Volume Title: The Effect of Education on Efficiency in Consumption

Volume Author/Editor: Robert T. Michael

Volume Publisher: NBER

Volume ISBN: 0-87014-242-9

Volume URL: http://www.nber.org/books/mich72-1

Publication Date: 1972

Chapter Title: Appendix B

Chapter Author: Robert T. Michael

Chapter URL: http://www.nber.org/chapters/c3521

Chapter pages in book: (p. 105 - 112)

Appendix B

THE PURPOSE OF this appendix is to consider the biases that might exist in the estimated coefficients and to discuss the relationship of these biases to the model's prediction of a positive correlation between the estimated income and education elasticities.

1. THE STRUCTURE

Suppose a household's expenditure on a market good, X, is a function of its permanent money income, Y_p, and its nonmarket efficiency, E_f:[1]

$$X = \alpha + \beta_Y Y_p + \beta_{E_f} E_f + u, \tag{B.1}$$

where u is an independent stochastic disturbance term. When fitted in the appropriate log form, β_Y is an estimate of the income elasticity of X and β_{E_f} is an estimate of the elasticity of expenditure on X with respect to nonmarket efficiency. We expect β_Y to be positive for most, if not all, of the market goods. The model developed here suggests that under certain specified assumptions, β_{E_f} will be positive, zero, or negative as β_Y is $\gtreqless$ unity.

Since neither Y_p nor E_f is directly observable, some proxy for each is used. We assume that total permanent consumption, C_p, is proportional to Y_p and that E_f is a function of the household's level of education, E, the age of its members, A, and a vector, V, of other factors including ability:

$$C_p = b_1 Y_p$$

or measured consumption, C, is

$$C = b_1 Y_p + C_t,$$

where C_t is transitory consumption, so

[1] The two other variables used in the empirical chapters, family size and region, add nothing of substance in the context of this appendix and will be ignored.

$$Y_p = (C - C_t)/b_1, \tag{B.2}$$

$$E_f = a_1 + a_2E + a_3A + a_4V, \tag{B.3}$$

where b_1, a_2, and $a_4 > 0$ and the sign of a_3 is ambiguous. Substituting (B.2) and (B.3) into (B.1),

$$X = (\alpha + a_1\beta_{E_f}) + (\beta_Y/b_1)C + (\beta_{E_f}a_2)E + (\beta_{E_f}a_3)A + (u + \beta_{E_f}a_4V - (\beta_Y/b_1)C_t)$$

or

$$X = \alpha^* + \beta^*_C C + \beta^*_E E + \beta^*_A A + w \tag{B.4}$$

where

$$w = u + \beta_{E_f}a_4V - \beta^*_C C_t.$$

2. THE ESTIMATING EQUATION

If equation (B-4) is estimated empirically by ordinary least squares, the estimated equation would be

$$X = a + b_C C + b_E E + b_A A + \epsilon \tag{B.5}$$

for each consumption item. In matrix notation

$$X = Y\beta + w \tag{B.6}$$

where X is an n by 1 column vector of the n observations on the expenditure item, β is a 3 by 1 vector of the regression coefficients, w is an n by 1 vector of error terms, and

$$Y = \begin{pmatrix} C_1 & E_1 & A_1 \\ C_2 & E_2 & A_2 \\ \vdots & \vdots & \vdots \\ C_n & E_n & A_n \end{pmatrix}, \tag{B.7}$$

expressing the variables in deviation form. By the use of ordinary least squares

$$\hat{\beta} = (Y'Y)^{-1}(Y'X) \tag{B.8}$$

and the expected value of $\hat{\beta}$ can be expressed as

$$\mathrm{E}(\hat{\beta}) = \beta + (Y'Y)^{-1}\mathrm{E}(Y'w). \tag{B.9}$$

The variance-covariance matrix of $\hat{\beta}$ is

$$\mathrm{var}\,(\hat{\beta}) = \sigma_w{}^2(Y'Y)^{-1}, \tag{B.10}$$

and Haitovsky has shown that the elements in this matrix can be expressed in terms of the partial correlations between the independent variables:[2]

$$\text{cov}\ (\beta_i \beta_j) = -\ \rho_{ij.l}\,\sigma\beta_i\sigma\beta_j, \tag{B.11}$$

where l is an index over other independent variables, and $\rho_{ij.l}$ is the partial correlation coefficient of independent variables i and j. Thus the covariation between, say, the income elasticity and the education elasticity ($l \equiv A$) is opposite in sign to the partial correlation of income and education. For $i = j$, (B.11) is an identity given the convention that the partial correlation coefficient of a variable with itself is minus one: $\rho_{ii.l} = -1$. So we may express the ijth element in $(Y'Y)^{-1}$ from (B.10) and (B.11) as

$$(Y'Y)_{ij}^{-1} = -\ \rho_{ij.l} k_{ij} \quad \text{where } k_{ij} = \frac{\sigma\beta_i\sigma\beta_j}{\sigma_w^{\ 2}} > 0;\ \text{all } i, j. \tag{B.12}$$

