Volume Title: Studies in Income and Wealth, Volume 15

Volume Author/Editor: Conference on Research in Income and Wealth

Volume Publisher: NBER

Volume ISBN: 0-870-14170-8

Volume URL: http://www.nber.org/books/unkn52-1

Publication Date: 1952

Chapter Title: A Method of Comparing Incomes of Families Differing in Composition

Chapter Author: Milton Friedman

Chapter URL: http://www.nber.org/chapters/c9763

Chapter pages in book: (p. 9 - 24)

Part I

A Method of Comparing Incomes of Families Differing in Composition

MILTON FRIEDMAN
University of Chicago

This paper was written in 1935 when I was a member of a group working under the direction of Hildegarde Kneeland at the National Resources Committee on plans for the Consumer Purchases Study and for the analysis of the consumption data it was expected to yield. It seemed to me that the defect pointed out in Section D 4 might well be fatal and that the method should not be published unless this defect could either be removed or shown not to be crucial and even then, not until the method could be tried out on some actual data. I planned to do further work on the problem but I never did and so the paper remained buried in my files.

I must confess that my original view still seems to me correct. The exhumation of the paper at this time is at the insistence of Dorothy Brady, to whom I happened to show it and who argues that, whether or not I am right about the significance of the defect in the method, the method itself is sufficiently suggestive to justify publication.

I have left the paper in its original form except for a few minor verbal changes, and the addition of two paragraphs (the present final paragraph in Section D 3 and in Section E).

A The Problem

The analysis of factors affecting consumption patterns and the derivation of criteria for ranking families by their relative economic status are two of the chief aims of consumption research. One of the most important factors affecting a family's consumption pattern is the culture complex prevailing within the socio-economic group of which it is a member. Since groups differ widely in their habits of expenditure, analysis of the effect of other factors must be carried on separately for each group. Similarly, we can compare the economic well-being of families only within a given socio-economic group for which it can be assumed that tastes are roughly the same.

But even within a given socio-economic group the solution of these two problems is greatly complicated by the difficulty of separating the influence of income from that of family composition. Two families with the same income but a different number of members will have different consumption patterns and cannot be considered equally well off.

B Possible Approaches

1 Segregation of Family Types

The difficulty can be treated in two ways. The first is to separate families for which there are data into rigidly defined family types and analyze the relation between income and expenditure separately for each family type. The results can then be compared for different family types and the influence of family type determined. This method has the advantage that it is direct and relatively simple. It has the disadvantages that the number of cases of each family type will be small and that it does not furnish criteria for ranking families by economic status.

2 Quantification of Family Composition

The second method is to reduce family composition to a quantitative basis by obtaining a measure of the economic size of families. Family size can then be treated as a quantitative variable in the same way as income, and its influence can be measured and/or eliminated by various statistical techniques. The great advantage of this method is that it enables the total number of families to be used in studying the relation between income and expenditures and provides a criterion for comparing the economic status of families. This paper is concerned with a method for obtaining a measure of the economic size of the family.

The number of persons in the family is not, of course, a satisfactory measure, since from the viewpoint of expenditures the different members cannot be considered equivalent. A family of two adults and two young children is probably 'better off' than a family of four adults if both have the same total income. Similarly, for certain categories of consumption a female can spend less and still be as 'well off' as a male. To adjust for these differences it is desirable to set up ratios of equivalence among the consumption requirements of individuals of different age and sex so as to express the total requirements of a family in terms of a common unit. Such ratios of equivalence form an 'ammain' scale. They will, of course, differ for the several categories of consumption. It is evident, for example, that while an adult female might consume only three-fourths the food of an 'equally well off' adult male, she might spend more on clothing.

