SUGGESTED METHODS FOR ESTIMATING ADULT EQUIVALENCE SCALES

R. McShane*

When trying to obtain an estimate of the demand function of households for certain food items, it is often convenient to express the demographic structure (size, age and sex composition) of individual households in terms of so many consumer (e.g. an adult male) equivalents. In many studies such conversions have been made by using adult equivalence scales derived from nutrionists estimates of the consumption requirements of human beings at various ages. However these scales, by their nature, are based on estimates of how much a person of a given age and sex should consume, rather than on how much the same person actually does consume in real life. In an endeavour to overcome this difficulty a method is suggested in this article whereby a scale can be calculated from the observed behaviour of households.

In the demand for many foodstuff items there exist strong reasons for expecting that a household's size, age and sex composition will exert an influence. In order to facilitate, and simplify the analysis used to determine the way in which other factors influence the demand for such commodities, it is often convenient to firstly take account of (or eliminate) the influence of these demographic characteristics on their consumption. The usual method used is to define a household's demographic structure in terms of consumer equivalents. In this method, family members of a given age and sex are expressed as a proportion of some "unit consumer" (e.g. an adult male) and the influence of household size and composition are amalgamated into one composite variable, which is then employed to express consumption in terms of so many ounces per adult equivalent.

The adult equivalence scales most commonly used have been derived from nutrionists' estimates of the consumption requirements of human beings at various ages. However such scales, by their nature, are based on estimates of how much a person of a given age and sex should consume, rather than on how much the same person actually does consume in real life. There have been, however, a number of alternative methods suggested[1] which, by using information obtained from household studies

* University of Newcastle.

[1] E. Sydenstricker and W. King, "The Measurement of Relative Economic Status of Families", *Journal of the American Statistical Association*, Vol. 17 (1921); M. Friedman, "A Method of Comparing Incomes of Families Differing in Composition", *Studies in Income and Wealth* (New York: National Bureau of Economic Research, 1952); J. A. C. Brown, "The Consumption of Food in Relation to Household Composition and Income", *Econometrica*, Vol. 22, No.1, (1954); Prais and H. S. Houthakker, *The Analysis of Family Budgets* (Harvard: Cambridge University Press, 1955).

and knowledge of the factors likely to influence household consumption of the particular item, enable an adult equivalence scale to be calculated from actual behaviour.

Of these methods, the one which has been most widely employed was that developed by Prais and Houthakker[2] from views expressed by Sydenstricker and King in an article published in 1921. Their method of calculation essentially rests on the assumption that a household's consumption of a particular item per adult equivalent[3] is primarily a function of income per adult equivalent[4], and that the form of the relationship between them is known. To illustrate their procedure let n_{ir} = total number of equivalent adults for the ith commodity in the rth family,

then

$$n_{ir} = \sum_{=i}^{z} \gamma_{ij} b_{jr} \tag{1}$$

where

b_{jr} = number of persons in the rth family and the jth age group.

γ_{ij} = the coefficient of adult equivalence for the ith commodity and the jth age group.

Z = total number of age groups.

Now denoting by m the value of income per adult-equivalent (on the income scale) they have

$$V_i / \sum_{j=1}^{z} \gamma_{ij} b_{jr} = f_i(m) \tag{2}$$

where

V_i represents the consumption of the ith commodity,

f_i represents the particular function holding between the variables.

Equation (2) may be transformed to

$$V_i / f_i(m) = \sum_{j=1}^{z} \gamma_{ij} b_{jr} \tag{3}$$

They then suggest that the parameters $\gamma_{1i} \ldots \gamma_{zi}$ be estimated by taking the regression of the composite variable $V_i/f_i(m)$ on the variables b_{1r} to b_{zr}.

---

[2] For a complete explanation of the logic behind their method see Prais and Houthakker, *op. cit.*, pp. 125–145.

[3] Derived from the equivalent adult size of the household in terms of the particular commodity.

[4] Based on the equivalent adult size of the household, which is derived from an income scale reflecting the proportion of income allocated to units on the basis of age and sex.

To further illustrate the procedure, consider the particular case of where the income elasticity of demand for the ith commodity is thought to be constant, that is

$$V_i / \sum_{j=1}^{z} \gamma_{ij} b_{jr} = A_i m^{\eta_i} \tag{4}$$

Thus in this case the regression equation becomes

$$V_i m^{\eta_i} = A_i \sum_{j=i}^{z} \gamma_{ij} b_{jr} \tag{5}$$

The actual method of estimation suggested by Prais and Houthakker procedes as follows:

(i) take some value of η_i.

