

Cowles Foundation Paper 26c

Reprinted from
ECONOMETRICA, Journal of the Econometric Society, Vol. 15, No. 4, October, 1947
The University of Chicago, Chicago 37, Illinois, U.S.A.

FAMILY EXPENDITURES AND THE MARGINAL
PROPENSITY TO CONSUME*

By TRYGVE HAAVELMO

1. INTRODUCTION

THE elasticities of expenditure (for a group of consumer goods or for all consumer goods) with respect to income, as obtained from family-budget data, frequently differ quite considerably from those obtained from time series of per capita income and per capita expenditures (eliminating effects of changes in prices). Several explanations have been offered. They seem to run along two different lines, one leading to the conclusion that the problem is "merely a problem of aggregation," another to the somewhat negative result that cross-section studies "have no meaning."

Those who hold that the problem is one of aggregation argue as follows: The results of cross-section studies of family expenditures indicate that the expenditure curves (family expenditure as a function of family income) are not linear. The total (or average) expenditure of the community will, therefore, depend not only upon total (or average) income but also upon the higher moments of the income distribution. Even if these higher moments should remain constant under variations in the average income the resulting functional relationship between average expenditure and average income would not, in general, be of the same form as the corresponding family-expenditure function. The argument then seems to imply that the "true" elasticities of expenditure with respect to income could be obtained from family-budget data when the sample is sufficiently large, while the observable "macro-" relation between average income and average expenditure might yield only biased results if used for this purpose.

The reasoning of those who question the meaning of relationships derived from cross-section studies is perhaps somewhat more vague. Their general argument is that "people are different," and that, therefore, a comparison of incomes and expenditures as *between* different individuals or families is not the same thing as comparing successive changes in income and expenditure for the same individual or family. These ideas would seem to lead to the conclusion that if there is some observable degree of stability in the relation between average expenditure and average income of the community this result derives from the persistence of a certain "constellation of types of people" in the community.

* This will be reprinted in Cowles Commission Papers, New Series, No. 26.

I do not think there need be any real disagreement between these two lines of thinking. I believe, in fact, that both of them contain useful ideas that can be rationalized and integrated into a more general explanation of the connection between relationships obtained from cross-section studies and the relationships obtained from the corresponding averages for the community as a whole. The purpose of the present note is to offer some suggestions in this respect. I shall discuss the problems involved only with reference to the particular example of family expenditures, but the results might perhaps be of some interest in connection with other types of cross-section studies.

2. GENERALIZATION OF THE FAMILY-EXPENDITURE FUNCTIONS

It seems fairly obvious that the consumer expenditures of a household or a family depend on many other factors besides income and prices. Such expenditures depend no doubt upon the size of the family, the age of the various members of the family, geographic location (e.g., whether city or rural), etc. The expenditures probably also depend on the amount of wealth or liquid assets of the family, the family income in the past, etc. Suppose there are n such variable characteristics, x_1, x_2, \dots, x_n , say, that affect the consumer expenditures of a family besides its (real) income and the prices paid. (One of these x 's may be interpreted as a random residual, a "catch-all" for numerous other factors that it is not possible to specify explicitly.) These ideas are of course not new by any means. In practically all the budget studies that have been published considerable effort is made to "eliminate the effect of other factors," by transforming the income and expenditure data to an "equivalent-adults" basis, by subdividing the families into "rural" and "urban," or according to occupational groups, etc.

Let c denote the total (real) consumer expenditure of a family. Let, further, y denote the (real) income of the family, and p_1, p_2, \dots, p_s , the existing (real) prices of the various consumer goods. Our generalized expenditure function for the individual family could then be written as

$$(2.1) \quad c = f(y, p_1, p_2, \dots, p_s, x_1, x_2, \dots, x_n).$$

The prices paid may depend, to some extent, upon the income level and the other characteristics of the family. For the sake of simplicity we shall here assume that this is not the case, i.e., we shall assume that the expenditure function for the i th family can be written

$$(2.1') \quad c_i = f(y^{(i)}, p_1, p_2, \dots, p_s, x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}).$$

Assuming a consumer-expenditure function of this general type means

accepting the point of view that "people are different" but that the differences can be described explicitly by the value of a set of observable characteristics, x_1, x_2, \dots, x_n .

