```

The syntax is the same as the `regress` command, except that after the comma the user has to specify the variable defining the two groups (in our case `poor`). The first block of output (table 12.1) reports the mean values of y for the two groups, and the difference between them. It then shows the contribution attributable to the gaps in endowments (E), the coefficients (C), and the interaction (CE). In this application, the gap in endowments accounts for the great bulk of the gap in outcomes.

Table 12.1 First Block of Output from `decompose`

Mean prediction high (H):	-1.442
Mean prediction low (L):	-1.861
Raw differential (R) {H-L}:	0.419
- due to endowments (E):	0.406
- due to coefficients (C):	-0.082
- due to interaction (CE):	0.095

Source: Authors.

The second block of output (table 12.2) shows how the explained and unexplained portions of the outcome gap vary depending on the decomposition used. The first and second columns correspond to the Oaxaca decomposition in equations 12.4' and 12.5', where $D = 0$ and $D = 1$, respectively. The third and fourth columns correspond to Cotton's and Reimers' decompositions, where the diagonal of D equals 0.5 and $f_{NP} = 0.562$ (in our case), respectively. The final column labeled "*" is Neumark's decomposition. Whatever decomposition is used, it is clearly the difference in the mean values of the x 's that accounts for the vast majority of the difference in malnutrition between poor and nonpoor children in Vietnam. Differences in the effects of the determinants play a tiny part in explaining malnutrition inequalities.

Table 12.2 Second Block of Output from `decompose`

D:	0	1	0.5	0.562	*
Unexplained (U) {C+(1-D)CE}:	0.014	-0.082	-0.034	-0.038	-0.032
Explained (V) {E+D*CE}:	0.406	0.501	0.454	0.458	0.451
% unexplained {U/R}:	3.2	-19.5	-8.1	-9.1	-7.5
% explained (V/R):	96.8	119.5	108.1	109.1	107.5

Source: Authors.

⁵From within Stata, give the command `findit decompose` and follow the links. Another ado file—`oaxaca`—is available for Stata version 8.2 and later. This has all the functions of `decompose` with the important addition of providing standard errors for the contributions.

The third block of output (table 12.3) allows the user to see how far gaps in individual x 's contribute to the overall explained gap. For example, focusing on the final column corresponding to Neumark's decomposition, we see that the gaps in the two demographic variables actually *favor* the poor, whereas the gaps in the remaining variables all *disfavor* the poor. Of the latter, it is the gap in household consumption that accounts for the bulk of the explained gap. It is not so much the correlates of poverty (poor water and sanitation, low educational levels) that account for malnutrition inequalities between poor and nonpoor children in Vietnam—it is poverty itself, in the form of lack of purchasing power.

Table 12.3 Third Block of Output from decompose

Variables	explained: D =						
	E(D=0)	C	CE	1	0.5	0.543	*
lnage	-0.027	0.282	0.005	-0.022	-0.024	-0.024	-0.024
Sex	-0.004	0.038	0.002	-0.002	-0.003	-0.003	-0.003
safwtr	0.029	0.005	0.004	0.033	0.031	0.031	0.033
oksan	-0.008	0.016	0.056	0.048	0.02	0.022	0.036
schmom	0.029	-0.103	-0.035	-0.006	0.012	0.01	0.009
lnpcexp	0.387	0.551	0.064	0.45	0.419	0.421	0.4
_cons	0	-0.87	0	0	0	0	0
Total	0.406	-0.082	0.095	0.501	0.454	0.458	0.451

Source: Authors.

The fourth and final block of output (table 12.4) gives the coefficient estimates, means, and predictions for each x for each group, the "high group" in this case being the nonpoor and the "low group" being the poor.