Even for the consumption of a particular category of goods by a given socio-economic group the ammain scale will depend on the criterion of equivalent requirements that we adopt – the economic, nonnormative one of equal actual expenditures or the normative, non-economic one of equal requirements based on some assumed standard of need. According to the former criterion the ratio of equivalence between a male and a female for say food, will depend on the ratio between the actual expenditures for the food of a male and a female who are members of the same family. According to the normative criterion, on the other hand, the ratio of equivalence between a male and a female for food is based on the ratio of the amounts of food, measured in some such unit as calories, necessary to keep the two equally well nourished in terms of some medical standard. Similarly, it would be possible to set up equivalence ratios for housing on the basis of an accepted standard of housing adequacy. Such scales may be of great use in investigations into social welfare or in social planning concerned with standards for the distribution of social income. They are a meaningful basis for comparing actual expenditures by families with different composition only to the extent that people actually conform to, i.e., behave in accordance with, the assumed standards of need and that it is possible to adjust the physical ratios for differences in prices. Since neither condition is generally fulfilled it is necessary to set up ammain scales for the latter purpose on the basis of the economic criterion mentioned above; i.e., to equate requirements of individuals on the basis of actual expenditures rather than on the basis of a standard of need. This is the point of view adopted here.

C Existing Procedures for Derivation of Ammain Scales

1 FOR ITEMS CONSUMED BY THE INDIVIDUAL

For certain categories of consumption, for example, clothing, a family's expenditures can be allocated to the members and ammain scales can be obtained directly from data for many families on the expenditures on behalf of the members of each family. One member, say an adult male of specified age, is selected as the unit of comparison; the amount expended for the particular category of consumption on behalf of each member of a particular family is then expressed as a ratio to the amount expended on behalf of the member taken as the unit of comparison; if these ratios do not depend on the number in the family or on the family income, they are averaged for particular ages and sexes; the result is an ammain scale. This is the procedure that has generally been used for consumption items that are bought on behalf of an individual. For example, Edgar Sydenstricker and W. I. King employ essentially this procedure in their excellent article, 'The Measurement of the Relative Economic Status of Families' (*Quarterly Publication of the American Statistical Association,* 1920-21, Vol. 17, 842-57) to obtain a scale for items of consumption other than food. The only departure from it is that they average the actual expenditure on behalf of the individual members of the family and then obtain ratios of equivalence, instead of performing the operations in reverse order.

This procedure has two basic difficulties: for many categories of consumption it is impossible to allocate the total family expenditure directly among the several members; the expenditure on behalf of a person of given age and sex is likely to depend on the number in the family, i.e., there are 'overhead costs' that do not vary with the number in the family.

2 FOR ITEMS CONSUMED BY THE FAMILY

To construct ammain scales for categories of consumption to which these objections apply, we would like to have data on the expenditure patterns of pairs of families that are 'equally well off' and whose family composition differs *only* in that one family has an additional member. From such data we could compute the additional expenditure on a particular consumption category that is necessary to leave a family as well off with an additional member of a given age and sex as it was before. From data for a sufficiently large number of such pairs of families we could construct ratios of equivalence that would take account not only of the age and sex of the individual but also of the number in the family. And in the computation of these scales it would not be necessary to allocate the total family expenditure among the members.

But we cannot obtain the necessary data unless we can tell when two families of different family composition are 'equally well off'. And we cannot tell this without some measure of the economic size of the family, i.e., without an ammain scale, even if we are willing to use a pecuniary measure of welfare. We thus seem to be involved in a vicious circle: we cannot get a satisfactory ammain scale unless we have one to begin with.

a *Sydenstricker and King solution*

In constructing a scale of equivalence ratios for food, which they call a 'fammain' scale (food for adult male maintenance), Sydenstricker and King break this circle by adopting an approximate scale, the Atwater scale based on dietary requirements. They use this scale to classify families by welfare and then employ a variant of the method discussed above to obtain corrections to the assumed scale. They use the corrected scale as a second approximation, and repeat the process for a third approximation. They correct the scale separately for sex and age rather than for the two variables simultaneously, and do not take account of the individual's order in the family, e.g., whether a child is the first or second in the family.

b *Suggested alternative solution*

Another way of breaking the vicious circle is to set up a rather circuitous definition of consumption equivalence, then use the ordinary method for the statistical treatment of interdependence, namely, multiple correlation. This procedure has never been applied to problems of this sort but seems to offer excellent possibilities.