If we were to register the values of y and the x 's for the various families we could construct a joint, multidimensional frequency table for these $n+1$ characteristics. We may think of this frequency table as representing a sample from an infinite population, described by a joint probability distribution of y and the x 's. If the sample is sufficiently large we may proceed as if we are dealing with an infinite population. To fix the ideas, let the joint probability density function of y and the x 's be

$$(2.2) \quad \Phi(y, x_1, x_2, \dots, x_n; \bar{y})$$

where \bar{y} is a variable parameter, viz., the *average family income* in the community. In what follows we shall assume that *the form Φ remains invariant, and that it contains no other variable parameters besides \bar{y} .*

Now, for any *given value of y* , the n x 's will have a joint probability distribution that in general will depend on y . E.g., there is obviously some correlation between family income and the age of the head of the family. Let

$$(2.3) \quad \omega(y; \bar{y})$$

be the marginal probability density function of y (i.e., the distribution of income). Then the *conditional* distribution, ψ , of the x 's for a *given value of y* is defined by

$$(2.4) \quad \psi(x_1, x_2, \dots, x_n; y; \bar{y}) = \frac{\Phi(y, x_1, x_2, \dots, x_n; \bar{y})}{\omega(y; \bar{y})} .$$

If the x 's and y are not stochastically independent this function will depend on y .

3. TWO PROBLEMS OF AGGREGATION

On the basis of the definitions and assumptions listed above one is led to consider the following two problems of aggregation: (1) To find the total, or average, level of consumer expenditure for all families having a *given income y* , regardless of the values of the characteristics x . (2) To find the total, or average, level of consumer expenditures of all families in the community. The first type of aggregation gives the ordinary income-expenditure function, or Engel curve. This function is obtained from family-budget data. It will, in general, depend also on the prices prevailing during the period when the budget inquiry is made. The second type of aggregation defines the "macro-" relation

between average consumer expenditure and average income as shown by the time series of these two variables. Also this function will, in general, depend on the prices p_1, p_2, \dots, p_s . We shall consider two cases: (a) a linear generalized expenditure function, and (b) a non-linear expenditure function.

(a) *Linear expenditure function*

Let us assume that the expenditure function (2.1) is a linear function of the following type

$$(3.1) \quad c = \alpha y + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \gamma,$$

where α and the β 's are constants, and where

$$(3.2) \quad \gamma = \gamma(p_1, p_2, \dots, p_s)$$

depends only upon the existing prices p_1, p_2, \dots, p_s .

From family-budget data we may calculate the average or *expected level of expenditure* at a given level of family income, y (and given levels of the average income, \bar{y} , and the prices, p_1, p_2, \dots, p_s). This function, which we denote by $E(c|y; \bar{y}; p_1, p_2, \dots, p_s)$, is determined by

$$(3.3) \quad \begin{aligned} & E(c|y; \bar{y}; p_1, p_2, \dots, p_s) \\ &= \iint \dots \int (\alpha y + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \gamma) \\ & \quad \cdot \psi(x_1, x_2, \dots, x_n; y; \bar{y}) dx_1 dx_2 \dots dx_n. \end{aligned}$$

(Here, and in similar formulae below the integrals are taken over the whole range of the distribution or, conventionally, from $-\infty$ to $+\infty$, for each of the variables of integration.)

From (3.3) it is seen that the expected value of c , given y , will be a function of y and p_1, p_2, \dots, p_s , and this function will depend on the parameter \bar{y} . *Obviously this function is not, in general, linear in y or, if it is, the coefficient of y will not in general be equal to α* , for the means $E(x_1|y; \bar{y}), \dots, E(x_n|y; \bar{y})$, will in general be functions of y and \bar{y} , depending on the manner in which y and \bar{y} enter into the distribution ψ . The ordinary family-expenditure function as given by (3.3) can, therefore, be written as

$$(3.4) \quad E(c|y; \bar{y}; p_1, p_2, \dots, p_s) = g(y; \bar{y}) + \gamma(p_1, p_2, \dots, p_s),$$

where g is some—perhaps complicated—function of the family income, y , and the average income, \bar{y} .