(ii) calculate the value of the variable on the L.H.S. of (5).

(iii) run the regression of the L.H.S. values on the variables on the R.H.S. and calculate the coefficient of determination (R^2).

(iv) continue the above steps, using various η_i values until the maximum coefficient of determination is obtained—this equation then provides an estimate of the adult equivalence scale.

The method suggested by Prais and Houthakker however suffers from a number of problems when given empirical application. These arise from its assumptions, and are summarized below:

(i) Even though it is likely that the prime determinant of a household's consumption is income, once its demographic structure is taken into account, it seems unrealistic to assume that income is the sole cause of variation between the consumption levels of households.

(ii) The second problem arises because the income scale is unknown and so, has to be guessed at before estimation can proceed. To avoid this problem most studies (including Prais and Houthakker's own) use income per person as the variable in the regression equation rather than income per adult equivalent. The underlying premise of this approach is that the size of a household in terms of income adult equivalents is closely approximated by its numerical size. However if, in fact, the numerical size does happen to over-state the true income adult size, it can be shown that this will lead to an overestimate of the particular commodities adult equivalent scale, by an amount that depends on the income elasticity of the commodity.

If it is considered that for some particular commodity its particular adult equivalence scale closely approximates the income scale, then this problem with the Prais and Houthakker approach can be overcome as shown by

the following example. In this example it is once again assumed that the income elasticity of demand is constant, that is

$$V_i/n_{ir} = A_i \left(\frac{Y_i}{n_{ir}}\right)^{\eta_i}$$

which is the following regression equation

(6)

$$\frac{V_i^{\frac{1}{1-\eta_i}}}{Y_i^{\frac{\eta_i}{1-\eta_i}}} = A_i^{\frac{1}{1-\eta_i}} \cdot \sum_{j=1}^{z} \gamma_{ij} b_{jr}$$

in which all the variables are observable.

(iii) The third difficulty with the method is that the form of the relationship between consumption per adult equivalent and income per adult equivalent is assumed known. In fact, in many cases the exact form of the relationship is not known, and if an inappropriate form is assumed, then a biased scale will result. Furthermore, if the correct functional form is a little more sophisticated than a double logarithmic or semi-logarithmic function, then the estimation procedure becomes complex and time consuming.

There is, however, an alternative procedure which permits adult equivalence scales to be calculated from cross-sectional data without requiring the assumptions of the Prais and Houthakker method. The steps required in this suggested method are outlined below.

The first step requires that the characteristics which significantly influence consumption of a particular good, besides size and composition, be determined, either by theoretical means or through the use of Contingency Table tests of independence. That is we have,

(7) $$C_{ri} = f(H_{ri}, N_{ri}, P_r^T, X_{ri} \ldots X_{rn})$$

where

C_{ri} = the rth household's consumption of the ith commodity during a given period.

H_{ri} = the psychological characteristics of the rth household, represented by the strength of its desire for the ith good.

N_{ri} = size of the rth household in terms of adult equivalents of the ith commodity.

P_r^T = a vector of prices confronting the rth household.

$X_{ri} \ldots X_{rn}$ = the other characteristics which are systematic in the influence they have on the consumption of the ith commodity by households.

To illustrate this procedure it shall be assumed at this stage that only one additional variable (X_{ri} – say religion) besides H, N and P^T's is considered to be significant in influencing household consumption of the ith commodity, that is

$$C_{ri} = f(H_{ri}, N_{ri}, P_r^T, X_{ri}) \tag{8}$$

The second step, then, entails separating the total sample survey into 'k' subsets, where 'k' is the number of categories of X_{ri} which produce significantly different consumption patterns. For example, in this case we may have 3 ($= k$) different religious categories which have different effects on consumption. Since all households face similar prices, it follows from equation (8) that any differences between the consumption levels of households within each of these 'k' subsets is due to differences in their psychological make up and size, that is

$$C_{ri}{}_s - C_{ti}{}_s = f(N_{ri} - N_{ti}, H_{ri} - H_{ti}) \tag{9}$$

$C_{ri}{}_s$ consumption of ith good by rth household in sth subset $s = 1 \ldots k$

Now the following procedure would produce realistic estimates of adult equivalence scales if the consumption level of each household within each of the 'k' categories was compared with that of another household within its category, whose consumption did not represent an abnormally strong like or dislike for the ith commodity[5]. However, it is virtually impossible to comment on the relative strength of a household's desire for a commodity. In order to minimize this danger it is suggested that the consumption of each household within a particular category be compared with the average consumption of a set of households within the category which have the same demographic structure. That is, it is safer to operate with the following function, instead of (9):

$$C_{ri}{}_s - C^*{}_{i}{}_s = f(N_{ri} - N^{**}{}_i, H^D) \quad s = 1 \ldots k \tag{10}$$

where

$C^*{}_{j}{}_s$ is the average consumption of households that have the same demographic structure and are of size N^*.