D Alternative Procedure

Since the procedure will be the same regardless of the consumption category for which the ammain scale is constructed, we may confine discussion to the problem of obtaining a scale for food requirements, which, following Sydenstricker and King, we may call a fammain scale.

1 Definitions

Let us, then, *define a fammain scale as a set of numbers (called fammains) assigned to families of different composition that have the property that the relation between expenditure on food per fammain and income per fammain is independent of family composition*. This definition, while somewhat different from those ordinarily formulated, does not depart from general notions of a fammain scale and the purpose it is intended to serve. Fammain scales are designed to eliminate the influence of family composition from certain characteristics of the data. The definition here suggested takes the relation between the expenditure on some consumption category and income as the characteristic to be freed from the influence of family composition. Further, the ordinary definition of consumption equivalence can be considered a corollary of the one given here. For if income per fammain is taken as the measure of welfare, the expenditure on food per fammain is the same for families who are equally well off, regardless of family composition.

We may select as the basic unit of our fammain scale the combination of husband and wife of certain ages with no dependents rather than as is ordinarily done, an adult male. Data are much more extensive and detailed for this type of family than for single person families. Further, it scarcely seems valid to use the same fammain scales for single person families and for other types of family. Consider the husband-wife combination selected as representing one fammain and from data for this family type determine the relation between expenditure on food per fammain (total food expenditure for this type of family) and income per fammain (total income). This can be done by fitting a curve by the method of least squares to the pairs of observations giving for each family total income and total food expenditure. For illustrative purposes, suppose the relation is linear and is given by the equation

$$(1) \qquad T = A + Br,$$

where T is the total expenditure on food and r the total income for families taken as the unit of comparison. We now wish to set up a fammain scale such that for all other family types this relation will hold (approximately) between income per fammain and food expenditure per fammain. That is

$$(2) \qquad \frac{T}{S} = A + B\frac{r}{S}$$

where S represents the number of fammains in a family.

2 SYMBOLIC REPRESENTATION OF SCALES

The first step is to represent our scale symbolically. This we can do in several ways. We may, for example, set up only a few classes, such as those used in the 1928 Federal Employees Study in five cities: adult male 15 years or over, adult female 15 years or over, children 11 through 14 years, children 7 through 10 years, children 4 through 6 years, and children

3 years or under; and use a different symbol to represent the ratio of the food requirements of each class of persons to the requirements of the husband-wife combination as a unit. But such a classification is too simple. It does not allow for sex differences below the age of 15 and assumes that the age of a person over 15 does not affect his consumption requirements. This can easily be remedied by using a much more detailed classification, making it conform, say, to that used by Sydenstricker and King who present equivalence ratios for each sex for all ages up to 80. For our purposes, however, even such a scale is not in sufficient detail since it does not take into account the order of the individual in the family, whether he is the first or second child, and so on. We could set up a separate scale for each child, thereby taking the number of individuals in the family into account. Such an extremely detailed scale is what we want. But its derivation presents enormous statistical difficulties, since if a separate symbol were assigned to each item in such a scale, there would be several hundred such symbols whose values it would be practically impossible to determine from the limited data available.

There is, however, a very simple method of overcoming this difficulty. Instead of assigning a discrete symbol to each item in the scale, we can assume that the ratio between the food requirements of say the first male child and the husband-wife combination, varies *continuously* with age and can be represented by a simple mathematical function of age. Thus we can represent the number of fammains, S_{1m}, for the first additional male member of the family by

$$S_{1m} = c_{1m} + d_{1m}y + e_{1m}y^2, \tag{3}$$

where y is the age of the person considered. The adoption of such a function means that we need only three constants, c_{1m}, d_{1m}, and e_{1m}, to represent the fammain scale for the first additional male instead of the eighty required by a scale of the Sydenstricker-King type. A similar function can be set up for additional males. In general terms this may be written

$$S_{im} = c_{im} + d_{im}y + e_{im}y^2, \tag{4}$$

where i stands for the order in the family attributed to the particular male considered. Similarly, we can set up functions representing the number of fammains for female members, which may be written

$$S_{if} = c_{if} + d_{if}y + e_{if}y^2. \tag{5}$$

No special significance is, of course, to be attached to the particular form of the function chosen for illustrative purposes. It will probably be found that other types of function are better suited to represent a fammain scale. Some light can be thrown on this question by fitting various types of curve to the scales that have already been computed. Since the scale will presumably have a maximum, i.e., consumption requirements will decrease after a certain age, a linear function cannot be employed.