That is, the sign of the ijth element in $(Y'Y)^{-1}$ is opposite the sign of $\rho_{ij.l}$.[3]

Writing out the three rows in (B.9) using (B.12):

$$\begin{aligned}
\text{E}(b_C) &= \beta_C^* + k_{CC}\text{E}(C'w) - \rho_{CE.A}k_{CE}\text{E}(E'w) - \rho_{CA.E}k_{CA}\text{E}(A'w)\\
\text{E}(b_E) &= \beta_E^* - \rho_{EC.A}k_{EC}\text{E}(C'w) + k_{EE}\text{E}(E'w) - \rho_{EA.C}k_{EA}\text{E}(A'w)\\
\text{E}(b_A) &= \beta_A^* - \rho_{AC.E}k_{AC}\text{E}(C'w) - \rho_{AE.C}k_{AE}\text{E}(E'w) + k_{AA}\text{E}(A'w).
\end{aligned} \tag{B.13}$$

To determine the direction of these effects we must know the partial correlation matrix and the sign of each of the expected value terms.

3. THE DIRECTION OF BIASES

We will assume that A is uncorrelated with u, C_t, and hence $\text{E}(A'w) = 0$ and the final term drops out of each line of (B.13). Writing out the remaining two terms,

[2] See Yoel Haitovsky, "On the Correlation Between Estimated Parameters in Linear Regressions," NBER, Mimeo., May 1969.

[3] From (B.10), var $(\hat{\beta})/\sigma_w^2 = (Y'Y)^{-1}$ and the ijth element would be

$$\frac{\text{cov}\ \beta_i\beta_j}{\sigma_w^{\ 2}} = -\ \frac{\rho_{ij.l}\ \sigma\beta_i\sigma\beta_j}{\sigma_w^{\ 2}} = -\ \rho_{ij.l}\, k_{ij}$$

If $i = j$, $k_{ij} = \sigma\beta_i^{\,2}/\sigma_w^{\,2} > 0$ and since $\rho_{ij.l} = -1$, the term is necessarily positive. This must be the case, obviously, since it is a variance term on the principal diagonal.

$$\mathrm{E}(C'w) = \mathrm{E}[(C'u) + \beta_E\, a_4(C'V) - \beta_C^*(C'C_t)]$$
$$\mathrm{E}(E'w) = \mathrm{E}[(E'u) + \beta_{E_f} a_4(E'V) - \beta_C^*(E'C_t)]. \qquad \text{(B.14)}$$

It seems reasonable to suppose $(E'u) = (E'C_t) = 0$ and that $(E'V) > 0$, since education is positively correlated with ability. Hence with $a_4 > 0$ and β_{E_f} having the sign of $(\beta_C^* - 1)$:

$$\mathrm{E}(E'w) = \beta_{E_f} a_4 \mathrm{E}(E'V) \gtreqless 0 \quad \text{as } \beta_C^* \gtreqless 1. \qquad \text{(B.15)}$$

Next consider E $(C'w)$. Since $C = C_p + C_t$, $(C'C_t) = (C_p'C_t) + (C_t'C_t) = n\sigma_{C_t}^2$ by the usual assumption that C_t and C_p are uncorrelated. Likewise, $(C'V) = (C_p'V) + (C_t'V) = (C_p'V) > 0$, since the permanent component in income or consumption is presumably positively related to ability. Finally, since u is the transitory expenditure on the item in question and is therefore unrelated to C_p but is related to C_t by the definition $\sum_g u_{g_i} = C_{t_i}$ (where g is an index over consumption items for the ith observation), $(C'u) = (C_t'u) > 0$, unless the transitory expenditure on one item is offset by a negative transitory expenditure on another item in the consumption basket. In fact, it may be reasonable to suppose that $(C_t'u) = 0$ for most items except those on which expenditures tend to be lumpy, e.g., durable goods. Since an expenditure on a durable—an automobile, a major appliance, a home, et cetera—is probably not offset within the period, for these durables it seems very likely that $(C_t'u) > 0$. Summarizing these effects on $(C'w)$:

$$\left.\begin{array}{lll} \text{"durables"} & \mathrm{E}(C'u) > 0 & \text{for durable items,} \\ \text{"ability"} & \beta_{E_f} a_4 \mathrm{E}(C'V) \gtreqless 0 & \text{as } \beta_C^* \gtreqless 1, \\ \text{"measurement"} & -\beta_C^* \mathrm{E}(C'C_t) < 0 & \text{for superior goods.} \end{array}\right\} \text{(B.16)}$$

In order to determine the direction of the biases on the estimated coefficients, we must also know the partial correlations between the independent variables. These are given in the following table for the 1960 BLS data used in Chapters 4 and 5.[4]

[4] The matrix includes the two additional explanatory variables, family size and region, despite their being omitted from the discussion in this appendix.