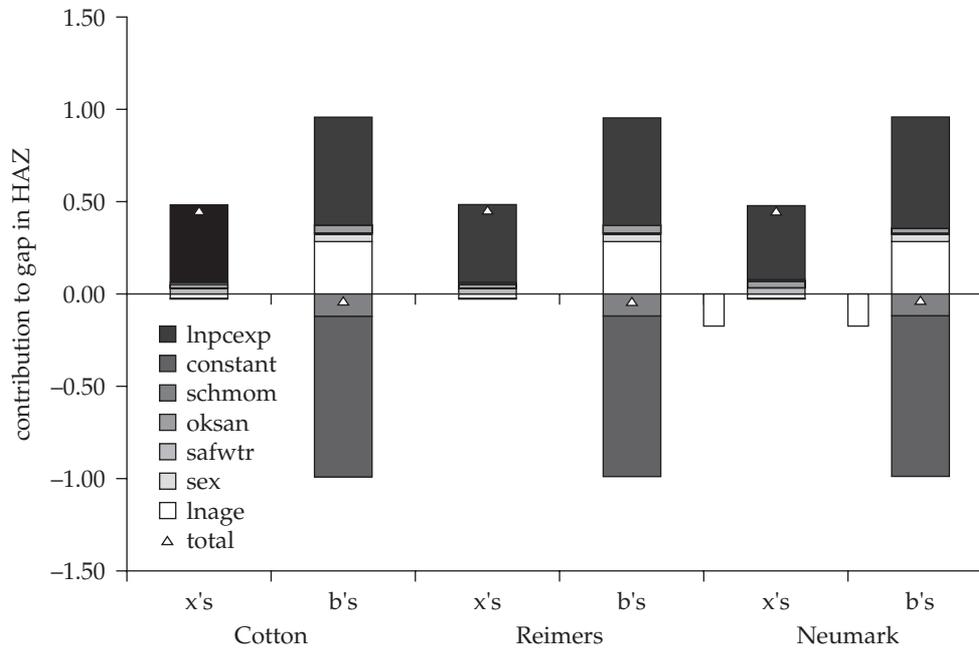
Table 12.4 Fourth Block of Output from decompose

Variables	High model		Low model		Pooled		
	Coef.	Mean	Pred.	Coef.	Mean	Pred.	Coef.
lnage	-0.321	4.021	-1.291	-0.392	3.952	-1.551	-0.354
Sex	-0.088	0.513	-0.045	-0.166	0.491	-0.081	-0.122
Safwtr	0.165	0.421	0.069	0.144	0.221	0.032	0.164
Oksan	0.195	0.313	0.061	-0.034	0.069	-0.002	0.147
Schmom	-0.003	7.696	-0.023	0.015	5.739	0.086	0.005
lnpcexp	0.544	7.99	4.348	0.467	7.162	3.346	0.483
_cons	-4.561	1	-4.561	-3.691	1	-3.691	-3.955
Total			-1.442			-1.861	

Source: Authors.

For the first Oaxaca decomposition (12.4'), columns 2 and 3 of table 12.3 allow us to identify how the gap in each of the β 's contributes to the overall unexplained gap. For the other decompositions, the contributions of the individual β 's can be found by taking the group difference in the variable specific predictions given in table

Figure 12.3 Contributions of Differences in Means and in Coefficients to Poor–Nonpoor Difference in Mean Height-for-Age z-Scores, Vietnam, 1998



Source: Authors.

12.4 and subtracting the explained part given in table 12.3 from this. A bar chart can then be presented, as in figure 12.3, showing the contribution of the difference in the means of each x and the difference in the coefficients on each x . As far as the means of the x 's are concerned, figure 12.3 tells us simply what we already knew from the second block of output: most of the explained part of the malnutrition gap is attributable to the gap in per capita consumption. The triangles in the chart, which indicate the overall contributions of the x 's and the β 's, also show us something else we already knew: that the bulk of the gap in malnutrition is from gaps in the x 's, not gaps in the β 's. Figure 12.3 makes clear that the unimportance overall of the unexplained portion is due to offsetting effects from different β 's. The poor have a higher intercept in the HAZ equation, but this is largely offset by the fact that the consumption effect is weaker for the poor.

Extensions

The framework above can be extended in a number of ways. One is to explain changes in gaps over time (Makepeace et al. 1999). Wagstaff and Nguyen (2003) use this framework to investigate why child survival continued to improve in Vietnam during the 1990s for the nonpoor but not for the poor.

Another extension would be to take selectivity into account. There are, in fact, two separate selectivity issues that might be explored. The first concerns sample selection. Consider the example of child malnutrition. Because a child's nutritional status influences its survival prospects, it also affects the probability that the child appears in the sample (Lee, Rosenzweig, and Pitt 1997). The resulting selection bias can be dealt with provided data are available to model the selection process. In the

example given, this would require fertility history data such that the probability of a child death before the survey date could be modeled as a function of household characteristics. The selection correction term—known as the inverse Mills ratio (IMR) (Wooldridge 2002)—can then be used to adjust the group mean difference in the outcome variable. In the `decompose` routine, this can be done with the option `lambda(varname)`, where `varname` would be that given to the IMR. A second selection problem concerns the selection into the poor and nonpoor groups—group assignment selection. Malnutrition in a child might itself reduce a household's living standards by, for example, keeping the mother at home to look after the child and preventing her from working or by reducing the amount of help the child can provide on the family farm (Ponce, Gertler, and Glewwe 1998). If this is the case, malnutrition may influence the selection of a child into poverty. If the sample selection issue is put aside, the group assignment problem can be dealt with by modeling the probability of being in one group rather than the other, and then using the selection correction terms to adjust the difference in group means. Again, this can be done in the `decompose` routine with the `lambda()` option.

