7-1970

Engel's Formula for Estimating the Costs of Producing an Individual: A Note

Elchanan Cohn
University of South Carolina - Columbia, feu00004@moore.sc.edu

Publication Info
Published in *Journal of Political Economy*, Volume 78, Issue 4, 1970, pages 778-782.

Engel's Formula for Estimating the Costs of Producing an Individual: A Note

Elchanan Cohn

Pennsylvania State University

In a survey article on the "Historical Roots of the Concept of Human Capital," B. F. Kiker presents Engel's formula for estimating the total costs of producing an individual through age x as follows:

$$C_x = c_0\{1 + x + k[x(x + 1)/2]\}, \qquad (1)$$

"where C_x is the total cost of producing a human being (neglecting interest, depreciation, and maintenance) through age x, c_0 denotes costs incurred up to the point of birth, and k is the annual percentage increase in cost" (Kiker 1966, p. 483).

Let c_x be the costs of production during age x (that is, from the $[x - 1]$st to the xth birthday). Then from (1),

$$c_x = c_0 + kc_0x \qquad (x = 0, 1, \ldots, n). \qquad (2)$$

Equation (2) implies that c_x follows an arithmetic progression over time with annual *marginal* costs $(c_x - c_{x-1})$ being kc_0 for all x. We may also note that Kiker's definition of k is correct only when the percentage increase in costs is computed relative to c_0. When we compute the percentage increase in costs relative to the previous year's costs we get

$$\frac{c_x - c_{x-1}}{c_{x-1}} = \frac{c_0k}{c_0 + kc_0(x - 1)} = \frac{k}{1 + k(x - 1)}, \qquad (3)$$

which is less than k whenever $x > 1$.

As Kiker points out, Engel's interest in introducing equation (1) was centered on measuring the value of humans. As such, the procedure is quite erroneous. However, such a formulation may be extremely useful for assessing the *costs* of human production (irrespective of its value). Indeed, if the costs of producing the "average" individual follow one or another

I would like to thank the referee of this *Journal* for his helpful comments and suggestions. Computational assistance was provided by Fred Chow.

version of Engel's formula, it may be of great significance to education economists.

However, it seems to me that equation (2) is objectionable on at least two grounds. First, c_x is alleged to be a function of c_0, the cost incurred up to the point of birth. But it is difficult to perceive why costs in year $x\ (=1, 2, \ldots, n)$ would bear any relation to costs of prenatal care. A more reasonable a priori hypothesis would predict $c_x\ (x = 2, 3, \ldots, n)$ on the basis of c_1. Retaining all of the other features of equation (2) we obtain

$$c_x = c_1 + kc_1(x - 1) \qquad (x = 1, 2, \ldots, n). \tag{4}$$

The second objection is related to the assumption implicit in Engel's formula that costs increase every year by a constant *absolute* sum. This implies that costs of producing (say) an eight-year-old boy differ from those of producing a seven-year-old boy by the same amount that they differ between the production of a two-year-old and one-year-old lad. It seems, however, that as a child grows costs of production increase at an increasing rate. First, it is to be emphasized that we are interested here in social, rather than private, costs. The former will include not just out-of-pocket costs of feeding, clothing, and sheltering the child but also public costs of education, parks and recreation, libraries, and so forth.* These would seem to increase quite rapidly with age. Also, one must take into account hidden opportunity costs, such as the earnings forgone by parents. As the mother extends her period of absence from the labor market, her forgone earnings increase markedly. Moreover, when the father, too, spends time with his child, his forgone opportunities are likewise almost certain to increase with time. In addition, the child's *own* opportunity costs of earnings increase very rapidly with age. This is not to deny the possibility that some cost items may be reduced over time. Medical care may be the prime example. Yet on the whole, the costs of production appear to be increasing rapidly with time, suggesting a non-linear relation between costs and age. Other questions involving the size of the family (scale economies), location in an urban or rural setting, and so on, cannot be explored within the scope of this note.