Partial Correlation Matrix Weighted by $\sqrt{}$Cell Size

	ln E	A	F	R
ln C	.662	.299	.865	−.441
ln E		−.755	−.733	.102
A			−.591	−.002
F				.376

Equation (B.13) can now be rewritten, defining k^*_{ij} *as* $|\rho_{ij.l}|k_{ij}$, which is always positive, and assigning the proper sign to each term from the partial correlation matrix:

$$\mathrm{E}(b_C) = \beta_C^* + k_{CC}\mathrm{E}(C'w) - k_{CE.A}\mathrm{E}(E'w)$$

$$\mathrm{E}(b_E) = \beta_E^* - k^*_{EC.A}\mathrm{E}(C'w) + k_{EE}\mathrm{E}(E'w) \qquad \text{(B.17)}$$

$$\mathrm{E}(b_A) = \beta_A^* - k^*_{AC.E}\mathrm{E}(C'w) + k^*_{AE.C}\mathrm{E}(E'w).$$

Several points can be made from equations (B.15)–(B.17). First, notice that the two "ability" biases work in opposite directions and may result in no net effect on the coefficients. The effect of $\mathrm{E}(C'V)$ on b_C is positive for a luxury, but the effect of $\mathrm{E}(E'V)$ for this case is negative;[5] similarly, the effects are opposite for necessities and for the education coefficient. Second, if we assume for the moment that $\mathrm{E}(C'V) = \mathrm{E}(E'V) = 0$ and focus on the measurement errors,

$$\mathrm{E}(C'w) = \mathrm{E}(C_t'u - \beta_C^* n\sigma_{C_t}{}^2),$$

so from (B.17)

$$\mathrm{E}(b_C) \gtreqless \beta_C^* \quad \text{as} \quad \left(\frac{\mathrm{cov}(C_t u)}{\sigma_{C_t}{}^2}\right) \gtrless \beta_C^*, \qquad \text{(B.18)}$$

which is equivalent to the statement made by Liviatan,[6] and

$$\mathrm{E}(b_E) \lesseqgtr \beta_E^* \quad \text{as} \quad \left(\frac{\mathrm{cov}(C_t u)}{\sigma_{C_t}{}^2}\right) \gtrless \beta_C^*. \qquad \text{(B.19)}$$

[5] Isolating these "ability" biases on $\mathrm{E}(b_C)$, from (B.17):

$$\mathrm{E}(b_C) - \beta_C^* = (\beta_{Ef}a_4/\sigma_w^2)(\sigma_{\beta_C}^2\mathrm{E}(C'V) - |\rho_{CE.l}|\sigma_{\beta_C}\sigma_{\beta_E}\mathrm{E}(E'V))$$

so the bias is positive, nonexistent, or negative as

$$(\sigma_{\beta_C}/\sigma_{\beta_E}|\rho_{CE.l}|) \gtreqless \mathrm{E}(E'V)/\mathrm{E}(C'V).$$

[6] Liviatan, "Errors in Variables," p. 338.

Since this covariance term is likely to be large for lumpy, durable expenditures, their income elasticities are likely to be biased upward, and their education elasticities biased downward.[7]

Notice here that the bias in one direction in the estimate of the income elasticity implies a bias in the opposite direction in the estimate of the education elasticity. And this is the case in general, given the positive partial correlation between income and education: an upward bias on the one coefficient is associated with a downward bias on the other coefficient.[8] Yet a positive relationship between these coefficients is implied by the theoretical model being tested here. If the income elasticities of durables, which tend to be luxuries according to estimates here and elsewhere, are biased upward and their education elasticities biased downward, these biases tend to impose a relationship on the coefficients opposite to that implied by the theory. Furthermore, since the weighted average of these elasticity estimates must be one and zero, respectively, if an upward bias exists in the luxuries, some other items, probably necessities, must be biased downward. Both of these, then, tend to impose a negative correlation on the elasticity estimates.[9]

Grouping the data, as was done for the empirical investigation in this study, should reduce these measurement error biases. The data were cross-classified by, among other things, measured income and education level, and group averages were used. This is the method suggested by Friedman for eliminating the transitory effects.[10] The

[7] From estimates of the relative biases on income coefficients, Liviatan states: "Upward biased elasticities are those of durables and clothing which are generally considered as 'unstable' or 'variable' items. On the other hand, rent, which can be considered as the most stable type of expenditure, exhibits the largest downward bias. These phenomena seem to be consistent with the analysis . . . which pointed out the relations between direction of bias and amount of random variability in a given expenditure item." Ibid., p. 343.