In addition to functions of types (4) and (5), similar functions will be needed for the husbands and wives who fall outside the age range used to define a fammain. These functions might be linear, for presumably they will relate only to ages greater than that at which individuals reach their maximum consumption requirements. We may represent these functions symbolically by

(6) $$S_h = c_h + d_h y,$$

and

(7) $$S_w = c_w + d_w y,$$

where h and w stand for husband and wife respectively.

The number of functions of types (4) and (5) needed will depend largely on the family types that are available for analysis and on the assumption concerning the influence of the number in the family. Four scales of type (4) and four of type (5) would take care of all six-member families and many seven- to ten-member families, and would require the estimation of 28 constants. An equally detailed scale of the Sydenstricker-King type would require about 640 constants.

The use of continuous functions to represent a fammain scale is not only convenient statistically but also seems more reasonable than the use of discrete scales, since food requirements might be expected to vary continuously with age.

3 STATISTICAL EVALUATION OF CONSTANTS

The next step is to determine the values of the constants for the equations selected. This can be done by using the requirement that the relation between expenditure on food per fammain and income per fammain shall be the same for all family types as expressed in equation (2),

$$\frac{T}{S} = A + B\frac{r}{S}.$$

This can be written in the form

(8) $$T - Br = AS.$$

In this equation S is the only term whose numerical value is unknown. The values of T and r are given by the data, and the values of A and B have been determined from families that have been defined as having one fammain.

For each family in the sample under consideration we express symbolically the number of fammains. For example, consider a family of husband and wife, who together count as one fammain, and who have two childen, a boy of 12 and a girl of 10. By using (4) and (5), we can express the total number of fammains for this family as follows:

$$(9) \qquad \begin{aligned} S &= 1 + a_{1m} + b_{1m}(12) + c_{1m}(12)^2 + a_{1f} + b_{1f}(10) + c_{1f}(10)^2 \\ &= 1 + a_{1m} + 12b_{1m} + 144c_{1m} + a_{1f} + 10b_{1f} + 100c_{1f}. \end{aligned}$$

Substituting this expression for S in equation (8) gives the following observation equation containing the unknown constants:

$$(10) \quad T - Br = A(1 + a_{1m} + 12b_{1m} + 144c_{1m} + a_{1f} + 10b_{1f} + 100c_{1f}).$$

This can be done for each family, yielding as many observation equations as there are families in our sample. As these equations are linear with respect to the unknown parameters, the method of least squares is directly applicable.

With a scale as complex and detailed as that described above, the statistical process of evaluating the constants involved will be extremely laborious. In general, however, it will not be necessary to have as many subclassifications with respect to sex and order in the family as were suggested. For certain consumption categories the order will be immaterial, and for others, the sex.

Once a scale has been obtained by this procedure it can be used to reduce the data to a per fammain basis. From these data for the whole sample a better estimate than equation (1) of the relation between expenditures on food per fammain and income per fammain can be obtained. If the new constants differ significantly from A and B in equation (1) it may be well to use them as the basis for the derivation of corrected constants for the fammain scales by the same procedure as that described above.

More generally, this process of successive approximation can itself be replaced by simultaneous determination. Instead of first computing the numerical values of A and B in equation (1), then estimating the other parameters in equations like (10), these two steps can be combined. A and B can be taken as unknown parameters in observation equations like (10), such equations included for families of the type taken as representing one fammain, and all the constants estimated simultaneously. Of course, the resulting equations will no longer be linear in the unknown parameters, thereby complicating the problem of statistical estimation.