A further extension is the case in which the relationship of interest is nonlinear. Examples include a probit model or a hazard model in the case of modeling child survival. In such cases, one option would be to work with the underlying latent variable that is linear in the covariates. Wagstaff and Nguyen (2003), for example, do their decomposition in terms of the negative of the log of the hazard rate.

The methods described above decompose the difference in the mean of an outcome variable between two groups. Group differences in other parameters of the distribution can also be of interest. For example, with respect to the example of child malnutrition, the difference in mean HAZ scores is arguably less interesting than the difference in the proportion of poor and nonpoor children that are stunted. The general Oaxaca approach can be extended to decompose differences in a full distribution of an outcome into the contribution of differences in the distributions of covariates, on the one hand, and differences in the effects of these covariates, on the other. For example, this can be done using quantile regression (Machado and Mata 2005). Apart from decomposing the full distribution, and not simply the mean, this approach has the advantage of allowing the effect of covariates to differ over the conditional distribution of the outcome. So, for example, one can allow for the possibility that income has a different marginal effect on the nutritional status of malnourished and well-nourished children. The approach has been used to explain the change in the distribution of HAZ scores in Vietnam between 1993 and 1998 (O'Donnell, López-Nicolás, and van Doorslaer 2005; O'Donnell, van Doorslaer, and Wagstaff 2006).

References

- Cotton, J. 1988. "On the Decomposition of Wage Differentials." *Review of Economics and Statistics* 70(2): 236–43.
- Glewwe, P., M. Gagnolati, and H. Zaman. 2000. *Who Gained from Vietnam's Boom in the 1990s?* Washington, DC: The World Bank.
- Jalan, J., and M. Ravallion. 2003. "Does Piped Water Reduce Diarrhea for Children in Rural India?" *Journal of Econometrics* 112(1): 153–73.
- Lee, L.-F., M. Rosenzweig, and M. Pitt. 1997. "The Effects of Improved Nutrition, Sanitation, and Water Quality on Child Health in High-Mortality Populations." *Journal of Econometrics* 77: 209–35.

- Machado, J. A. F., and J. Mata. 2005. "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression." *Journal of Applied Econometrics* 20(4): 445–65.
- Makepeace, G., P. Paci, P. H. Joshi, and P. Dolton. 1999. "How Unequally Has Equal Pay Progressed Since the 1970s." *Journal of Human Resources* 34(3): 534–56.
- Neumark, D. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." *Journal of Human Resources* 23(3): 279–95.
- Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14: 693–709.
- O'Donnell, O., A. López-Nicolás, and E. van Doorslaer. 2005. "Decomposition of Changes in the Distribution of Child Nutrition Using Quantile Regression: Vietnam 1993–98." Mimeo, University of Macedonia, Thessaloniki, Greece.
- O'Donnell, O., E. van Doorslaer, and A. Wagstaff. 2006. "Decomposition of Inequalities in Health and Health Care." In *Elgar Companion to Health Economics*, ed. A. M. Jones. Chichester, England: Edward Elgar.
- Ponce, N., P. Gertler, and P. Glewwe. 1998. "Will Vietnam Grow Out of Malnutrition?" In *Household Welfare and Vietnam's Transition*, ed. D. Dollar, P. Glewwe, and J. Litvack, 257–75. Washington, DC: World Bank.
- Reimers, C. W. 1983. "Labor Market Discrimination against Hispanic and Black Men." *Review of Economics and Statistics* 65(4): 570–79.
- Wagstaff, A., and N. N. Nguyen. 2003. "Poverty and Survival Prospects of Vietnamese Children under Doi Moi." In *Economic Growth, Poverty and Household Welfare: Policy Lessons from Vietnam*, ed. P. Glewwe, N. Agrawal, and D. Dollar. Washington, DC: World Bank.
- Wagstaff, A., E. van Doorslaer, and N. Watanabe. 2003. "On Decomposing the Causes of Health Sector Inequalities, with an Application to Malnutrition Inequalities in Vietnam." *Journal of Econometrics* 112(1): 219–27.
- Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