[8] This is related to Haitovsky's equation (see B.11), since, if $\rho_{CE.I} > 0$, the covariation between β_C and β_E is negative.

[9] If, instead, the income elasticities of luxuries were biased downward and those of necessities upward, then education elasticities of the former would be biased upward and those of the latter downward. In this case the biases alone would tend to produce empirical results that would "support" the model's prediction of a positive relationship between the elasticities across the items. It is fortuitous that the durables which are most likely to have income effects biased upward are also luxuries, for if the empirical results support the model it is in spite of, not as a result of, these measurement error biases.

[10] Friedman suggests, as a way of eliminating the influence of transitory factors affecting income, "to classify the families by measured income, to compute

same procedure was suggested by Liviatan. He shows that if measured income, Y, is highly correlated with permanent income and is not correlated with random elements in consumption, then grouping by measured income and using mean expenditures and mean total consumption eliminates the transitory bias. This is, in fact, the procedure used by Liviatan for much of his empirical work.[11]

Thus, the effect of the transitory components should be reduced by the grouping procedure used, although it need not be eliminated, especially in those cases in which the cell size is small.[12] There seems to be no reason to expect the grouping procedure to reduce also the influence of the two "ability" biases, but, as indicated above, these two work in opposite directions on all coefficients.[13]

4. THE TRADITIONAL ENGEL CURVE

Perhaps it should be stressed that some of the biases considered here result from our insistence on including as separate independent variables both the money income and the nonmarket efficiency of the household. Were we willing to combine these two into one variable

mean expenditures on an individual category of consumption and on all categories combined for each class. Under the relevant assumptions about correlations and mean transitory components of consumption, these means are estimates of the mean permanent components of the individual category and of total consumption. The relation between them is then an estimated relation between permanent components." Milton Friedman, *A Theory of the Consumption Function,* Princeton University Press for NBER, 1957, p. 207.

[11] See Nissan Liviatan, *Consumption Patterns in Israel,* Jerusalem, Falk Project for Economic Research in Israel, 1964, p. 78.

[12] Liviatan indicates that about fifty families per cell is "quite satisfactory" in this grouping method (ibid., p. 358). For my empirical work here, the 1960 data contained an average of nearly ninety households, although the 1950 data's average cell size was approximately fifty in Chapter 6 and is considerably smaller in Appendix D. Even for the 1960 data, since the cells were not of equal size, some had substantially fewer than fifty observations; hence the biases may not be fully eliminated.

[13] Notice that if V is uncorrelated with C_p but is correlated positively with education, then the resulting bias tends to support the model's prediction. But this particular circumstance might be considered legitimate supporting evidence in the sense that it indicates that ability or health or some other factor in V raises nonmarket efficiency ($a_4 > 0$). These effects should not, however, be attributed to formal schooling. But since these factors presumably also affect market efficiency, and hence C_p, the offsetting effects exist and the net effect may be negligible.

many of our problems would disappear. However, since the basic model suggests that human capital has an effect on consumption over and above its effect on money earnings or income, we wish to separate out these effects. If the model is accepted, it suggests that running Engel curves with only a money income variable, and omitting the productivity effect, creates a bias of another sort. This would be tantamount to fitting equation (B.1) while omitting the variable E_f. We know that the expected value of the coefficient from this simple regression $x = a + b_{Y_p}Y_p + e$, where $e = \beta_{E_f}E_f + u$, is

$$\mathrm{E}(b_{Y_p}) = \beta_Y + \beta_{E_f}b_{E_fY_p}, \qquad \text{(B.20)}$$

where $b_{E_fY_p}$ is the simple regression coefficient of E_f on Y_p, which is presumably positive in this case. Thus, our model suggests that this estimated coefficient b_{Y_p} is biased away from unity, since $\beta_Y \lesseqgtr 1$ implies $\beta_{E_f} \lesseqgtr 0$ from the model. That is, the income coefficient obtained in the traditional Engel curve includes both a money income effect and an efficiency effect, and so, for example, if human capital is correlated with E_f, but physical capital is not, then the estimated income coefficient would differ depending on which form of capital it resulted from. The income coefficient related to a return on financial or physical capital would be a "pure" income elasticity; the income coefficient which resulted from human capital would include the non-market efficiency effect and would tend to be further from unity.