One alternative formulation to (4) would be to assume a geometric progression of costs over time such that

$$\log c_x = \log c_1 + (\log k)(x - 1) \qquad (x = 1, 2, \ldots, n), \tag{5}$$

which can be written as

$$c_x = c_1 k^{x-1}. \tag{6}$$

Another possible reinterpretation of Engel's formula would assume that costs increase instantaneously in a manner consistent with

$$dc/c = k(dx/x). \tag{7}$$

Solving (7) we get

$$\int_{c_1}^{c_x} dc/c = k \int_{1}^{x} dx/x,$$

which implies that

$$\log c_x = \log c_1 + k(\log x), \qquad (8)$$

or

$$c_x = c_1 x^k \qquad (1 \leq x \leq n). \qquad (9)$$

Equations (6) and (9) provide, in my view, reasonable alternatives to (4). Which of these is the "best" formulation is basically an empirical question. Although a comprehensive investigation into the cost of producing an individual is beyond the scope of this note, some *suggestive* empirical results could still be presented.

Our data are based upon table 15 of the Dublin and Lotka classic (1946). Although table 15 is based on 1935–36 statistics, the relationships among the variables are what is important, and these, we believe, are still applicable today. Further, whatever the limitations of these data, some interesting patterns could still be noted.

Our first hypothesis was concerned with the use of c_1 rather than c_0 as one of the main parameters in Engel's formula. Equations (2′) and (4′) provide a useful comparison of equations (2) and (4). The results appear to corroborate our argument:[1]

$$c_x = 300 + \underset{(0.00101)}{0.03813}\, c_0 x \qquad (x = 0, 1, \ldots, 18); \; R^2 = 0.9209 \qquad (2')$$

$$c_x = 343 + \underset{(0.00044)}{0.02528}\, c_1(x - 1) \qquad (x = 1, 2, \ldots, 18); \; R^2 = 0.9819 \qquad (4')$$

Equations (2′) and (4′) are based on conditional least-squares estimation.[2] Equation (4′) provides a better fit to the data than does (2′), indicating that our first hypothesis could have both a priori and empirical foundations.

To test the second hypothesis, that is, that the relationship between costs and age is nonlinear, we estimated the following equations:

$$\log c_x = 2.5353 + \underset{(0.00014)}{0.00947}\,(x - 1) \qquad (x = 1, 2, \ldots, 18), \; R^2 = 0.9849 \qquad (5')$$

which can be rewritten as

$$c_x = 343 \cdot 1.022^{x-1}, \qquad (6')$$

[1] Data are from the first column of table 15 which lists the costs, exclusive of allowance for deaths and interest, "of bringing up a child to age 18 in families of five persons with annual incomes of $2,500" (Dublin and Lotka 1946, p. 57).

[2] The numbers in parentheses are the standard errors of the coefficients. The estimated coefficient in (2′), for example, is conditioned upon the assumption that $c_0 = 300$.