Thus, even when the generalized family-expenditure function is linear in y , the apparent marginal propensity to consume, as obtained from budget data, i.e., $\partial g/\partial y$, is in general not equal to α . In fact, $\partial g/\partial y$ will in general

depend also on the average level of income, \bar{y} , prevailing at the time when the budget inquiry was made.

The ordinary family-expenditure function, or Engel curve, given by (3.4) when (as in a budget study) the values of \bar{y} and p_1, p_2, \dots, p_s are held constant, would not show the effect upon expenditure of increasing the income of a given family. It would show the combined effect of the family getting more income *and*, at the same time, moving into a "milieu" where the x 's are, on the average, different.

Suppose that during the year, t_0 say, when the budget inquiry was made, the average income per family in the community was \bar{y}_{t_0} , and that γ was equal to $\gamma_{t_0} = \gamma(p_{1t_0}, p_{2t_0}, \dots, p_{st_0})$. We would then obtain the family-expenditure function $g(y; \bar{y}_{t_0}) + \gamma_{t_0}$. Regarded as a function of y only, this is an Engel curve. Suppose now that in a later year, t_1 say, the average income is $\bar{y}_{t_1} \neq \bar{y}_{t_0}$, and that the prices then are $p_{1t_1}, p_{2t_1}, \dots, p_{st_1}$. And suppose that we should attempt to calculate the average expenditure in the community during t_1 by aggregating the family-expenditure function $g(y; \bar{y}_{t_0}) + \gamma_{t_0}$ over all incomes during t_1 . The result would be given by the expression

$$(3.5) \quad \int g(y; \bar{y}_{t_0}) \omega(y; \bar{y}_{t_1}) dy + \gamma_{t_0}$$

But the average or expected expenditure, \bar{c}_{t_1} , in the whole community during t_1 is obviously given directly by the formulae

$$(3.6) \quad \begin{aligned} \bar{c}_{t_1} &= \int \int \dots \int (\alpha y + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \gamma_{t_1}) \\ &\quad \cdot \Phi(y, x_1, x_2, \dots, x_n; \bar{y}_{t_1}) dx_1 dx_2 \dots dx_n dy \\ &= \int \left[\int \int \dots \int (\alpha y + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \gamma_{t_1}) \right. \\ &\quad \left. \cdot \psi(x_1, x_2, \dots, x_n; y; \bar{y}_{t_1}) dx_1 dx_2 \dots dx_n \right] \omega(y; \bar{y}_{t_1}) dy \\ &= \int g(y; \bar{y}_{t_1}) \omega(y; \bar{y}_{t_1}) dy + \gamma_{t_1} \end{aligned}$$

where $\gamma_{t_1} = \gamma(p_{1t_1}, p_{2t_1}, \dots, p_{st_1})$. (3.6) is an aggregation process of the second type mentioned above. Obviously \bar{c}_{t_1} , as given by (3.6) will *not* in general be equal to the expression (3.5), even if the prices p_{it_1} and p_{it_0} were exactly the same during the two years, as long as, $\bar{y}_{t_1} \neq \bar{y}_{t_0}$.

Another question is whether the macrorelation between \bar{c}_t and \bar{y}_t

that is obtained by an aggregation process of the type (3.6) has the property that $\partial \bar{c}_t / \partial \bar{y}_t = \alpha$. From (3.6), setting $t_1 = t$, we obtain

$$(3.7) \quad \bar{c}_t = \alpha \bar{y}_t + \beta_1 \bar{x}_1(\bar{y}_t) + \beta_2 \bar{x}_2(\bar{y}_t) + \cdots + \beta_n \bar{x}_n(\bar{y}_t) + \gamma_t,$$

where $\bar{x}_i(\bar{y}_t)$ denotes the expected value of x_i and where $\gamma_t = \gamma(p_{1t}, p_{2t}, \cdots, p_{st})$. The means $\bar{x}_i(\bar{y}_t)$ will in general be functions of the average income \bar{y}_t . If the means $\bar{x}_i(\bar{y}_t)$ are independent of \bar{y}_t the partial derivative of \bar{c}_t with respect to \bar{y}_t in (3.7) will be constant and equal to α .