4 WEAKNESS IN SUGGESTED PROCEDURE

The chief weakness in the above procedure is the basic assumption that income per fammain measures economic welfare. For in deriving the scales for other categories of expenditure, say clothing, we shall have to assume that income per clothing ammain (let us call this the cammain scale) is the measure of economic welfare. And the number of cammains in a given family will ordinarily not equal the number of fammains. We would thus have as many measures of economic welfare as there are categories of consumption for which we could compute scales. It is true, of course, that from these several scales we could compute a weighted average scale (we may call this the ammain scale) which could be used as a single measure of economic welfare. But should we not then have defined the fammain scale originally in such a way as to render independent of family composition the relation between expenditure per *fammain* and income per *ammain?*

The basic consideration in not adopting this definition is entirely pragmatic: its adoption would render the statistical procedures too complicated. There is, further, some question which definition is preferable. Is it the income available for each food consumption unit, i.e., each fammain, that determines the expenditure for food on behalf of each such unit? Or is it the income available for each consumption unit (with respect to all items of expenditure), i.e., each ammain, that determines the expenditure for food on behalf of each food consumption unit? Our general vague knowledge of consumption habits does not provide a simple answer. We would expect that two families with similar tastes and with the same number of ammains and the same income per ammain but with different numbers of fammains would spend different amounts on food per fammain (or per ammain). But we would expect also that two families with similar tastes and with different numbers of ammains but with the same number of fammains and the same income per fammain would spend different amounts on food per fammain (or per ammain). That is, only in the exceptional cases where the number of ammains and the number of fammains are the same will families with similar tastes and the same income per ammain (or per fammain) spend the same amount on food per fammain and will the ordinary criteria of fammain or ammain scales be completely satisfied. And for this exceptional class of cases the definition that forms the basis of the statistical procedures suggested above will also be satisfied.

E OTHER APPLICATIONS OF THE CENTRAL IDEA

Although the preceding discussion is entirely in terms of family composition, it is clear that the central idea and general procedure are applicable to other problems of determining equivalent units. The problems of determining income levels that are equivalent for farm and nonfarm families, for families living in cities of different size, for families living in different regions, for families of different occupational status, and so on, are all logically identical with the problem of determining income levels that are equivalent for families of different size. In brief, these are all special cases of the general problem of computing cost of living indexes to determine when different economic units in different circumstances are 'equally well off'. The procedure will have the same central advantage for any of these special cases — that it is self-contained and independent of any assumed standard of 'need' — and the same chief defect — that it may yield different answers for different categories of consumption.

APPENDIX

Jean Mann Due, *University of Illinois**

To test the method, the summary of expenditures in 1945 among families in Portland, Oregon was obtained from the Bureau of Labor Statistics. Of the 192 records collected, those of single consumers, families whose members were not at home all year, and broken families were discarded, leaving 94. Because the sample was small, the expenditure pattern of the basic two person family was based on the data for families in which the husband was under 60 years of age.

The method for determining the scale of equivalence among different types of family was tested first with total expenditures for family living. For the two person families,[1] the regression of expenditures, including gifts and contributions, on money income was

$$y = 1171 + 0.45x. \tag{11}$$

Because a separate equation is needed for each variable introduced and because of the size of the sample and the calculations involved, it was decided to test the method by using only differences in ages of family members and omitting the influence of order in the family and sex. Age instead of order in the family or sex was chosen as a variable because it was thought to be a more important influence on family expenditures.

Data for each of the 68 families of more than two persons were substituted in the equation

$$T - Br = A(1 + a + by_i + cy_i^2), \tag{12}$$

where T is total family expenditures, r family income, A and B the parameters of the regression line (11), and y_i the age of the family member.

* The conclusions in the Appendix are taken from 'An Analysis of Measures of Equivalence Among Families of Varying Size and Composition', Master's Thesis, by Jean M. Mann (University of Illinois, Graduate College, 1950).

[1] The regression line for three person families was $y = 1382 + 0.49x$, for four person families $y = 889 + 0.6084x$, and for five or more person families $y = 1980 + 0.42x$.

The values of constants determined by least squares were $a = 0.17900132$; $b = 0.00194429$; $c = 0.00004238$.