H^D is the difference between the desire of the rth household and a 'so called' average or normal desire of the combined group for the ith commodity.

---

[5] The reason why it is necessary that the consumption level used as a basis for comparison should not represent an extreme dislike or desire for the "i" commodity is discussed later.

Now using the fact that it is possible to write a total number of equivalent adults for the *i*th commodity in the *r*th family by equation (1), it is possible to rewrite function (10) in the following form:

$$(11) \qquad C_{ri}{}_{s} - C^{*}{}_{i}{}_{s} = f\left[\sum_{j=1}^{z} \gamma_{ij}\,(b_{jr} - b_{*j}), H^{D}\right] \qquad s = 1 \ldots k$$

Assume for the moment that all households have the same desire for the *i*th commodity, then it seems reasonable to expect that a proportional relationship will hold between the difference in the size of households and the magnitude of the difference in their respective consumption levels, that is

$$(12) \qquad C_{ri}{}_{s} - C^{*}{}_{i}{}_{s} = \propto \left[\sum_{j=1}^{z} \gamma_{ij}\,(b_{jr} - b_{*j})\right]$$

If the above simplifying assumption is relaxed, it would be possible to include in equation (12) a term acknowledging the influence of the difference in the desires of households on the difference in their consumption, that is

$$(13) \qquad C_{ri}{}_{s} - C^{*}{}_{i}{}_{s} = \propto \left[\sum_{j=1}^{z} \gamma_{ij}\,(b_{jr} - b_{*j})\right] + f(H^{D}) \qquad s = 1 \ldots k$$

Now since the magnitude of differences in desires cannot usually be measured, equation (13) cannot be tested against observed data in its present form. However, it is reasonable to expect that the strength of the desire for a particular item of foodstuff by households to be independent and distributed approximately normally about some mean strength. It follows that, if the average desire of the households that were used as the comparison base closely approximates the population mean strength, then the variable reflecting the difference between the relative desires of households can be legitimately included in the error term. Thus equation (13) can be rewritten in a simple linear form which satisfies the requirements for least squares estimation, that is

$$(14) \qquad C_{ri}{}_{s} - C^{*}{}_{i}{}_{s} = \propto \left[\sum_{j=1}^{z} \gamma_{ij}\,(b_{jr} - b_{*j})\right] + e \qquad s = 1 \ldots k$$

Thus by running this regression on the households contained in the *s*th subset, an adult equivalence scale estimate can be obtained. It is to be noted that this procedure enables '*k*' estimates to be obtained.

The above suggestion for deriving an adult equivalence scale for a particular commodity has two advantages over the Prais-Houthakker method. In the first place it does not require *a priori* knowledge of the

exact functional form of the relationship between a household's consumption of the commodity and any of its characteristics. In other words the number of assumptions necessary for this procedure is minimal. Secondly, by enabling a number of different scales to be calculated, this procedure provides an indication of any change in the scales with different values of the variables that are significant in influencing the level of a household's consumption of the good. For example, it may be that the adult equivalence scale is different for households at different positions in the income distribution. If this is the case, then any single scale used in an aggregate study should be a weighted average of the scales holding at different income levels, with the weights being determined by the relative number of households contained in each of the income categories. Note that the possibility of bias, produced by the use of an abnormally high or low consumption level as the basis of comparison within any category, would be reduced by the averaging procedure.

The chief disadvantage of this method is that it requires information on a relatively large number of households in order to produce good estimates. The actual size requirement however is a function of the number of characteristics considered to be significant, the number of classifications within each of these, and also of the variability in the strength of the desire by households for the particular good. For example, if only one variable was considered significant and it took three values, then if the variability in the strength of desires by households was minor, approximately 300 to 400 households would be required. If the variability was large, the size of the sample would need to be around 600 to 700 households. However, if two factors were taken as significant and each had three possible values then, depending on the variability of the desire for the good by households, the sample would need to be at least 1,000 to 2,000. This problem with the size requirement may however be reduced a little in practice by dealing only with those variables which are considered to be dominant in influencing a household's consumption of the good and by assuming that the distributions of the other variables that have an affect on its consumption are similar in each of sample subsets examined.