The assumption that the means of the x 's are independent of \bar{y}_t , at least in the short run, seems reasonable in many cases. E.g., an increase in the average income cannot immediately affect the marginal age distribution of the population. The assumption that the means of the x 's do not depend much on the average income in the short run is, I think, in practice a much less dangerous assumption than assuming that the x 's and y are stochastically independent. But in general it would probably be necessary to retain some of the quantities \bar{x}_i as variables in the "macro-" expenditure function (3.7) in order to obtain an unbiased estimate of the coefficient α , although the omission of some of the quantities \bar{x}_i in this relation may be less serious than omitting the corresponding variables x_i in (3.1) when we are trying to estimate α from family-budget data.

(b) *Nonlinear expenditure function*

The results above may be generalized. Let the family-expenditure function have the general form (2.1). Then the ordinary expenditure function, the Engel curve, as obtained from budget data for the period t_0 is given by

$$(3.8) \quad \begin{aligned} E(c \mid y; \bar{y}_{t_0}; p_{1t_0}, p_{2t_0}, \cdots, p_{st_0}) \\ = \int \int \cdots \int f \psi(x_1, x_2, \cdots, x_n; y; \bar{y}_{t_0}) dx_1 dx_2 \cdots dx_n \\ = G(y; \bar{y}_{t_0}; p_{1t_0}, p_{2t_0}, \cdots, p_{st_0}), \end{aligned}$$

where G is some function the form of which will depend on f and ψ .

The "consumption function" as obtained from time series of the average consumer expenditure, \bar{c}_t , the average income, \bar{y}_t , and the prices $p_{1t}, p_{2t}, \cdots, p_{st}$, is the function

$$(3.9) \quad \begin{aligned} \bar{c}_t = \int \int \cdots \int f \Phi(y, x_1, x_2, \cdots, x_n; \bar{y}_t) dx_1 dx_2 \cdots dx_n dy \\ = F(\bar{y}_t; p_{1t}, p_{2t}, \cdots, p_{st}), \end{aligned}$$

where F is some function the form of which depends on f and Φ .

The ordinary expenditure function G , regarded as a function only of family income y , does not show the effect of a change in income only. It shows the combined effect of a family getting a different income and at the same time moving into a "milieu" where the x 's are, on the average, different. Furthermore, the value of G for a given value of y would change as the parameters \bar{y}_t and p_{it} change over time, i.e., the form of the Engel curve would ordinarily change from year to year, even when the form f remains absolutely unchanged.

Suppose that from a budget study during the period t_0 we have obtained the Engel curve (3.8). Then, in order to find the average family expenditure, \bar{c}_{t_1} , during some other period t_1 , it is not sufficient to find the weighted average of this Engel curve by means of the income distribution of the period t_1 . In general

$$(3.10) \quad \int G(y; \bar{y}_{t_0}; p_{1t_0}, p_{2t_0}, \dots, p_{st_0}) \omega(y; \bar{y}_{t_1}) dy \neq \bar{c}_{t_1}$$

if either the prices p_{it_1} and p_{it_0} differ, or if $\bar{y}_{t_1} \neq \bar{y}_{t_0}$, or both.

4. GENERAL REMARKS ON THE MEANING OF CROSS-SECTION STUDIES

In the preceding analysis we found that the Engel curves that are obtained from budget data cannot in general be used to estimate the effect upon consumer expenditures of a spontaneous change in all family incomes. For such a change means a change in the income distribution (e.g., a change in \bar{y}), and this again means a change in the joint distribution of income and the characteristics, x_i . This will in general alter the family-expenditure function G , as obtained from (3.8). The ordinary family-expenditure functions or Engel curves, as obtained from budget data, cannot be assumed to remain invariant under transformations of the income distribution, not even when the changes are of the simple, one-parameter type that were discussed above.

Similar remarks apply to many other types of cross-section studies, e.g., cross-section studies of firms for the purpose of measuring production functions. The conclusion is, I think, that in such studies one has to operate with a *much larger number of independent variables* than is commonly used, in order to make the individual units in a cross-section sample comparable. Otherwise the estimates of "elasticities," "marginal productivities," etc. that are obtained from cross-section studies might be useless for the purpose of estimating the effect of spontaneous changes in the independent variables considered.

*University Institute of Economics,
Oslo*