This parabola, which has a minimum at about 23 years, does not appear to be a rational result because normally an additional family member would be expected to increase family expenditures at an accelerating rate as he passed through the early ages, then at a relatively constant rate after maturity, probably decreasing again after retirement age. Inspection of the data indicated that one or two 'extreme' cases had affected the determination of the form of the relation.

To investigate these results further, graphs were drawn by plotting the equivalence ratios of three, four, and more than four person families against the sum of the ages of the family members other than husband and wife. The values of the three person families were scattered about a horizontal line indicating no correlation whatsoever between the equivalence ratio and the age of the third family member. Similar diagrams exhibited little relation between the equivalence ratio and the ages of additional members of families of more than three persons.

Examination of the regression lines for each size of family suggests that the slopes for all sizes of families may be the same. Accordingly, lines with slopes of 0.46 were fitted to the data for four and five person families. The deviations from these lines were not correlated with the ages, or sum of the ages, of family members other than husband and wife.

Confirmation of the finding that the age of additional family members was not an important variable in influencing total family expenditures among families of the same size was attempted using data from the Consumer Purchases Study;[2] expenditures were plotted against income for families with children of various ages. Three person families were used, with a child under 2 years, one 6-11 years, and one 16-29 years. These ages were sufficiently different for marked variations in expenditures to have been apparent if age was an important influence on family expenditures. Similarly, for four person families expenditures were plotted against income for families whose children were both under 5 years, both 16-29 years, and the older child 12-15 and the younger 6-15 years. In both cases there was no consistent relation between age and expenditures.

The conclusion was reached, therefore, that age is not an important variable in influencing total family expenditure in relation to income; the number of children in the family is significant but not their ages. This does not mean that the cost of supporting a 24 year old child is no greater than

[2] Day Monroe, Maryland Y. Pennell, Mary R. Pratt, and Geraldine S. DePuy, 'Family Spending and Saving as Related to Age of Wife and Number of Children', Department of Agriculture *Miscellaneous Publication 489* (1942).

that of supporting a 2 year old, but rather that adjustments for differences in expenditures on individual members occur within the different categories of consumption. For example, the husband and wife may have to make more adjustments among consumption items when the third member is 24 than when he is 2, but the total family expenditure pattern is unaffected by the age of the additional member *at the point of equivalence* established in this or some other manner.

To see if this conclusion would be true of food expenditures, which have generally been regarded as varying fairly directly with age, the same procedure was carried out using total food expenditures. This time the values of the constants turned out to be: $a = 0.24167785$; $b = 0.00074265$; $c = 0.00000711$.

These equivalence ratios for food expenditures increased continuously with age, a result that seems in conflict with the nutritional needs of persons of different ages.

The tests used for total money living expenditures were applied. When the sum of the ages of additional family members was plotted against the equivalence ratio for food for three, four, and more than four person families there was again no indication of correlation between the equivalence ratio and age. The food expenditures of families with children of different ages from North Central small cities in the Consumer Purchases Study were plotted against income, as were total expenditures; these varied widely but showed no correlation with the age of the additional family member. It would seem, therefore, that in the case of food as in the case of total expenditures the number in the family rather than the age of the members is significant for the pattern of family expenditure in relation to income.

The regression lines of food expenditures against income for the Portland families calculated for each size of family were almost parallel[3] and the intercepts increased fairly consistently with the size of family.[4] The parallel regression lines for total expenditures and for food expenditures satisfy the relation $T - Br = AS$ when the values of the equivalence ratios for various sizes of families are:

Size of Family	Total Expenditures	Food Expenditures
2	100	100
3	118	151
4	120	137
4+	125	184

[3] The four person family is an exception in the Portland sample.

[4] The regression equations of food expenditures on income for 2, 3, 4, and more than 4 person families were $y = 455 + 0.1082x$, $y = 687 + 0.1025x$, $y = 622 + 0.1146x$. and $y = 836 + 0.1084x$ respectively.

These results illustrate Mr. Friedman's criticism of the suggested method for determining the equivalence ratios. The scale values increase more rapidly with the number of persons when the relation of food expenditures to income is used as the criterion of equivalence than when the relation of total expenditures to income is used.