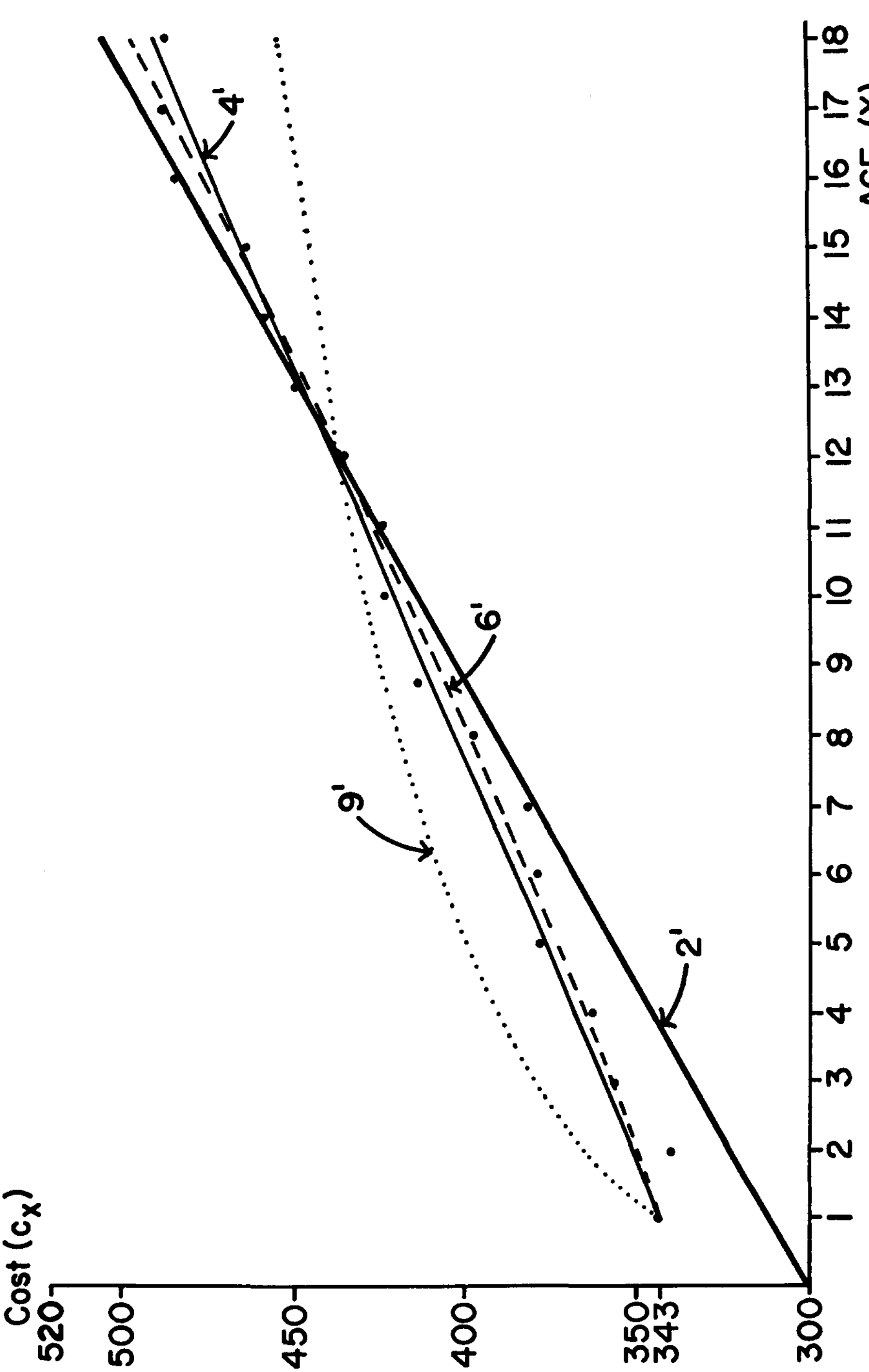
Cost (c_x)
520
500
450
400
350
343
300
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
AGE (X)
2'
4'
6'
9'

FIG. 1

and

$$\log c_x = 2.5353 + \underset{(0.00641)}{0.09711} \log x \qquad (1 \leq x \leq 18), \quad R^2 = 0.7398 \tag{8'}$$

or

$$c_x = 343 \cdot x^{0.09711}. \tag{9'}$$

Equations (5′) and (8′) suggest that while the geometric form (equation [6]) fits the data remarkably well, the alternative form (equation [9]) is statistically inferior. This is not to say, however, that such an hypothesis is necessarily wrong. With improved cost estimates we might find empirical grounds to support the hypothesis embodied in equation (9).

Figure 1 provides a visual illustration of the foregoing statistical analysis. The Dublin and Lotka figures are represented by the scatter of dots. The fitted relationships are also superimposed therein. From the figure, it might be observed that relation (6′) provides the closest fit to the data, although an extremely close fit is also provided by equation (4′). In conclusion, the data of the Dublin and Lotka study appear to confirm our first hypothesis (that is, that c_x should be related to c_1 rather than c_0) but do not provide ample evidence as to whether a geometric progression is superior to an arithmetic progression of costs.

References

Dublin, Louis I., and Lotka, Alfred J. *The Money Value of a Man*. Rev. ed. New York: Ronald, 1946.

Kiker, B. F. "The Historical Roots of the Concept of Human Capital." *J.P.E.* 74 (October 1966):481–99.